Roger W. Brockett

FINITE DIMENSIONAL LINEAR SYSTEMS

SERIES IN DECISION AND CONTROL

Ronald A. Howard, Editor

FINITE DIMENSIONAL LINEAR SYSTEMS

by Roger W. Brockett

DYNAMIC PROBABILISTIC SYSTEMS, Volume I: MARKOVIAN MODELS (in press)

by Ronald A. Howard

DYNAMIC PROBABILISTIC SYSTEMS, Volume II: SEMIMARKOVIAN AND DECISION PROCESSES (in press)

by Ronald A. Howard

OPTIMIZATION BY VECTOR SPACE METHODS

by David G. Luenberger

INTRODUCTION TO DYNAMIC PROGRAMMING

by George L. Nemhauser

OPTIMIZATION THEORY WITH APPLICATIONS

by Donald A. Pierre

FINITE DIMENSIONAL LINEAR SYSTEMS

ROGER W. BROCKETT

JOHN WILEY AND SONS, INC.

NEW YORK · LONDON · SYDNEY · TORONTO

Library of Congress Catalogue Card Number: 74-100326

SBN 471 10585 6

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To my parents

PREFACE

This book is based on a one-semester course on dynamical systems given in the Electrical Engineering Department at the Massachusetts Institute of Technology over the last five years. The students have been mostly electrical engineers in their first year of graduate school, but some students in aeronautics, economics, and mathematics have also attended. The topics covered form the core for advanced work in such fields of study as optimal control, estimation, stability, electrical networks, and the control of distributed systems. My objective in this course is to provide a solid foundation for learning about dynamical systems rather than to develop any special expertise. The particular subject matter chosen for discussion has been picked not only because of the ease with which it can be applied in specific system-theoretic problems, but also because of its importance in engineering analysis in general. Thus in one way or another, many of the standard topics of applied mathematics are touched upon—solubility of linear systems of equations, ordinary differential equations, calculus of variations, and basic ideas of vector analysis. These are discussed in the context of linear systems. The prerequisites are modest; the book can be read by students who are familiar with basic mathematical notation and who have a little experience with vector-matrix manipulation, provided that they have the patience for some detailed argumentation. To make the book more widely accessible, considerable background material on linear algebra is included in summary form.

The exposition is, I believe, in keeping with the spirit of modern engineering mathematics. First, I do not hesitate to use a few well-known, difficult theorems that are not proved here; however, when this is done, it is my intention to point out explicitly what theorem is being used, where it can be found, and why the hypotheses are fulfilled. Secondly, I have deliberately restricted the generality in such a way as to make it possible to give arguments which I believe to be complete and correct and which at the same time require a minimum of mathematical background. To illustrate, in most cases I have assumed continuity where square integrability in the sense of Lebesgue would be enough. Restatement of the results in an L_2 setting should be easy for those having an elementary knowledge of one-dimensional Lebesgue integration. In the construction of the proofs, I have avoided, as much as possible, appealing to Jordan Normal forms and other theorems for which no useful infinite dimensional analogue exists. Thus the reader familiar with Hilbert space theory should have little trouble carrying over many of the arguments to an infinite dimensional setting.

Because this is intended as a text, the book contains a number of nontrivial examples together with a large nunber of exercises. The reader who requires additional motivation should read these as part of the text. Regarding the Exercises, I should point out that some of these are fairly difficult. A student should not feel as if all solutions will follow immediately from the text. I think that attempting a difficult problem is likely to be more educational than completing a routine collection of drill problems. I have used the exercises to introduce some additional definitions and results that have been, or appear to be, applicable to the problems of interest here.

Since giving the student reliable intuition rather than a knowledge of specific facts is the main goal, I have attempted to develop the theory of linear systems in a systematic way, making as much use as possible of vector ideas. The basic facts about the solution of simultaneous linear equations, and one's intuition about them, are made to serve as the basis for building intuition about linear differential equations. However, despite the strong similarities between the development of ideas here and that found in linear algebra courses (especially as evidenced by Halmos' book, *Finite Dimensional Vector Spaces*), this is not a book on linear alegbra. Students should have some previous experience in this area. The order of presentation has been determined mainly by taking into account the dynamical aspects of the material with the assumption that linear algebra is a tool.

I have experimented a great deal with the development and choice of topics. The final result is, of course, a compromise. To understand this compromise, I suggest grouping the topics discussed into three categories: those that are necessary to make contact with previous work, those that are immediately useful, and those that will be needed for advanced work. Thus certain material on frequency response is brought in to help bridge the gap between introductory subjects based on the Laplace transform and the approach taken here. On the other hand, most of the material in the chapters on least squares and stability can be viewed as being immediately useful. Finally, topics such as the McMillan degree, Floquet theory, and certain other material that would be used in advanced courses on optimal control, estimation theory, stability theory, and network synthesis, are introduced. I am aware that the inclusion of the material on least squares is not standard in courses on linear system theory at this level. I think it should be. Without something like it which not only is useful, but at the same time draws in a nontrivial way on the material developed, the whole subject lacks a focal

point. (No pun intended!) Moreover, I think that least squares theory would be extensively taught at this level if it were more widely appreciated that it can be done without recourse to either the standard machinery of the calculus of variations or the maximum principle.

It is perhaps wise to point out in advance that although the emphasis here is on "time domain" methods, it does not follow that I feel transform techniques should be abandoned altogether. On the contrary, transform methods certainly play an important role in this subject since in a great many cases they are natural and effective. However, frequently the use of vector differential equations, together with elementary analysis, yields results that are difficult, if not impossible, to obtain using transform techniques alone.

Some specific points should be noted:

1. Transform techniques are effectively limited to linear time-invariant differential equations. Although special types of time-varying problems can be treated using transforms, these techniques do not form a basis for a *systematic* study of time-varying equations.
2. In the treatment of problems involving the integrals of quadratic forms, the Parseval–Plancherel relation makes transform techniques quite effective—if the equations are time-invariant and the interval of interest is infinite. Otherwise, vector space methods are generally superior.
3. Stability of constant linear equations is, of course, primarily a complex variable problem. However, when the equations are time-varying, the Routh–Hurwitz test is irrelevant, and time domain methods must be used.

The book contains more material than I have ever covered in a single semester. Generally omitted were Sections 19, 23, 26, 27, and 33 to 35. With these deletions the material can be covered in one semester even with some allowances being made for varying backgrounds in linear algebra. I found that it was highly desirable to schedule an additional hour each week devoted exclusively to working out in detail physical problems that illustrate the theory. These sessions were optional from the students' point of view and were only recommended for those having difficulty in visualizing the engineering significance. I have included as examples here some of the more successful of the problems discussed.

This book is intended as a text, not a historical account or a personalized research monograph. The vast majority of the material is well-known to workers in the field, although surprisingly little of it has made its way into textbook form. Accordingly, accurate acknowledgment is sometimes difficult, and my remarks in the notes and references should not be taken too seriously. The origin of some results is in dispute, and in other cases there is considerable difference of opinion as to what is a basic contribution and what is a minor embellishment. I have not tried to make the reference list complete in any sense. The items that are listed have all been of direct use to me in one way or another in learning the subject or preparing the manuscript. I strongly encourage students to use the references to get perspective on the field and to see this material developed from other points of view. At the same time I have been unable to acknowledge properly all sources, and I apologize in advance to those authors whose relevant work is not given credit.

I am indebted to many people who have made suggestions as to the content and organization of the earlier drafts. In particular I profited from the criticism of R. K. Brayton, T. Fortmann, L. A. Gould, I. B. Rhodes, F. C. Schweppe, D. L. Snyder, and R. N. Spann, all of whom taught from one version or another of the manuscript and each of whom in their own way influenced the final result. M. Athans participated in the early development of the course for which these notes were written and in this way also contributed. I want to particularly acknowledge Jan Willems who made many suggestions and listened to countless arguments. Naturally the students themselves provided valuable feedback. I would like to mention R. Canales, J. Davis, H. Geering, J. Gruhl, and R. Skoog in particular as having made a substantial contribution. For typing and retyping the manuscript countless times I want to thank A. Brazer, K. Erlandson, J. Gruber, and especially M. Stanton. The final thanks go to my wife Carolann for her constant encouragement and moral support.

ROGER W. BROCKETT

Cambridge, Massachusetts
October, 1969

CONTENTS

1
LINEAR DIFFERENTIAL EQUATIONS

In this chapter we discuss linear ordinary differential equations. For the most part we will not find it necessary to assume that these equations are also time invariant, although this is an important special case. We are interested in exposing the structure of the solutions and, to a lesser extent, developing actual solution techniques. The presentation relies heavily on vector space methods. This is entirely in keeping with modern methods of computation. The first section covers briefly the background from linear algebra needed in Chapter 1. This policy is repeated in the first section of the remaining three chapters. Taken together these four sections are intended as a convenient reference to make the book more nearly self contained.

1. LINEAR INDEPENDENCE AND LINEAR MAPPINGS

The theory of linear dynamical systems is so completely entwined with the study of basic linear algebra that any attempt to relegate the latter to appendices is, in our view, out of the question. On the other hand, excellent books devoted entirely to the many facets of the subject already exist and there is no need to duplicate here material that is widely available. Our policy will be to steer a middle course. At the start of each chapter, we give some background in those aspects of linear algebra that are most germane and try to point out in the subsequent sections the most informative relationships. In this section we discuss the algebraic structure of the set R^n of all n-tuples of real numbers,* and linear transformations of sets of n-tuples into sets of n-tuples.

By R^n we mean the set of all objects of the form $(x_1, x_2, \ldots, x_n)$ with the x_i real numbers. This is a specific example of a finite dimensional vector space. We define the following operations for members of R^n.

(a) The *sum of two n-tuples*, $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ is $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$.

(b) The *product of an n-tuple* by a real scalar a is $(ax_1, ax_2, \ldots, ax_n)$. With these two definitions understood, the set R^n is called *Cartesian n-space*.

* (1, 3, 5) is an example of a 3-tuple, the reader who is unfamiliar with this idea might think in terms of vectors in ordinary geometry.

Its elements may be called *n-tuples* or *vectors*. We use **0** for the n-tuple $(0, 0, \ldots, 0)$.

Theorem 1. *If* $\mathbf{x}$, $\mathbf{y}$ *and* $\mathbf{z}$ *belong to* R^n *and if* a *and* b *are real scalars, then*

(i) $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$
(ii) $\mathbf{0}+\mathbf{x}=\mathbf{x}$
(iii) $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$
(iv) $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$
(v) $a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}$
(vi) $(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}$
(vii) $(ab)\mathbf{x}=a(b\mathbf{x})$
(viii) $1\cdot\mathbf{x}=\mathbf{x}$

These are all easy to verify using the properties of the real numbers. We will not give a proof. Our motivation for including them here is that it is exactly these eight properties that are used in the general definition of an abstract real vector space. That is, any set of objects **v** together with a definition of sum and a definition of scalar multiplication which satisfies these eight conditions is called a *real vector space*.

Examples. In this book we need to examine 3 particular vector spaces.

(i) R^n as defined above.

(ii) Let $C^m[t_0, t_1]$ denote the set of m-tuples whose elements are continuous functions of time defined on the interval $t_0 \leqslant t \leqslant t_1$. We write the elements of $C^m[t_0, t_1]$ as column vectors, that is

$$\mathbf{u}=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \quad \text{and} \quad \mathbf{u}(t)=\begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{bmatrix}$$

It is easy to verify that the set $C^m[t_0, t_1]$ is a vector space provided we define $a\mathbf{u}$ and $\mathbf{u}+\mathbf{v}$ as

$$a\mathbf{u}=\begin{bmatrix} au_1 \\ au_2 \\ \vdots \\ au_m \end{bmatrix}; \qquad \mathbf{u}+\mathbf{v}=\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_m+v_m \end{bmatrix}$$

(iii) Let $R^{m\times n}$ denote the set of all m by n arrays of real numbers arranged in the format

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

One easily verifies that this set is a linear vector space provided addition and scalar multiplication are defined by

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

$$a\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} aa_{11} & aa_{12} & \cdots & aa_{1n} \\ aa_{21} & aa_{22} & \cdots & aa_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ aa_{m1} & aa_{m2} & \cdots & aa_{mn} \end{bmatrix};$$

Such arrays are called *matrices*.

The vectors $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 1)$ which we label $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, taken together as a set of n n-tuples, form what is called the *standard basis* for R^n. It is obvious that any n-tuple $\mathbf{x}$ can be expressed as $\mathbf{x}=a_1\mathbf{e}_1+a_2\mathbf{e}_2+\cdots+a_n\mathbf{e}_n$. Given an arbitrary collection of n or fewer vectors in R^n, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ we denote the subset of R^n which can be expressed as $a_1\mathbf{x}_1+a_2\mathbf{x}_2+\cdots a_k\mathbf{x}_k$ for some choice of a_i the *subspace spanned by* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$. If $k<n$, then the subspace spanned by $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k$ is not the whole space. However, if $k=n$, it may be. If $k=n$ and the space spanned by $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is the whole space, then we say that the collection of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ forms a *basis* for R^n.

A set of n-tuples $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is called *linearly independent* if $a_1\mathbf{x}_1+a_2\mathbf{x}_2, \ldots, a_n\mathbf{x}_k=\mathbf{0}$ implies that all the a_i are zero. The following theorem is fundamental.

Theorem 2. *Let* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ *and* $\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_j$ *be sets of* n*-tuples. If the set* $\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_j$ *is linearly independent, and if both sets span the same subspace of* R^n, *then* $k \geq j$ *and* $k=j$ *if and only if the collection* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ *is linearly independent.*

One implication of this theorem is that no fewer than j vectors can span the same space as $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_j$ if this collection is linearly independent. This, like most of the results from linear algebra which we use, is "obvious" from a geometrical point of view if the n-tuples are thought of as vectors in a 3-dimensional space.

We now consider linear mappings of the elements of one Cartesian space into another. If L denotes a rule which assigns to every element of R^n an element of R^m, in symbols: $L : R^n \to R^m$, then we say that L is a *linear transformation* if for $\mathbf{x}$ and $\mathbf{y}$ in R^n and all scalars we have

(i) $L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$
(ii) $L(a\mathbf{x}) = aL(\mathbf{x})$

The same definition of linearity is used for mappings defined on any real vector space.

The linear mappings of R^n into R^m can all be described by a set of simultaneous linear equations. Let $\mathbf{x}$ belong to R^n and $\mathbf{z}$ belong to R^m, then any linear mapping $L : R^n \to R^m$ can be described by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= z_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= z_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots &\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= z_m \end{aligned}$$

This notation is very clumsy and many times obscures the basic simplicity of linear mappings. However, it is comforting to know that any linear mapping of R^n into R^m can be expressed in this pedestrian way.

Let L be a linear mapping of R^n into R^m. We say that an n-tuple $\mathbf{x}_1$ lies in the *null space* of L if $L(\mathbf{x}_1) = \mathbf{0}$. We say that an m-tuple $\mathbf{z}_1$ lies in the *range space* of L if there exists an n-vector $\mathbf{x}_2$ such that $L(\mathbf{x}_2) = \mathbf{z}_1$. If there exists a set of k linearly independent n-tuples $x_1, x_2, \ldots, x_k$ such that the m-tuples $L(\mathbf{x}_1), L(\mathbf{x}_2), \ldots, L(\mathbf{x}_k)$ are all zero, and if there exists no set of $k + 1$ linearly independent n-tuples $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}$ such that $L(\mathbf{x}), L(\mathbf{x}_2), \ldots, L(\mathbf{x}_{k+1})$ are all zero, then we say that the *null space of L has dimension k* or that the *nullity of L* is k. If there exists a set of j n-tuples $\mathbf{x}_1, \mathbf{x}_2, \ldots, x_j$ such that the set of $L(\mathbf{x}_1), \mathbf{L}(\mathbf{x}_2), \ldots, L(\mathbf{x}_j)$ is linearly independent and if there exists no set of $j + 1$ n-tuples $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{j+1}$ such that the $L(\mathbf{x}_1), L(\mathbf{x}_2), \ldots, L(\mathbf{x}_{j+1})$ is linearly independent, then we say the *range space of L has dimension j*, or that the *rank of L* is j.

Matrices are introduced so that one can do the arithmetic associated with linear mappings. One of the common arithmetical operations is to compute the effect of two successive linear mappings. To carry this out we need an appropriate definition of matrix multiplication. If $\mathbf{A}$ belongs to $R^{m \times n}$ and $\mathbf{B}$ belongs to $R^{n \times p}$, then we define $\mathbf{AB}$, the *product* of $\mathbf{A}$ and $\mathbf{B}$, as that member $\mathbf{C}$ or $R^{m \times p}$ whose ijth element c_{ij} is given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

It must be observed that in general, $\mathbf{AB}$ is not equal to $\mathbf{BA}$ and indeed, $\mathbf{BA}$ need not even be defined for some $\mathbf{A}$'s and $\mathbf{B}$'s for which $\mathbf{AB}$ makes good sense. This failure of matrix multiplication to be commutative is one of the things that makes the subject interesting.

Two matrices have special significance: $\mathbf{0}$ (*the zero matrix*) is used to denote a matrix all of whose elements are zero and $\mathbf{I}$ (*the identity matrix*) is used to denote a square ($m = n$) matrix whose iith (diagonal) elements are 1 and whose off-diagonal elements are zero. Notice that $\mathbf{0} + \mathbf{A} = \mathbf{A}$ for all $\mathbf{A}$ and $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$ for all square matrices $\mathbf{A}$.

Matrices can be used to describe linear mappings of one finite dimensional vector space into a second finite dimensional vector space in the following way. Agree, henceforth, to write all n-tuples as column vectors,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then regarding $\mathbf{x}$ as an element of $R^{n \times 1}$ we know how to interpret the product $\mathbf{Ax}$ for any $\mathbf{A}$ in $R^{m \times n}$. The equation $\mathbf{Ax} = \mathbf{z}$ is, upon closer examination, simply shorthand notation for the set of simultaneous equations encountered earlier.

Frequently it is convenient to represent a matrix as being composed of two or more submatrices. If $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{C}$ are matrices with the same number of rows then we denote by $(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{C})$ the following matrix

$$(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{C}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} & b_{11} & b_{12} & \cdots & b_{1n} & \cdots & c_{11} & c_{12} & \cdots & c_{1p} \\ a_{21} & a_{22} & \cdots & a_{2m} & b_{21} & b_{22} & \cdots & b_{2n} & \cdots & c_{21} & c_{22} & \cdots & c_{2p} \\ \cdots & & & & & & & & & & & & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qm} & b_{q1} & b_{q2} & \cdots & b_{qn} & \cdots & c_{q1} & c_{q2} & \cdots & c_{qp} \end{bmatrix}$$

On the other hand, if $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are matrices with the same number of columns, we denote by $(\mathbf{A}; \mathbf{B}; \ldots ; \mathbf{C})$ the matrix

$$(\mathbf{A}; \mathbf{B}; \ldots; \mathbf{C}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \cdots & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{m} \\ b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \cdots & & & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nq} \\ \cdots & & & \cdots \\ c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \cdots & & & \cdots \\ c_{p1} & c_{p2} & \cdots & c_{pq} \end{bmatrix}$$

Thus commas are used to denote partitions according to columns and semicolons are used to denote partitions according to rows.

Partitions of the form

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

will be used to denote the matrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} & b_{11} & b_{12} & \cdots & b_{1q} \\ a_{21} & a_{22} & \cdots & a_{2p} & b_{21} & b_{22} & \cdots & b_{2q} \\ \cdots & & & & & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} & b_{m1} & b_{m2} & \cdots & b_{mq} \\ c_{11} & c_{12} & \cdots & c_{1p} & d_{11} & d_{12} & \cdots & d_{1q} \\ c_{21} & c_{22} & \cdots & c_{2p} & d_{21} & d_{22} & \cdots & d_{2q} \\ \cdots & & & & & & & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{np} & d_{n1} & d_{n2} & \cdots & d_{nq} \end{bmatrix}$$

Two square-partitioned matrices

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{bmatrix}$$

are said to be *partitioned conformably* if the submatrices **A**, **D**, **E**, and **H** are square with the dimensions of **A** and **D** being equal to those of **E** and **H** respectively. In this case

$$\mathbf{MN} = \begin{bmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{bmatrix}$$

and

$$\mathbf{NM} = \begin{bmatrix} \mathbf{EA} + \mathbf{FC} & \mathbf{EB} + \mathbf{FD} \\ \mathbf{GA} + \mathbf{HC} & \mathbf{GB} + \mathbf{HD} \end{bmatrix}$$

We immediately transfer our definitions of rank, range space, and null space from linear transformations to matrices. We say that $\mathbf{x}_1$ belongs to the *null space* of **A** if $\mathbf{A}\mathbf{x}_1 = \mathbf{0}$; we say $\mathbf{z}_1$ belongs to the *range space* of **A** if there exists an $\mathbf{x}_2$ such that $\mathbf{A}\mathbf{x}_2 = \mathbf{z}_1$. Finally, the *nullity* of A is the number of linearly independent vectors in the null space of **A** and the *rank* of **A** equals the number of linearly independent vectors in the range space of **A**.

Associated with any element of **A** of $R^{m \times n}$ is an element of $R^{n \times m}$ called the *transpose* of **A** in symbols: $\mathbf{A}'$, such that if a_{ij} is the entry in the ith row and jth column of **A**, then a_{ji} is the entry in the ith row and jth column of $\mathbf{A}'$. A matrix is said to be *symmetric* if it equals its transpose.

The transpose of a column vector **x** belonging to $R^{n \times 1}$ is a row vector $\mathbf{x}'$ belonging to $R^{1 \times n}$. Using the general definition of matrix multiplication we see that $\mathbf{y}'\mathbf{x}$ is defined for column vectors **x** and **y** of the same dimension and that $\mathbf{y}'\mathbf{x}$ is a scalar

$$\mathbf{y}'\mathbf{x} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

This product *viewed as a product between two elements of* R^n is called the *standard inner product* on R^n. Cartesian space, with this definition of inner product, is called the *standard Euclidean space*. We use the symbol E^n to denote this space.

The natural setting for much of the analysis in this book is an inner product space. Abstractly, an *inner product* in a real vector space X is a mapping (written as $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle$) of† $X \times X$ into R^1 such that

(i) $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \mathbf{x}_2, \mathbf{x}_1 \rangle$
(ii) $\langle \mathbf{x}_1, a\mathbf{x}_2 + b\mathbf{x}_3 \rangle = a\langle \mathbf{x}_1, \mathbf{x}_2 \rangle + b\langle \mathbf{x}_1, \mathbf{x}_3 \rangle$
(iii) $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle \geqslant 0$; $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle = 0$ if and only if $\mathbf{x}_1 = \mathbf{0}$.

A vector space with an inner product is called an *inner product space*. A finite dimensional inner product space is called a *Euclidean space*.

Theorem 3. *The standard inner product on* R^n *satisfies the conditions required of an inner product.*

By using the standard inner product in R^n meaning can be given to some additional geometric terms. The *length* of a vector in E^n is $\sqrt{\mathbf{x}'\mathbf{x}}$ and is written as $\|\mathbf{x}\|$. Two vectors, **x** and **y** are called *perpendicular* if and only if $\mathbf{x}'\mathbf{y} = 0$; the zero vector is perpendicular to all nonzero vectors.

Three different inner product spaces are used in this book:

(i) $E^n = R^n$ with $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}'\mathbf{y}$

(ii) $C_*^m[t_0, t_1] = C^m[t_0, t_1]$ with $\langle \mathbf{x}, \mathbf{y} \rangle = \int_{t_0}^{t_1} \mathbf{x}'(t)\mathbf{y}(t)\, dt$

(iii)* $E^{n \times m} = R^{n \times m}$ with $\langle \mathbf{X}, \mathbf{Y} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} y_{ij} = \operatorname{tr} \mathbf{X}'\mathbf{Y}$

The next problem we want to treat is that of solving linear equations of the form $L(\mathbf{x}) = \mathbf{z}$ where **z** is given. However, before we can treat this problem we need to define the adjoint transformation associated with a given linear transformation from one inner product space into another. Adjoint transformations can be defined for linear transformations without reference to an inner product, but when one is available, it helps to use it.

If X is an inner product space with inner product $\langle\ ,\ \rangle_X$ and if Y is

† $X \times X$ is the Cartesian product of X with itself. See, e.g., reference [1].
* If X is a square matrix then tr X (*trace* of X) is the sum of its diagonal elements.

an inner product space with inner product $\langle\ ,\ \rangle_y$ T is the *adjoint* of a linear transformation L if for all $\mathbf{x}$ and $\mathbf{y}$

$$\langle \mathbf{y}, L(\mathbf{x})\rangle_y = \langle T(\mathbf{y}), \mathbf{x}\rangle_x$$

We use the symbol L^* for the adjoint of L. It is possible for the spaces X and Y to be quite different. Keep in mind that the mapping L takes X into Y and the mapping L^* takes Y into X.

Examples. (i) Consider a mapping of E^n into E^m defined by $\mathbf{Ax} = \mathbf{y}$. Then the adjoint mapping is given by $\mathbf{A}'\mathbf{y} = \mathbf{x}$ because

$$\langle \mathbf{y}, \mathbf{Ax}\rangle = \mathbf{y}'A\mathbf{x} = (\mathbf{A}'\mathbf{y})'\mathbf{x} = \langle \mathbf{A}'\mathbf{y}, \mathbf{x}\rangle$$

(ii) Let $\mathbf{B}$ be an n by m matrix of continuous functions. We can regard

$$L(\mathbf{u}) = \int_{t_0}^{t_1} \mathbf{B}(\sigma)\mathbf{u}(\sigma)\, d\sigma$$

as defining a linear mapping of $C_m^*[t_0, t_1]$ into E^n. By definition of $C_*^m[t_0, t_1]$ and E^n we have

$$\langle \mathbf{u}_1, \mathbf{u}_2\rangle = \int_{t_0}^{t_1} \mathbf{u}_1'(\sigma)\mathbf{u}_2(\sigma)\, d\sigma$$

$$\langle \mathbf{x}_1, \mathbf{x}_2\rangle = \mathbf{x}_1'\mathbf{x}_2$$

A short calculation verifies that $L^*(\mathbf{x}) = \mathbf{B}'\mathbf{x}$.

(iii) Let $\mathbf{T}$ be such that the operator $L(\mathbf{u}) = \mathbf{y}$ defined by

$$\mathbf{y}(t) = \int_{t_0}^{t_1} \mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\, d\sigma$$

defines a mapping of $C_m^*[t_0, t_1]$ into $C_q^*[t_0, t_1]$. By definition we have for $\mathbf{y}$ in $C_q^*[t_0, t_1]$

$$\langle \mathbf{y}, L(\mathbf{u})\rangle = \int_{t_0}^{t_1}\int_{t_0}^{t_1} \mathbf{y}'(t)\mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\, d\sigma\, dt$$

$$= \int_{t_0}^{t_1}\int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{T}'(\sigma, t)\mathbf{y}(\sigma)\, d\sigma\, dt$$

provided the interchange of integration is valid. This gives an expression for $L^*(\mathbf{y}) = \mathbf{u}$ of the form

$$\mathbf{u}(t) = \int_{t_0}^{t_1} \mathbf{T}'(\sigma, t)\mathbf{y}(\sigma)\, d\sigma$$

(iv) If X is the space of n by n real matrices and if $\mathbf{A}$ is also an n by n matrix then

$$L(\mathbf{X}) = \mathbf{A}'\mathbf{X} + \mathbf{XA}$$

maps $R^{n\times n}$ into $R^{n\times n}$. We make $R^{n\times n}$ an inner product space with $\langle \mathbf{X}_1, \mathbf{X}_2\rangle = \operatorname{tr} \mathbf{X}_1'\mathbf{X}_2$. Then the adjoint of L is

$$L^*(\mathbf{Y}) = \mathbf{AY} + \mathbf{YA}'$$

We now have enough material on hand to state the basic facts about the solubility of linear algebraic equations.

Theorem 4. *The vector equation* $\mathbf{Ax} = \mathbf{z}$ *with* $\mathbf{A}$ *and* $\mathbf{z}$ *given has a solution if and only if any one of the following three equivalent conditions is satisfied:*

(i) $\mathbf{z}$ *lies in the range space of* $\mathbf{A}$

(ii) $\mathbf{z}$ *is perpendicular to every vector in the null space of* $\mathbf{A}'$

(iii) *the ranks of* $\mathbf{A}$ *and the augmented matrix* $(\mathbf{A}, \mathbf{z})$ *are the same.*

Moreover, if $\mathbf{x}_1$ *is any particular solution, then any other solution is of the form* $\mathbf{x}_1 + \mathbf{x}_2$ *where* $\mathbf{x}_2$ *lies in the null space of* $\mathbf{A}$.

The above theorem is concerned with solving $\mathbf{Ax} = \mathbf{z}$ with $\mathbf{z}$ specified. If $\mathbf{x}$ and $\mathbf{z}$ are n-tuples so that $\mathbf{A}$ belongs to $R^{n\times n}$, then one can ask if a solution exists for all $\mathbf{z}$. If so, there exists a matrix $\mathbf{A}^{-1}$ called the *inverse* of $\mathbf{A}$ such that $\mathbf{x} = \mathbf{A}^{-1}\mathbf{z}$. We assume the reader is familiar with the idea of the *determinant* of a square matrix. We will use $\det \mathbf{A}$ to indicate the determinant of $\mathbf{A}$. If a square matrix has a nonzero determinant we call it *nonsingular*; otherwise it is *singular*. Associated with every square matrix $\mathbf{A}$ is a *resolvent matrix* $(\mathbf{I}s - \mathbf{A})^{-1}$ which is viewed as a function of the complex variable s. The values of s for which $(\mathbf{I}s - \mathbf{A})$ does not have an inverse are called the *eigenvalues* of $\mathbf{A}$ and the equation $\det(\mathbf{I}s - \mathbf{A}) = 0$ which defines the eigenvalues is called the *characteristic equation* of $\mathbf{A}$. The polynomial $p(s) = \det(\mathbf{I}s - \mathbf{A})$ is called the *characteristic polynomial.*

Exercises

1. Show that the product of the eigenvalues of a matrix equal the determinant. Show that the sum of the eigenvalues of a matrix equals the sum of the elements on the main diagonal (= trace).
2. Mappings of $X \times X$ into the scalars which satisfy all inner product conditions except $\langle \mathbf{x}, \mathbf{x}\rangle \geqslant 0$ are sometimes called *pseudo-inner products*. The set of 4-tuples R^4 with the pseudo-inner product $\langle \mathbf{x}, \mathbf{y}\rangle$ defined as $x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$ is called *Minkowski space*. In this space a *Lorentz transformation* is any transformation which preserves inner products; i.e. any T such that $\langle \mathbf{Tx}, \mathbf{Ty}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. If the product of the eigenvalues of $\mathbf{T}$ is 1 (as opposed to -1), then $\mathbf{T}$ is called a *proper Lorentz transformation*. Show that every proper Lorentz transformation has a real eigenvector $\mathbf{x}_e$ such that $\langle \mathbf{x}_e, \mathbf{x}_e\rangle = 0$. Show that in contrast rotations

(i.e. transformations such that $\langle \mathbf{Tx}, \mathbf{Ty}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$) in the Euclidean Space E^4 do not necessarily possess any real eigenvalues.

3. Show that the space of n by n matrices is a vector space. Show that if $\mathbf{M}$ and $\mathbf{N}$ are n by n matrices then

$$L(\mathbf{X}) = \mathbf{MX} + \mathbf{XN}$$

defines a linear transformation from $R^{n\times n}$ to $R^{n\times n}$

4. Suppose $\mathbf{K}(t)$ is singular for all t, then is

$$\int_0^T \mathbf{K}(t)\,dt \qquad \text{singular?}$$

Hint: Let $\mathbf{K}(t) = \begin{bmatrix}\sin t\\ \cos t\end{bmatrix}[\sin t \cos t]$ and calculate $\int_0^{2\pi} \mathbf{K}(t)\,dt$.

5. Ordinarily one does not use division notation when dealing with matrices because $\mathbf{A}/\mathbf{B}$ might be interpreted as $\mathbf{AB}^{-1}$ or $\mathbf{B}^{-1}\mathbf{A}$ and these are not necessarily the same. Why is the notation $(\mathbf{I}+\mathbf{A})/(\mathbf{I}-\mathbf{A})$ unambiguous?
6. Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ be n by n matrices. Compute the inverse of

$$\mathbf{M} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix}$$

in terms of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$. (State clearly any additional assumptions needed.)

7. Show that $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.
8. Show that if $\mathbf{A}$ and $\mathbf{B}$ are nonsingular then $\mathbf{AB}$ is nonsingular and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
9. Prove Theorem 1.
10. Prove Theorem 2.
11. Prove Theorem 3.
12. A subset of a real vector space v which is itself a vector space is called a *subspace* of v. A vector space v is said to be the *direct sum* of two subspaces u and w, in symbols: $v = u \oplus w$, if every vector $\mathbf{x}$ in v can be uniquely decomposed as

$$\mathbf{x} = \mathbf{u} + \mathbf{w}$$

with $\mathbf{u}$ in u and $\mathbf{w}$ in w. Find examples of direct sum decompositions for R^n, $C^m(0, 1)$ and $R^{n\times m}$.

13. (Continuation). Let v be an inner product space and let u be a subspace. We define the *orthogonal complement* of u, in symbols: $u^\perp$, as

$$u^\perp = \{\mathbf{x} : \mathbf{x} \text{ in } v \text{ and } \langle \mathbf{x}, \mathbf{y}\rangle = 0 \text{ for } \mathbf{y} \text{ in } u\}$$

Show that $v = u \oplus u^\perp$.

14. The Linear transformation $L(u)$ defined by

$$L(u)(t) = \int_{-\infty}^{t} \eta(t-\sigma)u(\sigma)\,d\sigma$$

is assumed to map $C_*^1(-\infty, \infty)$ into itself. Compute the adjoint of L.

2. UNIQUENESS OF SOLUTIONS GIVEN THE STATE: LINEARIZATION

The concept of the state of a physical system is of central importance when it comes to applying mathematical methods of analysis. The intuitive idea of what one means by the statement, "The *state* of a system is $\mathbf{x}$," is that by knowing $\mathbf{x}$, one knows all that is needed to be known about the system in order to determine the future behavior assuming that future stimuli can be observed. This informal statement can be taken as a guide to the way the term state is used in the literature. Previous attempts to be more precise and at the same time maintain generality do not seem to have been very successful, and we will not pursue the subject in purely abstract terms. It is convenient, however, to talk about state and state variables in connection with physical systems provided they are described by a set of ordinary differential equations. In this context one has little trouble being precise. If we have a set of simultaneous differential equations, $(y_j^{(i)} = d^i y_j/dt^i)$

$$\begin{aligned}
g_1(y_1, y_1^{(1)}, \ldots, y_1^{(n)}, y_2, y_2^{(1)}, \ldots, y_2^{(n)}, \ldots, y_m^{(n)}) &= 0\\
g_2(y_1, y_1^{(1)}, \ldots, y_1^{(n)}, y_2, y_2^{(1)}, \ldots, y_2^{(n)}, \ldots, y_m^{(n)}) &= 0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
g_m(y_1, y_1^{(1)}, \ldots, y_1^{(n)}, y_2, y_2^{(1)}, \ldots, y_2^{(n)}, \ldots, y_m^{(n)}) &= 0
\end{aligned}$$

then we say that a particular set of initial conditions characterize the state of the physical system if by knowing that set of initial data one can, in principle, construct a unique solution for the system.

By reducing all differential equations involving 2nd and higher derivatives to first order vector equations one achieves a notational and conceptual simplicity which is otherwise unavailable. Since this reduction can ordinarily be done with little effort we usually take as our starting point, a system of first order equations such as $(\dot{x}_i = dx_i/dt)$

$$\begin{aligned}
\dot{x}_1(t) &= f_1[x_1(t), x_2(t), \ldots, x_n(t), t]\\
\dot{x}_2(t) &= f_2[x_1(t), x_2(t), \ldots, x_n(t), t]\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\dot{x}_n(t) &= f_n[x_1(t), x_2(t), \ldots, x_n(t), t]
\end{aligned} \qquad \text{(NL)}$$

Higher order scalar equations which are solved for the highest derivative, such as

$$y^{(n)}(t) = f[y(t), y^{(1)}(t), \ldots, y^{(n-1)}(t), t]$$

can be put in this form by letting $x_i = y^{(i-1)}$ for $i = 1, 2, \ldots, n$ to get

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= x_3(t) \\ &\cdots\cdots \\ \dot{x}_n(t) &= f[x_1(t), x_2(t), \ldots, x_n(t)\ t] \end{aligned}$$

Clearly this idea can be extended to simultaneous higher order equations as well, and thus the first order formulation includes a great many cases of interest.

For convenience we usually prefer to write simultaneous equations as a single *vector* differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), t]; \qquad \mathbf{x} = \text{column vector}$$

Assuming for a moment that this equation has a unique solution passing through $\mathbf{x}_0$ at time t_0, we express its value at a later time t as $\boldsymbol{\phi}(t, \mathbf{x}_0, t_0)$. The solution $\boldsymbol{\phi}$ clearly satisfies the *composition rule*

$$\boldsymbol{\phi}(t, \mathbf{x}_0, t_0) = \boldsymbol{\phi}(t, \boldsymbol{\phi}(t_1, \mathbf{x}_0, t_0), t_1)$$

Since the expression on the right simply says that the solution from t_0 to t is the composition of the solution from t_0 to t_1 and the solution from t_1 to t. The set of values which the solutions may take on is called the *state space* for the given differential equation (or physical system).

In spite of the generality suggested by the equation (NL), we are primarily interested in linear differential equations and, for the time being, we demand that they be homogeneous besides. This being the case, we can write the equation (NL) as

$$\begin{aligned} \dot{x}_1(t) &= a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \cdots + a_{1n}(t)x_n(t) \\ \dot{x}_2(t) &= a_{21}(t)x_1(t) + a_{22}(t)x_2(t) + \cdots + a_{2n}(t)x_n(t) \\ &\cdots\cdots\cdots \\ \dot{x}_n(t) &= a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \cdots + a_{nn}(t)x_n(t) \end{aligned} \qquad \text{(L}'\text{)}$$

By adopting the vector-matrix notation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

we can express the differential equations (L′) as

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\,\mathbf{x}(t) \qquad \text{(L)}$$

Two distinct questions about this equation require attention.

(i) Given an initial vector $\mathbf{x}_0$ and a time t_0, does there *exist* a solution that passes through $\mathbf{x}_0$ at time t_0?

(ii) If there exists a solution passing through $\mathbf{x}_0$ at time t_0, is it *unique*?

The uniqueness question is the easier of the two and we dispose of it at this time.

Theorem 1. *If* $\mathbf{A}$ *is an n by n matrix whose elements are continuous functions of time defined on the interval* $t_0 \leqslant t \leqslant t_1$, *then there is at most one solution of* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *which is defined on the interval* $t_0 \leqslant t \leqslant t_1$ *and takes on the value* $\mathbf{x}_0$ *at* $t = t_0$.

Proof. We will obtain a proof by contradiction. Assume that $\mathbf{x}_1$ and $\mathbf{x}_2$ are two distinct solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ and that $\mathbf{x}_1(t_0) = \mathbf{x}_2(t_0) = \mathbf{x}_0$. Then if we let $\mathbf{z}(t) = \mathbf{x}_1(t) - \mathbf{x}_2(t)$ it follows that $\dot{\mathbf{z}}(t) = \mathbf{A}(t)\mathbf{z}(t)$ with $\mathbf{z}(t_0) = \mathbf{0}$. Premultiplying this vector differential equation in $\mathbf{z}$ by $2\mathbf{z}'(t)$, gives the scalar equation

$$\begin{aligned} \frac{d}{dt}[\mathbf{z}'(t)\mathbf{z}(t)] &= 2\mathbf{z}'(t)\mathbf{A}(t)\mathbf{z}(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} 2z_i(t)a_{ij}(t)z_j(t) \\ &\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} \|\mathbf{z}(t)\|\, 2 \max_{ij} |a_{ij}(t)| \cdot \|\mathbf{z}(t)\| \\ &\leqslant \|\mathbf{z}(t)\|^2 \cdot 2n^2 \max_{ij} |a_{ij}(t)| \end{aligned}$$

Letting η denote the coefficient of $\|\mathbf{z}(t)\|^2$ in this last expression let us write

$$\frac{d}{dt}(\|\mathbf{z}(t)\|^2) - \eta(t)\,\|\mathbf{z}(t)\|^2 \leqslant 0$$

If this inequality is multiplied by the positive integrating factor,

$$\rho(t) = \exp\left[-\int_{t_0}^{t} \eta(\sigma)\, d\sigma\right]$$

then the result can be expressed as

$$\frac{d}{dt}[\rho(t)\,\|\mathbf{z}(t)\|^2] \leqslant 0$$

Integrating this gives for all $t_0 \leqslant t \leqslant t_1$

$$\rho(t)\,\|\mathbf{z}(t)\|^2 - \rho(t_0)\,\|\mathbf{z}(t_0)\|^2 \leqslant 0$$

or, since $\mathbf{z}(t_0) = \mathbf{0}$, $\rho(t)\,\|\mathbf{z}(t)\| \leqslant 0$. Since $\rho(t)$ is positive, this means that $\mathbf{z}(t) = \mathbf{0}$; hence $\mathbf{x}_1(t) = \mathbf{x}_2(t)$ for all $t_0 \leqslant t \leqslant t_1$. ∎

In order to prove that a solution exists, we will actually solve the differential equation, although not in closed form. Even innocent-looking linear equations such as $\ddot{x}(t) + (a + b \sin t)x(t) = 0$ cannot be solved in terms of elementary functions, hence we will need to use an iterative scheme. This is the subject of the next section. We conclude this section with some comments on linearization.

Suppose a given set of nonlinear equations is known to have a solution $\boldsymbol{\phi}$ corresponding to a particular set of initial data and certain parameters $\mathbf{u}$ in the equations. If the initial data and the parameters are changed slightly then it is to be expected that the solution will also change slightly. Assuming that the right side of the equation governing $\boldsymbol{\phi}$ is differentiable with respect to $\mathbf{x}$ and $\mathbf{u}$ we determine this change to first order accuracy by the following technique. If $\boldsymbol{\phi}$ satisfies $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), t, \mathbf{u}(t)]$ then expand $\mathbf{f}$ in a Taylor series about the solution $\boldsymbol{\phi}(t, \mathbf{x}_0, t_0)$ which corresponds to the initial data $\mathbf{x}_0$ at time t_0 and the choice $\mathbf{u}_0$ for the parameter. This gives

$$\begin{aligned}&\mathbf{f}[\boldsymbol{\phi}(t, \mathbf{x}_0, t_0) + \delta\mathbf{x}(t), t, \mathbf{u}_0(t) + \delta\mathbf{u}]\\ &\qquad = \mathbf{f}[\boldsymbol{\phi}(t, \mathbf{x}_0, t_0), t, \mathbf{u}_0(t)] + \mathbf{A}(t)\delta\mathbf{x}(t) + \mathbf{B}(t)\delta\mathbf{u}(t) + \text{higher order terms}\end{aligned}$$

where

$$a_{ij}(t) = \left.\frac{\partial f_i}{\partial x_j}\right|_{\boldsymbol{\phi}(t, \mathbf{x}_0, t_0), \mathbf{u}_0(t), t}$$

$$b_{ij}(t) = \left.\frac{\partial f_i}{\partial u_j}\right|_{\boldsymbol{\phi}(t, \mathbf{x}_0, t_0), \mathbf{u}_0(t), t}$$

Setting the first two terms in the Taylor series equal to the derivative and using the fact that $\boldsymbol{\phi}$ is a solution gives

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) + \mathbf{B}(t)\delta\mathbf{u}(t) + \text{higher order terms}$$

The equation

$$\dot{\mathbf{z}}(t) = \mathbf{A}(t)\mathbf{z}(t) + \mathbf{B}(t)\mathbf{v}(t)$$

is called the *linearized equation* about the solution $\boldsymbol{\phi}$ or *the equation of first variation.*

Example. (Satellite Problem) One example which we will frequently return to in this book is that of a point mass in an inverse square law force field.

The motion of a unit mass is governed by a pair of second order equations in the radius r and the angle θ (Fig. 1).

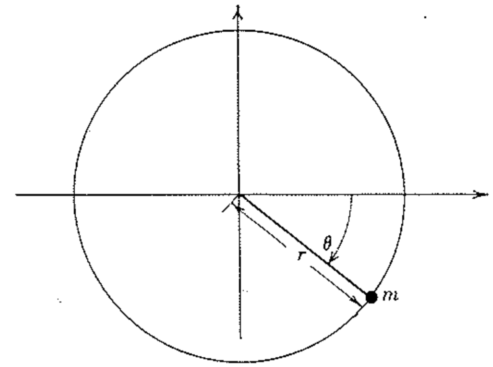


Figure 1. *Illustrating the problem of controlling a point mass in an inverse square law force field.*

If we assume that the unit mass (say a satellite) has the capability of thrusting in the radial direction with a thrust u_1 and thrusting in the tangential direction with a thrust u_2, then we have

$$\ddot{r}(t) = r(t)\dot{\theta}^2(t) - \frac{k}{r^2(t)} + u_1(t)$$

$$\ddot{\theta}(t) = -\frac{2\dot{\theta}(t)\dot{r}(t)}{r(t)} + \frac{1}{r(t)}u_2(t)$$

If $u_1(t) = u_2(t) = 0$, these equations admit the solution

$$\begin{aligned} r(t) &= \sigma \qquad (\sigma \quad \text{constant}) \\ \theta(t) &= \omega t \qquad (\omega \quad \text{constant}) \end{aligned} \qquad \sigma^3\omega^2 = k$$

That is, circular orbits are possible. If we let x_1, x_2, x_3 and x_4 be given by $x_1 = r - \sigma$, $x_2 = \dot{r}$, $x_3 = \sigma(\theta - \omega t)$, $x_4 = \sigma(\dot{\theta} - \omega)$, and normalize σ to 1, then it is easy to see that the linearized equations of motion about the given solution are

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \dot{x}_3(t)\\ \dot{x}_4(t)\end{bmatrix} = \begin{bmatrix}0 & 1 & 0 & 0\\ 3\omega^2 & 0 & 0 & 2\omega\\ 0 & 0 & 0 & 1\\ 0 & -2\omega & 0 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\\ x_4(t)\end{bmatrix} + \begin{bmatrix}0 & 0\\ 1 & 0\\ 0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}u_1(t)\\ u_2(t)\end{bmatrix}$$

Exercises

1. Find two different solutions of the nonlinear equation $\dot{x}(t) = \sqrt{x(t)}$ which pass through $x = 0$ at $t = 0$.

2. The *Euler equations* for the angular velocities of a rigid body are

$$I_1\,\dot\omega_1 = (I_2 - I_3)\omega_2\,\omega_3 + u_1$$
$$I_2\,\dot\omega_2 = (I_3 - I_1)\omega_1\,\omega_3 + u_2$$
$$I_3\,\dot\omega_3 = (I_1 - I_2)\omega_1\,\omega_2 + u_3$$

Here $\boldsymbol{\omega}$ is the angular velocity in a body fixed coordinate system coinciding with the principal axes, $\mathbf{u}$ is the applied torque, and I_1, I_2, and I_3 are the principal moments of inertia. If $I_1 = I_2$, we call the body *symmetrical*. In this case, linearize these equations about the solution $\mathbf{u} = 0$,

$$\begin{bmatrix}\dot\omega_1\\ \dot\omega_2\\ \dot\omega_3\end{bmatrix} = \begin{bmatrix}\cos\,\omega_0(I_2 - I_3)/I_1\\ \sin\,\omega_0(I_2 - I_3)/I_1\\ \omega_0\end{bmatrix}$$

3. A vector valued function of a vector $\mathbf{x}$ is said to satisfy a *Lipschitz condition* with respect to $\mathbf{x}$ if there exists a k such that

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)\| \leqslant k\,\|\mathbf{x}_1 - \mathbf{x}_2\|$$

for all $\mathbf{x}_1$ and $\mathbf{x}_2$. Show that for a given $\mathbf{x}_0$ there is at most one solution of the nonlinear equation $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)]$ passing through $\mathbf{x}_0$ if $\mathbf{f}$ satisfies a *Lipschitz condition.*

4. Matrices of the form

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-1}\end{bmatrix}$$

are said to be *companion matrices.* Show that if x satisfies the scalar nth order equation

$$x^{(n)}(t) + p_{n-1}(t)x^{(n-1)}(t) + \cdots + p_1(t)x^{(1)}(t) + p_0(t)x(t) = 0$$

and if the vector $\mathbf{x}$ is defined by $x_i = x^{(i-1)}$, then $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ with $\mathbf{A}$ being in companion form. Show that if $\phi_1(t), \phi_2(t); \ldots, \phi_n(t)$ are n solutions of the nth order equation, then the vectors

$$\boldsymbol{\phi}_1 = \begin{bmatrix}\phi_1\\ \phi_1^{(1)}\\ \vdots\\ \phi_1^{(n-1)}\end{bmatrix}; \qquad \boldsymbol{\phi}_2 = \begin{bmatrix}\phi_2\\ \phi_2^{(1)}\\ \vdots\\ \phi_2^{(n-1)}\end{bmatrix} \cdots \boldsymbol{\phi}_n = \begin{bmatrix}\phi_n\\ \phi_n^{(1)}\\ \vdots\\ \phi_n^{(n-1)}\end{bmatrix}$$

satisfy the vector equation. Finally, show that the matrix

$$\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \quad \boldsymbol{\phi}_2 \quad \cdots \quad \boldsymbol{\phi}_n]$$

satisfies the matrix equation $\dot{\mathbf{X}}(t) = \mathbf{A}(t)\,\mathbf{X}(t)$.

5. Show that there exists $x_1, x_2, \ldots, x_n$ such that the nth order equation

$$x^{(n)}(t) + p_{n-1}t^{-1}x^{(n-1)}(t) + p_{n-2}\,t^{-2}x^{(n-2)}(t) + \cdots$$
$$+ p_1 t^{-n+1}x^{(1)}(t)x + p_0\,t^{-n}x(t) = 0$$

is converted to $\dot{\mathbf{x}}(t) = t^{-1}\mathbf{A}\mathbf{x}(t)$ with $\mathbf{A}$ constant.

6. Electrical circuits will occasionally be used as examples. The basic building blocks are resistors, inductors and capacitors which we represent as shown in Fig. 2.

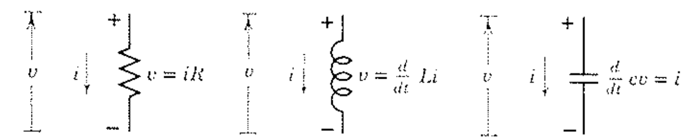


Figure 2

The resistance R, the inductance L and the capacitance C may depend on time or other parameters. Write the equations for the capacitor voltages in the circuit in Fig. 3.

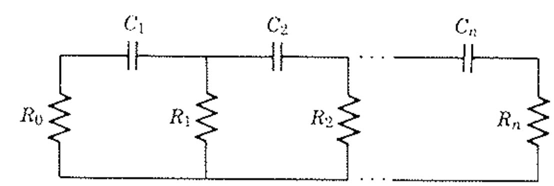


Figure 3

7. In the electrical network shown below the boxes contain nonlinear capacitors for which the voltage v and charge q are related by

$$v = \begin{cases} q^2, & q \geqslant 0\\ 0, & q \leqslant 0\end{cases}$$

There is a current source of strength $i_s = -(1/\sqrt{2}\,\sin\,(t + \pi/4)$ applied as shown in Fig. 4*a*. Show that for suitable values of $q(0)$ this network admits the solution

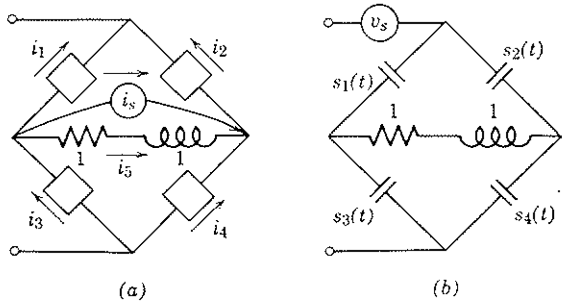


Figure 4 *(a) Nonlinear network; (b) Linearized network.*

$$i_1(t) = i_4(t) = -\tfrac{1}{2}\cos t$$
$$i_2(t) = i_3(t) = -\tfrac{1}{2}\cos t$$
$$i_5(t) = (1\sqrt{2})\sin(t+\pi/4)$$

Linearize the equation of motion about this solution. Call the linearized capacitance

$$c(v) = \frac{1}{2\sqrt{v}}$$

and call the reciprocal of the capacitance the *susceptance*. Show that upon linearization one gets the equivalent network shown in Fig. 4*b* with the susceptances being given by

$$s_1(t) = s_4(t) = 1 + \sin t$$
$$s_2(t) = s_3(t) = 1 - \sin t$$

and the voltage source having the value $\frac{1}{2}\,[1+\sin^2 t]$

8. Three chemical species, S_1, S_2 and S_3 are present in a reaction with reaction rate constants k_{ij}; that is, S_i turns into S_j at a rate $k_{ji}x_i$, where x_1, x_2 and x_3 are the concentrations of the species. Find the equations of evaluation and place them in vector-matrix form.

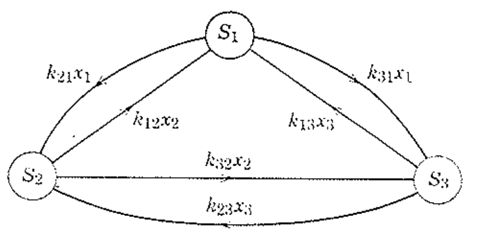


Figure 5

9. A very useful method for obtaining estimates on the behavior of the solutions of differential equations is the *Gronwall-Bellman inequality*. Assuming that $x(t) > 0$ show that the implicit inequality in ϕ

$$\phi(t) \leqslant \psi(t) + \int_a^t x(s)\phi(s)\,ds$$

implies the explicit inequality

$$\phi(t) \leqslant \psi(t) + \int_a^t x(s)\psi(s)\exp\left[\int_s^t x(u)\,du\right] ds$$

Hint: Let $r(t) = \int_a^t x(s)\phi(s)\,ds$ and show that

$$\dot{r}(t) - x(t)r(t) \leqslant x(t)\psi(t)$$

10. For $\mathbf{x}$ an n-vector and $\mathbf{A}$ an n by n matrix, establish the inequality $\|\mathbf{Ax}\| \leqslant n^2 \max_{ij} |a_{ij}| \cdot \|\mathbf{x}\|$. Show that this estimate can be improved on using the vector inequality $|\mathbf{y}'\mathbf{x}| \leqslant \|\mathbf{y}\|\,\|\mathbf{x}\|$ a total of n times.
11. Show that $\sin t$ is a solution of the nonlinear equation

$$\ddot{x}(t) + (4/3)x^3(t) = -\tfrac{1}{3}\sin 3t$$

provided the initial data is chosen properly. What is the linearized equation for motion about the given solution?

3. THE TRANSITION MATRIX

One way to compute the inverse of a given n by n matrix $\mathbf{A}$ is to solve the n linear equations $\mathbf{Az} = \mathbf{b}_1$, $\mathbf{Az} = \mathbf{b}_2$, ..., $\mathbf{Az} = \mathbf{b}_n$ with

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix}, \ldots, \mathbf{b}_n = \begin{bmatrix}0\\0\\0\\\vdots\\1\end{bmatrix}$$

and simply arrange the solutions $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$ in a matrix whose columns are the solution vectors. That is, make the definition $[\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n] = \mathbf{Z}$. Then $\mathbf{AZ} = \mathbf{I}$ so $\mathbf{A}^{-1} = \mathbf{Z}$. Something quite analogous can be used in solving the differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) \qquad \text{(L)}$$

Assume one can solve (L) for $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, subject to the boundary conditions,

$$\mathbf{x}_1(t_0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{x}_2(t_0) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{x}_n(t_0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Let us arrange the solution vectors, whose value at time t we denote by $\boldsymbol{\phi}_i(t, t_0)$, in a square matrix; thus

$$\boldsymbol{\Phi}(t, t_0) = [\boldsymbol{\phi}_1(t, t_0), \boldsymbol{\phi}_2(t, t_0), \ldots, \boldsymbol{\phi}_n(t, t_0)]$$

Then clearly

$$\frac{d}{dt}\boldsymbol{\Phi}(t, t_0) = \mathbf{A}(t)\boldsymbol{\Phi}(t, t_0), \qquad \boldsymbol{\Phi}(t_0, t_0) = \mathbf{I}$$

It is trivial to verify that just as $\mathbf{z}_0 = \mathbf{Z}\mathbf{b}_0$ solves $\mathbf{Az} = \mathbf{b}_0$, $\boldsymbol{\Phi}(t, t_0)\mathbf{x}_0$ solves $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ for $\mathbf{x}(t_0) = \mathbf{x}_0$. We now show that under reasonable assumptions on the smoothness of $\mathbf{A}$, there actually exists such a $\boldsymbol{\Phi}$. The proof involves showing convergence of a matrix valued sequence of time functions.

Recall that a sequence $x_1, x_2, x_3 \ldots$ of scalar valued functions of time defined on the interval $t_0 \leqslant t \leqslant t_1$ is said to *converge* on that interval if there exists a function x defined on $t_0 \leqslant t \leqslant t_1$ such that for every given t in the interval, the sequence of numbers $\{x_i(t)\}$ converges to $x(t)$. The sequence of functions is said to *converge uniformly* if there exists a function x such that for any given $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that for all $i \geqslant N$

$$\sup_{t_0 \leqslant t \leqslant t_1} |x_i(t) - x(t)| \leqslant \varepsilon;$$

A series of scalar valued functions of time $x_1 + x_2 + x_3 \cdots$ defined on the interval $t_0 \leqslant t \leqslant t_1$ is said to *converge* if the sequence of partial sums converges. The series is said to *converge uniformly* if the sequence of partial sums converges uniformly, and it is said to *converge absolutely* if the series remains convergent when every term is replaced by its absolute value. If $E_{ij}(\mathbf{M})$ denotes the ijth element, then we say that a sequence of matrices $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \ldots$, whose elements depend on time, converges uniformly on the interval $t_0 \leqslant t \leqslant t_1$ if each of the scalar sequences $E_{ij}(\mathbf{M}_1)$, $E_{ij}(\mathbf{M}_2)$, $E_{ij}(\mathbf{M}_3), \ldots$ converges uniformly. Series convergence for matrices is defined analogously.

Theorem 1. *If* $\mathbf{A}$ *is a square matrix whose elements are continuous functions of time on the interval* $t_0 \leqslant t \leqslant t_1$ *and if the sequence of matrices* $\mathbf{M}_k$ *is defined*

recursively by

$$\mathbf{M}_0 = \mathbf{I}$$

$$\mathbf{M}_k = \mathbf{I} + \int_{t_0}^{t} \mathbf{A}(\sigma)\mathbf{M}_{k-1}(\sigma)\, d\sigma$$

then the sequence of matrices $\mathbf{M}_0, \mathbf{M}_1 \ldots \mathbf{M}_k \ldots$ *converges uniformly on the given interval. Moreover, if the limit function is denoted by* $\boldsymbol{\Phi}$, *then for* $t_0 \leqslant t \leqslant t_1$

$$\frac{d}{dt}\boldsymbol{\Phi}(t, t_0) = \mathbf{A}(t)\boldsymbol{\Phi}(t, t_0); \qquad \boldsymbol{\Phi}(t_0, t_0) = \mathbf{I} \tag{T}$$

and the solution of (L) *which passes through* x_0 *at* $t = t_0$ *is* $\boldsymbol{\Phi}(t, t_0)\mathbf{x}_0$.

Proof. In order to prove that the sequence of $\mathbf{M}_k$ converges we need to show that the scalar sequence consisting of the 1, 1 elements converges, the scalar sequence consisting of the 1, 2 elements converges, etc. We recall that a series of continuous scalar functions $x_1 + x_2 + x_3 + \cdots$ defined on a closed interval $t_0 \leqslant t \leqslant t_1$, converges absolutely and uniformly on the interval if there exists a sequence of positive constants c_i such that for all t in the interval $|x_i(t)| \leqslant c_i$ and the series $c_1 + c_2 + \cdots$ converges.*

Define η as the maximum absolute value of any entry in $\mathbf{A}(t)$, that is,

$$\eta(t) = \max_{ij} |a_{ij}(t)|$$

Let $\gamma(t)$ be the integral of $\eta(t)$

$$\gamma(t) = \int_{t_0}^{t} \eta(\sigma)\, d\sigma$$

Using again the notation $E_{ij}(\ \ )$ to denote the ijth element of the matrix, we have for n by n matrices $\mathbf{A}$ and $\mathbf{B}$ the obvious inequality

$$|E_{ij}(\mathbf{AB})| \leqslant n \max_{ij} |E_{ij}(\mathbf{A})| \max_{ij} |E_{ij}(\mathbf{B})|$$

Using this we obtain the following estimate for all i and j

$$\begin{aligned} &E_{ij}[\mathbf{M}_k(t, t_0) - \mathbf{M}_{k-1}(t, t_0)] \\ &\quad = E_{ij}\left[\int_{t_0}^{t} \mathbf{A}(\sigma_1) \int_{t_0}^{\sigma_1} \mathbf{A}(\sigma_2) \cdots \int_{t_0}^{\sigma_{k-1}} \mathbf{A}(\sigma_k)\, d\sigma_k \cdots d\sigma_2\, d\sigma_1\right] \\ &\quad \leqslant \int_{t_0}^{t} \int_{t_0}^{\sigma_1} \cdots \int_{t_0}^{\sigma_{k-1}} n^{k-1}\eta(\sigma_1)\eta(\sigma_2) \cdots \eta(\sigma_k)\, d\sigma_k \cdots d\sigma_2\, d\sigma_1 \\ &\quad = \frac{n^{k-1}\gamma^k(t)}{k!} \end{aligned}$$

* This is the Weierstrass M-test. See, for example, Fulks, *Advanced Calculus*, page 364.

Thus we see that each term in the sum

$$E_{ij}\,\mathbf{M}_0(t, t_0) + \sum_{k=1}^{\infty} E_{ij}[\mathbf{M}_k(t, t_0) - \mathbf{M}_{k-1}(t, t_0)]$$

is less than the corresponding term in the sum

$$1 + \gamma(t) + \frac{n\gamma^2(t)}{2!} + \frac{n^2\gamma^3(t)}{3!} + \cdots$$

However, the latter converges for all t to $1 - \dfrac{1}{n} + \dfrac{e^{n\gamma(t)}}{n}$ so each element of the matrix series must converge as well.

To show that the limit $\mathbf{\Phi}$ actually satisfies the differential equation we differentiate term by term to get

$$\begin{aligned}\frac{d}{dt}\mathbf{\Phi}(t, t_0) &= \frac{d}{dt}\left[\mathbf{I} + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\,d\sigma_1 + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\int_{t_0}^{\sigma_1}\mathbf{A}(\sigma_2)\,d\sigma_2\,d\sigma_1 + \cdots\right] \\ &= \mathbf{A}(t)\left[\mathbf{I} + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\,d\sigma_1 + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\int_{t_0}^{\sigma_2}\mathbf{A}(\sigma_2)\,d\sigma_2\,d\sigma_1 \cdots\right] \\ &= \mathbf{A}(t)\,\mathbf{\Phi}\,(t,t_0)\end{aligned}$$

Since the original series is uniformly convergent and since the series obtained by term by term differentiation is also uniformly convergent, term by term differentiation actually yields the derivative* and we see, that $\mathbf{\Phi}$ *satisfies the matrix differential equation* of the theorem statement. It remains only to show that $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ satisfies equation (L). Clearly $\mathbf{\Phi}(t_0, t_0)\mathbf{x}_0 = \mathbf{x}_0$ since $\mathbf{\Phi}(t_0, t_0) = \mathbf{I}$. Thus our tentative solution satisfies the correct initial condition. Taking the derivatives of $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ with respect to time gives

$$\frac{d}{dt}[\mathbf{\Phi}(t, t_0)x_0] = \frac{d}{dt}[\mathbf{\Phi}(t, t_0)]\mathbf{x}_0 = \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0 \quad \blacksquare$$

The expression

$$\mathbf{\Phi}(t, t_0) = \mathbf{I} + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\,d\sigma_1 + \int_{t_0}^{t}\mathbf{A}(\sigma_1)\int_{t_0}^{\sigma_1}\mathbf{A}(\sigma_2)\,d\sigma_2\,d\sigma_1 \cdots \qquad \text{(PB)}$$

often is called the *Peano-Baker series* for the solution of the matrix equation (T).

Corollary 1. *If A is a scalar (one by one matrix), then the Peano-Baker series can be summed and*

$$\mathbf{\Phi}(t, t_0) = \exp\left[\int_{t_0}^{t}\mathbf{A}(\sigma)\,d\sigma\right]$$

* See, for example, Fulks, *Advanced Calculus*, page 370, Theorem 15.3g.

Proof. The proof relies on the fact that since $\mathbf{A}$ is a scalar

$$\mathbf{A}(\sigma_{i-1})\int_{t_0}^{\sigma_{i-1}}\mathbf{A}(\sigma_i)\,d\sigma_i = \int_{t_0}^{\sigma_{i-1}}\mathbf{A}(\sigma_i)\,d\sigma_i\,\mathbf{A}(\sigma_{i-1})$$

Using this fact it is easy to verify by successive differentiation that the kth term in Peano-Baker series can be expressed as

$$\frac{1}{k!}\left[\int_{t_0}^{t}\mathbf{A}(\sigma)\,d\sigma\right]^k$$

which is the kth term in the series expansion for the exponential. $\blacksquare$

This answer could also be obtained by using the "*integrating factor method*" usually found in introductory books on differential equations.

Corollary 2. *If $\mathbf{A}$ is a real constant, n by n matrix, then the Peano-Baker series is*

$$\mathbf{\Phi}(t, t_0) = \mathbf{I} + \mathbf{A}(t - t_0) + \mathbf{A}^2(t - t_0)^2/2! + \cdots$$

and this series converges uniformly and absolutely on any finite interval.

Proof. If $\mathbf{A}$ is a constant, then it can be removed from under the integrals and the Peano-Baker series becomes

$$\mathbf{\Phi}(t, t_0) = \mathbf{I} + \mathbf{A}\int_{t_0}^{t}d\sigma_1 + \mathbf{A}^2\int_{t_0}^{t}\int_{t_0}^{\sigma_1}d\sigma_2\,d\sigma_1 \cdots$$

Evaluating the integrals gives the series indicated. $\blacksquare$

The solution $\mathbf{\Phi}$ of the matrix differential equation (T) is called the *transition matrix** for equation (L).

The importance of the transition matrix lies in the fact that *all* solutions of equation (L) can be expressed in terms of the transition matrix as was brought out by the above theorem.

Example. (Simple harmonic motion). Consider a unit mass connected to a support through a spring whose spring constant is unity. (See Figure 1.) If x measures the displacement of the mass from equilibrium, then $\ddot{x}(t) + x(t) = 0$. Letting $x_1 = x$ and letting x_2 be the velocity of the mass gives

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

The associated transition matrix therefore satisfies

$$\begin{bmatrix}\dot{\phi}_{11}(t, t_0) & \dot{\phi}_{12}(t, t_0)\\ \dot{\phi}_{21}(t, t_0) & \dot{\phi}_{22}(t, t_0)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}\phi_{11}(t, t_0) & \phi_{12}(t, t_0)\\ \phi_{21}(t, t_0) & \phi_{22}(t, t_0)\end{bmatrix}$$

* In the older literature this is also called the *matrizant*.

with the initial condition $\Phi(t_0, t_0) = \mathbf{I}$. What is the physical interpretation of Φ in this case? The first column of Φ has as its first entry the position as a function of time which results when the mass is displaced by one unit and released at t_0 with zero velocity. The second entry in the first column is the corresponding velocity. The second column of Φ has as its first entry the position as a function of time which results when the mass is started from zero displacement but with unit velocity at $t = t_0$. The second entry in the second column is the corresponding velocity.

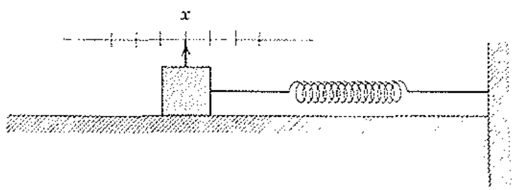


Figure 6. A physical model for simple harmonic motion.

The series for computing Φ in this case is easily summed because $\mathbf{A}$ is constant and in addition $\mathbf{A}^i = (-1)^{i+1}\mathbf{A}$ for i odd and $\mathbf{A}^i = (-1)^{i+1}\mathbf{I}$ for i even. A short calculation gives

$$\Phi(t, t_0) = \begin{bmatrix} \cos(t - t_0) & \sin(t - t_0) \\ -\sin(t - t_0) & \cos(t - t_0) \end{bmatrix}$$

Example. (Satellite Problem) In section 2 we introduced the equations of a unit mass in an inverse square law force field. These were then linearized about a circular orbit to get

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 \\ u_1(t) \\ 0 \\ u_2(t) \end{bmatrix}$$

If $u_1 = u_2 = 0$ then these equations are homogeneous and the results of the present section apply. Since the $\mathbf{A}$ matrix is constant it follows that Φ can be computed via Corollary 2. Using methods to be discussed in Section 5 this series can be summed to get

$$\Phi(t, 0) = \begin{bmatrix} 4 - 3\cos\omega t & \sin\omega t/\omega & 0 & 2(1 - \cos\omega t)/\omega \\ 3\omega\sin\omega t & \cos\omega t & 0 & 2\sin\omega t \\ 6(-\omega t + \sin\omega t) & -2(1 - \cos\omega t)/\omega & 1 & (-3\omega t + 4\sin\omega t)/\omega \\ 6\omega(-1 + \cos\omega t) & -2\sin\omega t & 0 & -3 + 4\cos\omega t \end{bmatrix}$$

Exercises

1. Verify that $\Phi(t, 0)$ for the 2-dimensional system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -2\delta \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is given by ($\omega = \sqrt{1 - \delta^2}$)

$$\begin{bmatrix} e^{-\delta t}\left[\cos\omega t + \dfrac{\delta}{\omega}\sin\omega t\right] & \dfrac{1}{\omega}e^{-\delta t}\sin\omega t \\ -\dfrac{1}{\omega}e^{-\delta t}\sin\omega t & e^{-\delta t}\left[\cos\omega t - \dfrac{\delta}{\omega}\sin\omega t\right] \end{bmatrix}$$

2. Show that every element of $\Phi(t, t_0)$ is nonnegative for all $t \geqslant t_0$ if $a_{ij}(t) \geqslant 0$ for all $i \neq j$ and all $t > t_0$. Conclude that $x_i(t) \geqslant z_i(t)$ for all i and all $t > t_0$ if $\mathbf{A}$ satisfies this condition with

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t); \quad \dot{\mathbf{z}}(t) = \mathbf{A}(t)\mathbf{z}(t); \quad x_i(t_0) \geqslant z_i(t_0) \quad \text{for all } i$$

3. Given that $\mathbf{A}$ is an n by n matrix in companion form (problem 4, section 2) find an expression for the transition matrix Φ associated with $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ in terms of the solutions of a suitable nth order scalar equation satisfying appropriate initial data.
4. Matrix differential equations are natural tools in some physical problems. For example, consider the problem of describing the orientation of one set of coordinate axes with respect to a second set of axes. (See Figure.) Say that the projection on the y_j axis of a unit vector along the x_i axis is a_{ij}. There are nine such direction cosines and we can arrange them as a matrix $\mathbf{A}$. If the x system is rotating about its x_1 axis with an angular velocity ω_1, about its x_2 axis with an angular velocity ω_2, and about its x_3 axis with an angular velocity ω_3, then $\mathbf{A}$ will change with time. Show that

$$\dot{\mathbf{A}}(t) = \Omega(t)\mathbf{A}(t); \qquad \Omega = \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{bmatrix}$$

5. Let $\mathbf{A}$ be a constant matrix. Find the transition matrix $\Phi(t, 0)$ for the time varying linear system

$$\dot{\mathbf{x}}(t) = f(t)\mathbf{A}\mathbf{x}(t)$$

where $f(\ )$ is a continuous function of t.

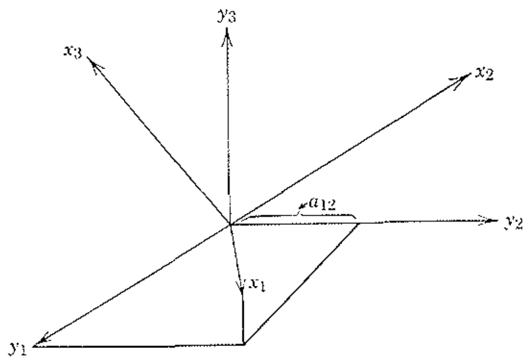


Figure 2

6. If $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$; $\mathbf{x}(0) = \mathbf{x}_0$ and $\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$; $\mathbf{p}(0) = \mathbf{p}_0$, show that $\mathbf{x}'(t)\mathbf{p}(t) = \mathbf{x}_0'\mathbf{p}_0$ for all t.
7. Given that $\mathbf{A}$ is a 2×2 constant matrix and given that

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

Suppose that if

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \qquad \text{then } \mathbf{x}(t) = \begin{bmatrix} e^{-3t} \\ -3e^{-3t} \end{bmatrix}$$

and if

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{then } \mathbf{x}(t) = \begin{bmatrix} e^{t} \\ e^{t} \end{bmatrix}$$

Determine the transition matrix for the system and the matrix $\mathbf{A}$.
8. Show that a square nonsingular matrix $\mathbf{\Phi}(\ ,\ )$, which depends on two arguments and is differentiable with respect to each, is a transition matrix if $\mathbf{\Phi}(t_0, t_0) = \mathbf{I}$ for all t_0 and the matrix

$$\left\{\frac{d}{dt}\mathbf{\Phi}(t, t_0)\right\}\mathbf{\Phi}^{-1}(t, t_0)$$

depends on t only.
9. Suppose that the boundary conditions for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ are specified in part at $t = t_0$ and in part at $t = t_1$. In particular, suppose

$$\mathbf{M}x(t_0) + \mathbf{N}\mathbf{x}(t_1) = \mathbf{b}$$

with Rank $(\mathbf{M}, \mathbf{N}) = \dim \mathbf{x}$. Show that for this *two point boundary value problem* there exists a unique solution if $\det[\mathbf{M} + \mathbf{N}\mathbf{\Phi}(t_1, t_0)]$ is nonzero.

10. *Conjecture*. If $w(t)$ is real and nonnegative for all $t > 0$ and if w is of the form $\mathbf{c}e^{\mathbf{A}t}\mathbf{b}$ then there exists a time-dependent matrix $\mathbf{A}(t)$ whose off-diagonal elements are nonnegative such that w is the 1,1-element in the transition matrix for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$, provided $w(0) = 1$.
11. As is well known, there is a one-to-one correspondence between the set of all 2 by 2 matrices of the form

$$\mathbf{M} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

and the complex members

$$z = (\sigma + j\omega)$$

Show that multiplications for these matrices, like multiplication for complex numbers is commutative. Let z and g be complex functions. What is the solution of

$$\dot{z}(t) = g(t)z(t)$$

Write this complex differential equation as a pair of real equations. Use this as a hint to find the transition matrix for

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} a(t) & b(t) \\ -b(t) & a(t) \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

4. THREE PROPERTIES OF $\mathbf{\Phi}$

The transition matrix specifies how the state of a dynamical system evolves in time. Because of its great importance its basic properties need to be studied with care. Three of the most important are collected here.

It is often essential to know when the transition matrix has an inverse or alternatively, when $\det \mathbf{\Phi}(t, t_0)$ is nonzero. Although it is possible to settle this question by use of the existence and uniqueness theorem we chose to resolve it by actually calculating $\det \mathbf{\Phi}$.

By the *trace* of a square matrix we mean the sum of its diagonal elements.

$$\text{trace } \mathbf{A} = \text{tr}(\mathbf{A}) = \sum_i^n a_{ii}$$

It is easily verified that even though $\mathbf{AB}$ is generally not equal to $\mathbf{BA}$ and indeed, even though $\mathbf{AB}$ and $\mathbf{BA}$ may be of different dimensions,

$$\text{tr } \mathbf{AB} = \text{tr } \mathbf{BA}$$

It follows that $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB})$ (exercise 1). It turns out that the trace of $\mathbf{A}$ and the determinant of $\mathbf{\Phi}$ are related by a curious formula whose origin is usually attributed to some combination of Abel, Jacobi, and Liouville.

Theorem 1. (Abel-Jacobi-Liouville) *If* $\mathbf{\Phi}$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$, *then*

$$\det \mathbf{\Phi}(t, t_0) = \exp\left[\int_{t_0}^{t} \operatorname{tr} \mathbf{A}(\sigma)\, d\sigma\right] \qquad \text{(AJL)}$$

and hence $\det \mathbf{\Phi}(t, t_0)$ *is nonzero if* $\int_{t_0}^{t} \operatorname{tr} \mathbf{A}(\sigma)\, d\sigma$ *is finite.*

Proof. Recall that if C_{ij} is the cofactor of the ij element of $\mathbf{\Phi}$, then for any j

$$\det \mathbf{\Phi} = \sum_{i=1}^{n} C_{ij}\phi_{ij}$$

Thus since ϕ_{ij} does not appear in C_{ij} we see that

$$\frac{\partial}{\partial \phi_{ij}} \det \mathbf{\Phi} = C_{ij}$$

Hence, using the chain rule we have

$$\frac{d}{dt} \det \mathbf{\Phi}(t, t_0) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{\partial}{\partial \phi_{ij}} \det \mathbf{\Phi}(t, t_0)\right]\frac{d}{dt}\phi_{ij}(t, t_0)$$

Using $\mathbf{C}(t, t_0)$ to denote the matrix of cofactors and using the definition of trace, this becomes

$$\frac{d}{dt} \det \mathbf{\Phi}(t, t_0) = \operatorname{tr} \mathbf{C}'(t, t_0)\dot{\mathbf{\Phi}}(t, t_0)$$

However, $\dot{\mathbf{\Phi}}(t, t_0) = \mathbf{A}(t)\mathbf{\Phi}(t, t_0)$, $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{CAB})$, and $\mathbf{\Phi}'\mathbf{C}' = \mathbf{I} \det \mathbf{\Phi}$ so

$$\frac{d}{dt} \det \mathbf{\Phi}(t, t_0) = \operatorname{tr} \mathbf{\Phi}(t, t_0)\mathbf{C}'(t, t_0)\mathbf{A}(t) = \det \mathbf{\Phi}(t, t_0) \operatorname{tr} \mathbf{A}(t)$$

An integration of this last differential equation with the boundary condition $\det \mathbf{\Phi}(t_0, t_0) = \mathbf{I}$ gives the equation (AJL). ▮

This result has an interesting geometrical interpretation. As is well known, the determinant of an n by n matrix can be interpreted as the volume in E^n contained in the parallelepiped generated by the columns in the matrix. Thus

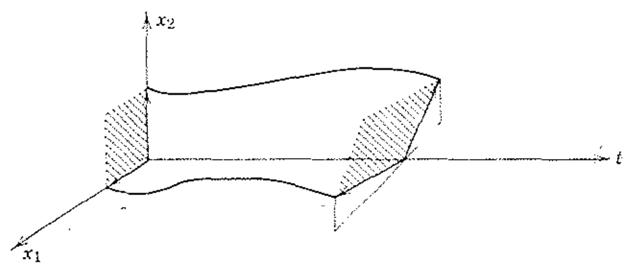


Figure 1. Illustrating the evolution of det $\mathbf{\Phi}$ *for a two dimensional equation.*

this formula gives an expression for how this volume evolves in time. (See figure.) Cases for which $\operatorname{tr} \mathbf{A}(t)$ vanishes identically are important in mechanics. In this case the volume is constant.

We note that as a consequence of the definition of $\mathbf{\Phi}(t, t_0)$ we have the following theorem which we call the *composition rule* for transition matrices.

Theorem 2. *If* $\mathbf{\Phi}$ *is a transition matrix, then it satisfies the functional equation**

$$\mathbf{\Phi}(t, t_0) = \mathbf{\Phi}(t, t_1)\mathbf{\Phi}(t_1, t_0)$$

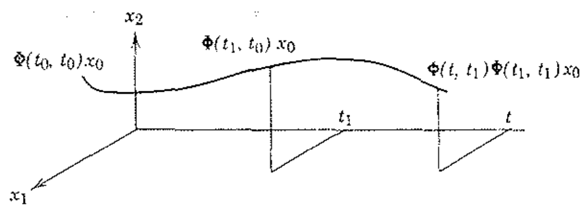


Figure 2. Illustration of the composition rule.

Proof. To see this it is helpful to draw a picture of what the transition matrix means in terms of solutions of equation (L). From the uniqueness theorem 2.1 we see that the two different ways of viewing the value of the state at t must give the same result.

Diagrams (see below) which illustrate alternative linear maps are called *commutative diagrams.* We will supplement the verbal statement of many of the harder results in this book by diagrams of this type.

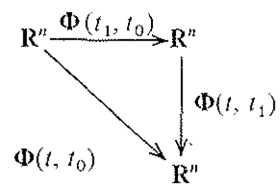


Figure 3. Diagramming the composition rule.

By this diagram we mean that the mapping of the state space into itself defined by $\mathbf{\Phi}(t, t_0)$ can be viewed alternatively by first applying $\mathbf{\Phi}(t_1, t_0)$ then mapping via $\mathbf{\Phi}(t, t_1)$.

The third property of $\mathbf{\Phi}$ of interest here concerns the behavior of $\mathbf{\Phi}$ under a change of coordinates. If the equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ has $\mathbf{\Phi}$ as a transition matrix and if a change of variable $\mathbf{z}(t) = \mathbf{P}(t)\mathbf{x}(t)$ is made, then what is the new

* An equation which relates the values of a function at different arguments is called a *functional equation.*

transition matrix? Assuming $\dot{\mathbf{P}}$ and $\mathbf{P}^{-1}$ exist, it follows that

$$\begin{aligned}\dot{\mathbf{z}}(t) &= \mathbf{P}(t)\dot{\mathbf{x}}(t) + \dot{\mathbf{P}}(t)\mathbf{x}(t)\\ &= [\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]\mathbf{z}(t)\end{aligned}$$

This leads to a third basic identity for transition matrices. Here and elsewhere we write $\Phi'_{\mathbf{A}}$ to identify the matrix which generates Φ if confusion is possible.

Theorem 3. *If* $\mathbf{P}$ *is differentiable and if* $\mathbf{P}^{-1}$ *exists, then*

$$\Phi_{\mathbf{A}}(t, t_0) = \mathbf{P}^{-1}(t)\Phi_{[\mathbf{PAP}^{-1} + \dot{\mathbf{P}}\mathbf{P}^{-1}]}(t, t_0)\mathbf{P}(t_0)$$

Like the composition law, this result is conveniently remembered in terms of a diagram. Figure 4 illustrates the fact that the mapping of $\mathbf{x}(t_0)$ into $\mathbf{x}(t)$ can be viewed in two ways. The simple way, via $\Phi_{\mathbf{A}}(t, t_0)$ and the complicated way involving going down to $\mathbf{z}(t_0)$, across to $\mathbf{z}(t)$, and back up to $\mathbf{x}(t)$.

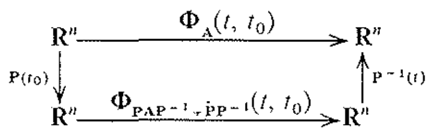


Figure 4. Diagramming the identity of Theorem 3.

Exercises

1. Show that if $\mathbf{A}$ and $\mathbf{B}$ are n by m and m by n, respectively, then $\operatorname{tr}\mathbf{AB} = \operatorname{tr}\mathbf{BA}$. Show that $\operatorname{tr}\mathbf{ABC} = \operatorname{tr}\mathbf{CAB}$.
2. Let $\mathbf{X}$ be an n by n matrix and let f be a scalar function of $\mathbf{X}$. Let $\nabla_X f$ denote an n by n matrix whose ijth element is the partial derivative of f with respect to x_{ij}. Show that if $\mathbf{A}$ is n by n, then

   (i) $\nabla_A \det \mathbf{A} = (\mathbf{A}')^{-1} \det \mathbf{A} = \mathit{Adj}\ \mathbf{A}$
   (ii) $\nabla_A \operatorname{tr}(\mathbf{A}'\mathbf{X}) = \mathbf{X}$

3. Show that to first order in ε

$$\det[I + \varepsilon\mathbf{A}(t)] = 1 + \varepsilon \operatorname{tr}\mathbf{A}(t)$$

   Use this to show that

$$\begin{aligned}\frac{d}{dt}\det\Phi(t, t_0) &= \lim_{h\to 0}\frac{1}{h}[\det\Phi(t+h, t_0) - \det\Phi(t, t_0)]\\ &= \lim_{h\to 0}\frac{1}{h}[\det[\Phi(t, t_0) + h\mathbf{A}\Phi(t, t_0)] - \det\Phi(t, t_0)]\\ &= [\operatorname{tr}\mathbf{A}(t)]\det\Phi(t, t_0)\end{aligned}$$

4. Given a scalar equation

$$x^{(n)}(t) + p_{n-1}(t)x^{(n-1)}(t) + \cdots + p_1(t)x^{(1)}(t) + p_0(t)x^{(0)}(t) = 0$$

   we have seen (Problem 4, section 2) that it can be represented in first order form with $\mathbf{A}$ in companion form. Show, with the help of problem 3, section 3, that if $x_i(t)$ is a solution which at t_0 has $n-1$ of its zeroth through n-1st derivatives zero and its ith derivative 1 then

$$\det\begin{bmatrix} x_1^{(0)}(t) & x_2^{(0)}(t) & \cdots & x_n^{(0)}(t)\\ x_1^{(1)}(t) & x_2^{(1)}(t) & \cdots & x_n^{(1)}(t)\\ \vdots & \vdots & & \vdots\\ x_1^{(n-1)}(t) & x_2^{(n-1)}(t) & \cdots & x_n^{(n-1)}(t)\end{bmatrix} = \exp\left[-\int_{t_0}^{t} p_{n-1}(\sigma)\,d\sigma\right]$$

   In this context the determinant on the left is called the *Wronskian* associated with the given equation and the given solutions.

5. MATRIX EXPONENTIALS

In the case where $\mathbf{A}$ is a constant the infinite series for the transition matrix takes the form

$$\Phi(t, t_0) = \mathbf{I} + \mathbf{A}(t - t_0) + \mathbf{A}^2(t - t_0)^2/2! + \cdots$$

In view of similarity between this series and the scalar series for e^{at}, it is logical to define the *matrix exponential** by

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2/2! + \cdots$$

From corollary 2 of the existence theorem, it follows that this series converges absolutely for all real (square) matrices $\mathbf{A}$. This definition is a useful one because it permits one to express the solution of the initial value problem

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t); \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{L}$$

very succinctly, i.e.,

$$\mathbf{x}(t) = e^{\mathbf{A}(t - t_0)}\mathbf{x}_0$$

However, we must not be deceived by the notation; evaluation of $e^{\mathbf{A}t}$ can be a difficult task.

The following theorem gives two very useful properties of matrix exponentials.

* Keep in mind that this notation, like most other notation used in mathematics, can be misleading. Certainly $10^{1/2}$ for the square root of 10 is not very informative if we insist that it is 10 multiplied by itself 1/2 times.

Theorem 1. *If* $\mathbf{A}$ *is given by*

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \quad \textit{then} \quad e^{\mathbf{A}} = \begin{bmatrix} e^{a_{11}} & 0 & \cdots & 0 \\ 0 & e^{a_{22}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e^{a_{nn}} \end{bmatrix}$$

If $\mathbf{P}$ *is any nonsingular* n *by* n *matrix, then*

$$e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P}$$

Proof. The first assertion is obvious from the definition of $e^{\mathbf{A}}$. To prove the second, it is enough to write out the series involved

$$\begin{aligned} e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} &= \mathbf{I} + \mathbf{P}^{-1}\mathbf{A}\mathbf{P} + \tfrac{1}{2}\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{P}^{-1}\mathbf{A}\mathbf{P} + \cdots \\ &= \mathbf{I} + \mathbf{P}^{-1}\mathbf{A}\mathbf{P} + \tfrac{1}{2}\mathbf{P}^{-1}\mathbf{A}^2\mathbf{P} + \cdots \\ &= \mathbf{P}^{-1}(\mathbf{I} + \mathbf{A} + \tfrac{1}{2}\mathbf{A}^2 + \cdots)\mathbf{P} \\ &= \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P} \quad \blacksquare \end{aligned}$$

The composition rule gives immediate verification of the identity

$$e^{\mathbf{A}t}e^{\mathbf{A}\sigma} = e^{\mathbf{A}(t+\sigma)}$$

None the less, it is generally false that $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}$ unless $\mathbf{AB} = \mathbf{BA}$. To see why, examine the series involved.

$$\begin{aligned} e^{\mathbf{A}}e^{\mathbf{B}} &= (\mathbf{I} + \mathbf{A} + \mathbf{A}^2/2! + \cdots)(\mathbf{I} + \mathbf{B} + \mathbf{B}^2/2! \cdots) \\ &= \mathbf{I} + \mathbf{A} + \mathbf{B} + \tfrac{1}{2}(\mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2) + \cdots \end{aligned}$$

On the other hand,

$$e^{\mathbf{A}+\mathbf{B}} = \mathbf{I} + \mathbf{A} + \mathbf{B} + \tfrac{1}{2}(\mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2) + \cdots$$

The difficulty is that we may have $2\mathbf{AB} \neq \mathbf{AB} + \mathbf{BA}$. There is a famous theorem which relates $e^{\mathbf{A}+\mathbf{B}}$ and $e^{\mathbf{A}}e^{\mathbf{B}}$ via the *commutator product* defined as $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. The essence of this result is that if the nonvanishing of $[\mathbf{A}, \mathbf{B}]$ is what prevents equality between $e^{\mathbf{A}}e^{\mathbf{B}}$ and $e^{\mathbf{A}+\mathbf{B}}$, then the true relationship should be expressible in terms of a power series involving $[\mathbf{A}, \mathbf{B}]$. Although we will make no direct use of the result here, it is included as a reminder that in general $e^{\mathbf{A}}e^{\mathbf{B}} \neq e^{\mathbf{A}+\mathbf{B}}$.

Theorem 2. (Baker, Campbell, Hausdorff) *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be square matrices. Then* $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{C}}$ *where* $\mathbf{C}$ *is given up to* 4th *order by*

$$\mathbf{C} = \mathbf{A} + \mathbf{B} + \tfrac{1}{2}[\mathbf{A}, \mathbf{B}] + \tfrac{1}{12}\{[[\mathbf{B}, \mathbf{A}], \mathbf{A}] + [[\mathbf{B}, \mathbf{A}], \mathbf{B}]\} + \cdots$$

The proof is an involved, but elementary, calculation and is left as an exercise.

The main importance of this fact for our work is that the formula

$$\mathbf{\Phi}(t, t_0) = \exp\left[\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma\right] \qquad \text{(generally not true)}$$

which one might be tempted to conjecture on the basis of corollary 1 of the existence theorem is generally false but does hold if for all t

$$\mathbf{A}(t) \int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma = \int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma \mathbf{A}(t)$$

The reason becomes apparent if the series definition is appealed to.

$$\begin{aligned} \frac{d}{dt}\exp\left[\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma\right] &= \frac{d}{dt}\left[\mathbf{I} + \int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma + \frac{1}{2}\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma \int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma + \cdots\right] \\ &= \mathbf{A}(t) + \frac{1}{2}\left[\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma \mathbf{A}(t) + \mathbf{A}(t)\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma\right] + \cdots \end{aligned}$$

If $\mathbf{A}$ and its integral commute, then the $\mathbf{A}(t)$ term can be brought out in front to give

$$\frac{d}{dt}\exp\left[\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma\right] = \mathbf{A}(t)\exp\left[\int_{t_0}^{t} \mathbf{A}(\sigma)\, d\sigma\right] \qquad \text{(if } \mathbf{A} \text{ and } \textstyle\int \mathbf{A}\, dt \text{ commute)}$$

Otherwise this is not necessarily true and the transition matrix is not necessarily the exponential of the integral of $\mathbf{A}$.

The *Laplace transform* provides an alternative approach to solving linear time-invariant equations and, as turns out, frequently leads to an explicit formula for the matrix exponential. Consider solving equation (L) using Laplace transform theory. Using the expression for the Laplace transform of a derivative, we get upon transforming (L)

$$s\hat{\mathbf{X}}(s) - \mathbf{A}\hat{\mathbf{X}}(s) = \mathbf{X}_0$$

where we use a circumflex to denote transformed variables. Solving for $\hat{\mathbf{X}}$ gives

$$\hat{\mathbf{X}}(s) = (\mathbf{I}s - \mathbf{A})^{-1}\mathbf{X}_0$$

Writing $(\mathbf{I}s - \mathbf{A})^{-1}$ as the adjoint over the determinant and using the complex inversion formula, we have

$$\mathbf{X}(t) = \frac{1}{2\pi i}\int_{-i\infty+\sigma}^{+i\infty+\sigma} \{\det(\mathbf{I}s - \mathbf{A})\}^{-1}\{Adj(\mathbf{I}s - \mathbf{A})\}e^{st}\, ds\mathbf{X}_0$$

where σ is larger than the real part of any pole of the integrand. Since the solution of equation (L) is known to be $e^{\mathbf{A}t}\mathbf{x}_0$ we see that

$$e^{\mathbf{A}t} = \frac{1}{2\pi i}\int_{-i\infty+\sigma}^{+i\infty+\sigma} \{\det(\mathbf{I}s - \mathbf{A})\}^{-1}\{Adj(\mathbf{I}s - \mathbf{A})\}e^{st}\, ds$$

The matrix $Adj(\mathbf{I}s - \mathbf{A})$ is a matrix of polynomials so the poles of the integrand are simply the zeros of $\det(\mathbf{I}s - \mathbf{A})$, that is, the eigenvalues of $\mathbf{A}$. Suppose $s_1, s_2 \ldots s_m$ denote the (distinct) eigenvalues of $\mathbf{A}$ and suppose s_i is repeated σ_i times. If $(\ )^{(k)}$ denotes the kth derivative, then from the Cauchy integral formula we obtain the following result.

Theorem 3. *The matrix exponential is given by*

$$e^{\mathbf{A}t} = \sum_{i=1}^{m}\sum_{k=0}^{\sigma_i - 1}\frac{1}{k!}\frac{1}{(\sigma_i - 1 - k)!}[(s - s_i)^{\sigma_i}(s\mathbf{I} - \mathbf{A})^{-1}]^{(\sigma_i - 1 - k)}t^k e^{st}\Big|_{s=s_i}$$

Examples. (i) *Newton's second law* for a point mass is $\ddot{x}(t) = f(t)$. Writing this in first order form gives

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}0\\ f(t)\end{bmatrix}$$

where x_1 is position and x_2 is velocity. In this case the series of $e^{\mathbf{A}t}$ contains only two terms since $\mathbf{A}^i = \mathbf{0}$ for $i \geqslant 2$. We have

$$\exp\left\{\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}t\right\} = \begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$$

The analysis of the effects of f will be postponed until section 6.

(ii) *Simple harmonic motion* is governed by the equation $\ddot{x}(t) + \omega^2 x(t) = 0$. If we let $x_1 = x$ and $\dot{x}_1 = \omega x_2$, then

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & \omega\\ -\omega & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

A quick calculation will verify that

$$\exp\left\{\begin{bmatrix}0 & \omega\\ \omega & 0\end{bmatrix}t\right\} = \begin{bmatrix}\cos\omega t & \sin\omega t\\ -\sin\omega t & \cos\omega t\end{bmatrix}$$

(iii) Using the fact that $\sigma\mathbf{I}$ commutes with any matrix, we can use the decomposition

$$\begin{bmatrix}\sigma & \omega\\ -\omega & \sigma\end{bmatrix} = \begin{bmatrix}\sigma & 0\\ 0 & \sigma\end{bmatrix} + \begin{bmatrix}0 & \omega\\ -\omega & 0\end{bmatrix}$$

and the rule $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$, valid for $\mathbf{AB} = \mathbf{BA}$, to verify that

$$\exp\left\{\begin{bmatrix}\sigma & \omega\\ -\omega & \sigma\end{bmatrix}t\right\} = \begin{bmatrix}e^{\sigma t} & 0\\ 0 & e^{\sigma t}\end{bmatrix}\cdot\begin{bmatrix}\cos\omega t & \sin\omega t\\ -\sin\omega t & \cos\omega t\end{bmatrix}$$

$$= \begin{bmatrix}e^{\sigma t}\cos\omega t & e^{\sigma t}\sin\omega t\\ -e^{\sigma t}\sin\omega t & e^{\sigma t}\cos\omega t\end{bmatrix}$$

(iv) If $\mathbf{A}$ and $\mathbf{B}$ are square matrices then an appeal to the series definition of the matrix exponential shows that

$$\exp\left\{\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}\end{bmatrix}t\right\} = \begin{bmatrix}e^{\mathbf{A}t} & \mathbf{0}\\ \mathbf{0} & e^{\mathbf{B}t}\end{bmatrix}$$

Exercises

1. Let $\mathbf{A}$ and $\mathbf{B}$ be constant n by n matrices. Obtain the first few terms in a power series for

$$\mathbf{P}(t) = \frac{d}{dt}e^{[\mathbf{A}+\mathbf{B}t]}$$

In what cases can the power series be expressed in terms of exponentials?

2. Show that if $\mathbf{A}$ and $\mathbf{B}$ are constant square matrices, then the transition matrix for the time-varying equation

$$\dot{\mathbf{x}}(t) = e^{-\mathbf{A}t}\mathbf{B}e^{\mathbf{A}t}\mathbf{x}(t)$$

is

$$\mathbf{\Phi}(t, t_0) = e^{-\mathbf{A}t}e^{(\mathbf{A}+\mathbf{B})(t-t_0)}e^{\mathbf{A}t_0}$$

3. If y is given by

$$y(t) = \sum_{k=1}^{n} a_k \sin 2\pi kt + b_k \cos 2\pi kt$$

Find column vectors $\mathbf{c}$ and $\mathbf{b}$ and a square matrix $\mathbf{A}$ such that

$$y(t) = \mathbf{c}'e^{\mathbf{A}t}\mathbf{b}$$

4. If $\mathbf{A}$ is a constant n by n matrix, and if it has an inverse, then show that

$$\int_0^t e^{\mathbf{A}\sigma}\, d\sigma = \mathbf{A}^{-1}e^{\mathbf{A}t}\Big|_0^t$$

5. Use the fact that

$$e^{\mathbf{A}t}e^{\mathbf{A}\sigma} = e^{\mathbf{A}(t+\sigma)}$$

to show that

$$\cos t \cos \sigma - \sin t \sin \sigma = \cos(t + \sigma)$$

and

$$\sin t \cos \sigma + \sin \sigma \cos t = \sin(t + \sigma)$$

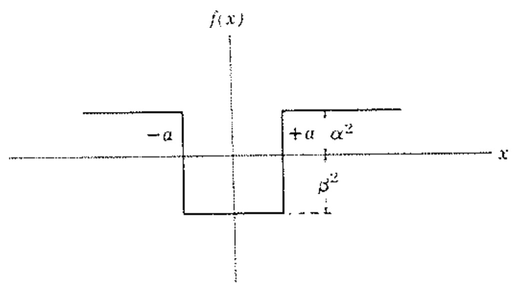
Hint: Use $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

6. One of the simplest problems in quantum mechanics is the one dimensional *potential well problem*. Here the wave function ψ satisfies

$$\frac{d^2\psi(x)}{dx^2} - \alpha^2\psi(x) = 0 \qquad |x| > a$$

$$\frac{d^2\psi(x)}{dx^2} + \beta^2\psi(x) = 0 \qquad |x| \leqslant a$$

Assume that α and β are real and assume that ψ and $d\psi/dx$ are continuous at $x = \pm a$. Determine conditions on α and β such that there exists a solution ψ which goes to zero at $\pm\infty$. Sketch ψ as a function of x (see Figure).



7. Show that the solution of the vector differential equation

$$\begin{bmatrix} \dot y_1(t) \\ \dot y_2(t) \end{bmatrix} = \begin{bmatrix} 0 & k^2(t) \\ -k^2(t) & 0 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

is

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \cos r(t) & \sin r(t) \\ -\sin r(t) & \cos r(t) \end{bmatrix} \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix}$$

where $r(t) = \int_0^t k^2(\sigma)\, d\sigma$. Let x be defined by $x(t) = (1/k(t))y_1(t)$ where we now assume $k(t) > 0$ for all t. Show that y_1 satisfies the second order equation

$$\ddot y_1(t) - (2\dot k(t)/k(t))\dot y_1(t) + k^4(t)y_1(t) = 0$$

and that x satisfies the second order equation

$$\ddot x(t) + [\ddot k/k - (2\dot k^2/k^2) + (k^4)]x(t) = 0$$

Hence if $[\dot k/k]^2$ and $|\ddot k/k|$ are both much smaller than k^4, x is a good approximation to the solution of

$$\ddot z(t) + k^4(t)z(t) = 0$$

This technique, generally referred to as the *WKB method*, is widely used in finding approximate solutions of Schrödinger's equation.

8. Show that every element of $e^{\mathbf{A}t}$ is $\geqslant 0$ for all $t \geqslant 0$ if and only if $a_{ij} \geqslant 0$ for all $i \neq j$. [Hint: for the only if part, write $e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + t^2\mathbf{M}(t)$ and look at the ijth element.] (See Problem 2, Section 3.)
9. (Jacobi Bracket Identity) Show that

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = \mathbf{0}$$

10. The problem of determining, as a function of time, the orientation of a symmetrical rigid body in torque free spinning motion is equivalent to that of finding the transition matrix for the system $\dot{\mathbf{X}}(t) = \mathbf{\Omega}(t)\mathbf{X}(t)$ where

$$\mathbf{\Omega}(t) = \begin{bmatrix} 0 & \omega_3 & \omega_2 \cos \omega t \\ -\omega_3 & 0 & -\omega_2 \sin \omega t \\ -\omega_2 \cos \omega t & \omega_2 \sin \omega t & 0 \end{bmatrix}$$

Show that

$$\mathbf{X}(t) = \begin{bmatrix} \cos \omega t & \sin \omega t & 0 \\ -\sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos \eta t & \sin \eta t & 0 \\ -\sin \eta t & \cos \eta t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where $\eta \cos\theta = \omega_3 - \omega$ and $\eta \sin\theta = \omega_2$

The solution $\mathbf{X}$ is, of course, the direction cosine matrix. (See Problem 2, Section 2 and Problem 4, Section 3.)

11. (Polya) Let $\mathbf{X}$ be a square matrix. Show that the only differentiable solutions of the functional equation

$$\mathbf{X}(t + \sigma) = \mathbf{X}(t)\mathbf{X}(\sigma) : \mathbf{X}(0) = \mathbf{I}$$

are matrix exponentials. (It is also true that the only continuous solutions are matrix exponentials, but the result is harder to prove.)

12. A simple partial differential equation is the one dimensional homogeneous *diffusion equation* which we write as

$$\frac{\partial x(t, z)}{\partial t} = \alpha^2 \frac{\partial^2 x(t, z)}{\partial x^2}$$

Let the boundary conditions be $x(t, 0) = x(t, 1) = 0$. Assume a solution of the form

$$x(t, z) = \sum_{n=1}^{\infty} y_n(t) \sin 2\pi n z$$

Proceeding formally, show that

$$\dot{y}_n = -n^2\alpha^2 y_n$$

This is an infinite set of equations. Find an appropriate "transition matrix."

13. Let f be a piecewise linear function of one variable defined by

$$f(x) = \begin{cases} -x, & |x| \leqslant 1 \\ x - 2, & x > 1 \\ x + 2, & x < -1 \end{cases}$$

Show that the differential equation

$$\ddot{x}(t) + f[\dot{x}(t)] + x(t) = 0$$

admits exactly one periodic solution in addition to the periodic solution $x = 0$.

6. INHOMOGENEOUS LINEAR EQUATIONS

Inhomogeneous linear differential equations of first order take the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t) \tag{IL}$$

Inhomogeneous equations of second and higher order can be reduced to first order form by the same device used before so we need not explicitly consider higher order equations.

Again we approach a new topic via an analogy with linear algebraic equations. Recall that the most general solution of $\mathbf{Ax} = \mathbf{b}$ consists of two parts, the first being one particular solution and the second being an element of the null space of $\mathbf{A}$. Otherwise stated, it is the sum of a particular solution $\mathbf{x}_1$ and a homogeneous solution, $\mathbf{x}_2$. To make the analogy still more vivid, rewrite equation (IL) as

$$[D\mathbf{I} - \mathbf{A}(t)]\mathbf{x}(t) = \mathbf{f}(t), \qquad D = \frac{d}{dt} \tag{IL}$$

The point of departure with linear algebraic equations comes with the introduction of initial conditions. These effectively fix which homogeneous solution we must choose.

The form of the homogeneous solutions of equation (IL) are known. They are $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$. What is a particular solution? In the special case where $\mathbf{A}$ is identically zero we have

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t)$$

which can be integrated to give

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{f}(\sigma)\, d\sigma$$

Now, reduce the equation

$$\dot{x}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t)$$

to the above form by letting $\mathbf{z}(t) = \mathbf{\Phi}(t_0, t)\mathbf{x}(t)$ with $\mathbf{\Phi}$ being the transition matrix for $\mathbf{A}$. This eliminates the dependent variable on the left and gives

$$\dot{\mathbf{z}}(t) = \mathbf{\Phi}^{-1}(t, t_0)\mathbf{f}(t) = \mathbf{\Phi}(t_0, t)\mathbf{f}(t)$$

Hence,

$$\mathbf{z}(t) = \mathbf{z}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

and hence

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\left[\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma\right] = \mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Phi}(t, \sigma)\mathbf{f}(\sigma)\, d\sigma \tag{VC}$$

In the literature, this equation is sometimes called the *variation of constants formula*. It is of major importance in most of what follows. For the sake of easy reference we summarize the above development with a theorem.

Theorem 1. (Variation of Constants Formula). *If* $\Phi(t, t_0)$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$, *then the unique solution of* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t)$; $\mathbf{x}(t_0) = \mathbf{x}_0$ *is given by*

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

Corollary. *The solution of the inhomogeneous linear constant equation* $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{f}(t)$; $\mathbf{x}(0) = \mathbf{x}_0$ *is given by*

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\sigma)}\mathbf{f}(\sigma)\, d\sigma$$

Notice that uniqueness of solutions for equation (IL) is not a problem, for if $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are two solutions of equation (IL) satisfying the same initial data, then

$$[\dot{\mathbf{x}}(t) - \dot{\mathbf{y}}(t)] = \mathbf{A}(t)[\mathbf{x}(t) - \mathbf{y}(t)], \qquad \mathbf{x}(t_0) - \mathbf{y}(t_0) = \mathbf{0}$$

and we have already treated uniqueness for this problem.

Example. Newton's second law for a unit mass is $\ddot{x}(t) = f(t)$. We have previously expressed this as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ f(t) \end{bmatrix}$$

The variation of constants formula gives an integral form

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_1(0) + t x_2(0) \\ x_2(0) \end{bmatrix} + \int_0^t \begin{bmatrix} 1 & (t-\sigma) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ f(\sigma) \end{bmatrix} d\sigma$$

and hence an integral form of Newton's law

$$x(t) = x(0) + t\dot{x}(0) + \int_0^t (t-\sigma) f(\sigma)\, d\sigma$$

Example. (**Satellite problem**). We are now in a position to express the solution of the linearized equations describing the motion of a satellite in a near circular orbit. Using the corollary we have for the equations of Section 2

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} =$$

$$\begin{bmatrix} 4-3 & \sin \omega t/\omega & 0 & 2(1-\cos \omega t)/\omega \\ 3\omega \sin \omega t & \cos \omega t & 1 & 2 \sin \omega t \\ 6(-\omega t + \sin \omega t) & -2(1-\cos \omega t)/\omega & 1 & (-3\omega t + 4 \sin \omega t)/\omega \\ 6\omega(-1 + \cos \omega t) & -2 \sin \omega t & 0 & -3 + 4 \cos \omega t \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \\ x_4(0) \end{bmatrix}$$

$$+ \int_0^t \begin{bmatrix} \sin \omega(t-\sigma)/\omega \\ \cos \omega(t-\sigma) \\ -2(1-\cos \omega(t-\sigma))/\omega \\ -2 \sin \omega(t-\sigma) \end{bmatrix} u_1(\sigma) +$$

$$\begin{bmatrix} 2(1-\cos \omega(t-\sigma))/\omega \\ 2 \sin \omega(t-\sigma) \\ (-3\omega(t-\sigma) + 4 \sin \omega(t-\sigma))/\omega \\ -3 + 4 \cos \omega(t-\sigma) \end{bmatrix} u_2(\sigma)\, d\sigma$$

In the language of elementary differential equation theory, Theorem 1 gives an explicit formula for a *particular integral*. For constant coefficient equations with exponential or sinusoidal forcing functions, a simple technique based on complex variable notation is far more direct. The idea is based on the fact that the derivative of $e^{\alpha t}$ is a scalar multiple of $e^{\alpha t}$ for α real or complex. Hence it is possible to find a particular solution of

$$\dot{x}(t) = Ax(t) + b \exp(\alpha t)$$

having the form $x_0 e^{\alpha t}$ except in certain degenerate cases. In fact, it is easy to verify that a particular integral for the given equation is

$$x(t) = (I\alpha - A)^{-1} b e^{\alpha t}$$

provided the indicated inverse exists. Denote the square root of -1 by i. If $\alpha = i\omega$, then it can be verified by a short calculation that a particular integral for

$$\dot{x}(t) = Ax(t) + b \sin \omega t$$

is

$$x(t) = \mathrm{Re}(Ii\omega - A)^{-1} b \sin t + \mathrm{Im}(Ii\omega - A)^{-1} b \cos t$$

We will return to this in section 16.

Exercises

1. Suppose that u and y are scalars related by equations of the form

$$x^{(n)}(t) + p_{n-1}x^{(n-1)}(t) + \cdots + p_1 x^{(1)}(t) + p_0 x(t) = u(t)$$
$$y(t) = q_{n-1}x^{(n-1)}(t) + q_{n-2}x^{(n-2)}(t) + \cdots + q_1 x^{(1)}(t) + q_0 x(t)$$

with the p_i and q_i constant. Show that if $x(0) = x^{(1)}(0) = \cdots x^{(n-1)}(0) = 0$ then there exists a continuous function w such that

$$y(t) = \int_0^t w(t-\sigma)u(\sigma)\,d\sigma$$

2. Let x be an n vector and let $\mathbf{A}$ be an n by n matrix. Define $\cos \mathbf{A}t$ and $\sin \mathbf{A}\mathbf{t}$ as $(i = \sqrt{-1})$

$$\cos \mathbf{At} = \operatorname{Re} e^{i\mathbf{A}t}$$
$$\sin \mathbf{At} = \operatorname{Im} e^{i\mathbf{A}t}$$

(The convergence proof for $e^{\mathbf{A}}$ works with $\mathbf{A}$ complex). Derive a variation of constants type formula for the second order vector differential equation

$$\ddot{\mathbf{x}}(t) = -\mathbf{A}^2\mathbf{x}(t) + \mathbf{f}(t)$$

Do not assume $\mathbf{A}^{-1}$ exists.

3. Suppose u is a scalar and we have $\mathbf{x}$ given by

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{b}u(t);\ \mathbf{x}(t_0) = \mathbf{x}_0$$

If $u(t)$ is constrained to lie in the convex set $|u(t)| \leqslant 1$ for all t in the interval $t_0 \leqslant t \leqslant T$, then the set of values $\Sigma(T, \mathbf{x}_0, t_0)$ which $\mathbf{x}$ can take on at T is called the *set of reachable states*. Find the set of reachable states for a unit mass with a bounded force,

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = u(t); \qquad |u(t)| \leqslant 1, \quad t_0 = 0, \quad \mathbf{x}_0 = \mathbf{0}$$

Show that in general the set of reachable states is convex. (Recall a set is convex if $\mathbf{x}_1$ and $\mathbf{x}_2$ belonging to the set implies that $(1-\alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2$ also belongs to the set if $0 \leqslant \alpha \leqslant 1$.)
Hint: Use the Neyman-Pearson lemma* in the calculation of the set of reachable states.

* The Neyman-Pearson lemma is discussed in Reference [5].

4. Convert the differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}(t)\mathbf{x}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0$$

into the integral equation

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\sigma)}\mathbf{B}(\sigma)\mathbf{x}(\sigma)\,d\sigma$$

5. The one dimensional *driven wave equation* defined on $0 \leqslant z \leqslant 1$; $t > 0$ is

$$\left(\frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial t^2}\right)x(z,t) = u(z,t); \qquad x(z,0) = x_1(z); \qquad \frac{\partial x}{\partial t}(z,0) = x_2(z)$$

We assume boundary conditions at $z = 0$ and $z = 1$ of the form

$$x(0,t) = x(1,t) = 0$$

and express x and u as

$$x(t,z) = \sum_{n=1}^{\infty} x_n(t) \sin 2\pi nz$$
$$u(t,z) = \sum_{n=1}^{\infty} u_n(t) \sin 2\pi nz$$

Proceeding formally show that the x_n satisfy an infinite set of equations of the form

$$\begin{bmatrix}\dot{x}_n\\ \dot{y}_n\end{bmatrix} = \begin{bmatrix}0 & \sqrt{2\pi}\,n\\ -\sqrt{2\pi}\,n & 0\end{bmatrix}\begin{bmatrix}x_n\\ y_n\end{bmatrix} + \begin{bmatrix}0\\ (\sqrt{2\pi}\,n)^{-1}\end{bmatrix}u_n$$

Find a variation of constants formula.

7. ADJOINT EQUATIONS

Unfortunately the word adjoint has two different meanings in applied mathematics. The classical terminology refers to the transpose of the matrix of cofactors associated with square matrix $\mathbf{A}$ as the adjoint of $\mathbf{A}$, thus leading to the formula $\mathbf{A}^{-1} = \{\text{adjoint }(\mathbf{A})\}/\det \mathbf{A}$. On the other hand, in the theory of linear transformations on an inner product space, the adjoint transformation associated with a given transformation L is defined as that transformation L^* which makes the inner product of $\mathbf{y}$ and $L(\mathbf{x})$ equal to that of $\mathbf{x}$ and $L^*(\mathbf{y})$. (See Section 1.)

The definition of adjoint used in differential equation theory corresponds to the latter idea. Given a linear homogeneous differential equation in $\mathbf{x}$ defined on an inner product space, we will say that a linear homogeneous differential equation in $\mathbf{p}$ defined on the same inner product space is the *adjoint equation* associated with the given equation in $\mathbf{x}$, provided that for any initial data, the inner product of a solution $\mathbf{x}$ and a solution $\mathbf{p}$ is constant. A less restrictive concept of adjointness is discussed in the exercises.

Theorem 1. *The adjoint equation associated with* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is* $\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$.

Proof: Differentiating the inner product $\mathbf{p}'\mathbf{x}$ gives

$$\frac{d}{dt}\mathbf{p}'(t)\mathbf{x}(t) = \dot{\mathbf{p}}'(t)\mathbf{x}(t) + \mathbf{p}'(t)\dot{\mathbf{x}}(t)$$
$$= \mathbf{p}'(t)[\mathbf{A}(t) - \mathbf{A}(t)]\mathbf{x}(t) = 0 \quad \blacksquare$$

The study of variational problems, the general theory of the two point boundary value problems, and problems in the existence and uniqueness of periodic solutions depend on properties of the adjoint equation. In these applications the key property of the adjoint equation is that it propagates the solution of the original equation backwards in time. This is brought out by the following theorem.

Theorem 2. *If* $\Phi(t, t_0)$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$, *then* $\Phi'(t_0, t)$ *is the transition matrix for its adjoint equation,* $\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$.

Proof. Differentiate $\mathbf{I} = \hat{\Phi}(t, t_0)\Phi(t, t_0)$ to get

$$0 = \frac{d}{dt}\mathbf{I} = \frac{d}{dt}[\Phi^{-1}(t, t_0)\Phi(t, t_0)]$$
$$= \left[\frac{d}{dt}\Phi^{-1}(t, t_0)\right]\Phi(t, t_0) + \Phi^{-1}(t, t_0)\frac{d}{dt}\Phi(t, t_0)$$
$$= \left[\frac{d}{dt}\{\Phi^{-1}(t, t_0)\} + \Phi^{-1}(t, t_0)\mathbf{A}(t)\right]\Phi(t, t_0)$$

Since $\Phi(t, t_0)$ is nonsingular, this means

$$\frac{d}{dt}\Phi^{-1}(t, t_0) = -\Phi^{-1}(t, t_0)\mathbf{A}(t)$$

or

$$\frac{d}{dt}\{\Phi^{-1}(t, t_0)\}' = -\mathbf{A}'(t)\{\Phi^{-1}(t, t_0)\}'$$

Since the derivative of the transpose is the transpose of the derivative, this gives the desired result. $\blacksquare$

A differential equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is said to be *self-adjoint* if for all t, $\mathbf{A}(t) = -\mathbf{A}'(t)$. Such systems are found in the study of mechanics; perhaps

the harmonic oscillator

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

is the best known example. The transition matrix for self-adjoint systems has the interesting property of being an orthogonal matrix.*

Theorem 3. *If* Φ *is the transition matrix for a self-adjoint system, then* $\Phi'(t, t_0)\Phi(t, t_0) = \mathbf{I}$ *for all* t *and* t_0.

Proof. To verify that $\Phi(t, t_0)$ is orthogonal if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is self adjoint, we simply observe that since $\mathbf{A}(t) = -\mathbf{A}'(t)$

$$\frac{d}{dt}[\Phi'(t, t_0)\Phi(t, t_0)] = \Phi'(t, t_0)[\mathbf{A}'(t) + \mathbf{A}(t)]\Phi(t, t_0) = \mathbf{0}$$

However, by definition of the transition matrix, $\Phi(t_0, t_0) = \mathbf{I}$ so

$$\Phi'(t_0, t_0)\Phi(t_0, t_0) = \mathbf{I}$$

and the result is immediate. $\blacksquare$

Exercises

1. Consider the mapping of $C_*^n[t_0, t_1]$ into $C_*^n[t_0, t_1]$ defined by

$$\mathbf{y}(t) = \int_{t_0}^{t} \Phi(t, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

   Compute its adjoint. (Compare with examples in Section 1.)

2. Consider a mapping of $E^n \times C_*^n[t_0, t_1]$ into $E^n \times C_*^n[t_0, t_1]$ defined by the pair of equations

$$\mathbf{x}(t_1) = \Phi(t_1, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t_1} \Phi(t_1, \sigma)\mathbf{f}(\sigma)\, d\sigma$$
$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \Phi(t, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

   This mapping is generated by $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t)$. Write down a similar mapping for $\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t) + \mathbf{g}(t)$. Show that it is the adjoint linear map in the sense given in Section 1, provided we regard the inner product on $E^n \times C^n[t_0, t_1]$ as the sum of the inner products on E^n and $C_*^n[t_0, t_1]$.

3. Show that if $A(t) = -\mathbf{A}'(-t)$ for all t, then $\Phi(t, t_0)$ and its inverse have the same eigenvalues.

* A square matrix is said to be *orthogonal* if $\mathbf{A}'\mathbf{A} = \mathbf{I}$.

4. As we have defined it here, the concept of self adjointness is not coordinate-free in the sense that $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ may not be self adjoint whereas the change of variable $\mathbf{z} = \mathbf{P}\mathbf{x}$ will give a self adjoint equation. Give necessary and sufficient conditions on $\mathbf{A}$ for there to exist a constant matrix $\mathbf{P}$ such that $\dot{\mathbf{z}}(t) = \mathbf{P}\mathbf{A}(t)\mathbf{P}^{-1}\mathbf{z}(t)$ is self adjoint.

8. PERIODIC HOMOGENEOUS EQUATIONS: REDUCIBILITY

Perhaps the simplest class of time-varying differential equations are those for which the time variations are periodic. Even with this assumption in force there is a significant departure from the constant case, and explicit solutions are generally impossible. Periodicity of the $\mathbf{A}$ matrix however, implies certain structural properties of the associated transition matrix, and it is this aspect that we want to focus on here. In applications, periodic linear equations arise very frequently as the result of linearizing a nonlinear system about a periodic solution. The main result here (Theorem 1) is generally attributed to either Floquet or Liapunov or both.

A function is said to be *periodic of period T* if $f(t + T) = f(t)$ for all t. In particular, f equals a constant (even 0) is a periodic function. Clearly any function which is periodic of period T is also periodic of period $2T$, etc.; and thus without some qualification "the period" of a function is ambiguous. When we use period below, we mean *any* period.

Consider the periodic, homogeneous system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t);\ \mathbf{A}(t + T) = \mathbf{A}(t) \qquad \text{(LP)}$$

Let $\Phi(t, t_0)$ denote its transition matrix. Since any nonsingular matrix $\mathbf{C}$ can be expressed as an exponential* $\mathbf{C} = e^{\mathbf{R}}$, let us write

$$\Phi(T, 0) = e^{\mathbf{R}T} \quad \text{definition of } \mathbf{R}$$

and define $\mathbf{P}(t)$ via

$$\mathbf{P}^{-1}(t) = \Phi(t, 0)e^{-\mathbf{R}t}$$

Notice that $\mathbf{P}^{-1}(t + T)$ can be expressed as

$$\begin{aligned}\mathbf{P}^{-1}(t + T) &= \Phi(t + T, 0)e^{-\mathbf{R}(t+T)}\\ &= \Phi(t + T, T)\Phi(T, 0)e^{-\mathbf{R}T}e^{-\mathbf{R}t}\\ &= \Phi(t + T, T)e^{-\mathbf{R}t}\\ &= \Phi(t, 0)e^{-\mathbf{R}t}\\ &= \mathbf{P}^{-1}(t)\end{aligned}$$

* See, for example, Coddington and Levinson [20], Chapter 3. If $\mathbf{C} = \mathbf{M}^2$ then $\mathbf{R}$ can be taken real; see Erugin [22]. Special cases are easily proved using the Jordan normal form; see Section 12 and exercises there.

In going from the first line to the second we use the composition rule $\Phi(t, t_0) = \Phi(t, \sigma)\Phi(\sigma, t_0)$. The third line follows from the second using the definition of $\mathbf{R}$. The fourth is obtained using the fact that since the system is periodic of period T, $\Phi(t + T, t_0 + T) = \Phi(t, t_0)$. Hence we have $\mathbf{P}^{-1}(t)e^{\mathbf{R}t} = \Phi(t, 0)$ and $\Phi(0, \sigma) = \Phi^{-1}(\sigma, 0) = e^{-\mathbf{R}\sigma}\mathbf{P}(\sigma)$. Putting this together gives the following theorem (see Figure 1).

Theorem 1. (Floquet-Liapunov) *If* $\mathbf{A}(t + T) = \mathbf{A}(t)$, *then the associated transition matrix can be written as*

$$\Phi(t, t_0) = \mathbf{P}^{-1}(t)e^{\mathbf{R}(t-t_0)}\mathbf{P}(t_0); \qquad \mathbf{P}(t) = \mathbf{P}(t + T)$$

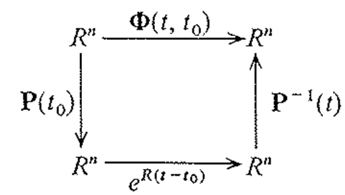


Figure 1 *The Floquet-Liapunov decomposition for periodic equations.*

We now consider a topic of broader interest embracing not only (LP), but nonperiodic equations as well. To begin, notice that in the development above $\mathbf{P}(t)$ is nonsingular so that the definition

$$\mathbf{P}(t)\mathbf{x}(t) = \mathbf{z}(t) \qquad \text{or} \qquad \mathbf{x}(t) = \mathbf{P}^{-1}(t)\mathbf{z}(t)$$

makes sense. The pleasant surprise is that $\mathbf{z}$ satisfies the *constant* equation

$$\dot{\mathbf{z}}(t) = \mathbf{R}\mathbf{z}(t)$$

This is an immediate consequence of Theorem 3 of Section 4.

The existence of special types of transformations which render a time varying system constant has been studied in detail by Liapunov and his followers. Their definitions are motivated by stability considerations, however, a discussion of them here is not completely out of place.

A transformation $\mathbf{z}(t) = \mathbf{L}(t)\mathbf{x}(t)$ is called a *Liapunov transformation* if

(i) $\mathbf{L}$ has a continuous derivative on the interval $(-\infty, \infty)$
(ii) $\mathbf{L}$ and $\dot{\mathbf{L}}$ are bounded* on the interval $(-\infty, \infty)$
(iii) there exists a constant m such that $0 < m \leqslant |\det \mathbf{L}(t)|$, for all t.

Linear differential equations which can be transformed into equations with

* The reader is cautioned that boundedness on $(-\infty, \infty)$ for scalars means that there exists a number M (not infinity!) such that for all $-\infty < t < \infty$ the absolute value of the function is less than or equal to M. A matrix is said to be bounded if each of its elements are.

constant coefficients by means of a Liapunov transformation are called *reducible*.

In light of our previous calculations, we see that Theorem 1 implies the following result.

Theorem 2. *The equation* (LP) *is reducible in the sense of Liapunov.*

Example. Let $e^{\mathbf{A}t}$ be periodic of period T and let $\mathbf{x}$ satisfy

$$\dot{\mathbf{x}}(t) = e^{\mathbf{A}t}\mathbf{B}e^{-\mathbf{A}t}\mathbf{x}(t)$$

If $\mathbf{z}$ is defined by

$$\mathbf{z}(t) = e^{-\mathbf{A}t}\mathbf{x}(t)$$

then

$$\dot{\mathbf{z}}(t) = (\mathbf{B} - \mathbf{A})\mathbf{z}(t)$$

Hence

$$\mathbf{z}(t) = e^{(\mathbf{B}-\mathbf{A})(t-t_0)}\mathbf{z}(t_0)$$

and

$$\mathbf{x}(t) = e^{\mathbf{A}t}e^{(\mathbf{B}-\mathbf{A})(t-t_0)}e^{-\mathbf{A}t_0}\mathbf{x}(t_0)$$

which is the decomposition called for by theorem 1.

Exercises

1. Show that if $\mathbf{z} = \mathbf{L}_1\mathbf{x}$ and $\mathbf{z} = \mathbf{L}_2\mathbf{x}$ are Liapunov transformations then $\mathbf{z} = \mathbf{L}_1\mathbf{L}_2\mathbf{x}$ is also.
2. Show that $\mathbf{z} = e^{\mathbf{A}t}\mathbf{x}$ is a Liapunov transformation if $\mathbf{A} = -\mathbf{A}'$.
3. Give necessary and sufficient conditions on $\mathbf{A}$ for $e^{\mathbf{A}t}$ to be periodic.
4. Show that there exists a Liapunov transformation which transforms the constant system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

into the constant system

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix}\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$

Generalize this result to n-dimensional equations.

5. Problem 10 of Section 5 gives a transition matrix for a periodic system. Find a decomposition of it in the form called by Theorem 1.
6. Show that if $\mathbf{A}$ is periodic of period T and that if for all t one has $\mathbf{A}(t) = -\mathbf{A}'(-t)$, then the eigenvalues of $\mathbf{\Phi}(t_0 + T, t_0)$ occur in reciprocal

pairs. That is, if ρ is an eigenvalue then so is ρ^{-1}. (See Problem 3, Section 7.)

7. Assume that the eigenvalues of $e^{\mathbf{A}}$ are distinct and that $\mathbf{A} = -\mathbf{A}'$. Suppose $\mathbf{B}$ is symmetric. Show that there exists $\varepsilon_0 > 0$ such that the eigenvalues of $e^{\mathbf{A}+\varepsilon\mathbf{B}}e^{\mathbf{A}-\varepsilon\mathbf{B}}$ all have magnitude 1 for $|\varepsilon| < \varepsilon_0$. (Use the results of Problem 6.)
8. Assume that $\mathbf{A}$ is periodic of period T and assume that $a_{ij}(t) \geqslant 0$ for all $i \neq j$ and all t. Show that $\mathbf{\Phi}(t_0 + T, t_0)$ has a real eigenvalue ρ_0 which is as large in magnitude as the magnitude of any other eigenvalue. (Hint: Use a version of the Perron-Frobenius theorem* and the results of Exercise 2 of Section 3. ($\mathbf{\Phi}$ is not necessarily irreducible.)
9. Let

$$\mathbf{A}(t) = \begin{bmatrix} -1 + \cos t & 0 \\ 0 & -2 + \cos t \end{bmatrix}$$

(a) Calculate the transition matrix $\mathbf{\Phi}(t, 0)$ for the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$$

(b) Since this system is periodic with period 2π, the transition matrix can be written

$$\mathbf{\Phi}(t, \sigma) = \mathbf{P}^{-1}(t)e^{\mathbf{R}(t-\sigma)}\mathbf{P}(\sigma)$$

Do so.

(c) Are all solutions of this system bounded? Do they decay to zero as $t \to \infty$?

10. Compute a real *logarithm* (any solution of $\mathbf{C} = e^{\mathbf{R}}$) for the following choices of $\mathbf{C}$.

(i) $\mathbf{C} = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}$ (ii) $\mathbf{C} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

(iii) $\mathbf{C} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ (iv) $\mathbf{C} = \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix}$

11. Show that the general solution of the periodic second order equation

$$\ddot{x}(t) + \eta(t)\dot{x}(t) + \gamma(t)x(t) = 0$$

$$\eta(t + T) = \eta(t); \qquad \gamma(t + T) = \gamma(t)$$

can be written as

$$x(t) = Ae^{\lambda_1 t}p_1(t) + Be^{\lambda_2 t}p_2(t)$$

* The Perron-Frobenius theorem is discussed in Reference [57].

with λ_1 and λ_2 real and p_1 and p_2 periodic of period T or else as

$$x(t) = Ae^{\lambda t}p_1(t) + Bt\, e^{\lambda t}p_2(t)$$

with λ real and p_1 and p_2 periodic of period T.

9. PERIODIC LINEAR INHOMOGENEOUS EQUATIONS

We are now interested in the structure of the solutions of inhomogeneous periodic equations of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t); \qquad \mathbf{A}(t+T) = \mathbf{A}(t); \qquad \mathbf{f}(t+T) = \mathbf{f}(t) \qquad \text{(LIP)}$$

When specialized to the case where $\mathbf{A}$ is constant and $\mathbf{f}$ is a constant vector times sin ωt these ideas play a key role in many aspects of system theory. The results of this section on the structure of solutions complement, but do not subsume or require the knowledge of, those of Floquet and Liapunov discussed in Section 8. The technique used to establish the results here will be repeated many times in what follows and bears some comment. It consists of two steps. The first is to convert the differential equation problem into a linear algebra problem using the variation of constants formula. The second is to invoke the basic results on the solubility of $\mathbf{Ax} = \mathbf{b}$. This sequence with a little juggling and interpretation accounts for a great many of the results in this book.

Theorem 1. *Let $\mathbf{\Phi}$ be the transition matrix associated with* $\mathbf{A}$. *The solution of* (LIP) *passing through $\mathbf{x}_0$ at time t_0 can be written as*

$$\mathbf{x}(t) = \mathbf{x}_p(t) + \mathbf{\Phi}(t, t_0)[\mathbf{x}_0 - \mathbf{x}_p(t_0)]$$

with $\mathbf{x}_p$ periodic of period T if and only if

$$\int_{t_0}^{t_0+T} \mathbf{p}'(\sigma)\mathbf{f}(\sigma)\, d\sigma = 0$$

for every n-vector $\mathbf{p}$ *which is periodic of period T and which satisfies the adjoint equation*

$$\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$$

Proof. Let us refer to the condition that the integral in the theorem statement vanishes for all periodic solutions of period T of the adjoint equation as the "orthogonality condition." We now show that the orthogonality condition holds if and only if the vector

$$\mathbf{v} = \int_{t_0}^{t_0+T} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

is perpendicular to every vector in the null space of $[\mathbf{\Phi}'(t_0, t_0+T) - \mathbf{I}]$. To establish this, recall that from Theorem 7.2 the solution of the adjoint equation which passes through $\mathbf{p}_0$ at $t = t_0$ can be written as $\mathbf{p}(t) = \mathbf{\Phi}'(t_0, t)\mathbf{p}_0$. If this is periodic, then $\mathbf{p}(t+T) = \mathbf{p}(t)$ and hence $\mathbf{p}_0$ satisfies $[\mathbf{\Phi}'(t_0, t_0+T) - \mathbf{I}]\mathbf{p}_0 = \mathbf{0}$. Vectors $\mathbf{p}_0$ in the null space of $[\mathbf{\Phi}'(t_0, t_0+T) - \mathbf{I}]$ give rise to periodic solutions. Therefore

$$\int_{t_0}^{t_0+T} \mathbf{p}_0' \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma = \int_{t_0}^{t_0+T} \mathbf{p}'(\sigma)\mathbf{f}(\sigma)\, d\sigma$$

and we see that indeed, the orthogonality condition is necessary and sufficient for $\mathbf{v}$ to be perpendicular to every vector in the null space of $[\mathbf{\Phi}'(t_0, t_0+T) - \mathbf{I}]$.

As an immediate consequence of this result, we claim that the equation

$$[\mathbf{\Phi}(t_0, t_0+T) - \mathbf{I}]\mathbf{x}_1 = \int_{t_0}^{t_0+T} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

has a solution if and only if the orthogonality condition is satisfied. Now bring in the variation of constants formula in the form

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\left[\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma\right]$$

If $\mathbf{x}(t_0+T)$ is to equal $\mathbf{x}(t_0)$ for any value of $\mathbf{x}(t_0)$, say $\mathbf{x}(t_0) = \mathbf{x}_1$, then $\mathbf{x}_1$ will satisfy

$$\mathbf{x}_1 = \mathbf{\Phi}(t_0+T, t_0)\left[\mathbf{x}_1 + \int_{t_0}^{t_0+T} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma\right]$$

Using the identity $\mathbf{\Phi}^{-1}(t_0+T, t_0) = \mathbf{\Phi}(t_0, t_0+T)$ together with a premultiplication by $\mathbf{\Phi}(t_0, t_0+T)$ gives the equivalent equation

$$[\mathbf{\Phi}(t_0, t_0+T) - \mathbf{I}]\mathbf{x}_1 = \int_{t_0}^{t_0+T} \mathbf{\Phi}(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

This has a solution if and only if the orthogonality condition is satisfied. At this point we can conclude that there exists an initial value $\mathbf{x}_1$ such that the solution passing through $\mathbf{x}_1$ at $t = t_0$ is periodic if and only if the orthogonality condition is satisfied.

The rest of the proof is easy. If there exists $\mathbf{x}_1$ such that the solution passing through $\mathbf{x}_1$ at $t = t_0$ is periodic, then let $\mathbf{x}_p$ be this solution. The solution passing through $\mathbf{x}_0$ at $t = t_0$ is

$$\mathbf{x}(t) = \mathbf{x}_p(t) + \mathbf{\Phi}(t, t_0)(\mathbf{x}_0 - \mathbf{x}_1)$$

and hence we have shown that if the orthogonality condition holds we can write the solution as the sum of a periodic part and a homogeneous solution.

On the other hand suppose the solution can be written as the sum of a periodic part and a homogeneous part. Then there exists a particular value of $\mathbf{x}_0$, namely $\mathbf{x}_0 = \mathbf{x}_1(t_0)$, such that the entire solution is periodic and hence the orthogonality condition must hold. ▮

There are a number of special cases of this theorem which are much simpler to interpret and use. We give two of these here as corollaries.

Corollary 1. *If the homogeneous equation* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *has no nontrivial periodic solution of period* T, *then the solution of* (LIP) *which passes through* $\mathbf{x}_0$ *at* $t = t_0$ *can be* uniquely *decomposed as*

$$\mathbf{x}(t) = \mathbf{x}_p(t) + \Phi(t, t_0)[\mathbf{x}_0 - \mathbf{x}_p(t_0)]$$

with $\mathbf{x}_p$ *periodic. Moreover,* $\mathbf{x}_p$ *is given by*

$$\mathbf{x}_p(t) = \Phi(t, t_0)[\Phi(t_0, t_0 + T) - \mathbf{I}]^{-1} \int_{t_0}^{t_0+T} \Phi(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma + \int_{t_0}^{t} \Phi(t, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

Proof. If the homogeneous equation has no periodic solution of period T, then $[\Phi(t_0 + T, t_0) - \mathbf{I}]\mathbf{x}_0 = \mathbf{0}$ has no nontrivial solution. Therefore $\det[\Phi(t_0 + T, t_0) - \mathbf{I}] \neq 0$. Using the identity valid for square matrices, $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$, we see that

$$\det[\Phi(t_0, t_0 + T) - \mathbf{I}] = \pm \det \Phi(t_0, t_0 + T)\det[\Phi(t_0 + T, t_0) - \mathbf{I}] \neq 0$$

As a result, $[\Phi(t_0, t_0 + T) - \mathbf{I}]\mathbf{x}_1 = \mathbf{v}$ can be solved uniquely for $\mathbf{x}_1$. Following through the algebra of the proof of the theorem gives

$$\mathbf{x}_1 = [\Phi(t_0, t_0 + T) - \mathbf{I}]^{-1} \int_{t_0}^{t_0+T} \Phi(t_0, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

and the representation given in the corollary follows immediately. ▮

Corollary 2. *If* $\mathbf{A}$ *is constant and has no eigenvalues with zero real parts, then the solution of* (LIP) *which passes through* $\mathbf{x}_0$ *at* $t = 0$ *can be expressed uniquely as*

$$\mathbf{x}(t) = \mathbf{x}_p(t) + e^{\mathbf{A}t}[\mathbf{x}_0 - \mathbf{x}_p(0)]$$

with $\mathbf{x}_p$ *periodic. Moreover,* $\mathbf{x}_p$ *is given by*

$$\mathbf{x}_p(t) = e^{\mathbf{A}t}[e^{-\mathbf{A}T} - \mathbf{I}]^{-1} \int_0^T e^{-\mathbf{A}\sigma}\mathbf{f}(\sigma)\, d\sigma + \int_0^t e^{\mathbf{A}(t-\sigma)}\mathbf{f}(\sigma)\, d\sigma$$

Proof. Identical to the previous proof with $\Phi(t, 0) = e^{\mathbf{A}t}$. ▮

Example. Consider a harmonic oscillator with a periodic drive.

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ f(t) \end{bmatrix}; \qquad f(t + 2\pi) = f(t)$$

What additional restrictions will be needed on f to insure the existence of a periodic solution? Since this system is self-adjoint we have

$$\begin{bmatrix} \dot{p}_1(t) \\ \dot{p}_2(t) \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} p_1(t) \\ p_2(t) \end{bmatrix}$$

and the orthogonality condition becomes

$$\int_0^{2\pi} \begin{bmatrix} \cos\sigma & \sin\sigma \\ -\sin\sigma & \cos\sigma \end{bmatrix} \begin{bmatrix} 0 \\ f(\sigma) \end{bmatrix} d\sigma = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This requires

$$\int_0^{2\pi} f(\sigma)\cos\sigma\, d\sigma = \int_0^{2\pi} f(\sigma)\sin\sigma\, d\sigma = 0$$

Hence we have a periodic solution if the Fourier series of f has no first harmonic.

Exercises

1. Consider the equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}$$

with $\mathbf{b}$ constant. Under what circumstances does there exist a constant solution for suitable $\mathbf{x}_0$? (The answer is not just $\det \mathbf{A} \neq 0$.)

2. Show that the equation

$$(D^2 + 1)(D + 1)x(t) = \sin nt; \left(D = \frac{d}{dt}\right)$$

has:

(i) no periodic solution for $n = 1$
(ii) a unique periodic solution of period $2\pi/n$ for $n = 2, 3, \ldots,$ and
(iii) a two parameter family of periodic solutions for $n = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$

3. Determine whether each of the following systems of equations has periodic solutions of period 2π, and if so, whether they are unique:

(a) $\dot{\mathbf{x}}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \cos t$

(b) $\dot{\mathbf{x}}(t) - \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \sin t$

(c) $\dot{\mathbf{x}}(t) = \begin{bmatrix} -5 & 3 \\ -6 & 4 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \sin t + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cos t$

(d) $\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

4. Consider

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u, \qquad y = \mathbf{c}'\mathbf{x}$$

with u a square wave

$$u(t) = \operatorname{sgn}(\sin t)$$

Decompose the solution in the form called for by Theorem 1 under a suitable hypothesis on $\mathbf{A}$.

5. Interpret the periodic equation (LIP) as a mapping of $E^n \times C_*^n[t_0, t_0 + T]$ into E^n according to

$$\mathbf{x}(t_0 + T) = \mathbf{\Phi}(t_0 + T, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t_0+T} \mathbf{\Phi}(t_0 + T, \sigma)\mathbf{f}(\sigma)\, d\sigma$$

Compute the adjoint of this mapping. Deduce Theorem 1 from the general principle which states $L(\mathbf{u}) = \mathbf{b}$ has a solution if and only if b is orthogonal to the null space of the adjoint of L. (See Exercises 1 and 2, Section 7.)

6. Let $\mathbf{A}$ be periodic of period T_1 and let $\mathbf{f}$ be periodic of period T_2. Show that if all solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ go to zero as t goes to infinity then there exists a vector $\mathbf{g}$ which is periodic of period T_2 and a matrix $\mathbf{P}$ which is periodic of period T_1 such that the general solution of

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(\mathbf{t}) + \mathbf{f}(t)$$

is expressible as

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\xi_0 + \mathbf{P}(t)\mathbf{g}(t)$$

10. SOME BASIC RESULTS OF ASYMPTOTIC BEHAVIOR

For linear time invariant systems and systems which are "close" to them in an appropriate sense, stability questions can be reduced to algebraic questions which, in turn, are easily resolved. In this section we give some results on linear time invariant systems and linear periodic systems. The conditions given by Theorems 1 and 3 are stated as sufficient conditions. They are also necessary but the proof is omitted. (See Exercise 4, Section 12.)

Theorem 1. *All solutions of the time invariant equation*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

approach zero as t approaches infinity if all the zeros of $\det(\mathbf{I}s - \mathbf{A})$ *lie in the half-plane* $\mathrm{Re}[s] < 0$. *All solutions are bounded for positive t if the zeros of* $\det(\mathbf{I}s - \mathbf{A})$ *lie in the half-plane* $\mathrm{Re}[s] \leqslant 0$ *and for* s_i *a zero with vanishing real part and multiplicity* $\sigma_i > 1$,*

$$[(s - s_i)^{\sigma_i}(\mathbf{I}s - \mathbf{A})^{-1}]^{(\sigma_i - 1 - k)}|_{s = s_i} = \mathbf{0}, \qquad k = 1, \ldots, \sigma_i - 1$$

Proof. Recall from Theorem 3, Section 5 that

$$e^{\mathbf{A}t} = \sum_{i=1}^{m} \sum_{k=0}^{\sigma_i + 1} \frac{1}{k!} \frac{1}{(\sigma_i - 1 - k)!} [(s - s_i)^{\sigma_i}(\mathbf{I}s - \mathbf{A})^{-1}]^{(\sigma_i - 1 - k)} t^k e^{st} \Big|_{s = s_i}$$

From this we see that $e^{\mathbf{A}t}$ is a sum of constant terms multiplied by $t^k e^{s_i t}$. Clearly if all the s_i have negative real parts then $e^{\mathbf{A}t}$ approaches zero since for any $\varepsilon > 0$

$$\lim_{t \to \infty} t^n e^{-\varepsilon t} = 0, \qquad n = 0, 1, 2, \ldots$$

On the other hand, if any of the eigenvalues of $\mathbf{A}$ have zero real parts then there is a term $\mathbf{A}_i e^{s_i t}$ which does not go to zero. Clearly then all solutions can be bounded even though one or more of the eigenvalues have zero real parts since convergence to zero is not required. On the other hand, if $\mathrm{Re}[s_i] = 0$ then $t^k e^{s_i t}$ is not bounded for $k \geqslant 1$ and hence we must require that the coefficients of these terms in the equation for $e^{\mathbf{A}t}$ vanish. But this is just what the residue condition demands.

To be more specific about a criterion for convergence to $\mathbf{0}$ we need to know under what circumstances the zeros of $\det(\mathbf{I}s - \mathbf{A})$ lie in the half-plane $\mathrm{Re}[s] < 0$. One set of necessary and sufficient conditions is given by the following theorem of Hurwitz.

Theorem 2. (Hurwitz) *A polynomial* $p(s) = p_n s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0$, *with real coefficients and* p_n *positive has all its zeros in the half-plane* $\mathrm{Re}[s] < 0$ *if and only if the n determinants*

$$\Delta_1 = p_{n-1};\ \Delta_2 = \det\begin{bmatrix} p_{n-1} & p_n \\ p_{n-3} & p_{n-2} \end{bmatrix} \Delta_3 = \det\begin{bmatrix} p_{n-1} & p_n & 0 \\ p_{n-3} & p_{n-2} & p_{n-1} \\ p_{n-5} & p_{n-4} & p_{n-3} \end{bmatrix}$$

$$\cdots \Delta_n = \det\begin{bmatrix} p_{n-1} & p_n & 0 & 0 & \cdots & 0 \\ p_{n-3} & p_{n-2} & p_{n-1} & p_n & \cdots & 0 \\ p_{n-5} & p_{n-4} & p_{n-3} & p_{n-2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & & p_0 \end{bmatrix} \tag{5}$$

are all positive.

We will not give a proof. See, for example, Kaplan [43].

* $[\ ]^{(p)}$ denotes pth derivative with respect to s.

These conditions can also be stated in terms of the positive definiteness of a symmetric matrix. In fact Hermite did so (as a special case of a somewhat more general problem 45 years before the paper of Hurwitz appeared).

If **A** is periodic then we know that the associated transition matrix can be expressed as

$$\mathbf{\Phi}(t, t_0) = \mathbf{P}^{-1}(t)e^{\mathbf{R}(t-t_0)}\mathbf{P}(t_0)$$

with **R** constant and **P** periodic. Thus the asymptotic behavior of solutions is determined by the eigenvalues of **R**. However $T\mathbf{R}$ is a logarithm of $\mathbf{\Phi}(t_0+T, t_0)$ where T is the period of **A** and as such, difficult to calculate. Moreover, for stability purposes it is entirely unnecessary to do so for if $T\mathbf{R}$ has as its eigenvalues λ_i then e^{λ_i} are the eigenvalues of $\mathbf{\Phi}(T + t_0, t_0)$. This gives immediately the following stability theorem for period systems.

Theorem 3. *All solutions of the periodic equation*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t); \qquad \mathbf{A}(t + T) = \mathbf{A}(t)$$

approach zero as t approaches infinity if the zeros $\det[\mathbf{I}s - \mathbf{\Phi}(t_0 + T, t_0)]$ *lie in the disk* $|s| < 1$. *All solutions are bounded for positive t if the zeros of* $\det[\mathbf{I}s - \mathbf{\Phi}(t_0 + T, t_0)]$ *lie in the disk* $|s| \leqslant 1$ *and for* s_i *a zero of magnitude* 1 *and multiplicity* $\sigma_i > 1$,

$$[(s - s_i)^{\sigma_i}\{\mathbf{I}s - \mathbf{\Phi}(t_0 + T, t_0)\}^{-1}]^{(\sigma_i - 1 - k)}|_{s=s_i} = \mathbf{0}, \qquad k = 1, 2, \ldots, \sigma_i - 1$$

To complete the picture we should give an algebraic criterion which is necessary and sufficient for the zeros of a polynomial to have all its zeros in the disk $|s| < 1$. Instead, we observe that if p is given by

$$p(s) = p_n s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0$$

then p has all its zeros in the disk $|s| < 1$ if and only if q as given by

$$q(s) = p\left[\frac{(1 + s)}{(1 - s)}\right] (1 - s)^n$$

has all its zeros in the half-plane $\mathrm{Re}[s] < 0$. Thus the Hurwitz criterion can be used here as well.

Examples. The zeros of real monic polynomials of degree 1, 2, 3, and 4 lie in the half-plane $\mathrm{Re}\, s < 0$ if and only if the following inequalities hold.

(a) $p(s) = s + a$; $\quad a > 0$
(b) $p(s) = s^2 + as + b$; $\quad a > 0, b > 0$
(c) $p(s) = s^3 + as^2 + bs + c$; $\quad a > 0, b > 0, c > 0, ab - c >$
(d) $p(s) = s^4 + as^3 + bs^2 + cs + d$; $\quad a > 0, b > 0, c > 0, d > 0$
$abc - c^2 - a^2 d > 0$

The zeros of real monic polynomials of degree 1, 2 and 3 lie inside the unit disk if and only if the following inequalities hold:

(a) $p(s) = s + a$; $\quad |a| < 1$
(b) $p(s) = s^2 + as + b$; $\quad |b| < 1, 1 + a + b > 0, 1 - a + b > 0$
(c) $p(s) = s^3 + as^2 + bs + c$; $\quad |c| < 1, 1 - a + b - c > 0, 3 - a - b + 3c > 0, 3 + a - b - 3c > 0, 1 + a + b + c > 0, (3 - b)^2 - (a - 3c)^2 > (1 + b)^2 - (a + c)^2$

Exercises

1. Find $f_0(n)$ for $n = 2, 3, \ldots,$ such that all solutions of
$$(D + 1)^n x(t) + fx(t) = 0$$
approach zero for all constant f in the interval $-1 \leqslant f \leqslant f_0(n)$.
Answer: $f_0(n) = (\sec \pi/n)^n$
2. Let $p(s)$ be a polynomial. Show that all the zeros of p are real and negative if and only if $p(s^2) + sp'(s^2)$ has all its zeros in the half-plane $\mathrm{Re}\, s < 0$. [Here $p' = (d/ds)p$.]
3. Let η and γ be periodic of period T. Show that all solutions of
$$\ddot{x}(t) + \eta(t)\dot{x}(t) + \gamma(t)x(t) = 0$$
are bounded for $t > 0$ if (i) the average value of η over one period is nonnegative and (ii) the transition matrix for
$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\gamma(t) & -\eta(t)\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$
has complex eigenvalues when evaluated at $t_0 + T, t_0$. Use this result and the result of Problem 6, Section 8 to show that there exists an $\varepsilon > 0$ such that all solutions of the given second order equation are bounded for $t > 0$ if the average value of η is nonnegative, $|\eta(t)| < \varepsilon$ for all t, and $\eta = 1$.
4. Consider the equation $\ddot{x}(t) + a^2(t)x(t) = 0$ with a positive. Let
$$\int_0^t a(\rho)\, d\rho = \sigma(t)$$
Show that
$$\frac{d^2x(\sigma)}{d\sigma^2} + a_1(\sigma)\frac{dx(\sigma)}{d\sigma} + x(\sigma) = 0$$
with a_1 suitably chosen. Use this and the results of the previous theorem to obtain a boundedness result for the original equation.

5. Show that the polynomial $p_n s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0$ has all its zeros in Re $s < 0$ if and only if the polynomial $p_0 s^n + p_1 s^{n-1} + \cdots + p_{n-1}s + p_n$ has all its zeros in Re $s < 0$. Use this to generate alternative forms of the Hurwitz conditions.
6. Let **A** and **B** be given by

$$\mathbf{A} = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$$

It is clear that $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ admits unbounded solutions. It is easy to verify that $e^{\mathbf{B}t}$ is a Liapunov transformation so that if **z** is related to **x** by $\mathbf{z}(t) = e^{\mathbf{B}t}\mathbf{x}(t)$ and **x** is unbounded, then **z** is also. The differential equation for **z** is

$$\dot{\mathbf{z}}(t) = [e^{\mathbf{B}t}\mathbf{A}e^{-\mathbf{B}t} + \mathbf{B}]\mathbf{z}(t)$$

Show that the eigenvalues of $e^{\mathbf{B}t}\mathbf{A}e^{-\mathbf{B}t} + \mathbf{B}$ are independent of t and have negative real parts. This shows that it is not possible to conclude that all solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ will be bounded just because the eigenvalues lie in Re $s < 0$. (Vinogradov, Markus-Yamabe, Rosenbrock.)
7. Closely related to first-order vector difference equations are vector difference equations of the form

$$\mathbf{x}_n = \mathbf{A}\mathbf{x}_{n-1}; \qquad \mathbf{x}_n = \mathbf{x}(n)$$

(a) Find the general solution of this equation.
(b) Show that all solutions of

$$\mathbf{x}_n = \mathbf{A}\mathbf{x}_{n-1}$$

tend to zero if all the eigenvalues of **A** have magnitude less than 1.
8. Let a be a Gauss-Markov process such that the mean of $a(t)$ is 0 and the variance of $a(t)$ is σ^2. Suppose the covariance, $E\{a(t)a(t+\tau)\}$ equals $\sigma^2 e^{-k|\tau|}$. Show that in the limit as t goes to infinity the expected value of the solution of $\dot{x}(t) = [a(t) - 1]x(t)$ goes to zero if $\sigma^2/k < 1$.

11. LINEAR MATRIX EQUATIONS

We found in Section 3 that the vector equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ could be most profitably studied by considering first the matrix equation defining the transition matrix. This interplay between vector equations and matrix equations reoccurs often. As it turns out, the most general linear matrix equation which arises in this book is

$$\dot{\mathbf{X}}(t) = \mathbf{A}_1(t)\mathbf{X}(t) + \mathbf{X}(t)\mathbf{A}_2(t) + \mathbf{F}(t) \qquad \text{(LM)}$$

A result completely analogous to the variation of constants formula holds for matrix equations of this form.

Theorem 1. (Matrix Variation of Constants Formula) *If* $\mathbf{\Phi}_1(t, t_0)$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}_1(t)\mathbf{x}(t)$ *and* $\mathbf{\Phi}_2(t, t_0)$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}_2'(t)\mathbf{x}(t)$, *then the solution of equation* (LM) *with the initial value* $\mathbf{X}(t_0)$ *is given by*

$$\mathbf{X}(t) = \mathbf{\Phi}_1(t, t_0)\mathbf{X}(t_0)\mathbf{\Phi}_2'(t, t_0) + \int_{t_0}^{t} \mathbf{\Phi}_1(t, \sigma)\mathbf{F}(\sigma)\mathbf{\Phi}_2'(t, \sigma)\, d\sigma$$

Proof. Differentiate this expression with respect to t and use the facts that $\dot{\mathbf{\Phi}}_1(t, t_0) = \mathbf{A}_1(t)\mathbf{\Phi}_1(t, t_0)$ and $\dot{\mathbf{\Phi}}_2(t, t_0) = \mathbf{A}_2'(t)\mathbf{\Phi}_2(t, t_0)$. This gives

$$\dot{\mathbf{X}}(t) = \mathbf{A}_1(t)\mathbf{\Phi}_1(t, t_0)\mathbf{X}(t_0)\mathbf{\Phi}_2'(t, t_0) + \mathbf{A}_1(t)\int_{t_0}^{t} \mathbf{\Phi}_1(t,\sigma)\mathbf{F}(\sigma)\mathbf{\Phi}_2'(t,\sigma)\, d\sigma$$
$$+ \mathbf{\Phi}_1(t, t_0)\mathbf{X}(t_0)\mathbf{\Phi}_2'(t, t_0)\mathbf{A}_2(t) + \int_{t_0}^{t} \mathbf{\Phi}_1(t, \sigma)\mathbf{F}(\sigma)\mathbf{\Phi}_2'(t, \sigma)\mathbf{A}_2(t)\, d\sigma + \mathbf{F}$$

Therefore **X** satisfies the differential equation. It is easily seen that it also satisfies the initial data. ▌

The matrix notation here is a possible source of confusion in that the matrix differential equation could have been written as a vector differential equation with an n^2-component vector.* Thus although we have offered a new proof for the variation of constants formula we have done so only because it is simpler to start afresh than to struggle with the notation required to apply the old proof.

We have already seen (Section 1) that the set of real n by n matrices is a vector space. The only difference between it and R^{n^2} is the way its elements are indexed. In Section 1 we have defined an inner product for this space.

$$\langle \mathbf{X}, \mathbf{P} \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\, p_{ij} = \operatorname{tr} \mathbf{P}'\mathbf{X} = \operatorname{tr} \mathbf{X}'\mathbf{P}$$

This choice is natural in the sense that regardless what index rule is chosen to convert an n by n matrix into an n^2-tuple, this gives the same value for the inner product as the standard inner product in R^{n^2} would.

We have defined the adjoint differential equation associated with a given homogeneous differential equation to be a homogeneous equation having the property that the inner product between its solutions and those of the original equation are a constant.

To satisfy the constraint

$$\frac{d}{dt}\operatorname{tr} \mathbf{X}'(t)\mathbf{P}(t) = \operatorname{tr} \dot{\mathbf{X}}'(t)\mathbf{P}(t) + \operatorname{tr} \mathbf{X}'(t)\dot{\mathbf{P}}(t) = 0$$

* For example, use lexographic ordering and write $(x_{11}, x_{12}, \ldots, x_{1n}, x_{21}x_{22}, \ldots, x_{nn})$

along the solutions of (LM) we require

$$0 = \mathrm{tr}[\dot{\mathbf{X}}'(t)\mathbf{P}(t) + \mathbf{X}'(t)\dot{\mathbf{P}}(t)] = \mathrm{tr}[\mathbf{X}'(t)\mathbf{A}_1'(t)\mathbf{P}(t) + \mathbf{A}_2'(t)\mathbf{X}'(t)\mathbf{P}(t) + \mathbf{X}'(t)\dot{\mathbf{P}}(t)]$$

Using the facts that $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ and $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}\,\mathbf{A} + \mathrm{tr}\,\mathbf{B}$ we see that

$$0 = \mathrm{tr}\{\mathbf{X}'(t)[\dot{\mathbf{P}}(t) + \mathbf{A}_1'(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}_2'(t)]\}$$

Thus if this is to hold for all $\mathbf{X}$, $\mathbf{P}$ must satisfy

$$\dot{\mathbf{P}}(t) = -\mathbf{A}_1'(t)\mathbf{P}(t) - \mathbf{P}(t)\mathbf{A}_2'(t) \qquad \text{(MA)}$$

This is the adjoint equation associated with equation (LM). This is summarized in the following theorem.

Theorem 2. *If the inner product between two matrices* $\mathbf{X}$ *and* $\mathbf{P}$ *is* $\mathrm{tr}\,\mathbf{P}'\mathbf{X}$, *then the adjoint differential equation associated with* $\dot{\mathbf{X}}(t) = \mathbf{A}_1(t)\mathbf{X}(t) + \mathbf{X}(t)\mathbf{A}_2(t)$ *is* $\dot{\mathbf{P}}(t) = -\mathbf{P}(t)\mathbf{A}_2'(t) - \mathbf{A}_1'(t)\mathbf{P}(t)$.

As an application of linear matrix equations, let us consider the problem of evaluating the integral

$$\eta = \int_{t_0}^{t_1} \mathbf{x}'(t)\mathbf{M}(t)\mathbf{x}(t)\,dt$$

for $\mathbf{x}$ satisfying a linear differential equation of first order, $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$. We would like an expression for η in terms of the initial value, $\mathbf{x}(t_0)$. Clearly one form of the answer is

$$\eta(\mathbf{x}_0) = \int_{t_0}^{t_1} \mathbf{x}'(t_0)\mathbf{\Phi}'(t, t_0)\mathbf{M}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0)\,dt$$

Thus η is a quadratic form in $\mathbf{x}(t_0)$. We adopt the notation

$$\mathbf{Q}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Phi}'(t, t_0)\mathbf{M}(t)\mathbf{\Phi}(t, t_0)\,dt$$

In order to calculate $\mathbf{Q}(t_0, t_1)$ it is not necessary to solve for $\mathbf{\Phi}(t, t_0)$ and integrate. It is possible to derive a linear differential equation for $\mathbf{Q}$ itself. In fact, replacing t_0 by t and differentiating with respect to t yields

$$\begin{aligned}\frac{d}{dt}\mathbf{Q}(t, t_1) &= \frac{d}{dt}\int_{t}^{t_1} \mathbf{\Phi}'(\sigma, t)\mathbf{M}(\sigma)\mathbf{\Phi}(\sigma, t)\,d\sigma \\ &= -\mathbf{A}'(t)\mathbf{Q}(t, t_1) - \mathbf{Q}(t, t_1)\mathbf{A}(t) - \mathbf{M}(t)\end{aligned}$$

The value of $\mathbf{Q}(t_1, t_1)$ is evidently zero so if we add to the differential equation a boundary condition to get

$$\frac{d}{dt}\mathbf{Q}(t, t_1) = -\mathbf{A}'(t)\mathbf{Q}(t, t_1) - \mathbf{Q}(t, t_1)\mathbf{A}(t) - \mathbf{M}(t); \qquad \mathbf{Q}(t_1, t_1) = \mathbf{0}$$

we obtain a direct method of evaluating η. Observe carefully that this is a differential equation with a boundary condition not at the initial time but at the final time.

The case where $\mathbf{A}$ and $\mathbf{M}$ are constant is of special importance. In this case the differential equation is time invariant

$$\frac{d}{dt}\mathbf{Q}(t, t_1) = -\mathbf{A}'\mathbf{Q}(t, t_1) - \mathbf{Q}(t, t_1)\mathbf{A} - \mathbf{M}$$

If there exists an equilibrium solution, i.e. a solution $\mathbf{Q}_1$ of

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \mathbf{M} = \mathbf{0}$$

then the differential equation can be rewritten in terms of $\mathbf{\Psi} = \mathbf{Q} - \mathbf{Q}_1$,

$$\frac{d}{dt}\mathbf{\Psi}(t, t_1) = -\mathbf{A}'\mathbf{\Psi}(t, t_1) - \mathbf{\Psi}(t, t_1)\mathbf{A}$$

From this we obtain immediately an explicit expression for the value of $\mathbf{Q}(t, t_1)$ namely

$$\mathbf{Q}(t, t_1) = \mathbf{Q}_1 - e^{\mathbf{A}'(t_1 - t)}\mathbf{Q}_1 e^{\mathbf{A}(t_1 - t)}$$

The equation $\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \mathbf{M} = \mathbf{0}$ is very important in the study of linear systems. The basic properties of it were studied by Liapunov in connection with stability questions. For our present purposes the following result is adequate although it is by no means the whole story.

Theorem 3. *If the eigenvalues of* $\mathbf{A}$ *have negative real parts, then* $\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{M}$ *can be solved for* $\mathbf{Q}$ *and the solution is unique. Moreover, under this same hypothesis, the solution,* $\mathbf{Q}_1$ *will be given by the convergent integral*

$$\mathbf{Q}_1 = \int_0^\infty e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\,dt$$

Proof. The integral is obviously convergent since it is a sum of terms of the form $t^i e^{\lambda_i t}$ with $\mathrm{Re}\,\lambda_i < 0$. The proof depends on the observation that $\frac{d}{dt}e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t} = \mathbf{A}'e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t} + e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\mathbf{A}$. To show that the integral satisfies the given equation substitute it in and use this identity, i.e.,

$$\begin{aligned}&\mathbf{A}'\int_0^\infty e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\,dt + \int_0^\infty e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\,dt\mathbf{A} \\ &\qquad = \int_0^\infty \frac{d}{dt}e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\,dt = e^{\mathbf{A}'t}\mathbf{M}e^{\mathbf{A}t}\Big|_0^\infty = -\mathbf{M}\end{aligned}$$

To show that this solution is unique, we observe that $L(\mathbf{Q}) = \mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A}$ can be viewed as a linear mapping of R^{n^2} into R^{n^2}. We have just shown that the range space of this mapping is n^2 dimensional by showing that a solution exists for all $\mathbf{M}$. Hence the rank is n^2, and the null space must be just the zero element. ▌

Exercises

1. Consider the differential equation

$$\ddot{x}(t) + 2\delta\dot{x}(t) + x(t) = 0$$

with the initial data $x(0) = 1$; $\dot{x}(0) = 0$. Evaluate

$$\eta = \int_0^\infty x^2(t)\,dt$$

Show that the value of δ which minimizes η is 1/2.

2. Generalize the above result to

$$x^{(n)}(t) + p_{n-1}x^{(n-1)}(t) + \cdots + p_1 x(t) + x(t) = 0$$

and the initial data $x(0) = 1$; $x^{(i)}(0) = 0$, $i = 1, 2, \ldots, n-1$. Find the coefficients which minimize

$$\eta = \int_0^\infty x^2(t)\,dt$$

(Anke, Babister, Bruckner, Parks).

3. Consider the time-invariant matrix equation

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t)$$

where the eigenvalues of $\mathbf{A}$ are assumed to be in $\mathrm{Re}[s] < 0$. Show that if $\mathbf{C} = \mathbf{C}'$ is constant then

$$\eta = \int_0^\infty \mathrm{tr}\,\mathbf{X}'(t)\mathbf{C}\mathbf{X}(t)\,dt = \mathrm{tr}\,\mathbf{X}'(0)\mathbf{Q}\mathbf{X}(0)$$

where $\mathbf{Q}'\mathbf{A} + \mathbf{A}'\mathbf{Q} = -\mathbf{C}$.
Hint: tr (= trace) is a linear operation i.e. $\mathrm{tr}(\mathbf{M} + \mathbf{N}) = \mathrm{tr}(\mathbf{M}) + \mathrm{tr}(\mathbf{N})$. Also $\mathrm{tr}(\mathbf{MN}) = \mathrm{tr}(\mathbf{NM})$.

4. Write out a vector equation equivalent to $\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t) + \mathbf{X}(t)\mathbf{A}(t)$ in the case where $\mathbf{A}$ and $\mathbf{X}$ are 2 by 2.

5. To evaluate quartic forms for linear constant systems

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

we can observe that dyad $\mathbf{x}\mathbf{x}'$ satisfies the linear equation

$$\frac{d}{dt}\mathbf{x}(t)\mathbf{x}'(t) + \mathbf{A}\mathbf{x}(t)\mathbf{x}'(t) + \mathbf{x}(t)\mathbf{x}'(t)\mathbf{A}'$$

If we write $\mathbf{x}(t)\mathbf{x}'(t)$ as an n^2 dimensional column vector, say

$$\mathbf{z} = (x_1x_1;\ x_1x_2;\ \cdots x_1x_n;\ x_2x_1;\ x_2x_2;\ \cdots x_2x_n;\ \cdots x_nx_n)$$

then the differential equation for $\mathbf{z}$ is linear and of the form

$$\dot{\mathbf{z}}(t) = \mathbf{B}\mathbf{z}(t)$$

Thus the quartic form in $\mathbf{x}$

$$\int_0^\infty \mathbf{z}'(t)\mathbf{M}\mathbf{z}(t)\,dt$$

can be evaluated by solving $\mathbf{B}'\mathbf{Q} + \mathbf{Q}\mathbf{B} = -\mathbf{M}$.
Follow through these steps to evaluate

$$\int_0^\infty y^4(t)\,dt$$

for

$$\ddot{y}(t) + 3\dot{y}(t) + 2y(t) = 0; \qquad y(0) = 1; \qquad \dot{y}(0) = 0$$

6. (*Parseval's Theorem*, special case). Show that if $\lambda > 0$ and $\gamma > 0$ then

$$\int_0^\infty (\alpha e^{-\lambda t})(\beta e^{-\gamma t})\,dt = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty}\left(\frac{\alpha}{s+\lambda}\right)\left(\frac{\beta}{\gamma - s}\right)ds$$

Show that if the eigenvalues of $\mathbf{A}$ lie in the half-plane $\mathrm{Re}\, s < 0$ then

$$\int_0^\infty e^{\mathbf{A}'t}\mathbf{C}e^{\mathbf{A}t}\,dt = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty}(-\mathbf{I}s - \mathbf{A}')^{-1}\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\,ds$$

7. Obtain conditions analogous to those given by Theorem 9-1 for the matrix equation

$$\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t) + \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t)$$

to have a periodic solution for suitable $\mathbf{X}(0)$ assuming $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are all periodic of period T.

8. Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices. Show that the matrix equation $\mathbf{AQ} + \mathbf{QB} = \mathbf{C}$ with $\mathbf{Q}$ regarded as unknown has a solution if and only if $\mathrm{tr}\,\mathbf{P}'\mathbf{C} = 0$ for every $\mathbf{P}$ which satisfies $\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{B}' = \mathbf{0}$.

9. If

$$\mathbf{A} = \begin{bmatrix} 5 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{and } \mathbf{M} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

does the equation

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{M}$$

have a unique solution?

10. Show that the Euler equations of Exercise 2, Section 2, can be written as a matrix equation in two ways

$$(\det \mathbf{I})\mathbf{I}^{-1}\dot{\Omega}I^{-1} = \mathbf{I}\boldsymbol{\Omega}^2 - \boldsymbol{\Omega}^2\mathbf{I} + \mathbf{N}$$
$$= (\det \mathbf{I})(\boldsymbol{\Omega}\mathbf{I}^{-1}\boldsymbol{\Omega}\mathbf{I}^{-1} - \mathbf{I}^{-1}\boldsymbol{\Omega}\mathbf{I}^{-1}\boldsymbol{\Omega}) + \mathbf{N}$$

where $\boldsymbol{\Omega}$, $\mathbf{I}$, and $\mathbf{N}$ are 3 by 3 matrices of the form

$$\boldsymbol{\Omega} = \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{bmatrix}; \quad \mathbf{I} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix}; \quad \mathbf{N} = \begin{bmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{bmatrix}$$

Show that the kinetic energy of a rigid body $\frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$ can be expressed as $-\frac{1}{4}(\det \mathbf{I})\mathrm{tr}(\boldsymbol{\Omega}\mathbf{I}^{-1})^2$. If $\mathbf{A}$ is an orthogonal matrix which satisfies $\dot{\mathbf{A}} = \boldsymbol{\Omega}\mathbf{A}$ then we can interpret $\mathbf{L} = \mathbf{A}'\mathbf{I}^{-1}\boldsymbol{\Omega}\mathbf{I}^{-1}\mathbf{A}(\det \mathbf{I})$ as the angular momentum. Show that $\dot{\mathbf{L}} = \mathbf{A}'\mathbf{N}\mathbf{A}$. Recall that $\mathbf{N}$ is expressed in the moving coordinate system.

11. There is an alternative approach to solving $\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{C}$ which is often more convenient when one is doing calculations. Define $\mathbf{B} \otimes \mathbf{A}$, the so-called *Kronecker product*, as a n^2 by n^2 dimensional matrix having the block form

$$\mathbf{B} \otimes \mathbf{A} = \begin{bmatrix} b_{11}\mathbf{A} & b_{12}\mathbf{A} & \cdots & b_{1n}\mathbf{A} \\ b_{21}\mathbf{A} & b_{22}\mathbf{A} & \cdots & b_{2n}\mathbf{A} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1}\mathbf{A} & b_{n2}\mathbf{A} & \cdots & b_{nn}\mathbf{A} \end{bmatrix}$$

It is clear that this multiplication, like the usual matrix multiplication is not commutative.

Convert the n by n matrix $\mathbf{Q}$ into a n^2 dimensional column vector $\mathbf{Q}_v$ by arranging the elements of $\mathbf{Q}$ in the order $q_{11}, q_{12}, \ldots, q_{1n}, q_{21}, q_{22}, \ldots, q_{2n}, \ldots, q_{n1}, q_{n2}, \ldots, q_{nn}$. Define a n^2 dimensional vector $\mathbf{C}_v$ in terms of the elements of $\mathbf{C}$ in a completely analogous way. Show that

$$(\mathbf{A}' \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}')\mathbf{Q}_v = -\mathbf{C}_v$$

For example, if $\mathbf{A}$ is 2 by 2 we are asserting that the matrix equation

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} + \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = -\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

is equivalent to the matrix equation

$$\begin{bmatrix} a_{11} + a_{11} & a_{21} & a_{21} & 0 \\ a_{12} & a_{22} + a_{11} & 0 & a_{21} \\ a_{12} & 0 & a_{11} + a_{22} & a_{21} \\ 0 & a_{12} & a_{12} & a_{22} + a_{22} \end{bmatrix} \begin{bmatrix} q_{11} \\ q_{12} \\ q_{21} \\ q_{22} \end{bmatrix} = -\begin{bmatrix} c_{11} \\ c_{12} \\ c_{21} \\ c_{22} \end{bmatrix}$$

NOTES AND REFERENCES

1. The material on linear algebra is completely standard. A good reference is Halmos [32]. A brief exposition of the basic facts (with no proofs) is contained in Abraham [1]. Gantmacher's two volumes [24] are very useful for reference purposes. They are a unique source for many concrete results on matrix theory. Greub's two volumes [28] are a very complete source for abstract linear algebra. For a less systematic, but quite enlightening presentation, see Bellman [6].
2. Uniqueness theorems can be found in most modern differential equation books. Birkhoff and Rota [8] is an excellent introductory book. Coddington and Levinson [20] and Lefschetz [53] are standard references.
3. Existence theorems for nonlinear equations are treated in differential equation books previously cited. The term "transition matrix" is not standard in mathematics. A *fundamental solution* is any nonsingular solution of $\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t)$, and many mathematical texts do not give a special name to the fundamental solutions which start at $\mathbf{I}$.
4. The behavior of $\boldsymbol{\Phi}$ under a change of coordinates in quite important although often it is not emphasized. The use of diagrams such as figure 3 is now quite common in mathematics. Abraham [1] makes especially good use of them. We will not hesitate to use them whenever they might make the ideas clearer.
5. Functions of matrices can be introduced in many ways. We refer the reader to Gantmacher [24] for some alternative points of view on the matrix exponential. In some settings $[\mathbf{A}, \mathbf{B}]$ is called the *Lie product*. It plays an extensive role in more advanced work. Pease [66] cites a number of applications. The Cauchy integral formula and related matters are discussed in Kaplan [43].
6. Following Coddington and Levinson [20] we use the term variation of constants in place of *variation of parameters*, or *method of undetermined coefficients*, as it is sometimes called. It is a pity that such a basic result has such an awkward collection of names.
7. The idea of an adjoint system is obvious in vector-matrix notation. On the other hand, the adjoint of a nth order scalar equation is much more difficult to appreciate, and also to compute. See, e.g. Coddington and Levinson [20].
8. Periodic theory and reducible equations are treated in most standard texts. They are treated extensively in Erugin [22]. Lefschetz [53] can also be consulted with profit.
9. These results are very important in the mathematical theory of oscillation for which Hale [31] is an excellent reference. Special cases are also of great practical importance (Bode plots, Nyquist Locus, etc. See Kaplan [43]).

10. For a proof of the Hurwitz criterion see Gantmacher [24], or Kaplan [43]. Lehnigk [54] has perhaps the most details on variations. Recently Hurwitz's original paper on this subject has been translated and reprinted [7]. Zadeh and Desoer [91] contains the Lienard-Chipart criterion which involves less calculation than the Hurwitz test as well as the results of Jury on conditions for a polynomial to have its zeros inside the unit disk.
11. Some results on linear matrix equations can be found in Gantmacher [24] and Bellman [6]. Viewing the space of n by n matrices as an inner product space is standard in Lie theory; see Pease [66]. The basic facts on $\mathbf{A}'\mathbf{Q} + \mathbf{QA} = -\mathbf{C}$ were discovered by Liapunov [55].

2
LINEAR SYSTEMS

The word system is used in many ways. In this book we mean by a *linear system* a pair of equations of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)$$

or more compactly,

$$\begin{bmatrix}\dot{\mathbf{x}}(t)\\ \mathbf{y}(t)\end{bmatrix} = \begin{bmatrix}\mathbf{A}(t) & \mathbf{B}(t)\\ \mathbf{C}(t) & \mathbf{D}(t)\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{u}(t)\end{bmatrix}$$

This object is more complicated than just a linear differential equation and there are many interesting questions which arise in the study of linear systems which have no counterpart in classical linear differential equation theory. We are interested in questions which relate to the dependence of $\mathbf{y}$ on $\mathbf{u}$ which these equations imply. The point of view one adopts is that $\mathbf{u}$ is something that can be manipulated and that $\mathbf{y}$ is a consequence of this manipulation. Hence $\mathbf{u}$ is called an *input*, $\mathbf{y}$ is called the *output*. Consistent with the terminology of the previous chapter, $\mathbf{x}$ is called the state.

The resulting theory has been applied quite successfully in two broad categories of engineering (i.e., synthesis) situations. The first is where the system is given and an input is to be found such that the output has certain properties. This is a typical problem in controlling dynamical systems. The second problem is where a given relationship between $\mathbf{u}$ and $\mathbf{y}$ is desired and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are to be chosen in such a way as to achieve this relationship. This is a typical problem in the synthesis of dynamical systems. We will consider some aspects of both questions.

We make an assumption at this point which is to hold throughout. The elements of the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are assumed to be continuous functions of time defined on $\infty < t < \infty$.

12. STRUCTURE OF LINEAR MAPPINGS

In Section 1 we discussed linear mappings and their representation by matrices. Here we continue the discussion, introducing some results on the structure of linear mappings. This discussion will motivate, to some extent, the material on the structure of linear systems which follows.

Given a set of objects $\{x_i\}$ we say that a binary relation $x_i \sim x_j$ (read: x_i equivalent to x_j) is an *equivalence relation* if it is

(i) *reflexive*: i.e. $x_i \sim x_i$
(ii) *symmetric*: i.e. $x_i \sim x_j$ if and only if $x_j \sim x_i$
(iii) *transitive*: i.e. $x_i \sim x_j$ and $x_j \sim x_k$ implies $x_i \sim x_k$.

The matrix $\mathbf{A}_1$ is said to be *equivalent* to the matrix $\mathbf{A}_2$ if there exist square nonsingular matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{PA}_1\mathbf{Q} = \mathbf{A}_2$. The reader should show that $\mathbf{A}_1$ is equivalent to itself; that $\mathbf{A}_1$ is equivalent to $\mathbf{A}_2$ if and only if $\mathbf{A}_2$ is equivalent to $\mathbf{A}_1$; and that if $\mathbf{A}_1$ is equivalent to $\mathbf{A}_2$ and $\mathbf{A}_2$ is equivalent to $\mathbf{A}_3$, then $\mathbf{A}_1$ is equivalent to $\mathbf{A}_3$. Thus this definition is reflexive, symmetric and transitive and hence an equivalence relation in the logical sense. The basic result on equivalent matrices is given by the following theorem.

Theorem 1. *If* $\mathbf{M}_1$ *belongs to* $R^{m\times n}$, *then there exist nonsingular matrices* $\mathbf{P}$ *in* $R^{m\times m}$ *and* $\mathbf{Q}$ *in* $R^{n\times n}$ *such that* $\mathbf{PM}_2\mathbf{Q} = \mathbf{M}_1$ *with* $\mathbf{M}_2$ *being an* m *by* n *matrix of the form*

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

with the number of 1*'s on the diagonal equal to the rank of* $\mathbf{M}_1$.

As a visual reminder of this theorem, we can construct a diagram to go along with the equations $\mathbf{y} = \mathbf{M}_1\mathbf{x} = \mathbf{PM}_2\mathbf{Qx}$

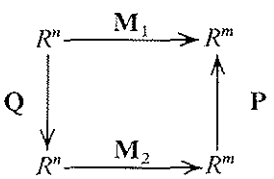


Figure 1. Diagramming the relationship between equivalent matrices.

As a direct consequence of this theorem, there becomes available some factorization results.

Corollary. *If* $\mathbf{M}$ *belongs to* $R^{m\times n}$ *and is of rank* r, *then there exist matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *belonging to* $R^{m\times r}$ *and* $R^{r\times n}$ *respectively, such that* $\mathbf{M} = \mathbf{PQ}$. *At the same time, there exist matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *and* $\mathbf{M}_2$ *belonging to* $R^{m\times m}$, $R^{n\times n}$, *and* $R^{r\times r}$ *such that* $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$, $\mathbf{M}_2$ *is nonsingular and* $\mathbf{M} = \mathbf{PM}_1\mathbf{Q}$ *with*

$\mathbf{M}_1$ *of the form*

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{M}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

A second important decomposition theorem exists for matrices which are symmetric. A matrix $\mathbf{A}_1$ is said to be *congruent* to a second matrix $\mathbf{A}_2$ if there exists a nonsingular matrix $\mathbf{P}$ such that $\mathbf{PA}_1\mathbf{P}' = \mathbf{A}_2$. Again it is easy to show that this definition is reflexive, symmetric and transitive. The following theorem is quite important in many applications. It describes the "simplest" matrix congruent to a given one.

Theorem 2. *If* $\mathbf{A}$ *is a symmetric matrix, then there exists a nonsingular matrix* $\mathbf{P}$ *such that* $\mathbf{A} = \mathbf{P}'\mathbf{SP}$ *where* $\mathbf{S}$ *is a diagonal matrix of plus ones, minus ones and zeros.*

$$\mathbf{S} = \begin{bmatrix} 1 \cdots 0 & 0 \cdots & & 0 \; 0 \cdots 0 \\ \vdots \; \ddots \; \vdots & & & \\ 0 \cdots 1 & 0 & & \cdots\cdots \\ 0 \cdots \cdot & -1 & & \cdots\cdots \\ \vdots & & \ddots & \\ 0 \cdots\cdots\cdots & & -1 & 0 \cdots \\ 0 \cdots\cdots\cdots & & \cdots & 0 \cdots \\ \vdots & & & \ddots \\ 0 \cdots\cdots\cdots & & \cdots & 0 \end{bmatrix}$$

Gantmacher [24] contains a proof.

The n-tuple of diagonal elements $(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$ is sometimes called the *signature n-tuple of* $\mathbf{A}$. The number of nonzero entries is equal to the rank of $\mathbf{A}$.

We say that a real symmetric matrix $\mathbf{Q}$ is *positive definite*, in symbols: $\mathbf{Q} > \mathbf{0}$, if for all real vectors $\mathbf{x} \neq \mathbf{0}$ we have $\mathbf{x}'\mathbf{Qx} > 0$. We say a real symmetric

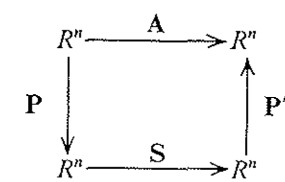


Figure 2. Diagramming the congruence relations.

matrix is *nonnegative definite* if for all real vectors, $\mathbf{x}$, we have $\mathbf{x}'\mathbf{Q}\mathbf{x} \geqslant 0$. As we define it, nonnegative definite includes positive definite as a special case. An important result on positive definiteness is the theorem by Sylvester.

Theorem 3. *An n by n symmetric matrix* $\mathbf{Q}$ *is positive definite if and only if its signature n-tuple consists of plus ones only. This will be the case if and only if the n quantities*

$$q_{11}; \det \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}; \det \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}; \cdots \det \mathbf{Q}$$

are all positive.

A very important criterion for linear independence of vectors can be stated in terms of the positive definiteness of a certain symmetric matrix. This is the *Gram criterion.* Recall that a set of real vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$, in E_n is called *linearly independent* if there exists no set of non-zero real numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$ such that

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_m \mathbf{x}_m = \mathbf{0}$$

If one wants to test a set of vectors for linear independence, a simple way is to form m equations for the α_i by premultiplying the above equation by $\mathbf{x}_1', \mathbf{x}_2'$, etc. to obtain

$$\begin{bmatrix} \mathbf{x}_1'\mathbf{x}_1 & \mathbf{x}_1'\mathbf{x}_2 & \cdots & \mathbf{x}_1'\mathbf{x}_m \\ \mathbf{x}_2'\mathbf{x}_1 & \mathbf{x}_2'\mathbf{x}_2 & \cdots & \mathbf{x}_2'\mathbf{x}_m \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{x}_m'\mathbf{x}_1 & \mathbf{x}_m'\mathbf{x}_2 & \cdots & \mathbf{x}_m'\mathbf{x}_m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The m by m matrix of inner products on the left will be called the *Gramian* associated with the given set of vectors. It is clearly symmetric since $\mathbf{x}_i'\mathbf{x}_j = \mathbf{x}_j'\mathbf{x}_i$; moreover it is nonnegative definite since an easy calculation shows

$$[\alpha_1, \alpha_2, \ldots, \alpha_m] \begin{bmatrix} \mathbf{x}_1'\mathbf{x}_1 & \mathbf{x}_1'\mathbf{x}_2 & \cdots & \mathbf{x}_1'\mathbf{x}_m \\ \mathbf{x}_2'\mathbf{x}_1 & \mathbf{x}_2'\mathbf{x}_2 & \cdots & \mathbf{x}_2'\mathbf{x}_m \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{x}_m'\mathbf{x}_1 & \mathbf{x}_m'\mathbf{x}_2 & \cdots & \mathbf{x}_m'\mathbf{x}_m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$
$$= \|\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_m \mathbf{x}_m\|^2$$

Clearly there exists no nonzero set of α's such that $\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 \cdots \alpha_m \mathbf{x}_m = \mathbf{0}$ if and only if the Gramian is positive definite. Observe that a set of m vectors of dimension n cannot be linearly independent if m exceeds n so that in this case the Gramian matrix is always singular.

If $\mathbf{P}$ is any nonsingular matrix and if $\mathbf{A}$ and $\mathbf{B}$ are two square matrices related by $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$ then $\mathbf{A}$ and $\mathbf{B}$ are said to be related by a *similarity transformation* and are called *similar matrices.* Using the fact that $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$ we see that $\det(\mathbf{I}s - \mathbf{A}) = \det \mathbf{P}^{-1}(\mathbf{I}s - \mathbf{P}^{-1}\mathbf{A}\mathbf{P})^{-1}\mathbf{P} = \det(\mathbf{I}s - \mathbf{P}^{-1}\mathbf{A}\mathbf{P})$, and hence similar matrices have the same characteristic equation. If $\mathbf{A}$ is a symmetric matrix then the eigenvalues of $\mathbf{A}$ (i.e. the zeros of $\det(\mathbf{I}s - \mathbf{A})$ are all real and the eigenvectors of $\mathbf{A}$ can be taken to be mutually perpendicular. By taking $\mathbf{P}$ to be a matrix where columns are the eigenvectors of $\mathbf{A}$ normalized to unit length, one obtains a matrix $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ which is similar to $\mathbf{A}$ and is both real and diagonal. For nonsymmetric matrices such a reduction is not generally possible. In the first place it is clear that if $\mathbf{D}$ is a diagonal matrix then any vector with only one nonzero entry is an eigenvector; using the fact that $\mathbf{D}\mathbf{x}_i = \lambda_i \mathbf{x}_i$ defines the eigenvalues it is likewise clear that each of the diagonal entries is an eigenvalue. Hence a real matrix $\mathbf{A}$ with complex eigenvalues cannot be similar to a real diagonal matrix. The following theorem illustrates to what extent an arbitrary matrix can be reduced by a similarity transformation.

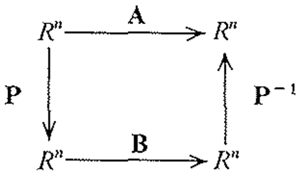


Figure 3. *Diagramming Similarity Transformations.*

Theorem 4. *If* $\mathbf{A}$ *is a real matrix then there exists a real nonsingular matrix* $\mathbf{P}$ *such that* $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ *where* $\mathbf{B}$ *is of the form*

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{B}_2 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \mathbf{B}_m & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{B}_{m+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \mathbf{B}_{m+2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \mathbf{B}_n \end{bmatrix}$$

with $\mathbf{B}_1$ *through* $\mathbf{B}_m$ *taking the form*

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{S}_i & \mathbf{I} & 0 & \cdots & 0 \\ 0 & \mathbf{S}_i & \mathbf{I} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \mathbf{S}_i \end{bmatrix} \qquad \mathbf{S}_i = \begin{bmatrix} \sigma_i & \omega_i \\ -\omega_i & \sigma_i \end{bmatrix}$$

and $\mathbf{B}_{m+1}$ *through* $\mathbf{B}_n$ *taking the form*

$$\mathbf{B}_j = \begin{bmatrix} s_j & 1 & 0 & \cdots & 0 \\ 0 & s_j & 1 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & s_j \end{bmatrix}$$

This is a very difficult result to prove, see Gantmacher [24].

An easy corollary of this theorem is the highly useful Cayley-Hamilton Theorem.

Theorem 5. *If* $\det(\mathbf{I}s - \mathbf{A}) = p(s)$ *then* $p(\mathbf{A}) = \mathbf{0}$. *That is if* $\det(\mathbf{I}s - \mathbf{A}) = s^n + p_{n-1}s^{n-1} + \cdots + p_0$ *then* $\mathbf{A}^n + p_{n-1}\mathbf{A}^{n-1} + \cdots + p_0\mathbf{I} = \mathbf{0}$.

It may happen that in addition to $p(\mathbf{A})$ being the zero matrix, there is a polynomial p_1 of degree less than the dimension of $\mathbf{A}$, such that $p_1(\mathbf{A}) = \mathbf{0}$. The polynomial of lowest degree having the property that $p_1(\mathbf{A}) = \mathbf{0}$ is called the *minimal polynomial* of $\mathbf{A}$.

Exercises

1. Given a collection of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ which is linearly independent, define a new set by

$$\mathbf{v}_1 = \mathbf{x}_1(\|\mathbf{x}_1\|)^{-1}$$
$$\mathbf{v}_2 = (\mathbf{x}_2 - \mathbf{v}_1(\mathbf{v}_1'\mathbf{x}_2))\,\|\mathbf{x}_2 - \mathbf{v}_1(\mathbf{v}_1'\mathbf{x}_2)\|^{-1}$$
$$\mathbf{v}_3 = (\mathbf{x}_3 - \mathbf{v}_1(\mathbf{v}_1'\mathbf{x}_3) - \mathbf{v}_2(\mathbf{v}_2'\mathbf{x}_3))\,\|\mathbf{x}_3 - \mathbf{v}_1(\mathbf{v}_1'\mathbf{x}_3) - \mathbf{v}_2(\mathbf{v}_2'\mathbf{x}_3)\|^{-1}$$
$$\cdots$$

Show that

$$\mathbf{v}_i'\mathbf{v}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

That is, show that $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)'(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \mathbf{I}$. This is called the *Gram-Schmidt orthonormalization process.*

2. Show that the eigenvalues of a symmetric matrix are real.
3. Show that an n by n symmetric matrix is not necessarily nonnegative definite if the n determinants in theorem 3 are nonnegative.
4. Use theorem 4 to show that all solutions of

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

do not go to zero if the conditions for convergence to $\mathbf{0}$ given by theorem 10-1 are violated. Show that all solutions do not remain bounded if the conditions for boundedness given by theorem 10-1 are violated.

5. Show that if $\mathbf{Q} = \mathbf{Q}'$ then there exists a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1} = \mathbf{P}'$ and $\mathbf{PQP}' = \mathbf{\Lambda}$ with $\mathbf{\Lambda}$ real and diagonal. Show that if $\mathbf{Q} = \mathbf{Q}' > \mathbf{0}$ then there exists a unique symmetric matrix $\mathbf{R} > \mathbf{0}$ such that $\mathbf{R}^2 = \mathbf{Q}$. $\mathbf{R}$ is called the *square root* of $\mathbf{Q}$. Show that any nonsingular matrix $\mathbf{T}$ can be written as $\mathbf{\Theta R}$ where $\mathbf{\Theta}' = \mathbf{\Theta}^{-1}$, i.e., $\mathbf{\Theta}$ is orthogonal and $\mathbf{R} = \mathbf{R}' > \mathbf{0}$. This is called the *polar form* of $\mathbf{T}$ because of the analogy with polar representation of complex numbers. Show that the polar form is unique except for the order of $\mathbf{\Theta}$ and $\mathbf{R}$.
6. Let $\mathbf{A}$ be an $n \times n$ matrix of the form

$$\mathbf{A} = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}$$

Show that

$$e^{\mathbf{A}t} = e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & 0 & & \ddots & t \\ & & & & 1 \end{bmatrix}$$

Hint: Let $\mathbf{A} = \lambda\mathbf{I} + \mathbf{Z}$ where

$$\mathbf{Z} = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}$$

and use the fact that (since $(\lambda\mathbf{I})\mathbf{Z} = \mathbf{Z}(\lambda\mathbf{I})$)

$$e^{\mathbf{A}t} = e^{(\lambda\mathbf{I}+\mathbf{Z})t} = e^{\lambda\mathbf{I}t}e^{\mathbf{Z}t}$$

7. If $\mathbf{A}$ is a symmetric matrix, show that its range space and null space are orthogonal in the sense that if $\mathbf{x}$ belongs to the null space of $\mathbf{A}$ and if $\mathbf{y}$ belongs to the range space of $\mathbf{A}$ then $\mathbf{x}'\mathbf{y} = 0$.
8. What we have done for linear independence of vectors we can equally well do for vector functions if we interpret scalar product and linear independence in a suitable way. Consider a set of vector valued continuous

functions $\phi_1, \phi_2, \ldots, \phi_m$ defined on the interval $t_0 \leqslant t \leqslant t_1$. We will say that these are independent if no nontrivial linear combination of them vanishes identically on $t_0 \leqslant t \leqslant t_1$. Let n be the dimension of the vectors. Given a set of m real constants $\alpha_1, \alpha_2, \ldots, \alpha_m$, the sum $\alpha_1\phi_1(t) + \alpha_2\phi_2(t) + \cdots + \alpha_m\phi_m(t)$ is identically zero if and only if

$$\int_{t_0}^{t_1} \|\alpha_1\phi_1(t) + \alpha_2\phi_2(t) \cdots \alpha_m\phi_m(t)\|^2\, dt = 0$$

Show that if we regard the scalar product of ϕ_i and ϕ_j as

$$\int_{t_0}^{t_1} \phi_i'(t)\phi_j(t)\, dt = w_{ij}$$

then the conditions on the α_i demanded by $\alpha_1\phi_1(t) + \alpha_2\phi_2(t) \cdots \alpha_m\phi_m(t) = \mathbf{0}$ are

$$[\alpha_1, \alpha_2, \ldots, \alpha_m]\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ w_{m1} & w_{m2} & \cdots & w_{mm} \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = 0$$

Show that $\mathbf{W}$ is symmetric and nonnegative definite. We will call it the *Gramian* associated with the set of ϕ_i and the interval $t_0 \leqslant t \leqslant t_1$. Show that a necessary and sufficient condition for linear independence of the ϕ_i is that $\mathbf{W}$ should be positive definite.

9. Show that a symmetric matrix $\mathbf{H}$ is positive definite if and only if $\operatorname{tr}(\mathbf{QH})$ is positive for every nonnegative definite matrix $\mathbf{Q} \neq \mathbf{0}$.
10. Show by means of the Cayley–Hamiltonian Theorem that there exist functions $\alpha(t)$ such that $e^{At} = \sum_{i=0}^{n-1} \alpha_i(t)\mathbf{A}^i$.

13. CONTROLLABILITY

Up until now, we have been concerned with the *analysis* of linear dynamics. That is, we have concentrated on how one goes from an implicit equation such as $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t)$ to an explicit expression for $\mathbf{x}$ as is given by the variation of constants formula. The forcing function $\mathbf{f}$ was regarded as given. On the other hand, in control theory the problem typically is to *find* the input which causes the state or the output to behave in a desired way.

The classical servomechanism problem is concerned with making certain variables associated with a dynamical system, say a telescope, follow a particular path. On the other hand, the standard problem in guidance of rockets is to make certain variables associated with a dynamical system take on particular values at one point in time. These two examples are typical

problems in control and may be thought of as motivating the results of this section. We will refer to the first as the *servomechanism controllability* problem and the second as the *ballistic controllability* problem.

To begin, lets consider a special case of the ballistic problem. Suppose the dynamics are given as

$$\dot{\mathbf{z}}(t) = \mathbf{B}(t)\mathbf{u}(t), \qquad \mathbf{z}(t_0) = \dot{\mathbf{z}}_0$$

Assume $\mathbf{B}$ and $\mathbf{z}_0$ are known and let the problem be that of finding $\mathbf{u}$ so as to insure that at $t = t_1 > t_0$, $\mathbf{z}$ takes on a certain value. In symbols: $\mathbf{z}(t_1) = \mathbf{z}_1$. An integration of the differential equation gives

$$\mathbf{z}(t_1) = \mathbf{z}_0 + \int_{t_0}^{t_1} \mathbf{B}(t)\mathbf{u}(t)\, dt$$

Hence, if $\mathbf{z}_1 - \mathbf{z}_0$ lies in the "range space" of the linear mapping

$$L(\mathbf{u}) = \int_{t_0}^{t_1} \mathbf{B}(t)\mathbf{u}(t)\, dt$$

then the desired transfer is possible and otherwise it is not.

To put this discussion on a more formal basis we need a precise definition. If $\mathbf{G}$ is an n by m matrix whose elements are continuous functions of time defined on the interval $t_0 \leqslant t \leqslant t_1$, then the mapping $L\colon C^m[t_0, t_1] \to R^n$ defined by

$$L(\mathbf{u}) = \int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{u}(t)\, dt$$

is linear. We say that $\mathbf{x}_1$ belongs to the *range space* of this mapping if there exists $\mathbf{u}_1$ in $C^m[t_0, t_1]$ such that

$$\mathbf{x}_1 = \int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{u}(t)\, dt$$

Lemma 1. *An n-tuple* $\mathbf{x}_1$ *lies in the range space of* $L(\mathbf{u}) = \int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{u}(t)\, dt$ *if and only if it lies in the range space of the matrix*

$$\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{G}'(t)\, dt$$

Proof. If $\mathbf{x}_1$ lies in the range space of $\mathbf{W}(t_0, t_1)$, then there exists $\boldsymbol{\eta}_1$ such that $\mathbf{W}(t_0, t_1)\boldsymbol{\eta}_1 = \mathbf{x}_1$.

Let $\mathbf{u}$ be given by $\mathbf{G}'\boldsymbol{\eta}_1$, then

$$\int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{u}(t)\, dt = \mathbf{W}(t_0, t_1)\boldsymbol{\eta}_1 = \mathbf{x}_1$$

On the other hand, if $\mathbf{x}_1$ does not lie in the range space then there exists an

$\mathbf{x}_2$ such that $\mathbf{W}(t_0, t_1)\mathbf{x}_2 = \mathbf{0}$ and $\mathbf{x}_2'\mathbf{x}_1 \neq 0$. Suppose, contrary to what we want to prove, that there exists a $\mathbf{u}_1$ such that

$$\int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{u}_1(t)\, dt = \mathbf{x}_1$$

then

$$\int_{t_0}^{t_1} \mathbf{x}_2' \mathbf{G}(t)\mathbf{u}_1(t)\, dt = \mathbf{x}_2'\mathbf{x}_1 \neq 0$$

But,

$$\mathbf{x}_2'\mathbf{W}(t_0, t_1)\mathbf{x}_2 = \int_{t_0}^{t_1} [\mathbf{x}_2'\mathbf{G}(t)][\mathbf{G}'(t)\mathbf{x}_2]\, dt = 0$$

Since $\mathbf{G}(t)$ is continuous we know that the vanishing of

$$\mathbf{x}_2'\mathbf{W}(t_0, t_1)\mathbf{x}_2 = \int_{t_0}^{t_1} \|\mathbf{G}'(t)\mathbf{x}_2\|^2\, dt$$

implies the vanishing of $\mathbf{G}'(t)\mathbf{x}_2$ and the contradiction of

$$\int_{t_0}^{t_1} \mathbf{x}_2'\mathbf{G}(t)\mathbf{u}_1(t)\, dt \neq 0 \quad \blacksquare$$

Corollary. *There exists a control* $\mathbf{u}$ *which transfers the state of the system* $\dot{\mathbf{z}}(t) = \mathbf{B}(t)\mathbf{u}(t)$ *from* $\mathbf{z}_0$ *at* $t = t_0$ *to* $\mathbf{z}_1$ *at* $t = t_1$ *if and only if* $\mathbf{z}_1 - \mathbf{z}_0$ *lies in the range space of* $\int_{t_0}^{t_1} \mathbf{B}(t)\mathbf{B}'(t)\, dt$. *If the transfer is possible then one particular control which actually drives the state to* $\mathbf{z}_1$ *at* $t = t_1$ *is* $\mathbf{u}(t) = \mathbf{B}'(t)\boldsymbol{\eta}$ *where* $\boldsymbol{\eta}$ *is any solution of*

$$\int_{t_0}^{t_1} \mathbf{B}(t)\mathbf{B}'(t)\, dt\boldsymbol{\eta} = \mathbf{z}_1 - \mathbf{z}_0$$

The extension of these results to dynamics of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

is the subject of theorem 1. (Note carefully the order of the arguments in $\boldsymbol{\Phi}$ and $\mathbf{W}$ in this theorem statement.)

Theorem 1. *There exists a* $\mathbf{u}$ *which drives the state of the system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

from the value $\mathbf{x}_0$ *at* $t = t_0$ *to the value* $\mathbf{x}_1$ *at* $t = t_1 > t_0$ *if and only if* $\mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$ *belongs to the range space of*

$$\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \boldsymbol{\Phi}(t_0, t)\mathbf{B}(t)\mathbf{B}'(t)\boldsymbol{\Phi}'(t_0, t)\, dt$$

Moreover, if $\boldsymbol{\eta}_0$ *is any solution of* $\mathbf{W}(t_0, t_1)\boldsymbol{\eta} = \mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$ *then* $\mathbf{u}$ *as given by* $\mathbf{u}(t) = -\mathbf{B}'(t)\boldsymbol{\Phi}'(t_0, t)\boldsymbol{\eta}_0$ *is one control which accomplishes the desired transfer.*

Proof. If we define $\mathbf{z}$ by $\mathbf{z}(t) = \boldsymbol{\Phi}(t_0, t)\mathbf{x}(t)$, then

$$\dot{\mathbf{z}}(t) = \boldsymbol{\Phi}(t_0, t)\mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{x}(t) = \boldsymbol{\Phi}(t, t_0)\mathbf{z}(t)$$

From the previous lemma we know that the set of values which $\mathbf{z}(t_1) - \mathbf{z}(t_0)$ can take on are those which lie in the range space of

$$\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \boldsymbol{\Phi}(t_0, t)\mathbf{B}(t)\mathbf{B}'(t)\boldsymbol{\Phi}'(t_0, t)\, dt$$

To accomplish the desired transfer, we require $\mathbf{z}(t_1) - \mathbf{z}(t_0) = \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1 - \mathbf{x}_0$. Hence, the desired transfer is possible and if and only if $[\mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1]$ lies in the range space of $\mathbf{W}(t_0, t_1)$. From the comments following the previous lemma we see that one particular control which accomplishes the transfer is defined by

$$\mathbf{u}(t) = -\mathbf{B}'(t)\boldsymbol{\Phi}'(t_0, t)\boldsymbol{\eta}$$

with $\boldsymbol{\eta}$ satisfying $\mathbf{W}(t_0, t_1)\boldsymbol{\eta} = \mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$. $\blacksquare$

It is instructive to interpret the quantity $\mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$ in terms of trajectories. Since $\boldsymbol{\Phi}$ describes the motion of $\mathbf{x}$ with no input, $\boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$ has the interpretation of a state $\mathbf{x}_1$ at $t = t_1$ propagated backwards in time to t_0. If $\mathbf{x}_0 - \boldsymbol{\Phi}(t_0, t_1)\mathbf{x}_1$ is zero, then no control need be applied to accomplish the transfer because the free motion passes through $\mathbf{x}_1$. (See figure.)

The matrix $\mathbf{W}$ plays an important role in the theory of forced linear systems. We call it the *controllability Gramian*. The following theorem describes some of its properties.

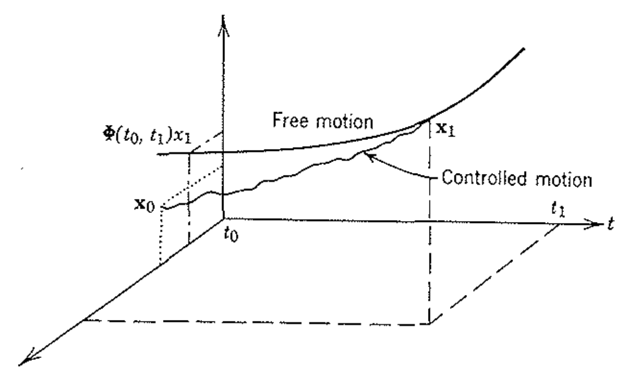


Figure 1. *Illustration of Theorem 1.*

Theorem 2. *The matrix* $\mathbf{W}$ *defined in the statement of Theorem 1 has the following properties*

(i) $\mathbf{W}(t_0, t_1)$ *is symmetric*
(ii) $\mathbf{W}(t_0, t_1)$ *is nonnegative definite for* $t_1 \geqslant t_0$.
(iii) $\mathbf{W}(t_0, t_1)$ *satisfies the linear matrix differential equation*

$$\frac{d}{dt}\mathbf{W}(t, t_1) = \mathbf{A}(t)\mathbf{W}(t, t_1) + \mathbf{W}(t, t_1)\mathbf{A}'(t) - \mathbf{B}(t)\mathbf{B}'(t); \qquad \mathbf{W}(t_1, t_1) = \mathbf{0}$$

(iv) $\mathbf{W}(t_0, t_1)$ *satisfies the functional equation*

$$\mathbf{W}(t_0, t_1) = \mathbf{W}(t_0, t) + \mathbf{\Phi}(t_0, t)\mathbf{W}(t, t_1)\mathbf{\Phi}'(t_0, t)$$

Proof. Property (i) follows immediately from the definitions. To establish (ii) we observe that for any constant vector $\boldsymbol{\eta}$ we have

$$\begin{aligned}\boldsymbol{\eta}'\mathbf{W}(t_0, t_1)\boldsymbol{\eta} &= \int_{t_0}^{t_1} \boldsymbol{\eta}'\mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\boldsymbol{\eta}\, d\sigma \\ &= \int_{t_0}^{t_1} \|\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\boldsymbol{\eta}\|^2\, d\sigma\end{aligned}$$

Hence $\boldsymbol{\eta}'\mathbf{W}(t_0, t_1)\boldsymbol{\eta} \geqslant 0$ for all real $\boldsymbol{\eta}$.

To establish (iii) it is enough to evaluate the derivative of $\mathbf{W}(t, t_1)$ with respect to t. Using Leibnitz's rule we have

$$\begin{aligned}&\frac{d}{dt}\int_t^{t_1} \mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t, \sigma)\, d\sigma \\ &\quad = -\mathbf{B}(t)\mathbf{B}'(t) + \int_t^{t_1} \frac{d}{dt}\{\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t, \sigma)\}\, d\sigma \\ &\quad = -\mathbf{B}(t)\mathbf{B}'(t) + \mathbf{A}(t)\int_t^{t_1} \mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t, \sigma)\, d\sigma \\ &\qquad + \int_t^{t_1} \mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t, \sigma)\, d\sigma\mathbf{A}'(t)\end{aligned}$$

Thus the differential equation has been verified. The boundary condition at $t = t_1$ is obvious from the definition of $\mathbf{W}$.

To prove (iv) we expand the integral defining $\mathbf{W}$ to get

$$\begin{aligned}\mathbf{W}(t_0, t_1) &= \int_{t_0}^{t} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\, d\sigma \\ &\quad + \int_t^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\, d\sigma \\ &= \mathbf{W}(t_0, t) + \mathbf{\Phi}(t_0, t)\int_t^{t_1} \mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t, \sigma)\, d\sigma\mathbf{\Phi}'(t_0, t)\end{aligned}$$

From which property (iv) follows using the definition of $\mathbf{W}(t, t_1)$. ▌

In the special case where $\mathbf{A}$ and $\mathbf{B}$ are time invariant it is possible to calculate the range space of $\mathbf{W}$ quite easily. Moreover, contrary to the general case, the range space does not depend on the arguments of $\mathbf{W}$ except in a trivial way. The following theorem expresses the situation.

Theorem 3. *For* $\mathbf{A}$ *and* $\mathbf{B}$ *constant and* $\mathbf{A}$ *n by n, the range space and the null space of* $\mathbf{W}(t_0, t)$ *for* $t > t_0$ *coincide with the range space and null space of the n by n matrix*

$$\mathbf{W}_T = [\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}][\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}]'$$

Moreover, for any vector $\mathbf{x}_0$ *and any* $t > t_0$,

$$\operatorname{Rank}[\mathbf{W}(t_0, t), \mathbf{x}_0] = \operatorname{Rank}[\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}), \mathbf{x}_0]$$

Proof. If $\mathbf{x}_1$ is in the null space of $\mathbf{W}(t_0, t_1)$, then

$$0 = \mathbf{x}_1'\mathbf{W}(t_0, t_1)\mathbf{x}_1 = \int_{t_0}^{t_1} \mathbf{x}_1' e^{\mathbf{A}(t_0 - \sigma)}\mathbf{BB}'e^{\mathbf{A}'(t_0 - \sigma)}\mathbf{x}_1\, d\sigma$$

Since the integrand is of the form $\|\mathbf{B}'e^{\mathbf{A}'(t_0-\sigma)}\mathbf{x}_1\|^2$ we can see that $\mathbf{x}_1'\mathbf{W}(t_0, t_1)\mathbf{x}_1$ can vanish only if, for all σ

$$\mathbf{B}'e^{\mathbf{A}'(t_0-\sigma)}\mathbf{x}_1 = \mathbf{0}$$

Expanding the left side of this equation in a Taylor series about $\sigma = t_0$ we see it demands that

$$\begin{aligned}\mathbf{B}'\mathbf{x}_1 &= \mathbf{0} \\ \mathbf{B}'\mathbf{A}'\mathbf{x}_1 &= \mathbf{0} \\ &\cdots\cdots \\ \mathbf{B}'(\mathbf{A}')^{n-1}\mathbf{x}_1 &= \mathbf{0}\end{aligned}$$

Thus, if $\mathbf{x}_1$ is in the null space of $\mathbf{W}(t_0, t_1)$ it is in the null space of $[\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}]'$ and hence in the null space of $\mathbf{W}_T$.

Now suppose $\mathbf{x}_1$ is in the null space of $\mathbf{W}_T$. By virtue of the Cayley-Hamilton theorem, we can express $e^{\mathbf{A}(t_0-\sigma)}$ as (Exercise 10, Section 12)

$$e^{\mathbf{A}(t_0-\sigma)} = \sum_{i=0}^{n-1} \alpha_i(t_0 - \sigma)\mathbf{A}^i$$

Hence

$$\mathbf{x}_1'\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \left[\sum_{i=0}^{n-1} \alpha_i(t_0 - \sigma)\mathbf{x}_1'\mathbf{A}^i\mathbf{B}\right]\mathbf{B}'e^{\mathbf{A}'(t_0-\sigma)}\, d\sigma$$

Now $\mathbf{x}_1'\mathbf{A}^i\mathbf{B} = \mathbf{0}$ ($i = 0, 1, \ldots, n-1$) since $\mathbf{x}_1$ is in the null space of $\mathbf{W}_T$. Therefore

$$\mathbf{x}_1'\mathbf{W}(t_0, t_1) = \mathbf{0}$$

and since $\mathbf{W}(t_0, t_1)$ is symmetric

$$\mathbf{W}(t_0, t_1)\mathbf{x}_1 = \mathbf{0}$$

so that $\mathbf{x}_1$ is in the null space of $\mathbf{W}(t_0, t_1)$. Using the fact that $\mathbf{W}(t_0, t_1)$ is a symmetric matrix and therefore that its null space and range space are orthogonal, we conclude that $\mathbf{W}_T$ and $\mathbf{W}(t_0, t_1)$ have the same range and null spaces. Clearly $(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})$ and $\mathbf{W}_T$ have the same range spaces so that the rank equality holds. ▮

We will call an n-dimensional linear constant system *controllable* if $(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})$ is of rank n. No attempt is made to define the term controllability in broader context.

Example. (Satellite Problem). The linearized equations for a unit mass with thrusters in an inverse square law force field have been given in Section 2. In the notation of the present section

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

An easy calculation gives

$$[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix}$$

As may be verified, this is of rank 4 so that the system is controllable.

On the other hand, if one of the inputs is inoperative is the system still controllable? Setting $u_2 = 0$ reduces $\mathbf{B}$ to $\mathbf{B}_1 = (0, 1, 0, 0)'$ and gives

$$(\mathbf{B}_1, \mathbf{AB}_1, \mathbf{A}^2\mathbf{B}_1, \mathbf{A}^2\mathbf{B}_1) = \begin{bmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix}$$

This matrix is only of rank 3. Setting $u_1 = 0$ reduces $\mathbf{B}$ to $\mathbf{B}_2 = (0, 0, 0, 1)'$ and gives

$$(\mathbf{B}_2, \mathbf{AB}_2, \mathbf{A}^2\mathbf{B}_2, \mathbf{A}^3\mathbf{B}_2) = \begin{bmatrix} 0 & 0 & 2\omega & 0 \\ 0 & 2\omega & 0 & -2\omega^3 \\ 0 & 1 & 0 & -4\omega^2 \\ 1 & 0 & -4\omega^2 & 0 \end{bmatrix}$$

This matrix is of rank 4. Since u_1 was radial thrust and u_2 was tangential

thrust we see that loss of radial thrust does not destroy controllability, whereas loss of tangential thrust does.

In some cases the entire state vector $\mathbf{x}$ is not of interest and some subset of its components is all that matters. The following theorem is then of interest. It generalizes Theorem 1 and although the generalization is not deep it helps fill in certain gaps.

Theorem 4. *There exists a* $\mathbf{u}$ *which drives the output of the system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

to $\mathbf{y}_1$ *at* $t = t_1 > t_0$ *if and only if* $\mathbf{y}_1 - \mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{x}_0$ *lies in the range space of* $\mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{W}(t_0, t_1)$.

Proof. Using theorem 1 we know that the set of states that can be reached at t_1 are those states expressible as $\mathbf{\Phi}(t_1, t_0)\mathbf{x}_0 + \mathbf{\Phi}(t_1, t_0)\mathbf{W}(t_0, t_1)\boldsymbol{\eta}$. Thus we know that the set of values $\mathbf{y}(t_1)$ can take on are those which are expressible as $\mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{x}_0 + \mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{W}(t_0, t_1)\boldsymbol{\eta}$. Hence in order to be able to reach a particular $\mathbf{y}_1$ it must happen that $\mathbf{y}_1 - \mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{x}_0$ lies in the range space of $\mathbf{C}(t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{W}(t_0, t_1)$. ▮

It is important not to read too much into the preceding theorem. It is a result about controllability in the ballistic sense and does not imply that the output $\mathbf{y}$ can be made to be a particular function, i.e. it does not imply that $\mathbf{y}(t)$ can be made to "follow a path". Questions about controllability in a servomechanism sense have been avoided thus far. The following informal discussion, which is based on the Laplace transform and hence is limited to time invariant systems, is intended to serve as an introduction.

Suppose we start with the time domain equations

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

and transform to get

$$\hat{\mathbf{y}}(s) = [\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\hat{\mathbf{u}}(s) + \mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{x}_0$$

If $\hat{\mathbf{y}}_1(s)$ is a given desired output, it is necessary to be able to solve

$$\hat{\mathbf{y}}_1(s) - \mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{x}_0 = [\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\hat{\mathbf{u}}(s)$$

for $\hat{\mathbf{u}}$ in order to generate it. Note however, that even if $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$ is invertible, some terms in the solution may correspond to delta functions when transformed back into the time domain. By requiring $\mathbf{y}_d$, defined as $\mathbf{y}_d(t) = \mathbf{y}_1(t) - \mathbf{C}e^{\mathbf{A}t}\mathbf{x}_0$, to have sufficiently many derivatives this problem can be avoided provided $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$ is invertible. The following theorem gives necessary and sufficient conditions for the invertibility of $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$.

Theorem 5. *Let* **A**, **B**, **C** *and* **D** *be n by n, n by m, m by n, and m by m, respectively. The determinant of the m by m matrix* $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$ *is identically zero if and only if the rank of the matrix*

$$\begin{bmatrix} \mathbf{D} & \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^{n-1}\mathbf{B} & \mathbf{CA}^{n}\mathbf{B} & \cdots & \mathbf{CA}^{2n}\mathbf{B} \\ \mathbf{0} & \mathbf{D} & \mathbf{CB} & \cdots & \mathbf{CA}^{n-2}\mathbf{B} & \mathbf{CA}^{n-1}\mathbf{B} & \cdots & \mathbf{CA}^{2n-1}\mathbf{B} \\ \mathbf{0} & \mathbf{0} & \mathbf{D} & \cdots & \mathbf{CA}^{n-3}\mathbf{B} & \mathbf{CA}^{n-2}\mathbf{B} & \cdots & \mathbf{CA}^{2n-2}\mathbf{B} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D} & \mathbf{CB} & \cdots & \mathbf{CA}^{n-1}\mathbf{B} \end{bmatrix}$$

is less than $(n+1)m$.

This theorem is not used in the sequel, and its proof is omitted.

Exercises

1. Show that the limit

$$\lim_{t\to\infty} \mathbf{W}(0, t) = \lim_{t\to\infty} \int_0^t \mathbf{\Phi}(0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(0, \sigma)\, d\sigma$$

exists if the transition matrix for $\dot{\mathbf{x}}(t) = -\mathbf{A}'(t)\mathbf{x}(t)$ approaches zero at an exponential rate and $\|\mathbf{B}(t)\|$ is bounded for $0 \leqslant t \leqslant \infty$.

2. Obtain controllability conditions analogous to those given by theorem 13-1 which apply to the matrix equation

$$\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t) + \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t)\mathbf{U}(t)\mathbf{D}(t)$$

3. Let $f(\;)$ be a continuously differentiable scalar function of a scalar argument. Assume $f'(y) \geqslant \varepsilon > 0$ for all y. Show that if $\det(\mathbf{b}, \mathbf{Ab}, \ldots, \mathbf{A}^{n-1}\mathbf{b}) \neq 0$ then for $t_1 > 0$ the equation

$$\int_0^{t_1} e^{-\mathbf{A}t}\mathbf{b} f(\mathbf{b}'e^{-\mathbf{A}'t}\mathbf{p})\, dt = \mathbf{x}$$

has a unique solution $\mathbf{p}$ for every given $\mathbf{x}$.

4. Suppose that for every constant k in the interval $0 \leqslant k \leqslant 1$

$$\det[\mathbf{b} + k\mathbf{c}, \mathbf{A}(\mathbf{b} + k\mathbf{c}), \ldots, \mathbf{A}^{n-1}(\mathbf{b} + k\mathbf{c})] \neq 0$$

does it follow that for all $t_1 > 0$ and all time-varying k satisfying $0 \leqslant k(t) \leqslant 1$, the matrix

$$\mathbf{W}(0, t_1) = \int_0^{t_1} e^{-\mathbf{A}t}[\mathbf{b} + k(t)\mathbf{c}][\mathbf{b} + k(t)\mathbf{c}]'e^{-\mathbf{A}'t}\, dt$$

will be positive definite?

5. Let $\mathbf{Q}$ be a positive definite matrix. When is the time invariant system

$$\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \dot{\mathbf{x}}_2(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{Q} \\ -\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix} \mathbf{u}(t)$$

controllable?

6. Show that the differential equation

$$\ddot{\chi}(t) + u(t)\dot{\chi}(t) + \chi(t) = 0$$

is controllable in the sense that given any χ_0, $\dot{\chi}_0$ and χ_1, $\dot{\chi}_1$ there exists a u and a $T > 0$ such that $\chi(T) = \chi_1$ and $\dot{\chi}(T) = \dot{\chi}_1$ provided $\chi_0^2 + \dot{\chi}_0^2$ and $\chi_1^2 + \dot{\chi}_1^2 \neq 0$.

7. Given the time invariant system

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{b}u(t); \qquad \mathbf{x}(0) = \mathbf{0}$$

and the constraint

$$u(t + \tfrac{1}{2}) = u(t)$$

Show that it is possible to select u such that $\mathbf{x}(1) = \mathbf{x}_1$ if and only if there exists a vector $\boldsymbol{\eta}$ such that

$$(e^{\mathbf{A}/2} + \mathbf{I})(\mathbf{b}, \mathbf{Ab}, \ldots, \mathbf{A}^{n-1}\mathbf{b})\boldsymbol{\eta} = \mathbf{x}_1$$

8. Specialize theorem 4 to the case where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant in such a way to obtain a generalization of Theorem 3.

9. Consider the vector differential equation

$$\dot{\mathbf{x}} = g(t)[\mathbf{Ax}(t) + \mathbf{Bu}(t)]$$

with g continuous and bounded

$$0 < \alpha \leqslant g(t) \leqslant \beta < \infty$$

Let $\mathbf{A}$ be n by n and $\mathbf{B}$ n by m. Both are constant. Show that if rank $(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})$ is n then given any $T > 0$ and any $\mathbf{x}_0$ and $\mathbf{x}_1$ there exists a control $\mathbf{u}$ which transfers the system from $\mathbf{x}_0$ at $t = 0$ to $\mathbf{x}_1$ at $t = T$.

Hint: Consider a change of time scale.

10. Calculate $\mathbf{W}(0, T)$ for the system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

11. Consider a harmonic oscillator with a drive u, satisfying the equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

Suppose we want to drive the system from the state $[1, 0]'$ to the state $[0, 0]'$ in 2π units of time.
(a) Does there exist a control u which makes this transfer?
(b) Now suppose u is to be a piecewise constant function of time of the form:

$$u(t) = \begin{cases} u_1; & 0 \leq t < 2\pi/3 \\ u_2; & 2\pi/3 \leq t < 4\pi/3 \\ u_3; & 4\pi/3 \leq t \leq 2\pi \end{cases}$$

Do there exist constants u_1, u_2 and u_3 such that we can make the transfer from $[1, 0]'$ at $t = 0$ to $[0, 0]'$ at time $t = 2\pi$?

12. If $\mathbf{A}$ and $\mathbf{B}$ are constant then

$$\mathbf{W}(0, t_1) = \int_0^{t_1} e^{-\mathbf{A}t}\mathbf{B}\mathbf{B}'e^{-\mathbf{A}'t}\, dt$$

From Theorem 2 we know that

$$\frac{d}{dt}\mathbf{W}(t, t_1) = \mathbf{A}\mathbf{W}(t, t_1) + \mathbf{W}(t, t_1)\mathbf{A}' - \mathbf{B}\mathbf{B}'; \qquad \mathbf{W}(t_1, t_1) = \mathbf{0}$$

If $\mathbf{Q}_0$ is an equilibrium solution of this equation, i.e. if

$$\mathbf{A}\mathbf{Q}_0 + \mathbf{Q}_0\mathbf{A}' - \mathbf{B}\mathbf{B}' = \mathbf{0}$$

Then show that $\Xi(t, t_1) = [\mathbf{W}(t, t_1) - \mathbf{Q}_0]$ satisfies the differential equation

$$\frac{d}{dt}\Xi(t, t_1) = \mathbf{A}\Xi(t, t_1) + \Xi(t, t_1)\mathbf{A}'$$

with the initial condition $\Xi(t_1, t_1) = -\mathbf{Q}_0$ and that as a result

$$\mathbf{W}(t, t_1) = \mathbf{Q}_0 - e^{-\mathbf{A}(t_1 - t)}\mathbf{Q}_0 e^{-\mathbf{A}'(t_1 - t)}$$

Remark: $\mathbf{A}\mathbf{Q} + \mathbf{Q}\mathbf{A}' = \mathbf{B}\mathbf{B}'$ may not have a solution $\mathbf{Q}_0$, here we are *making an explicit assumption* that it does.

13. Let $\mathbf{A}$ and $\mathbf{F}$ be constant $n \times n$ matrices. Let $\mathbf{b}$ be an n by 1 vector and let $\mathbf{G} = \mathbf{A} - \mathbf{F}$. Show that it is possible to drive the state of the system

$$\dot{\mathbf{x}}(t) = e^{\mathbf{F}t}\mathbf{A}e^{-\mathbf{F}t}\mathbf{x}(t) + e^{\mathbf{F}t}\mathbf{b}u(t)$$

from any state at $t = 0$ to $\mathbf{0}$ at $t = 1$ if and only if

$$\det(\mathbf{b}, \mathbf{G}\mathbf{b}, \ldots, \mathbf{G}^{n-1}\mathbf{b}) \neq 0$$

14. If $\mathbf{A}$ is n by n show that the range space of the two matrices

$$(\mathbf{B}, \mathbf{A}\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}) = \mathbf{L}_1$$
$$e^{\mathbf{A}}(\mathbf{B}, \mathbf{A}\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}) = \mathbf{L}_2$$

is the same.

15. Show that $\mathbf{W}$ satisfies the integral identity

$$\int_{t_0}^{t} \mathbf{\Phi}(t, \rho)\mathbf{W}(\rho, \rho + \delta)\mathbf{\Phi}'(t, \rho)\, d\rho = \int_{t_0}^{t_0 + \delta} \mathbf{\Phi}(t, \sigma)\mathbf{W}(\sigma, t - t_0 + \sigma)\mathbf{\Phi}'(t, \sigma)\, d\sigma$$

(Silverman-Anderson)

16. Consider a time-varying periodic system of the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \operatorname{sgn}(\sin t)\mathbf{D}\mathbf{x} + \mathbf{B}\mathbf{u}$$

where $\mathbf{x}$ is an n-vector and u is an m-vector. Show that a necessary and sufficient condition for the existence of a $\mathbf{u}$ defined on $0 \leq t \leq 2\pi$ such that any $\mathbf{x}_0$ can be driven to zero at 2π is that

$$\operatorname{rank}(\mathbf{B}, \mathbf{A}\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}, \mathbf{B}, \mathbf{D}\mathbf{B}, \ldots, \mathbf{D}^{n-1}\mathbf{B}) = n$$

17. Show that

$$\exp\begin{bmatrix} \mathbf{A} & \mathbf{B}\mathbf{B}' \\ \mathbf{0} & -\mathbf{A}' \end{bmatrix} t = \begin{bmatrix} e^{\mathbf{A}t} & e^{\mathbf{A}t}\mathbf{W}(0, t) \\ \mathbf{0} & e^{-\mathbf{A}'t} \end{bmatrix}$$

where $\mathbf{W}(0, t)$ is the controllability Gramian for $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$.

18. Let $\mathbf{b}$ be a column vector and let $\mathbf{A}$ be a real square constant matrix. Let $\mathbf{W}(0, T)$ be defined by Theorem 1 and factor $\mathbf{W}(0, T)$ as $\mathbf{D}\mathbf{D}'$. If we denote by ϕ_i the ith component of the vector $\mathbf{D}^{-1}e^{-\mathbf{A}t}\mathbf{b}$ then show

$$\int_0^T \phi_i(t)\phi_j(t)\, dt = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Thus the ϕ_i are orthogonal and normalized on the given interval. Also show that the least squares approximation to an arbitrary square integrable function ψ defined on $[0, T]$ using a linear combination of the ϕ_i is

$$\psi_a(t) = \mathbf{D}'e^{-\mathbf{A}'t}\mathbf{W}^{-1}(0, T)\int_0^T \psi(t)e^{-\mathbf{A}\sigma}\mathbf{b}\, d\sigma$$

19. Given that $\mathbf{W}_A(t_0, t_1)$ is the controllability Gramian for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ and given that $\mathbf{W}_B(t_0, t_1)$ is the controllability Gramian for

$$\dot{\mathbf{z}}(t) = [\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]\mathbf{z}(t) + \mathbf{P}(t)\mathbf{B}(t)\mathbf{u}(t)$$

show that

$$\mathbf{P}(t_0)\mathbf{W}_{\mathbf{B}}(t_0, t_1)\mathbf{P}'(t_0) = \mathbf{W}_{\mathbf{A}}(t_0, t_1)$$

20. Show that if rank $(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}) = n = \dim \mathbf{A}$ and that if the eigenvalues of $\mathbf{A}$ lie in the half-plane $\operatorname{Re} s < 0$ then the solution of $\mathbf{AW} + \mathbf{WA}' = -\mathbf{BB}'$ is positive definite and that the solution of $\mathbf{AW} + \mathbf{WA}' = \mathbf{BB}'$ is negative definite.
21. Show that the integral

$$\mathbf{W}(0, t) = \int_0^t \mathbf{N}e^{\mathbf{A}'t}\mathbf{BB}'e^{\mathbf{A}'t}\mathbf{N}'\,dt$$

is positive definite for all $t > 0$ if and only if

$$\mathbf{W}^* = \mathbf{N}(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})'\mathbf{N}'$$

is positive definite.
22. Show that the inverse of the controllability Gramian satisfies the nonlinear equation

$$\frac{d}{dt}[\mathbf{W}(t, t_1)]^{-1} = -\mathbf{A}'(t)[\mathbf{W}(t, t_1)]^{-1} - [\mathbf{W}(t, t_1)]^{-1}\mathbf{A}(t) + [\mathbf{W}(t, t_1)]^{-1}\mathbf{B}(t)\mathbf{B}'(t)[\mathbf{W}(t, t_1)]^{-1}$$

23. If the time invariant system

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{Bu}(t)$$

is controllable then show there exists a matrix $\mathbf{C}$ (depending on $\mathbf{b}_i$) such that

$$\dot{\mathbf{x}}(t) = (\mathbf{A} + \mathbf{BC})\mathbf{x}(t) + \mathbf{b}_i v(t)$$

is also controllable where $\mathbf{b}_i$ is any nonzero column of $\mathbf{B}$. (Heymann)
24. Let $\mathbf{A}$ be an n by n matrix. We say that a linear subspace of R^n is *invariant* under $\mathbf{A}$ if every vector $\mathbf{x}$ in that subspace has the property that $\mathbf{Ax}$ also belongs to that subspace. Show that $\det(\mathbf{b}, \mathbf{Ab}, \ldots, \mathbf{A}^{n-1}\mathbf{b}) \neq 0$ if and only if $\mathbf{b}$ does not belong to an invariant subspace ($\neq R^n$) of $\mathbf{A}$.

14. OBSERVABILITY

Observability questions relate to the problem of determining the value of the state vector knowing only the output $\mathbf{y}$ over some interval of time. This is a question of determining when the mapping of the state into the output associates a unique state with every output which can occur.

If we know $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$ and $\mathbf{u}(t)$ for $t_0 \leqslant t \leqslant t_1$, then clearly the output of the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

can be expressed as

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) + \mathbf{y}_1(t)$$

where $y_1(t)$ is the known quantity,

$$\mathbf{y}_1(t) = \int_{t_0}^t \mathbf{C}(t)\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{u}(\sigma)\,d\sigma$$

Thus if we want to ask questions about the determination of $\mathbf{x}(t_0)$, the starting state, based on the output $\mathbf{y}(t)$, we may as well consider the homogeneous system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\,\mathbf{x}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) \qquad (*)$$

provided the system and the input are known.

We see, therefore, that homogeneous problems lead to linear transformations of the form

$$L(\mathbf{x}_0)(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0 = \mathbf{H}(t)\mathbf{x}_0; \qquad t_0 \leqslant t \leqslant t_0$$

with $\mathbf{H}$ being continuous with respect to t. For these transformations we define the *null space* of this mapping as the set of all vectors $\mathbf{x}$ such that $\mathbf{H}(t)\mathbf{x}$ is identically zero on $t_0 \leqslant t \leqslant t_1$. A characterization of the null space is given by the following lemma.

Lemma 1. *Let* $\mathbf{H}$ *be an* m *by* n *matrix whose elements are continuous on the interval* $t_0 \leqslant t \leqslant t_1$. *The null space of the mapping* $L: R^n \to C^m[t_0, t_1]$ *defined by* $L(\mathbf{x}) = \mathbf{Hx}$ *coincides with the null space of*

$$\mathbf{M}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{H}'(t)\mathbf{H}(t)\,dt$$

Proof. If $\mathbf{M}(t_0, t_1)\mathbf{x}_0 = \mathbf{0}$, then $\mathbf{x}_0'\mathbf{M}(t_0, t_1)\mathbf{x}_0 = 0$ and so

$$\int_{t_0}^{t_1} \mathbf{x}_0'\,\mathbf{H}'(t)\mathbf{H}(t)\mathbf{x}_0\,dt = \int_{t_0}^{t_1} \|\mathbf{H}(t)\mathbf{x}_0\|^2\,dt = 0$$

Since $\mathbf{H}$ is assumed to be continuous this means $\mathbf{H}(t)\mathbf{x}_0$ vanishes for all t in the interval and so $\mathbf{Hx}_0 = \mathbf{0}$. On the other hand, if $\mathbf{Hx}_0$ is zero, then $\mathbf{H'Hx}_0$ is zero and hence its integral, which is $\mathbf{M}(t_0, t_1)\mathbf{x}_0$, is zero.

The following theorem is closely related to Theorem 1 of the previous section.

Theorem 1. *Suppose* $\mathbf{A}$, $\mathbf{C}$, *and* $\mathbf{y}$ *are given on the interval* $t_0 \leqslant t \leqslant t_1$ *together with*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

Then it is possible to determine $\mathbf{x}(t_0)$ *to within an additive constant vector* *which lies in the null space of* $\mathbf{M}(t_0, t_1)$ *where*

$$\mathbf{M}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\, dt$$

In particular, it is possible to determine $\mathbf{x}(t_0)$ *uniquely if* $\mathbf{M}(t_0, t_1)$ *is nonsingular. Moreover, it is impossible to distinguish, with a knowledge of* $\mathbf{y}$*, the starting state* $\mathbf{x}_1$ *from the starting state* $\mathbf{x}_2$ *if* $\mathbf{x}_1 - \mathbf{x}_2$ *lies in the null space of* $\mathbf{M}(t_0, t_1)$.

Proof. Starting with the differential equation we obtain

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0)$$

Premultiplication by $\mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)$ gives

$$\mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{y}(t) = \mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0)$$

Integrating this over $t_0 \leqslant t \leqslant t_1$ gives

$$\int_{t_0}^{t_1} \mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{y}(t)\, dt = \mathbf{M}(t_0, t_1)\mathbf{x}(t_0)$$

Since the left side is known the value of $\mathbf{x}(t_0)$ can be determined to within an additive constant vector $\mathbf{x}_1$ which lies in the null space of $\mathbf{M}(t_0, t_1)$. If $\mathbf{M}(t_0, t_1)$ is nonsingular then

$$\mathbf{x}(t_0) = \mathbf{M}^{-1}(t_0, t_1) \int_{t_0}^{t_1} \mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{y}(t)\, dt$$

Suppose that $\mathbf{x}_1 - \mathbf{x}_2$ lies in the null space of $\mathbf{M}(t_0, t_1)$. Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be the corresponding responses. Then

$$\begin{aligned}\int_{t_0}^{t_1} \|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|^2\, dt &= \int_{t_0}^{t_1} \|\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_1 - \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_2\|^2\, dt \\ &= (\mathbf{x}_1 - \mathbf{x}_2)' \int_{t_0}^{t_1} \mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\, dt(\mathbf{x}_1 - \mathbf{x}_2) \\ &= (\mathbf{x}_1 - \mathbf{x}_2)'\mathbf{M}(t_0, t_1)(\mathbf{x}_1 - \mathbf{x}_2) \\ &= 0\end{aligned}$$

Thus $\mathbf{y}_1 = \mathbf{y}_2$ on the given interval, hence $\mathbf{x}_1$ and $\mathbf{x}_2$ are indistinguishable. ▮

The matrix $\mathbf{M}(t_0, t_1)$ plays a role analogous to that of $\mathbf{W}(t_0, t_1)$ introduced in the previous section. We call it the *observability Gramian.*

Theorem 2. *The matrix* $\mathbf{M}$ *defined in the statement of Theorem* 1 *has the following properties*:

(i) $\mathbf{M}(t_0, t_1)$ *is symmetric*
(ii) $\mathbf{M}(t_0, t_1)$ *in nonnegative definite for* $t_1 \geqslant t_0$

(iii) $\mathbf{M}(t_0, t_1)$ *satisfies the linear matrix differential equation*

$$\frac{d}{dt}\mathbf{M}(t, t_1) = -\mathbf{A}'(t)\mathbf{M}(t, t_1) - \mathbf{M}(t, t_1)\mathbf{A}(t) - \mathbf{C}'(t)\mathbf{C}(t);\ \mathbf{M}(t_1, t_1) = \mathbf{0}$$

(iv) $\mathbf{M}(t_0, t_1)$ *satisfies the functional equation*

$$\mathbf{M}(t_0, t_1) = \mathbf{M}(t_0, t) + \mathbf{\Phi}'(t, t_0)\mathbf{M}(t, t_1)\mathbf{\Phi}(t, t_0)$$

Proof. Property (i) follows immediately from the definitions. To establish (ii) we observe that for any real constant vector $\boldsymbol{\eta}$ we have

$$\begin{aligned}\boldsymbol{\eta}'\mathbf{M}(t_0, t_1)\boldsymbol{\eta} &= \int_{t_0}^{t_1} \boldsymbol{\eta}'\mathbf{\Phi}'(t, t_0)\mathbf{C}'(t)\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\boldsymbol{\eta}\, dt \\ &= \int_{t_0}^{t_1} \|\mathbf{C}(t)\mathbf{\Phi}(t, t_0)\boldsymbol{\eta}\|^2\, dt \geqslant 0\end{aligned}$$

To obtain (iii) we merely evaluate the derivative of $\mathbf{M}(t, t_1)$ with respect to its first argument. Using Leibnitz's rule we have

$$\begin{aligned}\frac{d}{dt}\int_{t}^{t_1} &\mathbf{\Phi}'(\sigma, t)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t)\, d\sigma \\ &= -\mathbf{C}'(t)\mathbf{C}(t) + \int_{t}^{t_1} \frac{d}{dt}\mathbf{\Phi}'(\sigma, t)\mathbf{C}'(t)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t)\, d\sigma \\ &= -\mathbf{C}'(t)\mathbf{C}(t) - \mathbf{A}'(t)\int_{t}^{t_1} \mathbf{\Phi}'(\sigma, t)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t)\, d\sigma \\ &\quad - \int_{t}^{t_1} \mathbf{\Phi}'(\sigma, t)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t)\, d\sigma\mathbf{A}(t)\end{aligned}$$

A rearrangement of terms gives the differential equation. The boundary condition is obvious. To establish (iv) we expand the integral defining $\mathbf{M}$ as follows.

$$\begin{aligned}\mathbf{M}(t_0, t_1) &= \int_{t_0}^{t_1} \mathbf{\Phi}'(\sigma, t_0)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t_0)\, d\sigma \\ &= \int_{t_0}^{t} \mathbf{\Phi}'(\sigma, t_0)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t_0)\, d\sigma \\ &\quad + \int_{t}^{t_1} \mathbf{\Phi}'(\sigma, t_0)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t_0)\, d\sigma \\ &= \mathbf{M}(t_0, t) + \mathbf{\Phi}'(t, t_0)\mathbf{M}(t, t_1)\mathbf{\Phi}(t, t_0)\end{aligned}$$ ▮

We now state without proof the observability version of Theorem 3 of Section 13. The proof is easily constructed along the lines of the proof of that theorem.

Theorem 3. *For* $\mathbf{A}$ *and* $\mathbf{C}$ *constant and* $\mathbf{A}$ *n by n, the range space and null space of* $\mathbf{M}(t_0, t_1)$ *for* $t_1 > t_0$ *correspond with the range space and null space of*

$$\mathbf{M}_T = [\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1}]'[\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1}]$$

An n-dimensional linear constant system is called *observable* if the matrix $[\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1}]$ is of rank n.

Example. (**Satellite Problem**) As a continuation of the example of the previous section, we now examine the observability of the linearized equations of a particle near a circular orbit in an inverse square law force field. Assuming that the distance from the center of the force field and the angle can both be measured we have in the notation of Section 2

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}; \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}$$

with y_1 being a radial measurement and y_2 being an angle measurement.

$$(\mathbf{C}; \mathbf{CA}; \mathbf{CA}^2; \mathbf{CA}^3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -2\omega & 0 & 0 \\ 0 & -\omega^2 & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix}$$

This matrix is of rank 4 so that the system is observable. To minimize the measurements required we might consider not measuring y_2. This gives $\mathbf{C}_1 = (1, 0, 0, 0)$ and

$$(\mathbf{C}_1; \mathbf{C}_1\mathbf{A}; \mathbf{C}_1\mathbf{A}^2; \mathbf{C}_1\mathbf{A}^3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \end{bmatrix}$$

which is of rank 3. If y_1 is not measured then $\mathbf{C}_2 = (0, 0, 1, 0)$ and

$$(\mathbf{C}_2; \mathbf{C}_2\mathbf{A}; \mathbf{C}_2\mathbf{A}^2; \mathbf{C}_2\mathbf{A}^3) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix}$$

This matrix is of rank 4. Hence the system is observable from angle measurements alone but not radial measurements alone. It is physically clear that radial measurements alone will not determine a circular orbit. The other conclusion is not so obvious.

Exercises

1. Consider a linear time invariant set of equations $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$; $\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$. Suppose $\mathbf{y}$ can be measured only at particular points in time, say 0, T, $2T, \ldots$. Under what circumstances will sufficiently many such samples serve to determine $\mathbf{x}(0)$ uniquely?
2. Show that if $\mathbf{A}$ has its eigenvalues in the half-plane Re $s < 0$ and

$$\mathbf{Q} = \int_0^\infty e^{\mathbf{A}'t}\mathbf{C}'\mathbf{C}e^{\mathbf{A}t}\,dt$$

then $\mathbf{x}'\mathbf{Q}\mathbf{x}$ is positive for all $\mathbf{x} \neq \mathbf{0}$ if and only if the rank of $(\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1}) = n$ where n is the dimension of $\mathbf{A}$.
3. The variable x is known to satisfy the differential equation

$$\ddot{x}(t) + x(t) = 0$$

and the value of x is known at $t = \pi, 2\pi, 3\pi, \ldots$, etc. Can $x(0)$ and $\dot{x}(0)$ be determined from this data?
4. Suppose that the rank of $(\mathbf{b}, \mathbf{Ab}, \ldots, \mathbf{A}^{n-1}\mathbf{b})$ is n with $\mathbf{A}$ being n by n. Show that if $\mathbf{c}_1(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} = \mathbf{c}_2(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$ then $\mathbf{c}_1 = \mathbf{c}_2$.
5. Show that the linear constant system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$

is controllable if and only if the system

$$\dot{\mathbf{x}}(t) = -\mathbf{A}'\mathbf{x}(t)$$
$$\mathbf{y}(t) = \mathbf{B}'\mathbf{x}(t)$$

is observable.

15. WEIGHTING PATTERNS AND MINIMAL REALIZATIONS

In addition to the possibility of describing a relationship between an input $\mathbf{u}$ and a response $\mathbf{y}$ by the system

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{C}(t) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} \tag{S}$$

there exists the possibility of a description by an integral equation of the form

$$\mathbf{y}(t) = \mathbf{y}_0(t) + \int_{t_0}^{t} \mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\,d\sigma$$

In fact, if $\mathbf{\Phi}$ is the transition matrix associated with system (S) then it implies

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{C}(t)\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)\mathbf{u}(\sigma)\, d\sigma$$

The matrix

$$\mathbf{T}(t, \sigma) = \mathbf{C}(t)\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma) \qquad \text{(WP)}$$

associated with the system (S) is called the *weighting pattern*. A given matrix $\mathbf{T}(\ ,\ )$ is said to be *realizable* as the weighting pattern of a linear finite dimensional dynamical system if there exist matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ such that equation (WP) holds for all pairs (t, σ) with $\dot{\mathbf{\Phi}}(t, t_0) = \mathbf{A}(t)\mathbf{\Phi}(t, t_0)$ and $\mathbf{\Phi}(t_0, t_0) = \mathbf{I}$. The triplet $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ is then called a *realization* of $\mathbf{T}(\ ,\ )$.

It is apparent that if a given weighting pattern has one realization, then it has many. For example, if $\mathbf{P}(t)$ is nonsingular and differentiable for all t and if system (S) is one realization of $\mathbf{T}(\ ,\ )$ then in terms of $\mathbf{z}(t) = \mathbf{P}(t)\mathbf{x}(t)$ we have an alternative realization

$$\begin{bmatrix} \dot{\mathbf{z}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t) & \mathbf{P}(t)\mathbf{B}(t) \\ \mathbf{C}(t)\mathbf{P}^{-1}(t) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{z}(t) \\ \mathbf{u}(t) \end{bmatrix} \qquad \text{(S')}$$

It is instructive to look at the expression for $\mathbf{y}$ in a slightly different way. Since $\mathbf{\Phi}(t, \sigma) = \mathbf{\Phi}(t, t_0)\mathbf{\Phi}(t_0, \sigma)$ we can, assuming $\mathbf{x}_0 = \mathbf{0}$, write

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0) \int_{t_0}^{t} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)u(\sigma)\, d\sigma$$

An interpretation of the sequence of operations on $\mathbf{u}$ which yields $\mathbf{y}$ is obtained by looking at the top row of the following diagram. On the other hand, the same diagram for S′ is obtained by following the lower route. The main point is to see clearly the interrelationships between the mappings in the two *similar* realizations.

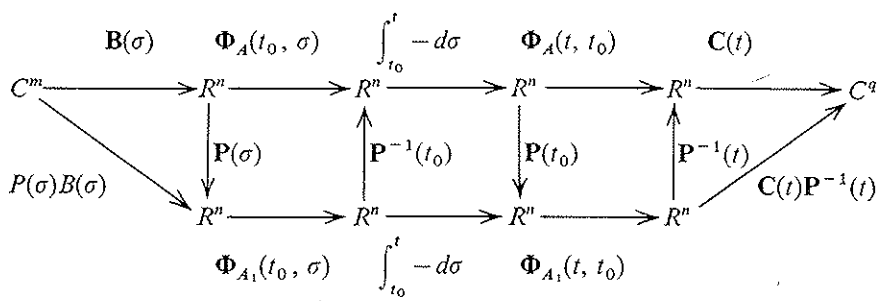


Figure 1. Illustrating alternative realizations. Here $\mathbf{A}_1 = \mathbf{PAP}^{-1} + \dot{\mathbf{P}}\mathbf{P}^{-1}$

Example. Given a time invariant realization $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ it is sometimes of interest to find a second realization $[\mathbf{F}, \mathbf{G}, \mathbf{H}]$ having the property that

$\mathbf{F}$ is symmetric, and $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are bounded. This can always be done. Let $\mathbf{P}(t) = e^{(\mathbf{A}'-\mathbf{A})t/2}$. Then $\mathbf{z}(t) = \mathbf{P}(t)\mathbf{x}(t)$ satisfies

$$\dot{\mathbf{z}}(t) = \tfrac{1}{2}\mathbf{P}(t)(\mathbf{A} + \mathbf{A}')\mathbf{P}'(t)\mathbf{z}(t) + \mathbf{P}(t)\mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{P}'(t)\mathbf{z}(t)$$

It is easy to verify that $\mathbf{F}(t) = \frac{1}{2}\mathbf{P}(t)(\mathbf{A} + \mathbf{A}')\mathbf{P}'(t)$, $\mathbf{G}(t) = \mathbf{P}(t)\mathbf{B}$ and $\mathbf{H} = \mathbf{C}\mathbf{P}'(t)$ meet the symmetry and boundedness constraints.

By focusing on the weighting pattern it is possible to define a type of equivalence relation between various systems with the same input-output characteristics assuming $\mathbf{x}(t_0) = \mathbf{0}$. Moreover, in the design of systems for control or communication purposes it is frequently much more natural to specify a weighting pattern than it is to specify the complete system (S). This point of view makes the following theorem important.

Theorem 1. *A given matrix* $\mathbf{T}(\ ,\ )$ *is realizable as the weighting pattern of a linear finite dimensional dynamical system if and only if there exists a decomposition valid for all t and σ of the form*

$$\mathbf{T}(t, \sigma) = \mathbf{H}(t)\mathbf{G}(\sigma)$$

with $\mathbf{H}$ *and* $\mathbf{G}$ *being finite dimensional matrices.*

Proof. Sufficiency: Suppose the factorization given in the theorem statement is possible. Consider the *degenerate realization* $[\mathbf{0}, \mathbf{G}, \mathbf{H}]$. That is, consider the realization

$$\dot{\mathbf{x}}(t) = \mathbf{G}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t)$$

Clearly this gives

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t_0) + \mathbf{H}(t)\int_{t_0}^{t} \mathbf{G}(\sigma)\mathbf{u}(\sigma)\, d\sigma$$
$$= \mathbf{H}(t)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\, d\sigma$$

Necessity: Suppose there exists a finite dimensional realization in the form of system (S). Then as we have seen

$$\mathbf{T}(t, \sigma) = \mathbf{C}(t)\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)$$

However, if t_0 is any constant we have from the composition law for transition matrices

$$\mathbf{\Phi}(t, \sigma) = \mathbf{\Phi}(t, t_0)\mathbf{\Phi}(t_0, \sigma)$$

Hence if we make the identification

$$\mathbf{H}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0); \qquad \mathbf{G}(\sigma) = \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma) \qquad (*)$$

we see that $\mathbf{H}(t)\mathbf{G}(\sigma) = \mathbf{T}(t, \sigma)$. ∎

It is clear that for a given weighting pattern there exist realizations having state vectors of different dimensions. A trivial example of this is the realization of $\mathbf{T}(\ ,\ )$ obtained from system (S) by adding a vector equation which does not affect the output; e.g.

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \\ & \qquad\qquad\qquad ; \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) \\ \dot{\mathbf{z}}(t) &= \mathbf{F}(t)\mathbf{z}(t) + \mathbf{G}(t)\mathbf{u}(t)\end{aligned}$$

Apparently there can be no upper bound on the dimension of a realization of $\mathbf{T}$. However, under reasonable assumptions there does exist a lower bound.

If system (S) realizes the weighting pattern $\mathbf{T}(\ ,\ )$ it will be called a *minimal realization* if there exists no other realization of $\mathbf{T}(t, \sigma)$ having a lower dimensional state vector. This minimum dimension is called the *order of the weighting pattern.*

Since a lack of controllability indicates a deficiency in the coupling between input and state vector and a lack of observability indicates a deficiency in the coupling between state and output it is reasonable to expect that minimality is related to these ideas. The following theorem shows that this is indeed the case. The proof gives a constructive method of minimizing the dimension of the state vector and is of independent interest.

Theorem 2. *The system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

is a minimal realization of $\mathbf{T}(t, \sigma) = \mathbf{C}(t)\mathbf{\Phi}(t, \sigma)\mathbf{B}(\sigma)$ *if and only if*

$$\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\, d\sigma$$

and

$$\mathbf{M}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Phi}'(\sigma, t_0)\mathbf{C}'(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}(\sigma, t_0)\, d\sigma$$

are both positive definite for some pair (t_0, t_1).

Proof. Sufficiency: We will show that if the realization is not minimal then the matrices $\mathbf{W}$ and $\mathbf{M}$ cannot both be positive definite for any t_1 and t_0. Assume the given realization is not minimal and that

$$\dot{\mathbf{z}}(t) = \mathbf{F}(t)\mathbf{z}(t) + \mathbf{G}(t)\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{H}(t)\mathbf{z}(t)$$

is a lower dimensional realization. For definiteness assume $\dim \mathbf{x} = n$ and $\dim \mathbf{z} = \nu < n$. Clearly the $\mathbf{z}$-realization of $\mathbf{T}(\ ,\ )$ gives rise, via equation (*), to a decomposition

$$\mathbf{T}(t, \sigma) = \mathbf{\Psi}^*(t)\mathbf{\Gamma}^*(\sigma)$$

with $\mathbf{\Psi}^*$ having ν columns and $\mathbf{\Gamma}^*$ having ν rows. Let $\mathbf{\Psi}(t, t_0) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)$ and let $\mathbf{\Gamma}(\sigma, t_0) = \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)$. Then

$$\mathbf{\Psi}(t, t_0)\mathbf{\Gamma}(\sigma, t_0) = \mathbf{\Psi}^*(t)\mathbf{\Gamma}^*(\sigma)$$

and a pre and post multiplication by $\mathbf{\Psi}'(t, t_0)$ and $\mathbf{\Gamma}'(\sigma, t_0)$ respectively gives

$$\mathbf{\Psi}'(t, t_0)\mathbf{\Psi}(t, t_0)\mathbf{\Gamma}(\sigma, t_0)\mathbf{\Gamma}'(\sigma, t_0) = \mathbf{\Psi}'(t, t_0)\mathbf{\Psi}^*(t)\mathbf{\Gamma}^*(\sigma)\mathbf{\Gamma}'(\sigma, t_0)$$

Integrating over the square $t_0 \leqslant t \leqslant t_1$, $t_0 \leqslant \sigma \leqslant t_1$ gives

$$\mathbf{M}(t_0, t_1)\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{\Psi}'(t, t_0)\mathbf{\Psi}^*(t)\, dt \int_{t_0}^{t_1} \mathbf{\Gamma}^*(\sigma)\mathbf{\Gamma}'(\sigma, t_0)\, d\sigma$$

Now from the dimension of $\mathbf{\Psi}^*$ and $\mathbf{\Gamma}^*$ we know the right side has rank less than or equal to $\nu < n$. Hence both $\mathbf{M}(t_0, t_1)$ and $\mathbf{W}(t_0, t_1)$ cannot be positive definite for any t_0 and t_1.

Necessity: Fix t_0 and t_1. Let $\mathbf{H}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)$ and let $\mathbf{G}(t) = \mathbf{\Phi}(t_0, t)\mathbf{B}(t)$. Since the matrices $\mathbf{W}(t_0, t_1)$ and $\mathbf{M}(t_0, t_1)$ are symmetric and positive semidefinite the congruence theorem asserts that there exist nonsingular matrices $\mathbf{P}$ and $\mathbf{Q}$ and signature matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ such that $\mathbf{S}_1^2 = \mathbf{S}_1$, $\mathbf{S}_2^2 = \mathbf{S}_2$ and

$$\mathbf{P}\mathbf{S}_1\mathbf{P}' = \int_{t_0}^{t_1} \mathbf{G}(t)\mathbf{G}'(t)\, dt = \mathbf{W}(t_0, t_1)$$

$$\mathbf{Q}'\mathbf{S}_2\mathbf{Q} = \int_{t_0}^{t_1} \mathbf{H}'(t)\mathbf{H}(t)\, dt = \mathbf{M}(t_0, t_1)$$

It is easy to show that for $t_0 \leqslant t \leqslant t_1$

$$\mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{G}(t) = \mathbf{G}(t)$$

and

$$\mathbf{H}(t)\mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q} = \mathbf{H}(t)$$

To verify the first of these we observe that

$$\begin{aligned}\int_{t_0}^{t_1} &[\mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{G}(t) - \mathbf{G}(t)][\mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{G}(t) - \mathbf{G}(t)]'\, dt \\ &= \mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{W}(t_0, t_1)\mathbf{P}'^{-1}\mathbf{S}_1\mathbf{P}' - \mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{W}(t_0, t_1) - \mathbf{W}(t_0, t_1)\mathbf{P}'^{-1}\mathbf{S}_1\mathbf{P}' \\ &\quad + \mathbf{W}(t_0, t_1) \\ &= \mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{P}\mathbf{S}_1\mathbf{P}'\mathbf{P}'^{-1}\mathbf{S}_1\mathbf{P}' - \mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{P}\mathbf{S}_1\mathbf{P}' - \mathbf{P}\mathbf{S}_1\mathbf{P}'\mathbf{P}'^{-1}\mathbf{S}_1\mathbf{P}' + \mathbf{P}\mathbf{S}_1\mathbf{P}' \\ &= \mathbf{P}\mathbf{S}_1\mathbf{P}' - \mathbf{P}\mathbf{S}_1\mathbf{P}' - \mathbf{P}\mathbf{S}_1\mathbf{P}' + \mathbf{P}\mathbf{S}_1\mathbf{P}' = \mathbf{0}\end{aligned}$$

A similar calculation works for $\mathbf{H}$.

Hence,

$$\mathbf{H}(t)\mathbf{G}(\sigma) = \mathbf{H}(t)\mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q}\mathbf{P}\mathbf{S}_1\mathbf{P}^{-1}\mathbf{G}(\sigma)$$

Using the corollary of the equivalence theorem (Theorem 1) of section 12, we can write $\mathbf{S}_2\mathbf{QPS}_1$ as $\mathbf{N}_1\mathbf{N}_2$ with the number of columns of $\mathbf{N}_1$ equal to the number of rows of $\mathbf{N}_2$ and both equal to the rank of $\mathbf{S}_2\mathbf{QPS}_1$. Clearly the rank of $\mathbf{S}_2\mathbf{QPS}_1$ equals the number of rows in $\mathbf{G}$ if and only if $\mathbf{S}_1$ and $\mathbf{S}_2$ are nonsingular. This is the case if and only if $\mathbf{M}(t_0, t_1)$ and $\mathbf{W}(t_0, t_1)$ are both nonsingular.

This shows that a reduction of the realization is possible for t and σ restricted by $t_0 \leqslant t \leqslant t_1$ and $t_0 \leqslant \sigma \leqslant t_1$. The minimum order of a weighting pattern restricted to a square $|t| \leqslant t_1$, $|\sigma| < t_1$ is clearly a monotone–increasing, integer–valued, function of t_1. Since it is bounded from above, it has a limit and there exists a finite t_1 such that on the square $|t| \leqslant t_1$ the order of the weighting pattern is maximum. Since t_0 and t_1 in the proof above are arbitrary, a reduction is possible if $W(t_0, t_1)$ and $M(t_0, t_1)$ are not positive definite for any t_0 and t_1. ▮

The necessity part of this proof provides an algorithm to reduce the dimension of a nonminimal realization in one step to a minimal realization. The various steps in the reduction procedure are illustrated in the following diagram. Here n_0 is the rank of $\mathbf{S}_2\mathbf{QPS}_1$.

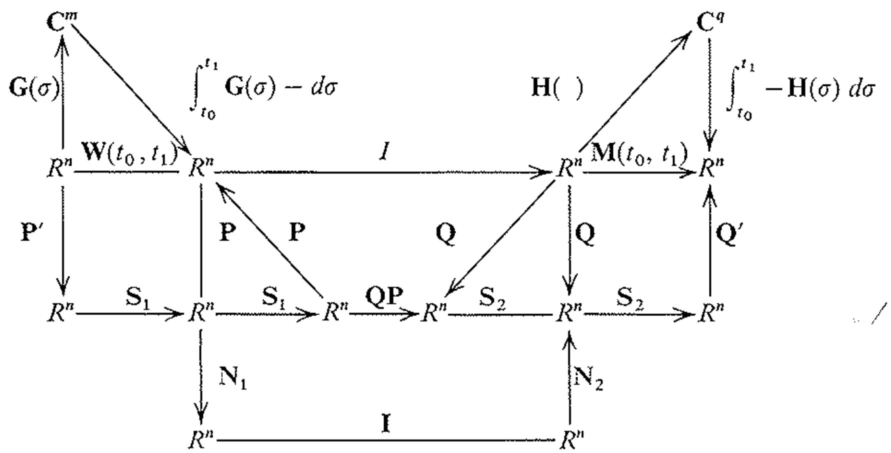


Figure 2. Illustrating the proof of Theorem 2.

We conclude this section with a specialization of this construction to the stationary case.

Theorem 3. *If* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *is a constant realization and if* $\mathbf{W}_T$ *and* $\mathbf{M}_T$ *have the meaning given in Sections* 13 *and* 14, *respectively, then the realization* $[\mathbf{0}, \mathbf{N}e^{-\mathbf{A}t}\mathbf{B}, \mathbf{C}e^{\mathbf{A}t}\mathbf{R}$ *is minimal where* $\mathbf{R}$ *and* $\mathbf{N}$ *satisfy*

$$\mathbf{RN} = \mathbf{Q}^{-1}\mathbf{S}_2\mathbf{QPS}_1\mathbf{P}^{-1}$$

with $\mathbf{Q}$ *and* $\mathbf{P}$ *defined by congruence decompositions*

$$\mathbf{PS}_1\mathbf{P}' = \mathbf{W}_T; \qquad \mathbf{Q}'\mathbf{S}_2\mathbf{Q} = \mathbf{M}_T$$

and $\mathbf{R}$ *and* $\mathbf{N}$ *being* q *by* n_0 *and* n_0 *by* m *where* n_0 *is the rank of* $\mathbf{Q}^{-1}\mathbf{S}_2\mathbf{QPS}_1\mathbf{P}^{-1}$.

Proof. The key step in the reduction part of the necessity proof above now becomes equivalent to establishing

$$\mathbf{PSP}^{-1}e^{-\mathbf{A}\sigma}\mathbf{B} = e^{-\mathbf{A}\sigma}\mathbf{B} \quad \text{and} \quad \mathbf{C}e^{\mathbf{A}t}\mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q} = \mathbf{C}e^{\mathbf{A}t}$$

for the given definitions of $\mathbf{P}$ and $\mathbf{Q}$. However this is equivalent, by the Cayley–Hamilton Theorem, to showing for $i = 0, 1, \ldots, n-1$

$$\mathbf{PSP}^{-1}\mathbf{A}^i\mathbf{B} = \mathbf{A}^i\mathbf{B} \quad \text{and} \quad \mathbf{CA}^i\mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q} = \mathbf{CA}^i$$

This will follow if for $i = 0, 1, \ldots, n-1$

$$(\mathbf{I} - \mathbf{PSP}^{-1})\mathbf{A}^i\mathbf{BB}'\mathbf{A}'^i(\mathbf{I} - \mathbf{PSP}^{-1})' = \mathbf{0}$$

and

$$(\mathbf{I} - \mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q})'\mathbf{A}'^i\mathbf{C}'\mathbf{CA}^i(\mathbf{I} - \mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q}) = \mathbf{0}$$

These matrices are nonnegative definite, hence they vanish if and only if their sum does. However, $\mathbf{W}_T$ is the sum, $\mathbf{BB}' + \mathbf{ABB}'\mathbf{A}' + \cdots + \mathbf{A}^{n-1}\mathbf{BB}'\mathbf{A}'^{n-1}$ and similarly for $\mathbf{M}_T$. Hence establishing the original equalities is equivalent to showing

$$(\mathbf{I} - \mathbf{PSP}^{-1})\mathbf{W}_T(\mathbf{I} - \mathbf{PSP}^{-1})' = \mathbf{0}$$

and

$$(\mathbf{I} - \mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q})'\mathbf{M}_T(\mathbf{I} - \mathbf{Q}^{-1}\mathbf{S}_2\mathbf{Q}) = \mathbf{0}$$

However, both follow immediately from the definitions. The remainder of the proof follows the necessity part of Theorem 2.

Exercises

1. Given a constant matrix $\mathbf{A}$ and a weighting pattern

$$\mathbf{T}(t, \sigma) = \mathbf{H}(t)\mathbf{G}(\sigma)$$

find a finite dimensional linear dynamical realization of the form

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

For example, if $T(t, \sigma) = \cos te^{\sigma}$, find a realization of the form

$$\dot{x}(t) = 2x(t) + b(t)u(t)$$
$$y(t) = c(t)x(t)$$

2. Theorem 1 leaves unanswered the question: "When does there exist a linear finite dimensional system such that a given pair, or set of pairs, of time functions are input-output pairs?" Can you show that there exists no finite dimensional linear system such that $\sin \omega t$ and $\cos \omega t$ are input-output pairs for all ω?
3. Consider a time varying system of the form

$$\ddot{x}(t) + a\dot{x}(t) + bx(t) = \dot{u}(t) + f(t)u(t)$$

with a and b constant and f differentiable. Find a first order vector differential equation representation for the system. Find the weighting pattern relating x and u.
4. The substitution $\dot{z} = \dot{x}^3$ converts the equation

$$\ddot{x} + \dot{x}/3 = u/\dot{x}^2; \qquad y = x$$

into the linear equation

$$\ddot{z} + \dot{z} = 3u$$

Is it possible to express x in terms of z and $\dot{z}$? That is, does there exist a g such that

$$\begin{aligned}\dot{z}_1 &= z_2\\ \dot{z}_2 &= -z_1 + 3u\\ y &= g(z_1, z_2)\end{aligned}$$

is a state representation?
5. Nonlinear realizations of weighting patterns are also sometimes useful. Using polar coordinates, verify that the linear system

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & -a\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}u(t); \qquad y(t) = bx_1(t) + cx_2(t)$$

and the nonlinear system

$$\begin{aligned}\dot{\theta}(t) &= -1 - a[\cos\theta(t)][\sin\theta(t)] + u(t)[\cos\theta(t)]/r(t)\\ \dot{r}(t) &= -ar(t)\sin^2\theta(t) + \sin\theta(t)u(t)\\ y(t) &= b[\cos\theta(t)]r(t) + c[\sin\theta(t)]r(t)\end{aligned}$$

realize the same weighting pattern.

16. STATIONARY WEIGHTING PATTERNS: FREQUENCY RESPONSE

Because of the important role played by constant realizations it is desirable to characterize their weighting patterns as completely as possible. There are two main questions which arise.

(i) Given a weighting pattern, does it have a time invariant realization?
(ii) Given a realization, is its weighting pattern time invariant?

Only the first question will be treated here although some results on the second are discussed in the exercises.

If a weighting pattern has a time invariant realization we call it *stationary*. It turns out that fairly simple necessary and sufficient conditions for a given weighting pattern to be stationary are available.

Theorem 3. *A given matrix* $\mathbf{T}(\ ,\ )$ *has a time invariant, finite dimensional realization*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

if and only if $\mathbf{T}(t, \sigma)$ *can be expressed as* $\mathbf{H}(t)\mathbf{G}(\sigma)$ *with* $\mathbf{H}(t)$ *and* $\mathbf{G}(t)$ *differentiable and*

$$\mathbf{T}(t, \sigma) = \mathbf{T}(t - \sigma, 0)$$

Proof: Necessity: Clearly for time invariant systems we have

$$\mathbf{T}(t, \sigma) = \mathbf{C}e^{\mathbf{A}(t-\sigma)}\mathbf{B} = (\mathbf{C}e^{\mathbf{A}t})(e^{-\mathbf{A}\sigma}\mathbf{B}) = \mathbf{T}(t - \sigma, 0)$$

Let $\mathbf{H}(t) = \mathbf{C}e^{\mathbf{A}t}$ and let $\mathbf{G}(\sigma) = e^{-\mathbf{A}\sigma}\mathbf{B}$. Both $\mathbf{H}$ and $\mathbf{G}$ are differentiable.

Sufficiency: From Theorem 15.2 we know that if a weighting pattern is realizable it has a factorization such that

$$\mathbf{T}(t - \sigma, 0) = \mathbf{T}(t, \sigma) = \mathbf{H}(t)\mathbf{G}(\sigma)$$

with

$$\mathbf{W}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{G}(\sigma)\mathbf{G}'(\sigma)\, d\sigma$$

and

$$\mathbf{M}(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{H}'(t)\mathbf{H}(t)\, dt$$

both positive definite for a suitable t_0 and t_1. Because $\mathbf{T}(t, \sigma) = \mathbf{T}(t - \sigma, 0)$ it follows that

$$\mathbf{0} = \frac{d}{dt}\mathbf{T}(t, \sigma) + \frac{d}{d\sigma}\mathbf{T}(t, \sigma) = \frac{d}{dt}[\mathbf{H}(t)\mathbf{G}(\sigma)] + \frac{d}{d\sigma}[\mathbf{H}(t)\mathbf{G}(\sigma)]$$

This gives

$$\left[\frac{d}{dt}\mathbf{H}(t)\right]\mathbf{G}(\sigma)\mathbf{G}(\sigma) + \mathbf{H}(t)\left[\frac{d}{d\sigma}\mathbf{G}(\sigma)\right]\mathbf{G}'(\sigma) = \mathbf{0}$$

Integrating this equation with respect to σ from t_0 to t_1 yields

$$\frac{d}{dt}[\mathbf{H}(t)]\mathbf{W}(t_0, t_1) + \mathbf{H}(t)\mathbf{W}_1(t_0, t_1) = \mathbf{0}$$

where

$$\mathbf{W}_1(t_0, t_1) = \int_{t_0}^{t_1} \left[\frac{d}{d\sigma}\mathbf{G}(\sigma)\right]\mathbf{G}'(\sigma)\, d\sigma$$

Clearly this gives

$$\mathbf{H}(t) = \mathbf{H}(0)e^{\mathbf{A}t}; \qquad \mathbf{A} = -\mathbf{W}_1(t_0, t_1)\mathbf{W}^{-1}(t_0, t_1)$$

on the other hand

$$\mathbf{H}(t)\mathbf{G}(\sigma) = \mathbf{H}(t-\sigma)\mathbf{G}(0)$$

and since $\mathbf{H}(t) = \mathbf{H}(0)e^{\mathbf{A}t}$

$$\mathbf{H}(t)\mathbf{G}(\sigma) = \mathbf{H}(0)e^{\mathbf{A}(t-\sigma)}\mathbf{G}(0)$$

The triplet $[\mathbf{A}, \mathbf{G}(0), \mathbf{H}(0)]$ is therefore a realization of $\mathbf{T}(t, \sigma)$. ▮

Given a constant nonminimal realization $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ we showed in section 15 that it is possible to minimize it without the calculation of any matrix exponentials or doing any integration. We now show that it is possible to get a *constant* minimal realization with a little additional work.

Theorem 2. *Suppose we are given a minimal realization* $[0, \mathbf{N}e^{-\mathbf{A}t}B, \mathbf{C}e^{\mathbf{A}t}\mathbf{R}]$ *which has a stationary weighting pattern. Then the constant realization* $[\mathbf{NAW}_T\mathbf{N}'(\mathbf{NW}_T\mathbf{N}')^{-1}, \mathbf{NB}, \mathbf{CR}]$ *is also a minimal realization. As before*

$$\mathbf{W}_T = (\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})'$$

Proof. The proof of this theorem proceeds along the lines of the proof of Theorem 16-1. Since the weighting pattern is stationary we have

$$\frac{d}{dt}\mathbf{C}e^{\mathbf{A}t}\mathbf{RN}e^{-\mathbf{A}\sigma}\mathbf{B} = \mathbf{CA}e^{\mathbf{A}t}\mathbf{RN}e^{-\mathbf{A}\sigma}\mathbf{B}$$

Now take partial derivatives of this with respect to σ and evaluate at $\sigma = 0$ to get

$$\dot{\mathbf{H}}(t)\mathbf{N}(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}) = \mathbf{H}(t)\mathbf{NA}(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})$$

Now post-multiply by $(\mathbf{B}, \mathbf{AB}, \ldots \mathbf{A}^{n-1}\mathbf{B})'\mathbf{N}'$ to get

$$\dot{\mathbf{H}}(t)\mathbf{NW}_T\mathbf{N}' = \mathbf{H}(t)\mathbf{NAW}_T\mathbf{N}'$$

This gives

$$\dot{\mathbf{H}}(t) = \mathbf{H}(t)\mathbf{NAW}_T\mathbf{N}'(\mathbf{NW}_T\mathbf{N}^*)^{-1}$$

The rest of the proof follows as in the proof of Theorem 1. ▮

Theorem 15.3 and the above result give a simple reduction procedure for rendering a constant realization minimal. This problem has been particularly stressed by Kalman.

Given that a system is linear and time-invariant there is a second characterization of the input-output properties which is of great importance in both theoretical and applied work.

The matrix $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ of rational functions in complex variables is called the *frequency response* of the linear time-invariant system

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{Bu}(t); \qquad \mathbf{y}(t) = \mathbf{Cx}(t) \tag{S}$$

Justification of this definition is not possible on purely mathematical grounds. It is apparent that $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ is just the Laplace transform of the weighting pattern. On the other hand, when one restricts the complex variable s to the line $s = i\omega$ with $i = \sqrt{-1}$ and $0 \leqslant \omega \leqslant \infty$ it is possible to give the frequency response an interpretation which is much more in keeping with its name.

Theorem 3. *Let* $\mathbf{u}$ *and* $\mathbf{y}$ *be related by the time-invariant system* (S). *Assume that* $\mathbf{A}$ *has no eigenvalues on the line* $\mathrm{Re}[s] = 0$. *Then if* $\mathbf{u}$ *is given by*

$$\mathbf{u}(t) = \mathbf{u}_0 \sin \omega t$$

there is for all ω *a unique periodic response* $\mathbf{y}$ *given by*

$$\mathbf{y}_p(t) = \mathrm{Re}\{\mathbf{C}(\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{Bu}_0\}\sin \omega t + \mathrm{Im}\{\mathbf{C}(\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{Bu}_0\}\cos \omega t$$

Proof. Examining the equation

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{Bu}_0 \sin \omega t \tag{*}$$

and using the corollary of Theorem 9-1, reveals that there exists a unique periodic solution for (*) because the adjoint system

$$\dot{\mathbf{p}}(t) = -\mathbf{A}'\mathbf{p}(t)$$

has no periodic solutions. (Observe $\lambda(\mathbf{A}) = -\lambda(-\mathbf{A}')$ and hence $-\mathbf{A}'$ has no purely imaginary eigenvalues.) Thus, if we can display one periodic solution of (*) we will know it is the only one. Try a solution of the form

$$\mathbf{x}(t) = \mathbf{x}_1 \sin \omega t + \mathbf{x}_2 \cos \omega t$$

Substituting into (*) gives

$$\mathbf{x}_1\omega \cos \omega t - \mathbf{x}_2\omega \sin \omega t = \mathbf{Ax}_1 \sin \omega t + \mathbf{Ax}_2 \cos \omega t + \mathbf{Bu}_0 \sin \omega t$$

or equivalently, the simultaneous equations

$$\mathbf{x}_1\omega - \mathbf{Ax}_2 = \mathbf{0}$$
$$-\mathbf{Ax}_1 - \mathbf{x}_2\omega = +\mathbf{Bu}_0$$

Multiplying the first by $i = \sqrt{-1}$ and adding it to the second gives

$$(\mathbf{x}_1 + i\mathbf{x}_2)i\omega - \mathbf{A}(\mathbf{x}_1 + i\mathbf{x}_2) = \mathbf{B}\mathbf{u}_0$$

or

$$(\mathbf{x}_1 + i\mathbf{x}_2) = (\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}_0$$

and the indicated inverse exists for all real ω. Thus the periodic solution $\mathbf{x}$ is given by

$$\mathbf{x}(t) = \mathrm{Re}(\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{B}u_0 \sin \omega t + \mathrm{Im}(\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{B}u_0 \cos \omega t$$

The expression for $\mathbf{y}_p$ follows if we premultiply by $\mathbf{C}$.

Thus the matrix $\mathbf{C}(\mathbf{I}i\omega - \mathbf{A})^{-1}\mathbf{B}$ can be viewed as a mapping which gives the output as a function of the frequency ω. From a knowledge of it one can, of course, determine $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ for all complex s. A common way of finding the frequency response for a complicated physical system is to let $\mathbf{u}$ be a sinusoid and to measure the resulting output when the transients have disappeared. Naturally this only works if all solutions of the homogeneous system go to zero as t approaches infinity, and must be repeated many times if a characterization over a range of frequencies is desired.

Example. The weighting pattern associated with the differential-difference equation

$$\dot{y}(t) = u(t-1)$$

is $f(t-\sigma)$, where $f(\tau)$ is 1 for τ larger than 1 and zero for $\tau < 1$. The frequency response is e^{-s}/s. There is, however, no finite dimensional realization of this weighting pattern. If there were to be one, there would necessarily be a time invariant realization since the weighting pattern depends only on the difference of the arguments. However, no function of the form $\mathbf{C}e^{\mathbf{A}\tau}\mathbf{B}$ can vanish for $0 \leqslant \tau \leqslant 1$ without vanishing identically everywhere (use Taylor Series), hence no realization exists.

Exercises

1. Find a finite dimensional linear dynamical system which yields a weighting pattern of the form

   (i) $\sin t \sin \sigma$

   repeat for

   (ii) $\sin(t - \sigma)$

   For which of these problems could a time invariant realization be found?

2. Show that a constant realization $[\mathbf{A}, \mathbf{B}, \mathbf{B}']$ with $\mathbf{A} = -\mathbf{A}'$ has a weighting pattern of the form

$$\mathbf{T}(t, \sigma) = \mathbf{G}(t)\mathbf{G}'(\sigma)$$

3. Show that weighting pattern associated with the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + e^{\mathbf{F}t}\mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}e^{-\mathbf{F}t}\mathbf{x}(t)$$

   is stationary if $\mathbf{FA} = \mathbf{AF}$. Calculate it if $\mathbf{FA} = \mathbf{AF}$.

4. Let $p(t)$ be a continuous function and $b(t)$ and $c(t)$ be differentiable. Show that the one dimensional system

$$\dot{x}(t) = p(t)x(t) + b(t)u(t); \qquad y(t) = c(t)x(t)$$

   has a time invariant weighting pattern if there exists a constant λ such that

$$\frac{\dot{c}(t)}{c(t)} + p(t) = \lambda = -\frac{\dot{b}(t)}{b(t)} + p(t)$$

   Under what additional hypothesis is this also a necessary condition?

5. Let $\mathbf{A}$ and $\mathbf{B}$ be infinitely differentiable functions of time and let $D = d/dt$. Show that if the weighting pattern of

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{C}(t) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

   is stationary then $\mathbf{C}(t)[D\mathbf{I} - \mathbf{A}(t)]^i\mathbf{B}(t)$ is constant for $i = 0, 1, 2, \ldots$.

6. Let $\mathbf{F}$ be an n by n matrix and let $\mathbf{C}e^{\mathbf{A}t}\mathbf{B}$ be n by n. Show that $e^{-\mathbf{F}t}\mathbf{C}e^{\mathbf{A}(t-\sigma)}\mathbf{B}e^{\mathbf{F}\sigma}$ is a stationary weighting pattern if and only if $\mathbf{F}^k\mathbf{C}\mathbf{A}^i\mathbf{B} = \mathbf{C}\mathbf{A}^i\mathbf{B}\mathbf{F}^k$ for all k and i positive integers.

7. Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be constant. Consider the system

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}(\sin t)\mathbf{u}(t) \\ \dot{\mathbf{z}}(t) &= \mathbf{A}\mathbf{z}(t) + \mathbf{B}(\cos t)\mathbf{u}(t) \\ \mathbf{y}(t) &= (\sin t)\mathbf{C}\mathbf{x}(t) + (\cos t)\mathbf{C}\mathbf{z}(t) \end{aligned}$$

   (a) Calculate the weighting pattern, and show that it is time invariant.
   (b) Calculate the frequency response function relating $\mathbf{y}$ and $\mathbf{u}$.
   Repeat for the system

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}(\sin t)\mathbf{u}(t) \\ \dot{\mathbf{z}}(t) &= \mathbf{A}\mathbf{z}(t) + \mathbf{B}(\cos t)\mathbf{u}(t) \\ \mathbf{y}(t) &= (\cos t)\mathbf{C}\mathbf{x}(t) - (\sin t)\mathbf{C}\mathbf{z}(t) \end{aligned}$$

8. Let η be defined on the interval $0 \leqslant t \leqslant \infty$ and let it be differentiable and positive. Show that the weighting pattern associated with

$$\dot{z}(t) = [\alpha - \dot{\eta}(t)/\eta(t)]z(t) + [1/\eta(t)]u(t)$$
$$y(t) = \eta(t)z(t)$$

is stationary.

9. Show that the weighting pattern relating i and v for the network shown is stationary if and only if the two resistors are constant.

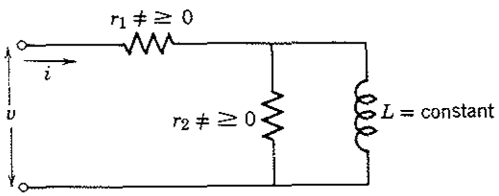


Figure 1.

10. Let L and C be constant in the network shown below. Under what circumstances is the weighting pattern relating i and v stationary?

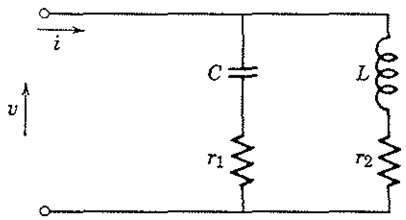


Figure 2.

11. In exercise 7 of section 2 the circuit shown below was introduced

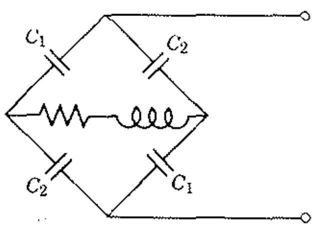


Figure 3.

where $C_1(t) = (1 + \sin t)^{-1}$ and $C_2(t) = (1 - \sin t)^{-1}$. Show that the differential equations relating the voltage v, the current i and the charge

q flowing through the inductor are

$$\ddot{q}(t) + \dot{q}(t) + q(t) = (\sin t)i(t)$$
$$\dot{v}(t) = 2i(t)(\sin t) - i(t)$$

Therefore the weighting pattern relating the voltage to the driving current is

$$T(t, \sigma) = 2(t - \sigma) - \sin t\omega(t - \sigma)\sin \sigma$$

where ω is the inverse Laplace transform of $(s^2 + s + 1)^{-1}$. Now consider putting two such circuits together as shown with $C_3(t) = (1 + \cos t)^{-1}$; $C_4(t) = (1 - \cos t)^{-1}$. In this case the weighting pattern of the circuit on the left is

$$T(t, \sigma) = 2(t - \sigma) - \cos t\omega(t - \sigma)\cos \sigma$$

To get the weighting pattern relating v to i we simply add the weighting patterns. This gives

$$T_1(t, \sigma) = 4(t - \sigma) + (\sin t \sin \sigma + \cos t \cos \sigma)\omega(t - \sigma)$$
$$= 4(t - \sigma) - \cos(t - \sigma)\omega(t - \sigma)$$

Hence conclude that the weighting pattern is stationary.

(Brockett–Skoog)

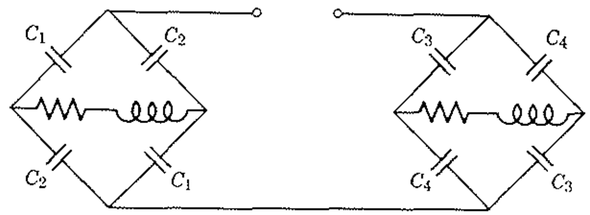


Figure 4.

12. If [**A**, **B**, **C**] is a (possible time varying) realization of a stationary weighting pattern, then show that $[\bar{\mathbf{A}}', \bar{\mathbf{C}}', \bar{\mathbf{B}}']$ realizes the transpose of the weighting pattern where $\bar{A}(t) = A(-t)$, $\bar{B}(t) = B(-t)$, and $\bar{C}(t) = C(-t)$.

(Brockett–Skoog)

17. REALIZATION THEORY FOR TIME-INVARIANT SYSTEMS

We have seen how the matrix $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ arises naturally in the study of time-invariant systems of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \tag{S}$$

Clearly this matrix has elements which are rational functions of s. Moreover, since $C(Is-A)^{-1}B$ has the character of $C(Is)^{-1}B$ as $|s|$ becomes large we see that $C(Is-A)^{-1}B$ approaches zero for large $|s|$ and, hence, that the degree of the denominator of any element of the matrix must exceed the degree of the corresponding numerator. Here we will show that subject to these two constraints *any* matrix of rational functions R can be written in this form. This is called finding a realization of R because if R is given as a desired frequency response, the decomposition $R(s) = C(Is-A)^{-1}B$ gives, via system (S), a first order vector differential equation which yields this frequency response.

To fix notation we assume R is a q by m matrix with elements r_{ij}. Initially we want to establish what was claimed above, regarding the decomposition of R.

Theorem 1. *Given any matrix of rational functions* R, *such that the degree of the denominator of each element exceeds the degree of the numerator of that element, there exist constant matrices* A, B *and* C *such that*

$$C(Is-A)^{-1}B = R(s)$$

Actually we will give two proofs of this because the proofs have independent interest and lead to useful definitions.

Proof 1. Let $p(s)$ be the monic least common multiple of the denominators of the $r_{ij}(s)$. Then $p(s)R(s)$ is a matrix of polynomials and can be written as

$$p(s)R(s) = (R_0 + R_1 s + \cdots + R_{r-1}s^{r-1})$$

where r is the degree of $p(s)$. Let 0_m be the m by m zero matrix and let I_m be the m by m identity matrix. Define A as an rm by rm matrix of the form

$$A = \begin{bmatrix} 0_m & I_m & 0_m & \cdots & 0_m \\ 0_m & 0_m & I_m & \cdots & 0_m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0_m & 0_m & 0_m & \cdots & I_m \\ -p_0 I_m & -p_1 I_m & -p_2 I_m & \cdots & -p_{r-1} I_m \end{bmatrix}$$

where

$$p(s) = s^r + p_{r-1}s^{r-1} + \cdots + p_1 s + p_0$$

and let B and C be given by

$$B = \begin{bmatrix} 0_m \\ 0_m \\ \vdots \\ 0_m \\ I_m \end{bmatrix} \qquad C = [R_0, R_1, \ldots, R_{r-2}, R_{r-1}]$$

Our claim is that with these definitions, $C(Is-A)^{-1}B = R(s)$.

To verify this observe that $(Is-A)^{-1}B$ is the solution $\hat{X}$ of the equation

$$(Is-A)\hat{X} = B$$

If we partition $\hat{X}$ conformably with A and take $A\hat{X}$ to the right, this equation becomes

$$\begin{bmatrix} s\hat{X}_1 \\ s\hat{X}_2 \\ \vdots \\ s\hat{X}_{r-1} \\ s\hat{X}_r \end{bmatrix} - \begin{bmatrix} 0_m & I_m & 0_m & \cdots & 0_m \\ 0_m & 0_m & I_m & \cdots & 0_m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0_m & 0_m & 0_m & \cdots & I_m \\ -p_0 I_m & -p_1 I_m & -p_2 I_m & \cdots & -p_{r-1} I_m \end{bmatrix} \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \\ \vdots \\ \hat{X}_{r-1} \\ \hat{X}_r \end{bmatrix} = \begin{bmatrix} 0_m \\ 0_m \\ \vdots \\ 0_m \\ I_m \end{bmatrix}$$

Thus we see that

$$s\hat{X}_i = \hat{X}_{i+1}; \qquad i = 1, 2, \ldots, r-1$$

and hence from the equation for $s\hat{X}_r$ it follows that

$$(s^r + p_{r-1}s^{r-1} + \cdots p_1 s + p_0)\hat{X}_1 = I_m$$

Since $C(Is-A)^{-1}B = C\hat{X}$ we see using $s\hat{X}_i = \hat{X}_{i+1}$

$$C(Is-A)^{-1}B = [1/p(s)][R_0 + R_1 s + \cdots R_{r-1}s^{r-1}] = R(s)$$

The realization of R given by these equations will be called the *standard controllable realization of* R.

Proof 2. Let $p(s)$ be defined as above and let r be its degree. This time we expand R about $|s| = \infty$ to get

$$R(s) = L_0 s^{-1} + L_1 s^{-2} + L_2 s^{-3} + \cdots$$

Let 0_q be the q by q zero matrix and let I_q be the q by q identity matrix. Define A as an rq by rq matrix of the form

$$A = \begin{bmatrix} 0_q & I_q & 0_q & \cdots & 0_q \\ 0_q & 0_q & I_q & \cdots & 0_q \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0_q & 0_q & 0_q & \cdots & I_q \\ -p_0 I_q & -p_1 I_q & -p_2 I_q & \cdots & -p_{r-1} I_q \end{bmatrix}$$

and let B and C be given by

$$B = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_{r-2} \\ L_{r-1} \end{bmatrix} \qquad C' = \begin{bmatrix} I_q \\ 0_q \\ \vdots \\ 0_q \\ 0_q \end{bmatrix}$$

To verify that this gives a realization, we observe that the expansion of $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ about $|s| = \infty$ is given by

$$\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} = \mathbf{CB}s^{-1} + \mathbf{CAB}s^{-2} + \mathbf{CA}^2\mathbf{B}s^{-3} + \cdots$$

Since for our choice of $\mathbf{C}$ and $\mathbf{A}$ we have

$$\begin{aligned}\mathbf{C} &= (\mathbf{I}_q, \mathbf{0}_q, \ldots, \mathbf{0}_q, \mathbf{0}_q)\\ \mathbf{CA} &= (\mathbf{0}_q, \mathbf{I}_q, \ldots, \mathbf{0}_q, \mathbf{0}_q)\\ &\cdots\cdots\cdots\cdots\\ \mathbf{CA}^{r-1} &= (\mathbf{0}_q, \mathbf{0}_q, \ldots, \mathbf{0}_q, \mathbf{I}_q)\end{aligned}$$

it follows from the definition of $\mathbf{B}$ that

$$\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} = \mathbf{L}_0 s^{-1} + \mathbf{L}_1 s^{-2} + \mathbf{L}_2 s^{-3} + \cdots + \mathbf{L}_{r-1}s^{-r} + \cdots$$

Thus the first r terms in the expansion of $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ agree with those of $\mathbf{R}$. To complete the argument, observe that

$$\det(\mathbf{I}s - \mathbf{A}) = \det\begin{bmatrix} s\mathbf{I}_r - \mathbf{A}_0 & \mathbf{0}_r & \cdots & \mathbf{0}^r \\ \mathbf{0}_r & s\mathbf{I}_r - \mathbf{A}_0 & \cdots & \mathbf{0}^r \\ \cdots & \cdots & \cdots & \mathbf{0}_r \\ \mathbf{0}_r & \cdots & \cdots & s\mathbf{I}_r - \mathbf{A}_0 \end{bmatrix}$$

where

$$\mathbf{A}_0 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -p_0 & -p_1 & \cdot & \cdots & -p_{r-1} \end{bmatrix}$$

therefore $p(s)$ is the minimal polynomial of the matrix $\mathbf{A}$. Hence,

$$p(\mathbf{A}) = \mathbf{0}.$$

Since $p(s)\mathbf{R}(s)$ is a polynomial matrix, the coefficient of s^{-1} must vanish in the expansion

$$p(s)\mathbf{R}(s) = (s^r + p_{r-1}s^{r-1} + \cdots + p_1 s + p_0)(\mathbf{L}_0 s^{-1} + \mathbf{L}_1 s^{-2} + \mathbf{L}_2 s^{-3} + \cdots)$$

Thus we see that

$$p_0\mathbf{CB} + p_1\mathbf{CAB} + \cdots + p_{r-1}\mathbf{CA}^{r-1}\mathbf{B} + \mathbf{CA}^r\mathbf{B} = \mathbf{C}p(\mathbf{A})\mathbf{B} = \mathbf{0}$$

and

$$p_0\mathbf{L}_0 + p_1\mathbf{L}_1 + \cdots + p_{r-1}\mathbf{L}_{r-1} + \mathbf{L}_r = \mathbf{0}$$

This shows that agreement of the first r terms in the expansion (which we have) insures that the $r + 1$ term will agree. If we repeat this, now using the fact that the coefficient of s^{-2} must also be zero then we obtain

$$p_0\mathbf{CAB} + p_1\mathbf{CA}^2\mathbf{B} + \cdots p_{r-1}\mathbf{CA}^r\mathbf{B} + \mathbf{CA}^{r+1}\mathbf{B} = \mathbf{0}$$
$$p_0\mathbf{L}_1 + p_1\mathbf{L}_2 + \cdots p_{r-1}\mathbf{L}_r + \mathbf{L}_{r+1} = \mathbf{0}$$

which implies $\mathbf{CA}^{r+1}\mathbf{B} = \mathbf{L}_{r+1}$. Repeating this we obtain $\mathbf{CA}^i\mathbf{B} = \mathbf{L}_i$ for all nonnegative integers. ▮

This proof leads to a second definition. The realization of $\mathbf{R}(s)$ defined by the choice of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in Proof 2 will be called the *standard observable realization of* $\mathbf{R}$.

By way of justification of our terminology we observe that for the standard controllable realization the controllability condition given in Theorem 13-3 is satisfied i.e.

$$(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{r-1}\mathbf{B}) = \begin{bmatrix} \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m & \mathbf{I}_m \\ \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{I}_m & \sim \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_m & \mathbf{I}_m & \cdots & \sim & \sim \\ \mathbf{I}_m & \sim & \cdots & \sim & \sim \end{bmatrix}$$

is clearly of rank rm. Likewise for the standard observable realization we have the observability condition of Theorem 14-3 fulfilled since

$$(\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{r-1}) = \begin{bmatrix} \mathbf{I}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \mathbf{0}_q \\ \mathbf{0}_q & \mathbf{I}_q & \cdots & \mathbf{0}_q & \mathbf{0}_q \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{I}_q & \mathbf{0}_q \\ \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \mathbf{I}_q \end{bmatrix}$$

is of rank rq. Notice we have truncated both the controllability and observability matrices since we stopped at $\mathbf{A}^{r-1}$ instead of going on to $\mathbf{A}^{mr-1}$ or $\mathbf{A}^{qr-1}$. These additional terms would certainly not decrease the rank however.

Example. (Kalman) To relate the ideas of controllability, observability, minimality and frequency response in one concrete physical problem consider the electrical network shown. Consider the voltage v as being the input and

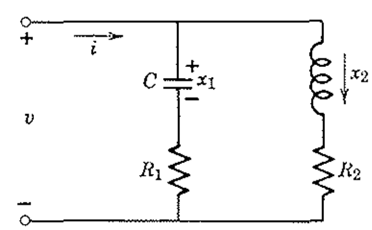


Figure 1. *A constant resistance network for $R_1 = R_2$ and $R_1^2 C = L$.*

consider the current i as being the output. Let x_1, the voltage across the capacitor, and x_2, the current through the inductor, be the state variables. The equations of motion are

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}-1/R_1C & 0\\ 0 & -R_2/L\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}1/R_1C\\ 1/L\end{bmatrix}v(t)$$

$$i(t) = [-1/R_1 \quad 1]\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + (1/R_1)v(t)$$

The frequency response is $\mathbf{c}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} + 1/R_1$ which comes out to be

$$r(s) = \frac{(R_1^2C - L)s + (R_1 - R_2)}{(Ls + R_2)(R_1^2Cs + R_1)} + \frac{1}{R_1}$$

The matrices $(\mathbf{b};\ \mathbf{Ab})$ and $(\mathbf{c};\ \mathbf{cA})$ are

$$(\mathbf{b};\ \mathbf{Ab}) = \begin{bmatrix}1/R_1C & -1/(R_1C)^2\\ 1/L & -R_2/L^2\end{bmatrix}$$

$$(\mathbf{c};\ \mathbf{cA}) = \begin{bmatrix}-1/R_1 & 1\\ 1/R_1^2C & -R_2/L\end{bmatrix}$$

This system

(i) is uncontrollable if $R_1R_2C = L$,
(ii) is unobservable if $R_1R_2C = L$,
(iii) has a constant impedance if $R_1 = R_2$ and $R_1^2C = L$.

If condition (iii) is in force, then the network is said to be a *constant resistance network*. Such networks are quite useful in electrical filter design.

Exercises

1. Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ be a minimal realization of a symmetric frequency response $\mathbf{R}$. (That is, $\mathbf{R}(s) = \mathbf{R}'(s)$ for all s.) Show that $[\mathbf{A}', \mathbf{C}', \mathbf{B}']$ is also a minimal realization of $\mathbf{R}$ and show that there exists a unique, nonsingular, symmetric matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{A}'$, $\mathbf{P}^{-1}\mathbf{B} = \mathbf{C}'$ and $\mathbf{CP} = \mathbf{B}'$. (Youla and Tissi)
2. Show that if $[\mathbf{A}, \mathbf{b}, \mathbf{c}]$ is a realization such that $\mathbf{c}[\mathrm{Adj}(\mathbf{I}s - \mathbf{A})]\mathbf{b} = s^2 + 3s + 2$ and $\det(\mathbf{I}s - \mathbf{A}) = s^3 + 3s^2 - s - 3$, then the realization is not minimal.
3. Two matrices with distinct eigenvalues commute if and only if they have the same set of eigenvectors. Show that $\mathbf{A}$ and $\mathbf{bc}$ do not commute if the system

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{b}u; \qquad y = \mathbf{cx}$$

is minimal. What are the implications for the analysis of the system

$$\dot{\mathbf{x}} = (\mathbf{A} - f(t)\mathbf{bc})\mathbf{x} \qquad (\mathbf{A}, \mathbf{b} \text{ and } \mathbf{c} \text{ constant})$$

obtained by adding time-varying feedback, i.e., do $\mathbf{A} - f(t)\mathbf{bc}$ and its time integral ever commute?
4. Consider a dynamical system

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{b}u; \qquad y = \mathbf{cx} + u$$

show that the system

$$\dot{\mathbf{x}} = (\mathbf{A} - \mathbf{bc})\mathbf{x} + \mathbf{b}y; \qquad u = -\mathbf{cx} + y$$

is *inverse* to the original system in the sense that

$$[\mathbf{c}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} + 1][1 - \mathbf{c}(\mathbf{I}s - \mathbf{A} + \mathbf{bc})^{-1}\mathbf{b}] = 1$$

5. Assume that the single-input system

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{b}u(t)$$
$$\mathbf{y}(t) = \mathbf{Cx}(t)$$

is controllable (but not necessarily observable). Show that there exists a similarity transformation which puts this in standard controllable form. Explain why your proof will not work for a multiple-input system.
6. Write down the conditions on a, b, c and d for the polynomials $s^3 + as^2 + bs + 1$ and $s^2 + cs + d$ not to have common factors.
Hint: Use observability criterion.
7. Consider the network shown below which is made up of n sections of

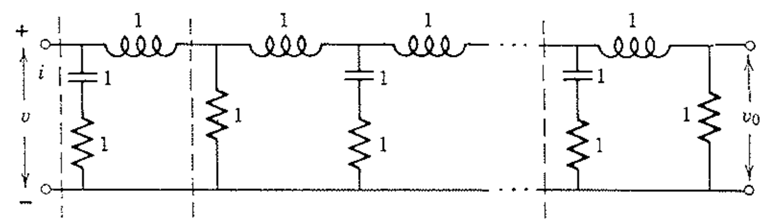


Figure 2.

the constant resistance type each having unit resistance. Calculate the frequency response $\hat{\imath}/\hat{v}$. Calculate the frequency response $\hat{v}_0/\hat{v}$.

18. McMILLAN DEGREE

Given a stationary weighting pattern $\mathbf{T}(\ ,\ )$ or a frequency response $\mathbf{R}$, it is interesting to characterize, as completely as possible, all the time invariant, minimal realizations of it. In particular, we know that all controllable and observable realizations use a state vector of the same dimension; how can this dimension be computed?

In general the dimension of $\mathbf{x}$ in either the standard controllable realization or standard observable realization is not minimal since this would be the case only if the realization simultaneously satisfied a controllability and observability condition. To get at the question of minimal realization we find it convenient to make the following definition.

If $\mathbf{R}$ is a matrix of rational functions each element of which has a denominator whose degree exceeds its numerator, then we will say that k is the *McMillan degree* of $\mathbf{R}$ if $\mathbf{R}$ has a realization $\mathbf{C}(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{B}$ with $\mathbf{A}$ being k by k and no realizations with $\mathbf{A}$ of dimension r by r with $r<k$ exists.

This leads the way to the following important result on linear constant realization.

Theorem 1. *Given any q by m matrix of rational functions $\mathbf{R}$ such that the degree of the denominator of each element exceeds the degree of the numerator of that element, and given that $\mathbf{R}$ has the expansion*

$$\mathbf{R}=\mathbf{L}_0 s^{-1}+\mathbf{L}_1 s^{-2}+\mathbf{L}_2 s^{-3}+\cdots$$

*the McMillan degree k of $\mathbf{R}$ is given by $k=\text{rank}\ \mathbf{H}_{r-1}$ where**

$$\mathbf{H}_i=\begin{bmatrix}\mathbf{L}_0 & \mathbf{L}_1 & \mathbf{L}_2 & \cdots & \mathbf{L}^i\\ \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 & \cdots & \mathbf{L}^{i+1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \mathbf{L}_i & \mathbf{L}_{i+1} & \mathbf{L}_{i+2} & \cdots & \mathbf{L}_{2i}\end{bmatrix}$$

with r being the degree of the least common multiple of the denominators of $\mathbf{R}$.

Proof. First observe that for any realization

$$\mathbf{H}_{r-1}=(\mathbf{C};\mathbf{CA};\ldots;\mathbf{CA}^{r-1})(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{r-1}\mathbf{B})$$

Since the rank of the product of two matrices cannot exceed the rank of either and since neither $(\mathbf{C};\mathbf{CA},\ldots,\mathbf{CA}^{r-1})$ nor $(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{r-1}\mathbf{B})$ can have rank greater than dim $\mathbf{x}$, we see that any realization must be of dimension $n\geqslant \text{rank}\ \mathbf{H}_{r-1}$.

We now show that there is at least one realization of dimension $n=\text{rank}\ \mathbf{H}_{r-1}$. We will obtain it by reducing the n-dimensional standard observable representation of $\mathbf{R}$ via the technique used in the proof of Theorem 15-2. In fact, if

$$\dot{\mathbf{x}}(t)=\mathbf{Ax}(t)+\mathbf{Bu}(t);\qquad \mathbf{y}(t)=\mathbf{Cx}(t)$$

* Matrices whose jkth element depends only on $j+k$ are called *Hankel matrices.* Accordingly, partitioned matrices whose jkth block depends only on $j+k$ are called *generalized Hankel matrices.*

is the standard observable realization of $\mathbf{R}$ then by Theorem 13-3, rank $\mathbf{W}(0,t_1)$ for $t_1>0$ equals rank $(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{n-1}\mathbf{B})$. The reduction to a controllable realization will eliminate all but rank $(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{n-1}\mathbf{B})$ components of the state vector. However, the second phase of the reduction is trivial since we have observability already. Thus, the final state vector has a dimension equal to rank $(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{n-1}\mathbf{B})$. For the standard observable representation, however, one easily sees that $(\mathbf{C};\mathbf{CA};\ldots;\mathbf{CA}^{r-1})=\mathbf{I}$ and so

$$\mathbf{H}_{r-1}=(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{r-1}\mathbf{B})\ \blacksquare$$

When proving results which depend only on the weighting pattern or frequency response it is often convenient to work with a particular realization such as the standard controllable realization. We now show that all time-invariant minimal realizations are equivalent in the sense that we may pass from any one to any other one using a nonsingular time invariant coordinate transformation.

Theorem 2. *Given any two time-invariant minimal realizations of the same frequency response, say*

$$\dot{\mathbf{x}}(t)=\mathbf{Ax}(t)+\mathbf{Bu}(t);\qquad \mathbf{y}(t)=\mathbf{Cx}(t)\qquad (*)$$

and

$$\dot{\mathbf{z}}(t)=\mathbf{Fz}(t)+\mathbf{Gu}(t);\qquad \mathbf{y}(t)=\mathbf{Hz}(t)$$

there exists a constant nonsingular matrix $\mathbf{P}$ such that

$$\mathbf{PAP}^{-1}=\mathbf{F};\quad \mathbf{PB}=\mathbf{G};\qquad \mathbf{CP}^{-1}=\mathbf{H}$$

Proof. Since the two systems realize the same frequency response they have the same weighting pattern and so for all τ, t, and σ

$$\mathbf{C}e^{\mathbf{A}(\tau+t-\sigma)}\mathbf{B}=\mathbf{H}e^{\mathbf{F}(\tau+t-\sigma)}\mathbf{G}$$

Pre and post multiply by $e^{\mathbf{A}'\tau}\mathbf{C}'$ and $\mathbf{B}'e^{-\mathbf{A}'\sigma}$ respectively, to get

$$e^{\mathbf{A}'\tau}\mathbf{C}'\mathbf{C}e^{\mathbf{A}\tau}e^{\mathbf{A}t}e^{-\mathbf{A}\sigma}\mathbf{BB}'e^{-\mathbf{A}'\sigma}=e^{\mathbf{A}'\tau}\mathbf{C}'\mathbf{H}e^{\mathbf{F}\tau}e^{\mathbf{F}t}e^{-\mathbf{F}\sigma}\mathbf{GB}'e^{-\mathbf{A}'\sigma}$$

An integration of this with respect to τ and σ over the square $0\leqslant\tau\leqslant T$, $0\leqslant\sigma\leqslant T$ gives with the obvious definition of $\mathbf{M}_1$ and $\mathbf{W}_1$

$$\mathbf{M}(0,T)e^{\mathbf{A}t}\mathbf{W}(0,T)=\mathbf{M}_1e^{\mathbf{F}t}\mathbf{W}_1$$

using the fact that the $[\mathbf{A},\mathbf{B},\mathbf{C}]$ realization is minimal we see that $\mathbf{M}(0,T)$ and $\mathbf{W}(0,T)$ are invertible. Hence

$$e^{\mathbf{A}t}=\mathbf{M}^{-1}(0,T)\,\mathbf{M}_1e^{\mathbf{F}t}\mathbf{W}_1\mathbf{W}^{-1}(0,T)$$

Since the [A, B, C] and the [F, G, H] realizations are minimal it follows that A and F are of the same dimension. At $t = 0$ we have

$$\mathbf{I} = [\mathbf{M}^{-1}(0, T)\mathbf{M}_1][\mathbf{W}_1\mathbf{W}^{-1}(0, T)]$$

so the pre and post multipliers are inverses of each other. We denote these by $\mathbf{P}^{-1}$ and $\mathbf{P}$ respectively. Using Theorem 1 of Section 5 we have

$$e^{\mathbf{A}t} = \mathbf{P}^{-1}e^{\mathbf{F}t}\mathbf{P} = e^{\mathbf{P}^{-1}\mathbf{F}\mathbf{P}t}$$

Thus we see that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{F}\mathbf{P}$.

To show that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{G}$ we write the weighting pattern identity as

$$\mathbf{C}e^{\mathbf{A}t}\mathbf{B} = \mathbf{H}e^{\mathbf{F}t}\mathbf{G}$$

and premultiply by $e^{\mathbf{A}'t}\mathbf{C}'$. An integration gives

$$\int_0^T e^{\mathbf{A}'t}\mathbf{C}'\mathbf{C}e^{\mathbf{A}t}\, dt\mathbf{B} = \int_0^T e^{\mathbf{A}'t}\mathbf{C}'\mathbf{H}e^{\mathbf{F}t}\, dt\mathbf{G}$$

Hence

$$\mathbf{B} = \mathbf{M}^{-1}(0, T)\mathbf{M}_1\mathbf{G} = \mathbf{P}^{-1}\mathbf{G}$$

Finally, write the weighting pattern identity as

$$\mathbf{C}e^{-\mathbf{A}t}\mathbf{B} = \mathbf{H}e^{-\mathbf{F}t}\mathbf{G}$$

and postmultiply by $\mathbf{B}'e^{-\mathbf{A}'t}$. An integration gives

$$\mathbf{C} = \mathbf{H}\mathbf{W}_1\mathbf{W}^{-1}(0, T) = \mathbf{H}\mathbf{P} \quad \blacksquare$$

The substance of this result is that if [A, B, C] and [F, G, H] are two constant minimal realizations of the same weighting pattern then not only are they of the same dimension, but in addition there exists a constant nonsingular matrix P such that the following diagram commutes

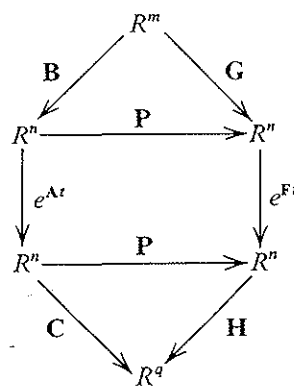


Figure 1. *The generation of all constant minimal realizations from a given minimal realization.*

Example. (Satellite Problem) The problem of controlling and observing a mass near a circular orbit in an inverse square law force field has been examined in Sections 2, 3, 6, 13, and 14. Let's now represent the linearized equations in terms of the frequency response. If we regard x_1 and x_3 as the output vector then (see previous sections)

$$\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B} = \begin{bmatrix} \dfrac{1}{s^2 + \omega^2} & \dfrac{2\omega}{s(s^2 + \omega^2)} \\[2ex] \dfrac{-2\omega}{s(s^2 + \omega^2)} & \dfrac{s^2 - 3\omega^2}{s^2(s^2 + \omega^2)} \end{bmatrix}$$

Notice that the dynamics may appear to be 2nd, 3rd, or 4th order depending on exactly what is regarded as the input and what is regarded as the output. The frequency response matrix makes it clear that the system is not controllable from u_1 since the McMillan degree of the first column is 3 not 4. Likewise it is apparent that the system cannot be observable from y_1 since the McMillan degree of the first row is 3 not 4.

Exercises

1. Suppose that $\mathbf{R}(s)$ has a partial fraction expansion of the form

$$\mathbf{R}(s) = \sum_{i=1}^{r} \mathbf{Z}_i(s + \lambda_i)^{-1}$$

   Show that the McMillan degree is equal to the sum of the ranks of the $\mathbf{Z}_i$.

2. Suppose that $\mathbf{R}(s)$ has the partial fraction expansion

$$\mathbf{R}(s) = \sum_{i=1}^{n} \sum_{j=1}^{\sigma_i} \mathbf{Z}_{ij}(s + \lambda_i)^{-j}$$

   Let $\mathbf{H}_i$ be given by

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{Z}_{i1} & \mathbf{Z}_{i2} & \cdots & \mathbf{Z}_{i\sigma_i} \\ \mathbf{Z}_{i2} & \mathbf{Z}_{i3} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{Z}_{i\sigma_i} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

   Show that the McMillan degree of $\mathbf{R}$ equals the sum of ranks of the $\mathbf{H}_i$. (*Hint:* Consider the special case where $n = 1$ and $\lambda_1 = 0$.)

3. Determine the McMillan degree of

$$\mathbf{R}(s) = \begin{bmatrix} \dfrac{s+1}{s^2 + 2s + 1} & \dfrac{s}{s^2 + 1} \\[2ex] \dfrac{1}{s + 2} & \dfrac{2}{s^2 + 3s + 2} \end{bmatrix}$$

   Find a minimal realization.

4. Let r be a scalar rational function of s with the property that at $|s| = \infty$, $sr(s) = a > 0$. Suppose the poles and zeros of r are real, simple, and interlace; i.e. if $p_1, p_2, \ldots, p_n$ are the values of s at which r has its poles and $z_1, z_2, \ldots, z_{n-1}$ are the values of z at which r has zeros, then with suitable ordering of the p_i and z_i

$$p_1 < z_1 < p_2 \cdots p_{n-1} < z_{n-1} < p_n$$

Show that r has a realization $[\mathbf{A}, \mathbf{b}, \mathbf{b}']$ with $\mathbf{A}$ real and diagonal and $\mathbf{b}$ real. Also show that if $\mathbf{Q}$ is any real symmetric matrix then the poles and zeros of $\mathbf{b}'(\mathbf{I}s - \mathbf{Q})^{-1}\mathbf{b}$ are real and interlace. (*Hint:* As a lemma show that $p + kq$ has real zeros for all $-\infty < k < \infty$ if and only if the poles and zeros of q/p are real and interlace.)

5. Let $r(s)$ be as in Problem (4). Show that the associated Hankel matrix $\mathbf{H}_{r-1}$ as given by Theorem 1 is necessarily positive definite.
6. If we denote by $\delta(\mathbf{R}_i)$ the McMillan degree of $\mathbf{R}_i$ then if $\mathbf{R}_1 + \mathbf{R}_2$ is defined, show that $|\delta(\mathbf{R}_1) - \delta(\mathbf{R}_2)| \leqslant \delta(\mathbf{R}_1 + \mathbf{R}_2) \leqslant \delta(\mathbf{R}_1) + \delta(\mathbf{R}_2)$ and if $\mathbf{R}_1\mathbf{R}_2$ is defined, show that

$$|\delta(\mathbf{R}_1) - \delta(\mathbf{R}_2)| \leqslant \delta(\mathbf{R}_1\mathbf{R}_2) \leqslant \delta(\mathbf{R}_1) + \delta(\mathbf{R}_2).$$

7. Show that if $\mathbf{T}(t, \sigma) = \mathbf{T}(t - \sigma, 0)$ and if $\mathbf{T}(t, \sigma) = \mathbf{H}(t)\mathbf{G}(\sigma)$ then $\mathbf{T}(t, \sigma)$ is analytic in t and σ.

19. FEEDBACK

A central problem in classical control theory is that of evaluating the effects obtained by replacing the input by a linear combination of the input and the output. We want to study some aspects of this question here and will return to it again in Section 34 where the feedback results of Nyquist will be discussed.

We will say that *two realizations* $[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]$ and $[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]$ are in the *same feedback equivalence class* if $\mathbf{B}_1 = \mathbf{B}_2 = \mathbf{B}$, $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}$ and for some matrix $\mathbf{K}$

$$\mathbf{A}_1 - \mathbf{A}_2 = \mathbf{BKC}$$

We will say that *two weighting patterns* $\mathbf{T}_1$ and $\mathbf{T}_2$ *are in the same feedback equivalence class* if they have realizations which are in the same feedback equivalence class.

Notice that if the realizations $[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]$ and $[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]$ are in the same feedback equivalence class and if the realizations $[\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]$ and $[\mathbf{A}_3, \mathbf{B}_3, \mathbf{C}_3]$ are in the same feedback equivalence class, then $[\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]$ and $[\mathbf{A}_3, \mathbf{B}_3, \mathbf{C}_3]$ are also in the same equivalence class since $\mathbf{B} = \mathbf{B}_1 = \mathbf{B}_2 = \mathbf{B}_3$, $\mathbf{C} = \mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}_3$ and the fact that $\mathbf{A}_1 - \mathbf{A}_2 = \mathbf{BK}_{12}\mathbf{C}$ and $\mathbf{A}_2 - \mathbf{A}_3 = \mathbf{BK}_{23}\mathbf{C}$ implies $\mathbf{A}_1 - \mathbf{A}_3 = \mathbf{B}(\mathbf{K}_{12} + \mathbf{K}_{23})\mathbf{C}$. This transitivity is more difficult to prove for the weighting pattern definition since not all realizations are required to be in the same feedback equivalence class.

Theorem 1. *Consider two realizations in the same feedback equivalence class,* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *and* $[(\mathbf{A} - \mathbf{BKC}), \mathbf{B}, \mathbf{C}]$.

(i) *If* $\mathbf{\Phi}_1$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *and* $\mathbf{\Phi}_2$ *is the transition matrix for* $\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{B}(t)\mathbf{K}(t)\mathbf{C}(t)]\mathbf{x}(t)$ *then*

$$\mathbf{\Phi}_2(t, t_0) = \mathbf{\Phi}_1(t, t_0) - \int_{t_0}^{t} \mathbf{\Phi}_1(t, \sigma)\mathbf{B}(\sigma)\mathbf{K}(\sigma)\mathbf{C}(\sigma)\mathbf{\Phi}_2(\sigma, t_0)\, d\sigma \qquad \text{(i)}$$

(ii) *If* $\mathbf{T}_1$ *is the weighting pattern of the realization* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *and* $\mathbf{T}_2$ *is the weighting pattern of* $[(\mathbf{A} - \mathbf{BKC}), \mathbf{B}, \mathbf{C}]$ *then*

$$\mathbf{T}_2(t, t_0) = \mathbf{T}_1(t, t_0) - \int_{t_0}^{t} \mathbf{T}_1(t, \sigma)\mathbf{K}(\sigma)\mathbf{T}_2(\sigma, t_0)\, d\sigma \qquad \text{(ii)}$$

(iii) *Assume the realizations* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *and* $[(\mathbf{A} - \mathbf{BKC}), \mathbf{B}, \mathbf{C}]$ *are time invariant. If* $\mathbf{R}_1$ *is the frequency response of* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ and $\mathbf{R}_2$ *is the frequency response of* $[\mathbf{A} - \mathbf{BKC}, \mathbf{B}, \mathbf{C}]$ *then*

$$\mathbf{R}_2 = \mathbf{R}_1 - \mathbf{R}_1\mathbf{K}\mathbf{R}_2 \qquad \text{(iii)}$$

Proof. (i) Using the variations of constants formula we can write the solution of the matrix equation

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) - \mathbf{B}(t)\mathbf{K}(t)\mathbf{C}(t)\mathbf{X}(t)$$

as

$$\mathbf{X}(t) = \mathbf{\Phi}_1(t, t_0)\mathbf{X}(t_0) - \int_{t_0}^{t} \mathbf{\Phi}_1(t, \sigma)\mathbf{B}(\sigma)\mathbf{K}(\sigma)\mathbf{C}(\sigma)\mathbf{X}(\sigma)\, d\sigma$$

Letting $\mathbf{X}(t_0) = \mathbf{I}$ gives equation (i).

(ii) The weighting patterns $\mathbf{T}_1$ and $\mathbf{T}_2$ are given by

$$\mathbf{T}_1(t, \sigma) = \mathbf{C}(t)\mathbf{\Phi}_1(t, \sigma)\mathbf{B}(\sigma)$$
$$\mathbf{T}_2(t, \sigma) = \mathbf{C}(t)\mathbf{\Phi}_2(t, \sigma)\mathbf{B}(\sigma)$$

Pre and postmultiplication of equation (i) by $\mathbf{C}(t)$ and $\mathbf{B}(\sigma)$ respectively gives

$$\mathbf{C}(t)\mathbf{\Phi}_2(t, \sigma)\mathbf{B}(\sigma) = \mathbf{C}(t)\mathbf{\Phi}_1(t, \sigma)\mathbf{B}(\sigma) - \int_{\sigma}^{t} \mathbf{C}(t)\mathbf{\Phi}_1(t, \rho)\mathbf{B}(\rho)\mathbf{K}(\rho)\mathbf{C}(\rho)\mathbf{\Phi}_2(\rho, \sigma)\mathbf{B}(\sigma)\, d\rho$$

This gives equation (ii).

(iii) In terms of Laplace transform quantities, the above equation in the case $\Phi_1(t, \sigma) = e^{A(t-\sigma)}$ becomes

$$C(Is - A + BKC)^{-1}B = C(Is - A)^{-1}B - C(Is - A)^{-1}BKC(Is - A + BKC)^{-1}B \quad \blacksquare$$

In the case where $C(Is - A)^{-1}B$ is a scalar equation (iii) has a special significance. Let the transfer function associated with $[A, b, c]$ be written as q/p. Then

$$r_2 - q/p = -r_2 q/p$$

or

$$r_2 = q/(p + kq)$$

This is the important formula "forward gain over one plus loop gain" on which so much of classical control theory is based. It has an immediate and nontrivial matrix interpretation, namely

$$c(Is - A)^{-1}b = \frac{\det(Is - A + bc) - \det(Is - A)}{\det(Is - A)}$$

To verify this observe that the poles of $c(Is - A)^{-1}b$ are the zeros of $\det(Is - A)$ and the poles of $c(Is - A + bc)^{-1}b$ are, from the equation for r_2, the zeros of $p + q$.

Theorem 2. *Let* $[A_1, B, C]$ *and* $[A_2, B, C]$ *be two realizations in the same feedback equivalence class. If* W_1 *and* W_2 *satisfy*

$$\frac{d}{dt} W_1(t, t_1) = A_1(t)W_1(t, t_1) + W_1(t, t_1)A_1'(t) - B(t)B'(t); \; W_1(t_1, t_1) = 0$$

$$\frac{d}{dt} W_2(t, t_1) = A_2'(t)W_2(t, t_1) - W_2(t, t_1)A_2(t) - B(t)B'(t); \; W_2(t_1, t_1) = 0$$

Then their range spaces coincide for all t *and* t_1. *If* M_1 *and* M_2 *satisfy*

$$\frac{d}{dt} M_1(t, t_1) = -A_1'(t)M(t, t_1) - M(t, t_1)A_1(t) - C'(t)C(t); \; M_1(t_1, t_1) = 0$$

$$\frac{d}{dt} M_2(t, t_1) = -A_2'(t)M(t, t_1) - M(t, t_1)A_2(t) - C'(t)C(t); \; M_2(t_1, t_1) = 0$$

then their range spaces coincide for all t *and* t_1.

Proof. We can establish this result without calculation. We know from Theorem 13-1 that the range space of $W_1(t, t_1)$ is the set of states at time t from which we can drive the system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t); \qquad y(t) = C(t)x(t) \qquad \text{(S)}$$

to the origin at time t_1. This set is clearly the same for the system

$$\dot{x}(t) = [A(t) - B(t)K(t)C(t)]x(t) + B(t)u(t); \qquad y(t) = C(t)x(t) \qquad \text{(SF)}$$

because if a control u_0 drives the solution of system (S) to the origin along a trajectory x_0 then the control $u_1 = u_0 + BKCx_0$ drives the solution of equation (SF) to zero along exactly the same trajectory. Since the same transfers are possible for systems (S) and (SF) the range space of W_1 and W_2 must coincide.

This result says that the range space of W_1 and W_2 coincide if $A_1 - A_2$ is expressible as BK_1C. Since M and W satisfy the same type of equation we can use this result. Thus the range space of the solutions M_1 and M_2 coincide if $-A_1' + A_2'$ is expressible as $C'K_2B'$. Taking transposes and letting K_1 go into $-K_2'$ gives the result. $\blacksquare$

Corollary. *If two realizations are in the same feedback equivalence class they are either both minimal or both nonminimal.*

Proof. This is an immediate consequence of Theorem 15-2 and the theorem above.

An important final result on time invariant systems is given in Theorem 3. Existing proofs are a little too difficult for inclusion here.

Theorem 3. *If* $[A, B, C]$ *is a time invariant realization and* C *is* n *by* n *and nonsingular, then there exists a constant* K *such that* $\det(Is - A + BKC)$ *has its* n *zeros in any preassigned configuration, provided rank* $(B, AB, \ldots, A^{n-1}B) = n$.

It is appropriate to conclude our discussion of feedback by introducing the block diagram notation which is used so effectively in systems engineering. These are schematics for showing the innerconnection of systems and are particularly effective in dealing with feedback. The basic idea is to reduce all operations to addition and multiplication and to describe these operations by the symbols in figure 1.

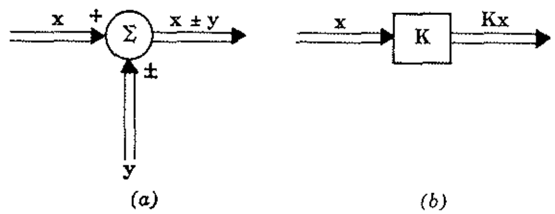


Figure 1. Block diagrams for addition (subtraction) and multiplication.

It is common practice to allow the multiplication to be either in the time domain or the frequency domain. Hence it is common to view the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t)$$

as a block diagram with $s^{-1}\mathbf{I}$ used for integration. (See figure 2.)

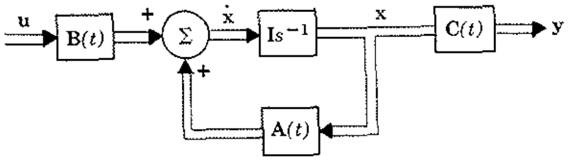


Figure 2. Basic System Configuration.

We will find it useful in our treatment of least squares to express the optimal control $\mathbf{u}_0$ as being generated by a certain time varying matrix operating on the state. The result will then be a configuration of the form shown in figure 3. It is the latter configuration to which the theorems of this section primarily apply.

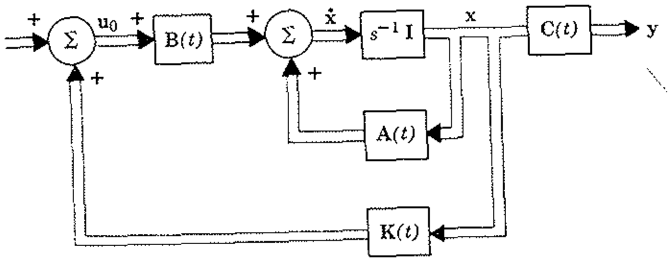


Figure 3. Showing a feedback control.

Exercises

1. Let r be a scalar valued function of a complex variable s. We call the subset of the complex plane given by

$$L = \{\sigma + i\omega : \text{Im } r(\sigma + i\omega) = 0\}$$

the *root locus* of r. We also divide L into two parts

$$L^+ = \{\sigma + i\omega : \text{Im } r(\sigma + i\omega) = 0 \text{ and } \text{Re } r(\sigma + i\omega) \geqslant 0\}$$
$$L^- = \{\sigma + i\omega : \text{Im } r(\sigma + i\omega) = 0 \text{ and } \text{Re } r(\sigma + i\omega) \leqslant 0\}$$

and call L^+ the *positive gain root-locus* and call L^- the *negative gain root-locus*. Show that if $\sigma + i\omega$ lies on the root locus of $\mathbf{c}'(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$

then for some real k, $\sigma + i\omega$ is an eigenvalue of $(\mathbf{A} - \mathbf{b}\mathbf{c}'k)$. Sketch the root locus of r if $r(s) = 1/(s^2 + 1)$ and also $r(s) = s/(s^2 + 1)$.

2. Show that the McMillan degrees of $\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ and $\mathbf{C}(\mathbf{I}s - \mathbf{A} + \mathbf{BKC})^{-1}\mathbf{B}$ are the same.
3. Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ be a time invariant minimal realization. Show that if $\mathbf{K}(t)$ is bounded for $-\infty < t < \infty$ then the Controllability Gramian for $[\mathbf{A} - \mathbf{BK}, \mathbf{B}, \mathbf{C}]$ can be bounded from below uniformly with respect to t_0 by

$$\mathbf{W}(t_0, t_0 + \delta) \geqslant \varepsilon\mathbf{I}$$

for suitable positive ε and δ.
(Silverman-Anderson)

NOTES AND REFERENCES

12. These topics are standard in linear algebra. Most often the normal form for similarity is given in complex form with the eigenvalues on the diagonal. The form given here is easily obtained from that one. The references for Section 1 are appropriate here as well.
13. The paper of Kalman, Ho, and Narendra [42] is a basic reference. The $(\mathbf{B}, \mathbf{AB}, \ldots \mathbf{A}^{n-1}\mathbf{B})$ condition appears in early work on optimal control see e.g. LaSalle [50] and Pontryagin *et al* [67]. We use the term "controllability Gramian" to suggest that the role **W** plays is very much like the conventional Gramian defined in Section 12. Theorems on output controllability appear in Kreindler and Sarachik [47] and Brockett and Mesarovic [17]. The latter reference contains a statement of Theorem 5.
14. Observability in the sense we study it here was introduced by Kalman and used in a series of papers on various aspects of system theory (see Kalman [36], [39], and [40]).
15. We follow the terminology of Weiss and Kalman [77] and Youla [87] in designating **T** as the "weighting pattern," the term *impluse response* is more common. Kalman [36] shows that separability implies realizability. In our development we follow Youla's excellent paper [87] which gives the minimality construction used here. Youla's paper also gives additional results beyond what we have included here.
16. The idea of characterizing a system by Laplace transform and/or its frequency response is standard. The idea of introducing complex numbers to solve for the periodic solution is popularly attributed to the electrical engineer Steinmetz. In our proof of Theorem 1 we follow Youla [87].

17–18. Two important early papers in this area are those of Gilbert [26] and Kalman [39]. These papers relate controllability and observability to minimality and discuss the question of the dimension of the minimal realization. Gilbert points out that in the special case where the denominators of **G** have no common factors a minimal realization can be obtained by elementary methods (partial fractions). The term McMillan degree is used frequently in Electrical Network Theory. It is inspired by a fundamental paper of McMillan [58] on synthesis questions. The relationship between the McMillan degree and Hankel Matrices has been pointed out recently by Youla and Tissi [90] and Ho and Kalman [34].

19. These questions have not received much attention in the literature. Some results along these lines are contained in Brockett [14] and Morgan [60]. See also Silverman and Anderson [73]. The result on the relocation of poles by feedback (Theorem 3) is easy in the case where **B** has one column (see, e.g. [14]) but is much harder in the general case. Popov [68] contains a proof. (See also Wonham [81].) Feedback equivalence classes are discussed in Popov [68].

3
LEAST SQUARES THEORY

The subject of optimal control is concerned with the calculation of a time function **u** which, subject to certain constraints, drives the state of a system $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)$ from an initial value $\mathbf{x}_0$ to some target set and minimizes a loss functional of the form $\int_{t_0}^{t_1} L[\mathbf{x}(t), \mathbf{u}(t)]\, dt + v[\mathbf{x}(t_1)]$. A central result in this area, the maximum principle of Pontryagin, is beyond the scope of this book. However, there is a class of problems of this type which can be treated in a completely satisfactory way using elementary methods. This class, which is characterized by quadratic loss functions and linear systems, is perhaps understood better and used more often than any other in the entire field of optimal control.

The systematic use of least squares methods in systems engineering began around 1940 with the work of Wiener on filtering. The part of his work which was most rapidly incorporated into the engineering literature considered linear time invariant systems and infinite time intervals. For these problems the spectral factorization methods proposed by Wiener were ideal. More recently there has been a considerable effort devoted to unifying his approach with the older methods based on calculus of variations. In this chapter we consider both points of view. This is necessary because neither by itself is really capable of conveying all the essential ideas.

By a *least squares problem* we mean a problem for which one is given

i) $\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{N}(t) \\ \mathbf{N}'(t) & \mathbf{L}(t) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$

(ii) $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{x}(t_0) = \mathbf{x}_0$

(iii) the set to which $\mathbf{x}(t_1)$ must belong.

and asked to find a control function **u** defined on the interval $t_0 \leqslant t \leqslant t_1$ such that **u** minimizes η and meets the boundary conditions. Without loss of generality we can let **L** and **Q** be symmetric.

Although these are problems in the calculus of variations, it is not necessary to view them as such. It is possible to obtain the minimizing function in a direct way using the linearity of the equations of motion and elementary properties of quadratic forms. In fact one can obtain somewhat more information about the answer using direct arguments than one does by a routine application of variational principles. For example, one obtains using a direct argument an expression for the amount by which η is increased when the minimizing control is replaced by an arbitrary second control. We also get automatically conditions under which there actually exists a solution to the minimization problem.

It is important to distinguish between two types of answers to these problems. If we find the optimal $\mathbf{u}$ as a function of $\mathbf{A}, \mathbf{B}, \mathbf{N}, \mathbf{L}, \mathbf{Q}\ \mathbf{x}_0$, the set which $\mathbf{x}(t_1)$ must belong, and t, but not $\mathbf{x}(t)$, then we say we have an *open loop* solution of the least squares problem. On the other hand, if we express the optimal $\mathbf{u}$ as a linear function of the optimal $\mathbf{x}$, $\mathbf{u}(t) = \mathbf{F}(t)\mathbf{x}(t)$, then we say we have a *closed loop* solution of the least squares problem. Of course, this distinction never occurs in the classical calculus of variations. It is important in system theory, however, because open loop and closed loop implementations of an optimal control behave differently in the presence of uncertainty and because open loop and closed loop solutions are implemented in a significantly different way.

In this chapter all controls are assumed to be continuous functions of time. When we claim that a control minimizes an integral we mean that among all controls which are continuous it gives a minimum. In no case are there controls which are discontinuous and known to be superior for the problems we treat.

20. MINIMIZATION IN INNER PRODUCT SPACES

Before considering the main question of this chapter, which concerns choosing functions so as to minimize integrals, it will be helpful to consider carefully some basic results on the minimization of scalar valued functions of many variables. In this section all the results depend on having an inner product space. Recall that in Section 1 we defined inner product spaces and gave several examples. As indicated there, we are most interested in E^n and $C_*^m[t_0, t_1]$, but to avoid duplication of effort we work with an abstract inner product space.

The very first extremization question which arises in an inner product space is the subject of *Schwartz's inequality*.

Theorem 1. *If* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are vectors in an inner product space then*

$$(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle)^2 \leqslant \langle \mathbf{x}_1, \mathbf{x}_1 \rangle \langle \mathbf{x}_2, \mathbf{x}_2 \rangle \qquad (*)$$

and equality holds if and only if $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are proportional.*

Proof. Let μ and λ be real numbers. Then

$$0 \leqslant \langle \mu\mathbf{x}_1 + \lambda\mathbf{x}_2, \mu\mathbf{x}_1 + \lambda\mathbf{x}_2 \rangle = \mu^2 \langle \mathbf{x}_1, \mathbf{x}_1 \rangle + 2\mu\lambda \langle \mathbf{x}_1, \mathbf{x}_2 \rangle + \lambda^2 \langle \mathbf{x}_2, \mathbf{x}_2 \rangle$$

This can be written as

$$[\mu, \lambda] \begin{bmatrix} \langle \mathbf{x}_1, \mathbf{x}_1 \rangle & \langle \mathbf{x}_1, \mathbf{x}_2 \rangle \\ \langle \mathbf{x}_1, \mathbf{x}_2 \rangle & \langle \mathbf{x}_2, \mathbf{x}_2 \rangle \end{bmatrix} \begin{bmatrix} \mu \\ \lambda \end{bmatrix} \geqslant 0$$

A necessary and sufficient condition for this matrix to be nonnegative definite is that the Schwartz inequality hold. Clearly equality holds if $\mathbf{x}_1 = a\mathbf{x}_2$. To prove it does not hold otherwise, notice that if equality holds then

$$[\langle \mathbf{x}_1, \mathbf{x}_1 \rangle + \langle \mathbf{x}_2, \mathbf{x}_2 \rangle] \langle \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2 \rangle = [\langle \mathbf{x}_1, \mathbf{x}_1 \rangle - \langle \mathbf{x}_2, \mathbf{x}_2 \rangle]^2$$

Hence if $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle = \langle \mathbf{x}_2, \mathbf{x}_2 \rangle$ we see that equality demands that $\mathbf{x}_1 = \mathbf{x}_2$. On the other hand, if $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle \neq \langle \mathbf{x}_2, \mathbf{x}_2 \rangle$ we can obtain a contradiction by scaling $\mathbf{x}_2$. ▮

If we regard $\mathbf{x}_2$ as being fixed, it is possible to view Schwartz's inequality as answering the question of how to minimize $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle$ subject to the constraint $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = a$. The answer is to let $\mathbf{x}_1$ be proportional to $\mathbf{x}_2$ and to pick the constant of proportionality so as to meet the constraint. The following theorem is equivalent to Theorem 1.

Theorem 1′. *Let* $\mathbf{x}_2 \neq \mathbf{0}$ *be a given vector in an inner product space and let* a *be a given scalar. Then the value of* $\mathbf{x}$ *which minimizes* $\langle \mathbf{x}, \mathbf{x} \rangle$ *subject to the constraint* $\langle \mathbf{x}, \mathbf{x}_2 \rangle = a$ *is* $\mathbf{x}_1 = a\langle \mathbf{x}_2, \mathbf{x}_2 \rangle^{-1}\mathbf{x}_2$.

Proof. First notice that $\mathbf{x}_1$ meets the constraint. If $\mathbf{x}$ is any other solution of $\langle \mathbf{x}, \mathbf{x}_2 \rangle = a$ then

$$\langle \mathbf{x}, \mathbf{x}_2 \rangle = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = a$$

thus

$$\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$$

multiplying by $a\langle \mathbf{x}_2, \mathbf{x}_2 \rangle^{-1}$ and completing the square gives

$$\langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}_1, \mathbf{x}_1 \rangle = \langle \mathbf{x}_1 - \mathbf{x}, \mathbf{x}_1 - \mathbf{x} \rangle \geqslant 0 \ ▮$$

The geometrical picture which goes along with this idea is that of finding the vector $\mathbf{x}$ joining the origin and a hyperplane* normal to $\mathbf{x}_2$ such that the length of $\mathbf{x}$ is a minimum. The basic device is geometric and consists of constructing a line which is perpendicular to a given line and also passes through a particular point. This construction is important in inner product spaces as a method for determining the minimum distance between a point and a hyperplane which does not contain that point.

Recall that if X is an inner product space with inner product $\langle\ ,\ \rangle_x$ and Y is an inner product space with inner product $\langle\ ,\ \rangle_y$ then in Section 1 we

* A *hyperplane* is the set of all points of the form $\{\mathbf{x} : \langle \mathbf{x}, \mathbf{b} \rangle = c\}$ for some vector $\mathbf{b}$ and some real number c.

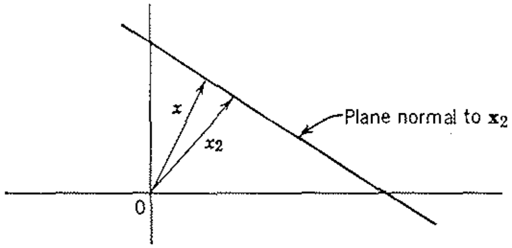


Figure 1. Illustrating Schwartz's inequality in E^2.

agreed to say that a linear transformation T is the adjoint transformation associated with a linear transformation L if for all $\mathbf{x}$ and $\mathbf{y}$

$$\langle \mathbf{y}, L(\mathbf{x})\rangle_y = \langle T(\mathbf{y}), \mathbf{x}\rangle_x$$

We used the symbol L^* for the adjoint of L. Again we emphasize that it is possible for the spaces X and Y to be quite different; the mapping L takes X into Y and the mapping L^* takes Y into X.

The following theorem generalizes Theorem 1′ and although it is by no means the most general result of its type, it is sufficient for our purposes.

Theorem 2. *If $L: X \to Y$ is a linear mapping from an inner product space X into a finite dimensional inner product space Y and if the composite map $LL^*: Y \to Y$ is invertible, then $L(\mathbf{x}) = \mathbf{y}_0$ has a solution*

$$\mathbf{x}_0 = L^*(LL^*)^{-1}(\mathbf{y}_0)$$

Moreover, if $\mathbf{x}_1$ is any other solution of $L(\mathbf{x}) = \mathbf{y}_0$ then

$$\langle \mathbf{x}_1, \mathbf{x}_1\rangle \geqslant \langle \mathbf{x}_0, \mathbf{x}_0\rangle$$

Proof. Let $LL^*(\mathbf{y}_1) = \mathbf{y}_0$; such a $\mathbf{y}_1$ exists because LL^* is invertible. Then $\mathbf{x}_0 = L^*(\mathbf{y}_1)$. Clearly $L\mathbf{x}_0 = LL^*(LL^*)^{-1}\mathbf{y}_0 = \mathbf{y}_0$ so $\mathbf{x}_0$ is a solution. If $\mathbf{x}_1$ is any other solution, then $L(\mathbf{x}_1) - L(\mathbf{x}_0) = \mathbf{0}$ and so

$$\langle \mathbf{y}_1, L(\mathbf{x}_0)\rangle = \langle L^*(\mathbf{y}_1), \mathbf{x}_0\rangle = \langle \mathbf{x}_0, \mathbf{x}_0\rangle = \langle \mathbf{x}_0, \mathbf{x}_1\rangle$$

This implies immediately that

$$\langle \mathbf{x}_0 - \mathbf{x}_1, \mathbf{x}_0\rangle = 0$$

Completing the square gives

$$\langle \mathbf{x}_1, \mathbf{x}_1\rangle - \langle \mathbf{x}_0, \mathbf{x}_0\rangle = \langle \mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_1 - \mathbf{x}_0\rangle \geqslant 0 \blacksquare$$

As a memonic device, the reader may find the following diagram helpful. The spirit of this picture is that although it is impossible to pass directly up from Y to X, because any $\mathbf{y}$ has many inverse images in X, it is possible to

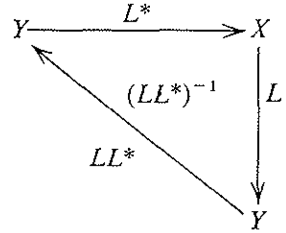


Figure 2. Diagramming the result of Theorem 2

identify the inverse image which minimizes $\|\mathbf{x}\|$ by passing from Y to X via $L^*(LL^*)^{-1}$.

Example. Let $\mathbf{z}$ be an m-vector and let $\mathbf{A}$ be an m by n matrix of rank m. Our problem is to find an n-vector $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{z}$ and $\|\mathbf{x}\|$ is a minimum. Since $\mathbf{A}$ is of rank m, it follows that the m by m matrix $\mathbf{AA}'$ is invertible, and so by the previous theorem

$$\mathbf{x}_0 = \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{z}$$

is the solution with the smallest length. As a slight generalization, let $\mathbf{Q} = \mathbf{N}'\mathbf{N}$ be given and ask for the solution of $\mathbf{Ax} = \mathbf{z}$ which minimizes $\mathbf{x}'\mathbf{Qx}$. If $\mathbf{N}$ is nonsingular (as it will be if $\mathbf{Q}$ is positive definite and not just nonnegative definite), then we can let $\mathbf{Nx} = \mathbf{y}$ and proceed as before. The best $\mathbf{x}$ is then

$$\mathbf{x}_0 = \mathbf{Q}^{-1}\mathbf{A}'(\mathbf{AQ}^{-1}\mathbf{A}')^{-1}\mathbf{z}$$

In addition to these results which describe minimization of $\langle \mathbf{x}, \mathbf{x}\rangle$ subject to linear constraints, there are many problems of interest where one wants to minimize the sum of a quadratic and a linear term. We will use a direct method to prove an elementary but immediately useful result.

Theorem 3. *If $\mathbf{Q} = \mathbf{Q}'$ is positive definite, then for all $\mathbf{x}$ in R^n*

$$\eta = \mathbf{x}'\mathbf{Qx} + 2\mathbf{r}'\mathbf{x} + b \geqslant b - \mathbf{r}'\mathbf{Q}^{-1}\mathbf{r}$$

with equality being achieved if and only if $\mathbf{x}_0 = -\mathbf{Q}^{-1}\mathbf{r}$.

Proof. First notice that since $\mathbf{Q}$ is positive definite, $\mathbf{Q}^{-1}$ exists. If we add and subtract $\mathbf{r}'\mathbf{Q}^{-1}\mathbf{r}$ from η we "complete the square" and obtain

$$\eta = (\mathbf{x} + \mathbf{Q}^{-1}\mathbf{r})'\mathbf{Q}(\mathbf{x} + \mathbf{Q}^{-1}\mathbf{r}) + b - r'\mathbf{Q}^{-1}\mathbf{r}$$

Since $\mathbf{Q}$ is positive definite the first term has a minimum value of 0 and this is achieved only for $\mathbf{x} = -\mathbf{Q}^{-1}\mathbf{r}$. $\blacksquare$

Finally, we describe a result where $\langle \mathbf{x}, \mathbf{x}\rangle$ is to be minimized subject to constraints of the form

$$\langle \mathbf{x}, \mathbf{Qx}\rangle = 1$$

These problems can very often be treated using the Lagrange multiplier technique after existence of minima has been established; again we use a direct method.

Theorem 4. *Let* $\mathbf{Q}$ *be real and symmetric and let* $\lambda_{\min}(Q)$ *and* $\lambda_{\max}(Q)$ *denote its minimum and maximum eigenvalues, respectively. Then for all* $\mathbf{x}$ *in* R^n

$$\lambda_{\min}(\mathbf{Q}) \leqslant \frac{\mathbf{x}'\mathbf{Q}\mathbf{x}}{\mathbf{x}'\mathbf{x}} \leqslant \lambda_{\max}(\mathbf{Q})$$

and there exist values of $\mathbf{x}$ *such that the extremes are achieved.*

Proof. Consider the problem of extremizing $\mathbf{x}'\mathbf{Q}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$ over $\mathbf{x} \neq \mathbf{0}$. Setting the partial derivatives with respect to the components of $\mathbf{x}$ equal to zero gives

$$\mathbf{Q}\mathbf{x} - \mu(\mathbf{x})\mathbf{x} = 0; \qquad \mu(\mathbf{x}) = \mathbf{x}'\mathbf{Q}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$$

Since μ is a scalar we see this necessary condition for a minimum can only be achieved if $\mathbf{x}$ is an eigenvector and μ is an eigenvalue. Since $\mathbf{Q}$ is symmetric all the eigenvalues are real and there is a minimum and a maximum eigenvalue. From the definition of μ we see that the inequality in the theorem statement holds.

Exercises

1. Let $\mathbf{Q}$ be symmetric and nonnegative definite. Show that there exists a minimum for

$$\eta = \mathbf{x}'\mathbf{Q}\mathbf{x} + 2\mathbf{r}'\mathbf{x} + b$$

if $\mathbf{r}$ lies in the range space of $\mathbf{Q}$ and that no minimum exists otherwise. Is the minimizing value of $\mathbf{x}$ unique?

2. Consider the problem of heating steel billets which are passing through a furnace which is segmented into 5 zones. Suppose that the heating costs associated with each zone is proportional to the square of the temperature and suppose the temperature of a billet coming out of a zone is related to its temperature going in according to

$$T_{\text{out}} = T_{\text{in}} + (T_{\text{zone}} - T_{\text{in}})\tfrac{1}{2}$$

If the billets start at 10° and must be heated to 1000°, find the values of T_1, T_2, T_3, T_4 and T_5 which minimize the heating cost. (Inoue–Leftowitz).

3. Find the solution of the vector equation

$$\begin{bmatrix} 3 & 2 & 1 \\ 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which minimizes $x_1^2 + x_2^2 + x_3^2$.

4. Establish the identity

$$\mathbf{x}'\mathbf{Q}^{-1}\mathbf{x} = -\det\begin{bmatrix} 0 & \mathbf{x}' \\ \mathbf{x} & \mathbf{Q} \end{bmatrix}(\det \mathbf{Q})^{-1}$$

5. Show that if $\mathbf{A}$ and $\mathbf{B}$ are positive definite symmetric matrices then $\mathbf{A} - \mathbf{B}$ is positive definite if and only if $\mathbf{B}^{-1} - \mathbf{A}^{-1}$ is positive definite. Show by example that $\mathbf{A} - \mathbf{B}$ positive definite does not imply $\mathbf{A}^2 - \mathbf{B}^2$ positive definite.
6. Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric and positive definite. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$ listed in decreasing order and $s_1, s_2, \ldots, s_n$ are the eigenvalues of $\mathbf{A} + \mathbf{B}$ listed in decreasing order, then $s_i - \lambda_i \geqslant 0$.
7. Let $\mathbf{Q}$ be nonnegative definite. Show that $\mathbf{x}'\mathbf{Q}\mathbf{x}$ vanishes if and only if $\mathbf{Q}\mathbf{x}$ vanishes. (*Hint:* Consider $(\mathbf{x} + \lambda\mathbf{y})'\mathbf{Q}(\mathbf{x} + \lambda\mathbf{y})$ for $-\infty < \lambda < \infty$; $\mathbf{y}'\mathbf{Q}\mathbf{y} = 0$ and $\mathbf{Q}\mathbf{y} \neq 0$.)
8. Let $\mathbf{A}$ be an n by m matrix and assume that for all m by m matrices $\mathbf{C}$ the equation

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = \mathbf{C}$$

can be solved. Find the solution for which tr $\mathbf{Q}'\mathbf{Q}$ is a minimum.

9. Theorem 2 deals with underdetermined sets of linear equations; the problem is to choose an optimal solution from many possible ones. The complementary problem is, in the simplest case, where $\mathbf{A}\mathbf{x} = \mathbf{y}$ has no solutions, but it is desirable to find $\mathbf{x}$ so that $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|$ is a minimum. Show that if $\mathbf{A}$ is m by n and of rank n then the choice $\mathbf{x}_0 = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}$ minimizes $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|$.
10. Given a set of n pairs (x_i, t_i) find the function $f(t) = \alpha t + \beta$ (α and β constants to be determined) such that $\sum_{i=1}^{n} [f(t_i) - x_i]^2$ (use Problem 9).
11. Show that if the integrals

$$\int_0^\infty x^2(t)\, dt \qquad \text{and} \qquad \int_0^\infty \dot{x}^2(t)\, dt$$

converge then $x(t)$ approaches zero as t approaches infinity.
Hint:

$$\tfrac{1}{2}x^2(\sigma)\Big|_{t_0}^{t+t_0} = \int_{t_0}^{t+t_0} \dot{x}(t)x(t)\, dt \leqslant \sqrt{\int_{t_0}^{t+t_0} x^2(t)\, dt \int_{t_0}^{t+t_0} \dot{x}^2(t)\, dt}$$

Also show that if in addition $|\dot{x}(t)| \leqslant M|x(t)|$ for some M and all t then there exists $\gamma > 0$ and $\lambda > 0$ such that $|x(t)| \leqslant \gamma e^{-\lambda t}$.

12. Minimize $\mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{b}'\mathbf{x}$ subject to the constraint $\mathbf{C}\mathbf{x} = \mathbf{0}$ where $\mathbf{C}$ is m by n and of rank m.
13. Let $\mathbf{Q} = \mathbf{Q}'$ be positive definite with minimum and maximum eigenvalues λ_1 and λ_n respectively. Show that for $\|\mathbf{x}\| = 1$

$$1 \leqslant (\mathbf{x}'\mathbf{Q}\mathbf{x})(\mathbf{x}'\mathbf{Q}^{-1}\mathbf{x}) \leqslant \frac{1}{4}\left[\left(\frac{\lambda_1}{\lambda_n}\right)^{1/2} + \left(\frac{\lambda_n}{\lambda_1}\right)^{1/2}\right]^2$$

(Kantorovich)

21. FREE END-POINT PROBLEMS

In this book we will consider explicitly two types of least squares problems: those for which the terminal state $\mathbf{x}(t_1)$ is unconstrained, and those for which the terminal state is completely fixed. The first case is slightly simpler and will be considered in this section. It should be pointed out in advance that the problems considered here are only "solved" to the extent that the optimal trajectories and the optimal controls are computed in terms of a solution of a certain nonlinear matrix differential equation which can be explicitly evaluated only in exceptional cases. Even so, the results play a central role in modern system theory because they give a great deal of structural information and when numbers are desired they are easily translated into computer algorithms.

The problems of interest here are those for which the end point is free and the performance functional is of the form

$$\eta = \int_{t_0}^{t_1} \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{u}(t)\,dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$$

We assume that $\mathbf{L}$ is symmetric but not necessarily positive definite or even nonnegative definite although it will be in many control theory applications. Because $\mathbf{L}$ is not assumed to be nonnegative definite it is not generally true that a minimum for η exists. It turns out that the determination of conditions under which a minimum exists as well as the actual calculation of the minimizing control can be made to depend upon solving a first order matrix equation of the *Riccati* form, i.e. an equation of the form

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{K}(t) - \mathbf{L}(t)$$

Since this equation is nonlinear it is not clear that for a given initial condition matrix $\mathbf{K}_0$, a solution will exist. Moreover, even if solutions do exist for some times, they may fail to exist over longer intervals. To illustrate the typical difficulties which arise, consider $\dot{k}(t) = k^2(t) + 1$; $k(0) = 0$. Separating variables and solving gives $k(t) = \tan t$. The solution passes off to $+\infty$ at $t = \pi/2$ and

hence it does not exist on the interval $0 \leqslant t \leqslant 2$. In the statement of both the theorems of this section it is assumed that it is possible to find solutions for appropriate Riccati equations which do not pass off to infinity on the interval in question. Our notation for the solution of the Riccati equation will be $\mathbf{\Pi}$ and by $\mathbf{\Pi}(t, \mathbf{K}_1, t_1)$ we mean the value at time t of the solution which passes through $\mathbf{K}_1$ at $t = t_1$.

Our treatment rests on the following simple identity.

Lemma 1. *Let* $\mathbf{A}, \mathbf{B}$, *and* $\mathbf{K} = \mathbf{K}'$ *be given matrices. Suppose that* $\dot{\mathbf{K}}$ *exists on the interval* $t_0 \leqslant t \leqslant t_1$. *Then for* $\mathbf{x}$ *and* $\mathbf{u}$ *related by* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$,

$$0 = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)]\begin{bmatrix} 0 & \mathbf{B}'(t)\mathbf{K}(t) \\ \mathbf{K}(t)\mathbf{B}(t) & \dot{\mathbf{K}}(t) + \mathbf{A}'(t)\mathbf{K}(t) + \mathbf{K}(t)\mathbf{A}(t)\end{bmatrix}\begin{bmatrix}\mathbf{u}(t) \\ \mathbf{x}(t)\end{bmatrix} dt - \mathbf{x}'(t)\mathbf{K}(t)\mathbf{x}(t)\Big|_{t_0}^{t_1}$$

Proof. If $\mathbf{x}$ is any differentiable trajectory and if $\mathbf{K}$ is any differentiable matrix, then

$$\int_{t_0}^{t_1} \mathbf{x}'(t)\dot{\mathbf{K}}(t)\mathbf{x}(t) + \dot{\mathbf{x}}'(t)\mathbf{K}(t)\mathbf{x}(t) + \mathbf{x}'(t)\mathbf{K}(t)\dot{\mathbf{x}}(t)\,dt - \mathbf{x}'(t)\mathbf{K}(t)\mathbf{x}(t)\Big|_{t_0}^{t_1} = 0$$

The substitution of $\mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ for $\dot{\mathbf{x}}(t)$ establishes the lemma. ∎

It is not difficult to use this result to get a solution to the free end point least squares problem.

Theorem 1. *Let* $\mathbf{A}, \mathbf{B}, \mathbf{L} = L'$ *and* $\mathbf{Q} = \mathbf{Q}'$ *be given matrices. Suppose there exists on the interval* $t_0 \leqslant t \leqslant t_1$ *a solution* $\mathbf{\Pi}$ *of the differential equation*

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{K}(t) - \mathbf{L}(t); \qquad \mathbf{K}(t_1) = \mathbf{Q}$$

Then there exists a control $\mathbf{u}$ *which minimizes*

$$\eta = \int_{t_0}^{t_1} \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{u}(t)\,dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$$

for the system $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$; $\mathbf{x}(t_0) = \mathbf{x}_0$. *The minimum value of* η *is* $\mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{Q}, t_1)\mathbf{x}(t_0)$. *The minimizing control in closed loop form is*

$$\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{x}(t)$$

whereas in open loop form it is

$$\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{\Phi}(t, t_0)\mathbf{x}_0$$

with $\mathbf{\Phi}$ *being the transition matrix for*

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)]\mathbf{x}(t)$$

Proof. Since $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$, adding the identity of Lemma 1 to η gives

$$\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}(t) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1) + \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)]$$
$$\times \begin{bmatrix} \mathbf{0} & \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1) \\ \mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{B}(t) & \dot{\mathbf{\Pi}}(t, \mathbf{Q}, t_1) + \mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{A}(t) + \mathbf{A}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1) \end{bmatrix}$$
$$\times \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt - \mathbf{x}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{x}(t) \Big|_{t_0}^{t_1}$$

Combining these two integrals and using the differential equation for $\mathbf{\Pi}$ gives

$$\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1) \\ \mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{B}(t) & \mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt$$
$$+ \mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{Q}, t_1)\mathbf{x}(t_0)$$

$$= \int_{t_0}^{t_1} \|\mathbf{u}(t) + \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{x}(t)\|^2 \, dt + \mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{Q}, t_1)\mathbf{x}(t_0)$$

But this expression makes the optimal control obvious. The integrand is nonnegative. Clearly the minimum value of η is $\mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{Q}, t_1)\mathbf{x}(t_0)$ and the minimizing choice of $\mathbf{u}(t)$ is $-\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{x}(t)$. Hence for optimality, $\mathbf{x}$ satisfies

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)]\mathbf{x}(t)$$

If $\mathbf{\Phi}$ is the transition matrix for this equation, then $\mathbf{u}$ is given by $\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_1)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0)$. ∎

It is of some interest to visualize the closed loop optimal control as a feedback diagram. This is illustrated in Figure 1.

Example. Given that $x(0) = 1$ find x on the interval $0 \leqslant t \leqslant T$ such that

$$\eta = \int_0^T \dot{x}^2(t) + x^2(t)\, dt$$

is a minimum. To convert this into a control problem set $\dot{x} = u$. The Riccati

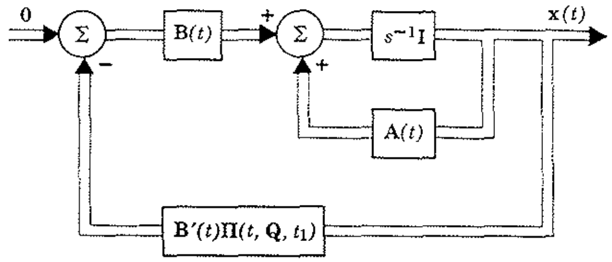


Figure 1. A feedback visualization of the free end-point problem.

equation is then

$$\dot{k}(t) = k^2(t) - 1$$

The solution which passes through zero at T is $\tanh(T - t)$. Hence the optimum x satisfies

$$\dot{x}(t) = -\tanh(T - t)x(t)$$

This integrates to give

$$x(t) = \cosh t - \left(\frac{\sinh T}{\cosh T}\right) \sinh t$$

$$u(t) = -\tanh(T - t)x(t)$$

The minimum value of η is $x(0)\pi(0, 0, T)x(0) = x^2(0)\tanh T$.

There is one special case in which the solution of the Riccati equation is directly interpretable in terms of the controllability Gramian. Suppose $\mathbf{L}$ is $\mathbf{0}$. Then the Riccati equation is

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{K}(t)$$

Now if $\mathbf{K}$ has an inverse we know that $\frac{d}{dt}\mathbf{K}^{-1}(t) = -\mathbf{K}^{-1}(t)\dot{\mathbf{K}}(t)\mathbf{K}^{-1}(t)$ and hence it follows from the above equation that

$$\frac{d}{dt}\mathbf{K}^{-1}(t) = \mathbf{K}^{-1}(t)\mathbf{A}'(t) + \mathbf{A}(t)\mathbf{K}^{-1}(t) - \mathbf{B}(t)\mathbf{B}'(t)$$

Notice that from Theorem 13.2 we know that $\mathbf{W}$ satisfies an identical linear matrix equation. The matrix variation of constants formula gives immediately

$$\mathbf{\Pi}(t, \mathbf{Q}, t_1) = [\mathbf{W}(t, t_1) + \mathbf{\Phi}(t, t_1)\mathbf{Q}^{-1}\mathbf{\Phi}'(t, t_1)]^{-1}$$

provided the indicated inverses exist. This calculation yields the following theorem.

Theorem 2. *Let* **A**, **B**, *and* $\mathbf{Q} = \mathbf{Q}'$ *be given matrices. Let* **W** *be the controllability Gramian for the system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

If the matrix

$$\mathbf{H}(t, t_1) = \mathbf{W}(t, t_1) + \mathbf{\Phi}(t, t_1)\mathbf{Q}^{-1}\mathbf{\Phi}'(t, t_1)$$

is invertible for all t in the interval $t_0 \leqslant t \leqslant t_1$ *then there exists a control which minimizes*

$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t)\,dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$$

As in all aspects of linear systems theory, the effects of changing coordinates needs to be considered in order to have a reasonably complete picture. For the least squares problem, what is the effect of changes of variable on the structure of the Riccati equation? In what coordinate system does it take the "simplest" form? We conclude this section with a study of these questions. Consider making the change of variable $\mathbf{z}(t) = \mathbf{P}(t)\mathbf{x}(t)$ with **P** nonsingular and differentiable. As we have seen (Section 4), the dynamics of Theorem 2 assume the form

$$\dot{\mathbf{z}}(t) = [\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]\mathbf{z}(t) + \mathbf{P}(t)\mathbf{B}(t)\mathbf{u}(t)$$

and η becomes

$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{z}'(t)\mathbf{P}'^{-1}(t)\mathbf{L}(t)\mathbf{P}^{-1}(t)\mathbf{z}(t)\,dt + \mathbf{z}'(t_1)\mathbf{P}'^{-1}(t_1)\mathbf{Q}\mathbf{P}^{-1}(t_1)\mathbf{z}(t_1)$$

Hence the Riccati equation for the new system is

$$\begin{aligned}\dot{\mathbf{K}}_1(t) = &-[\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]'\mathbf{K}_1(t) - \mathbf{K}_1(t)[\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) \\ &+ \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)] + \mathbf{K}_1(t)\mathbf{P}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{P}'(t)\mathbf{K}_1(t) - \mathbf{P}'^{-1}(t)\mathbf{L}(t)\mathbf{P}^{-1}(t)\end{aligned}$$

A particular case of interest is to let $\mathbf{P}(t) = \mathbf{\Phi}(t_0, t)$ in which case the linear terms in **K** vanish and the Riccati equation takes the form

$$\dot{\mathbf{K}}(t) = \mathbf{K}(t)\mathbf{G}(t)\mathbf{G}'(t)\mathbf{K}(t) - \mathbf{H}'(t)\mathbf{H}(t)$$

The entire set of transformational properties of the least squares problems can be read off the diagram in Figure 2. In this diagram **K**, $\dot{\mathbf{K}}$, etc. are to be regarded as linear transformations on R^n even though in some respects this is unnatural.

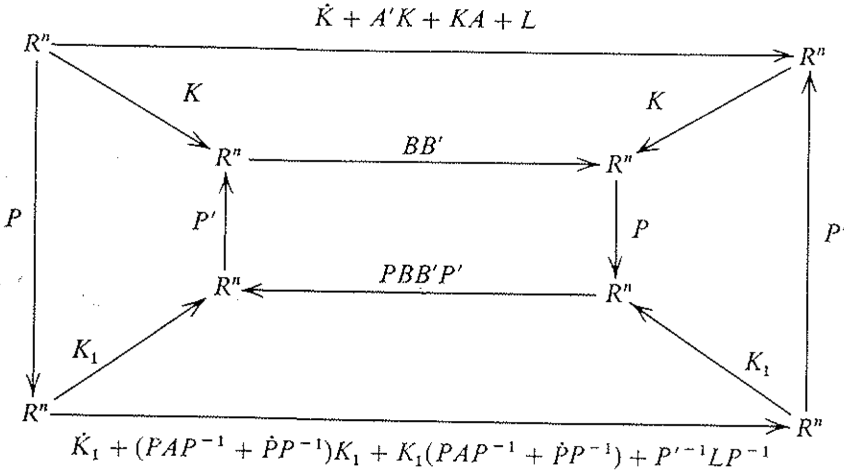


Figure 2. The effects of a transformation $\mathbf{z}(t) = \mathbf{P}(t)\mathbf{x}(t)$ *on the Riccati equation for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$; $\eta(t) = \mathbf{u}'(t)\mathbf{u}(t) + x'(t)\mathbf{L}(t)\mathbf{x}(t)$.

Exercises

1. Consider minimizing the quantity

$$\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{N}(t) \\ \mathbf{N}'(t) & \mathbf{L}(t) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt$$

for the system $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$. Show that the relevant Riccati equation is

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) - \mathbf{L}(t) + [\mathbf{N}(t) + \mathbf{B}'(t)\mathbf{K}(t)]'[\mathbf{N}(t) + \mathbf{B}'\mathbf{K}(t)]$$

and prove an analog of Theorem 1.

2. Show that if $\mathbf{\Pi}$ is a solution of the time invariant Riccati equation

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A} + \mathbf{K}(t)\mathbf{B}\mathbf{B}'\mathbf{K}(t) - \mathbf{L}$$

then $e^{-\mathbf{A}'t}\mathbf{\Pi}(t, \mathbf{Q}, t_1)e^{-\mathbf{A}t}$ satisfies the time varying Riccati equation

$$\dot{\mathbf{K}}(t) = \mathbf{K}(t)e^{\mathbf{A}t}\mathbf{B}\mathbf{B}'e^{\mathbf{A}'t}\mathbf{K}(t) - e^{-\mathbf{A}'t}\mathbf{L}e^{-\mathbf{A}t}$$

3. Show that there exists an initial state $\mathbf{p}(t_0)$ for the adjoint equation $\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$ such that the control u_0 which minimizes

$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t)\,dt + \mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$$

for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ is given by

$$\mathbf{u}_0(t) = -\mathbf{B}'(t)\mathbf{p}(t)$$

4. Consider the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

and suppose that we want to keep $\mathbf{x}(t)$ close to the trajectory $\mathbf{d}(t)$, where $\mathbf{d}(t)$ is a solution of the homogeneous equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$. To do this, we define the cost functional

$$J = \int_0^T \mathbf{u}'(t)\mathbf{u}(t) + [\mathbf{x}(t) - \mathbf{d}(t)]'\mathbf{L}(t)[\mathbf{x}(t - \mathbf{d}(t)]\, dt$$

Show that the control which minimizes this has the form

$$\mathbf{u}^*(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{g}(t)$$

What are $\mathbf{F}(t)$ and $\mathbf{g}(t)$?

5. Let $\mathbf{A}$ and $\mathbf{B}$ be constant matrices and let

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$

be a controllable system. Assume $\mathbf{A}$ has no eigenvalues with zero real parts. Show that if $\mathbf{K}$ is defined as

$$\mathbf{K} = \lim_{t\to\infty} [\mathbf{W}(0, t)]^{-1}$$

then $\mathbf{A} - \mathbf{BB}'\mathbf{K}$ has all its eigenvalues in Re $s < 0$.

6. A *pulse width modulator* is a device which operates on continuous functions whose amplitude is between plus and minus one to produce piecewise constant functions y according to the rule

$$y = \begin{cases} \operatorname{sgn} x(nT); & 0 \leqslant t - nT \leqslant Tx(nT) \\ 0; & Tx(nT) < t - nT \leqslant T \end{cases}$$

Show that for a suitable choice of β

$$\int_0^T ty^2(t)\, dt \leqslant \beta \int_0^T u^2(t) + \dot{u}^2(t)\, dt$$

22. FIXED END-POINT PROBLEMS

If the control $\mathbf{u}$ is required to drive a system to a particular state at the end of the interval $t_0 \leqslant t \leqslant t_1$ or if the state at time t_1 is required to belong to a certain set, then the methods of the previous section must be modified. We consider explicitly only the case where the terminal state is completely fixed, but various extensions are possible within the framework presented here. (See the Exercises.) As it turns out, the fixed end point problems are generally more difficult than those of the previous section in that not only is one required to solve a Riccati equation, but in addition a certain controllability Gramian must be evaluated. We begin with a simple case where L is zero.

Theorem 1. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be given matrices and let* $\mathbf{W}$ *be the controllability Gramian for the system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

If $\mathbf{u}_0$ *is any control of the form*

$$\mathbf{u}_0(t) = -\mathbf{B}'(t)\mathbf{\Phi}'(t_0, t)\xi$$

where ξ *satisfies*

$$\mathbf{W}(t_0, t_1)\xi = \mathbf{x}_0 - \mathbf{\Phi}(t_0, t_1)\mathbf{x}_1$$

then the control $\mathbf{u}_0$ *drives the system from* $\mathbf{x}_0$ *at* $t = t_0$ *to* $\mathbf{x}_1$ *at* $t = t_1$, *and if* $\mathbf{u}_1$ *is any other control which drives the system from* $\mathbf{x}_0$ *at* $t = t_0$ *to* $\mathbf{x}_1$ *at* $t = t_1$ *then*

$$\int_{t_0}^{t_1} \mathbf{u}_1'(t)\mathbf{u}_1(t)\, dt \geqslant \int_{t_0}^{t_1} \mathbf{u}_0'(t)\mathbf{u}_0(t)\, dt$$

Moreover, if $\mathbf{W}(t_0, t_1)$ *is nonsingular then*

$$\int_{t_0}^{t_1} \mathbf{u}_0'(t)\mathbf{u}_0(t)\, dt = [\mathbf{x}_0 - \mathbf{\Phi}(t_0, t_1)\mathbf{x}_1]'\mathbf{W}^{-1}(t_0, t_1)[\mathbf{x}_0 - \mathbf{\Phi}(t_0, t_1)\mathbf{x}_1]$$

Proof. By assumption $\mathbf{u}_1$ achieves the desired transfer and so

$$\mathbf{x}_1 = \mathbf{\Phi}(t_1, t_0)\left[\mathbf{x}(t_0) + \int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{u}_1(\sigma)\, d\sigma\right]$$

As was shown in the proof of Theorem 13-1, $\mathbf{u}_0$ also achieves the desired transfer. Thus

$$\mathbf{x}_1 = \mathbf{\Phi}(t_1, t_0)\left[\mathbf{x}(t_0) + \int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{u}_0(\sigma)\, d\sigma\right]$$

Subtracting these two different equations for $\mathbf{x}_1$ gives

$$\int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)[\mathbf{u}_1(\sigma) - \mathbf{u}_0(\sigma)]\, d\sigma = \mathbf{0}$$

The pre-multiplication of this equation by ξ' together with the definition of $\mathbf{u}_0$ gives

$$\int_{t_0}^{t_1} \mathbf{u}_0'(\sigma)[\mathbf{u}_1(\sigma) - \mathbf{u}_0(\sigma)]\, d\sigma = 0$$

This equality permits one to verify directly that

$$\int_{t_0}^{t_1} \mathbf{u}_1'(\sigma)\mathbf{u}_1(\sigma) - \mathbf{u}_0'(\sigma)\mathbf{u}_0(\sigma)\, d\sigma = \int_{t_0}^{t_1} [\mathbf{u}_1(\sigma) - \mathbf{u}_0(\sigma)]'[\mathbf{u}_1(\sigma) - \mathbf{u}_0(\sigma)]\, d\sigma$$

Since the right side is clearly nonnegative, this completes the proof of the optimality of $\mathbf{u}_0$.

Observe that in the event that $W(t_0, t_1)$ has an inverse (and this is probably the situation of greatest interest) one has

$$\begin{aligned}\int_{t_0}^{t_1} \mathbf{u}_0'(t)\mathbf{u}_0(t)\,dt &= \xi' \int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{B}'(\sigma)\mathbf{\Phi}'(t_0, \sigma)\,d\sigma\xi \\ &= \xi'\mathbf{W}(t_0, t_1)\xi \\ &= [\mathbf{x}_0 - \mathbf{\Phi}(t_0, t_1)\mathbf{x}_1]'\mathbf{W}^{-1}(t_0, t_1)[\mathbf{x}_0 - \mathbf{\Phi}(t_0, t_1)\mathbf{x}_1] \quad \blacksquare\end{aligned}$$

Notice that the proof gives additional information beyond what is given in the theorem statement; i.e. it gives an expression for the amount by which the right side of inequality

$$\int_{t_0}^{t_1} \mathbf{u}_1'(t)\mathbf{u}_1(t)\,dt \geqslant \int_{t_0}^{t_1} \mathbf{u}_0'(t)\mathbf{u}_0(t)\,dt$$

exceeds the left.

The structure of the optimal feedback system is shown in figure 1.

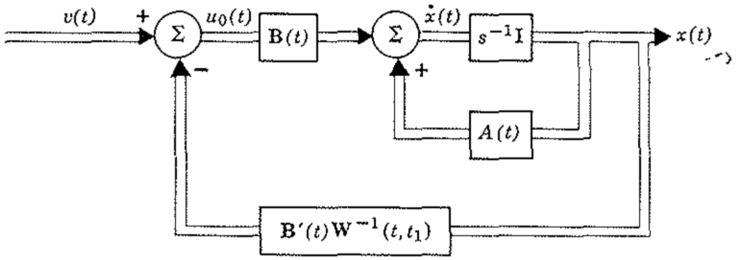


Figure 1. The optimal feedback control for a fixed end-point problem. Here $\mathbf{v}(t) = \mathbf{B}'(t)\mathbf{W}^{-1}(t, t_1)\mathbf{\Phi}(t, t_1)\mathbf{x}_1$

We could have used Theorem 2 of Section 20 to prove Theorem 1. To do so observe that it is necessary to solve

$$L(\mathbf{u}) = \int_{t_0}^{t_1} \mathbf{\Phi}(t_0, \sigma)\mathbf{B}(\sigma)\mathbf{u}(\sigma)\,d\sigma = [\mathbf{\Phi}(t_0, t_1)\mathbf{x}_1 - \mathbf{x}_0]$$

Now $L(\ )$ defines a linear mapping of $C_*^m[t_0, t_1]$ into E^n. Since $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \mathbf{x}_1'\mathbf{x}_2$ and

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \int_{t_0}^{t_1} \mathbf{u}_1'(\sigma)\mathbf{u}_2(\sigma)\,d\sigma$$

are the inner products in E^n and $C_*^m[t_0, t_1]$ respectively, the adjoint L^* maps E^n into $C_*^m[t_0, t_1]$ according to

$$L^*(\mathbf{p}) = \mathbf{B}'(t)\mathbf{\Phi}'(t_0, t)\mathbf{p}$$

The composite map LL^* is simply $\mathbf{W}(t_0, t_1)$. If this is invertible, then according to Theorem 20-2 the choice

$$\mathbf{u}_0(t) = -\mathbf{B}'(t)\mathbf{\Phi}'(t_0, t)\mathbf{W}^{-1}(t_0, t_1)\mathbf{p}$$

minimizes $\langle \mathbf{u}, \mathbf{u} \rangle$ for $L(\mathbf{u}) = \mathbf{p}$. The following diagram illustrates these ideas.

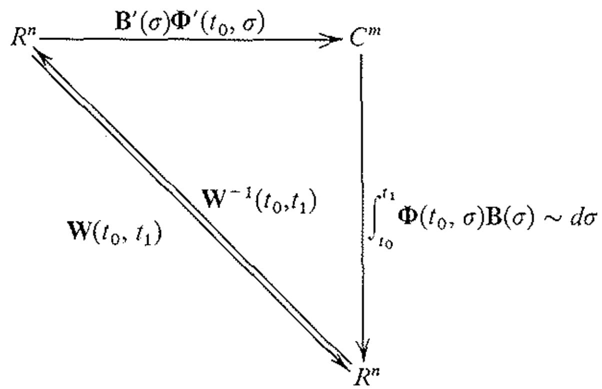


Figure 2. Illustrating the solution given by Theorem 1.

Example. The equation of motion for the electrical network shown in Figure 3 is

$$\frac{d}{dt}cv(t) = i(t); \qquad c = \text{value of capacitance}$$

where i is current out of the source. The energy dissipated in the resistor over the interval $0 \leqslant t \leqslant t_1$ is

$$d = \int_0^{t_1} ri^2(t)\,dt; \qquad r = \text{value of resistance}$$

Suppose we are to find the current i as a function of time such that $v(0) = v_0$, $v(t_1) = v_1$ and the energy dissipated in the resistor is a minimum. We want to minimize

$$\int_0^{t_1} i^2(t)\,dt$$

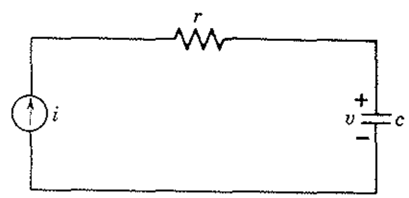


Figure 3. Minimizing the loss in charging a capacitor.

For this problem the controllability Gramian is $W(0, t_1) = t_1/c^2$ so that the optimum drive is

$$i_0(t) = +(c/t_1)(v_1 - v_0)$$

and hence is constant. If v_0 is 0 a short calculation shows that the ratio of the energy stored over the energy delivered in this case is $1/(1 + 2rc/t_1)$ and hence that this ratio limits the efficiency of this circuit as an energy storage device.

The last step on our assault on least squares problems is to treat those fixed end point problems which have an $\mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)$ term in the expression for η. This is only a matter of bringing to bear the right combination of preceding results. The final solution is however complicated to state.

Assuming dynamics of the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

and a penalty

$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt$$

the solution, again, centers around the Riccati equation

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{K}(t) - \mathbf{L}(t) \qquad \text{(R)}$$

However, now we have some freedom in the choice of the boundary conditions.

Theorem 2. *Assume that there exists a symmetric matrix* $\mathbf{K}_1$ *such that the solution* $\mathbf{\Pi}(t, \mathbf{K}_1, t_1)$ *of the matrix Riccati equation* (R) *exists on the interval* $t_0 \leqslant t \leqslant t_1$. *Then a differentiable trajectory* $\mathbf{x}$ *defined on the interval* $t_0 \leqslant t \leqslant t_1$ *minimizes*

$$\eta = \int_{t_0}^{t_0} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt$$

for the differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

and the boundary conditions $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{x}(t_1) = \mathbf{x}_1$ *if and only if it minimizes*

$$\eta_1 = \int_{t_0}^{t_1} \mathbf{v}'(t)\mathbf{v}(t)\, dt$$

for the differential equation

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1)]\mathbf{x}(t) + \mathbf{B}(t)\mathbf{v}(t)$$

and the boundary conditions $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{x}(t_1) = \mathbf{x}_1$. *Moreover, along any trajectory meeting the boundary condition*

$$\eta = \eta_1 + \mathbf{x}_0'\mathbf{\Pi}(t_0, \mathbf{K}_1, t_1)\mathbf{x}_0 - \mathbf{x}'(t_1)\mathbf{K}_1\mathbf{x}(t_1)$$

Proof. Since $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ adding the identity of Lemma 1 (Section 21) to η gives

$$\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}(t) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt + \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)]$$
$$\times \begin{bmatrix} \mathbf{0} & \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1) \\ \mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{B}(t) & \dot{\mathbf{\Pi}}(t, \mathbf{K}_1, t_1) + \mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{A}(t) + \mathbf{A}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1) \end{bmatrix}$$
$$\times \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt - \mathbf{x}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{x}(t)\Big|_{t_0}^{t_1}$$

Combining these two integrals and using the differential equation for $\mathbf{\Pi}$ gives

$$\eta = \int_{t_0}^{t_1} [\mathbf{u}'(t), \mathbf{x}'(t)] \begin{bmatrix} \mathbf{I} & \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1) \\ \mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{B}(t) & \mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1) \end{bmatrix} \begin{bmatrix} \mathbf{u}(t) \\ \mathbf{x}(t) \end{bmatrix} dt$$
$$+ \mathbf{x}(t_0)\mathbf{\Pi}(t_0, \mathbf{K}_1, t_1)\mathbf{x}(t_0)$$
$$= \int_{t_0}^{t_1} \|\mathbf{u}(t) + \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{x}(t)\|^2\, dt + \mathbf{x}(t_0)\mathbf{\Pi}(t_0, \mathbf{K}_1, t_1)\mathbf{x}(t_0)$$

Letting $\mathbf{v}(t) = \mathbf{u}(t) + \mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_1)\mathbf{x}(t)$ we obtain an equation in $\mathbf{x}$ and $\mathbf{v}$

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{K}_1, t_0)]\mathbf{x}(t) + \mathbf{B}(t)\mathbf{v}(t)$$

and a functional to minimize

$$\eta = \int_{t_0}^{t_1} \mathbf{v}'(t)\mathbf{v}(t)\, dt + \mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{K}_1, t_1)\mathbf{x}(t_0) - \mathbf{x}'(t_1)\mathbf{K}_1\mathbf{x}(t_1)$$

This completes the proof of equivalence. ▮

It is unfortunate that the ideas here tend to become obscured by the complicated formulas. The feedback diagram in Figure 4 illustrates the definition of $\mathbf{v}$.

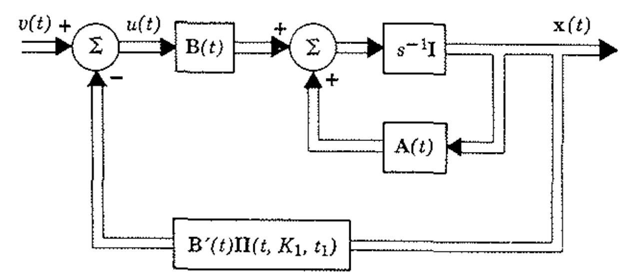


Figure 4. Illustrating the relationship between **u** *and* **v** *in Theorem 2.*

Exercises

1. Determine under what circumstances there exists a control $\mathbf{u}_0(t)$ which drives the system
$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
from $\mathbf{x}_0$ at t_0 to $\mathbf{x}_1$ at t_1 and minimizes
$$J = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + 2\mathbf{u}'(t)\mathbf{N}\mathbf{x}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt$$
Assume that the controllability Gramian $\mathbf{W}(t_0, t_1)$ is positive definite. (cf. Problem 21-1.) Find the control assuming it exists.
2. Let $\mathbf{C}$ be an m by n matrix of rank m. Find $\mathbf{u}$ such that the state of the system
$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t); \qquad \mathbf{x}(0) = \mathbf{x}_0$$
is driven to the null space of $\mathbf{C}$ at $t = 1$ and
$$\eta = \int_0^1 \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{u}(t)\, dt$$
is minimized.
3. Show that there exists an initial state for the adjoint equation
$$\dot{\mathbf{p}}(t) = -\mathbf{A}'(t)\mathbf{p}(t)$$
such that the control $\mathbf{u}$ which minimizes
$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t)\, dt$$
for transferring the state of
$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
from $\mathbf{x}_0$ at t_0 to $\mathbf{x}_1$ at t_1 is given by $\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{p}(t)$.
4. Find the minimum value of
$$\int_0^{\pi} \dot{x}^2(t)\, dt$$
with the constraints $x(\pi) = 0$ and $x(0) = 1$.
5. This and the following two problems consider classical inequalities associated with Rayleigh and Wirtinger which are of use in partial differential equations.

   Show that if x is differentiable and $x(0) = 0$, then
$$\int_0^{\pi/2} \dot{x}^2(t)\, dt \geq \int_0^{\pi/2} x^2(t)\, dt$$

   Hint: Replace $\pi/2$ by $\pi/2 - \varepsilon$ and carefully analyze the limiting case.
6. Show that if $x(0) = x(\pi) = 0$. Then
$$\int_0^{\pi} [\dot{x}(t)]^2\, dt \geq \int_0^{\pi} [x(t)]^2\, dt$$
Show that if $x(0) = \dot{x}(0) = x(\pi) = \dot{x}(\pi) = 0$ then
$$\int_0^{\pi} [\ddot{x}(t)]^2\, dt \geq \int_0^{\pi} [x(t)]^2\, dt$$
Is this the best inequality of its type?
7. Show that if $\int_0^{2\pi} x(t)\, dt = 0$ and $x(0) = x(2\pi) = 0$ then
$$\int_0^{2\pi} \dot{x}^2\, dt \geq \int_0^{2\pi} x^2(t)\, dt$$
8. Show that for $0 < T < \pi$ and x differentiable
$$\int_0^{T} \dot{x}^2(t) - x^2(t)\, dt \geq \left(\frac{T}{\sin^2 T}\right)[x(0), x(T)]\begin{bmatrix} -1 & -\cos T \\ -\cos T & 1 \end{bmatrix}\begin{bmatrix} x(0) \\ x(T) \end{bmatrix}$$
9. Let $\mathbf{A}$ be n by n and constant and let $\mathbf{b}$ be a constant n-vector. Suppose $\det(\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{n-1}\mathbf{b}) \neq 0$. Given a time function ψ defined on $0 \leq t \leq \sigma$ find $\mathbf{x}_0$ such that the following integral is minimized
$$q = \int_0^{\sigma} [\psi(t) - \mathbf{b}'e^{-\mathbf{A}'t}\mathbf{x}_0]^2\, dt$$
10. We have considered minimizing
$$\eta = \int_0^{t_1} u^2(t)\, dt$$
Subject to the constraint that u should drive the system
$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$
from $\mathbf{x}_0$ at $t = 0$ to $\mathbf{x}_1$ at $t = t_1$. The same methods work for
$$\eta = \int_0^{t_1} G(u)\, dt$$
provided G is a differentiable, convex function. Assume we express G as
$$G(u) = \int_0^{u} g(v)\, dv; \qquad g^{-1}(\ ) = f(\ )$$
and let $\mathbf{p}$ be the unique solution of (Exercise 3, Section 13)
$$\int_0^{t_1} e^{-\mathbf{A}t}\mathbf{b}f(\mathbf{b}'e^{-\mathbf{A}'t}\mathbf{p})\, dt = \mathbf{x}_0 - e^{-\mathbf{A}t_1}\mathbf{x}_1$$

Then show that the control

$$u(t) = -f(\mathbf{b}'e^{-A't}\mathbf{p})$$

accomplishes the transfer and that if v is any other control which accomplishes the transfer then

$$\int_0^{t_1} G[v(t)] - G[u(t)]\,dt = \int_0^{t_1} G[v(t)] - G[u(t)] + [u(t) - v(t)]g[u(t)]\,dt$$

Since G is convex the integrand on the left is nonnegative for all u and v and hence the given control is optimum.

11. Find u as a function of time such that the unstable system

$$\dot{x}(t) = x(t) + u(t)$$

is driven to zero and

$$\eta = \int_0^{\infty} u^4(t)\,dt$$

is a minimum. Find u as a function of x. That is, find a feedback control which steers the system along the optimal trajectory. (Use Problem 10.)

12. Let $\mathbf{A}$ and $\mathbf{b}$ be derived from the inverse square law force field problem with tangential thrusting and with $\omega = 1$.

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 \end{bmatrix}; \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Show that $\mathbf{W}(0, \sigma)$ is given by

$$\begin{aligned}
w_{11} &= (6\sigma - 8 \sin \sigma + 2 \sin 2\sigma) \\
w_{12} &= (-3 + 4 \cos \sigma - \cos 2\sigma) \\
w_{13} &= (3\sigma^2 + 2 \cos \sigma - 6\sigma \sin \sigma - 2 \cos \sigma) \\
w_{14} &= 2(-5\sigma + 7 \sin \sigma - \sin 2\sigma) \\
w_{22} &= 2\sigma - \sin 2\sigma \\
w_{23} &= (-\sigma \sin \sigma + 6\sigma \cos \sigma + 4\sigma - 2 \sin 2\sigma) \\
w_{24} &= (4 + 2 \cos 2\sigma - 6 \cos \sigma) \\
w_{33} &= (3\sigma^3 + 8\sigma - 4 \sin 2\sigma - 24 \sin \sigma + 24\sigma \cos \sigma) \\
w_{34} &= \left(-\frac{9}{2}\sigma^2 + 12\,\sigma \sin \sigma + 4 \cos 2\sigma - 4\right) \\
w_{44} &= 17\sigma + 4 \sin 2\sigma - 24 \sin \sigma
\end{aligned}$$

(Hempel and Tschauner)

13. Let $\mathbf{A}$ be n by n. Suppose that

$$[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}]$$

has rank n and $\mathbf{v}(t)$ is a known function of time. Find the control $\mathbf{u}(t)$ which drives the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{C}\mathbf{v}(t)$$

from $\mathbf{x}_0$ to $\mathbf{0}$ in T units of time and minimizes the quantity

$$J = \int_0^T \|\mathbf{u}(t)\|^2\,dt$$

14. Find the control u and the final time T such that the scalar system

$$\dot{x}(t) = u(t)$$

is driven from $x(0) = 0$ to $x(T) = 1$ while minimizing [over u and T] the quantity

$$J = \int_0^T u^2(t)\,dt + T$$

15. Consider the problem of driving x to the line shown and at the same time minimizing

$$\eta = \int_0^{t_1} u^2(t) + x^2(t)\,dt$$

where t_1 is the time when the line is reached and the dynamics are

$$\dot{x}(t) = u(t); \qquad x(0) = 1$$

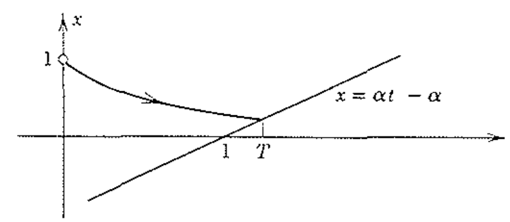


Figure 5

16. Find u such that the scalar system

$$\dot{x}(t) = -x(t) + u(t)$$

is driven from $x = 1$ at $t = 0$ to $x = 0$ at $t = 1$ and

$$\eta = \int_0^{1/2} u^2(t)\,dt + 2\int_{1/2}^{1} u^2(t)\,dt$$

is a minimum.

17. Consider the system $\dot{\chi}(t) = a\chi(t) + u(t)$ with u constrained by

$$u(t + t_1) = u(t)$$

The boundary conditions on χ are $\chi(0) = 1$, $\chi(2t_1) = 0$. Find the minimum value of

$$\eta_1 = \int_0^{2t_1} u^2(t)\, dt$$

Compare this result with that obtained without the constraint on $u(t)$. *Hint:* Consider the two dimensional system

$$\dot{x}_1(t) = ax_1(t) + u(t)$$
$$\dot{x}_2(t) = ax_2(t) + u(t)$$

with

$$\chi(t) = x_1(t) \quad \text{for} \quad 0 \leqslant t \leqslant t_1$$
$$\chi(t) = x_2(t - t_1) \quad \text{for} \quad t_1 \leqslant t \leqslant 2t_1$$

Note that it is necessary to have the condition $x_1(t_1) = x_2(0)$.

18. Find the control that transfers

$$\dot{x}(t) = b(t)u(t)$$

from the state $x(0) = 1$ to the state $x(1) = 0$ and minimizes

$$\eta = \int_0^1 u^2\, dt$$

19. Consider a system with two controls

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u_1 + \mathbf{c}u_2$$

The controls u_1 and u_2 are constrained by

$$\int_0^1 u_1^2(t)\, dt \leqslant 1; \qquad \int_0^1 u_2^2(t)\, dt \leqslant 1$$

Imagine that u_1 and u_2 have conflicting objectives. The object of u_1 is to minimize $\|x(1)\|$ while that of u_2 is to maximize $\|\mathbf{x}(1)\|$. We assume that u_1 and u_2 cannot observe each other directly but that u_1 and u_2 only know

$$\mathbf{x}(0) \quad \text{for} \quad 0 \leqslant t < \tfrac{1}{4}$$
$$\mathbf{x}(0), \mathbf{x}(\tfrac{1}{4}) \quad \text{for} \quad \tfrac{1}{4} \leqslant t < \tfrac{1}{2}$$
$$\mathbf{x}(0), \mathbf{x}(\tfrac{1}{4}), \mathbf{x}(\tfrac{1}{2}) \quad \text{for} \quad \tfrac{1}{2} \leqslant t < \tfrac{3}{4}$$
$$\mathbf{x}(0), \mathbf{x}(\tfrac{1}{4}), \mathbf{x}(\tfrac{1}{2}), \mathbf{x}(\tfrac{3}{4}) \quad \text{for} \quad \tfrac{3}{4} < t \leqslant 1$$

Find the best play for u_1 and u_2.

23. TIME-INVARIANT PROBLEMS

Specializing any result to the time invariant case is a pleasant task since some simplification and clarification is inevitable. In the case at hand the simplification is considerable. The solution of the Riccati equation can be obtained (conceptually at least) by computing a matrix exponential, partitioning it, and taking an inverse. To get the easiest case however it is necessary to let the time interval be infinite and restrict $\mathbf{L}$ to be nonnegative definite. In that situation the whole story is told by the solution of a set of algebraic equations. In the literature, this infinite time problem is referred to as the *regulator problem.* It is a useful starting point for the design of many types of systems.

Theorem 1. *If* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *is a constant minimal realization then there exists a real symmetric positive definite solution of the quadratic matrix equation*

$$\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{B}\mathbf{B}'\mathbf{K} = -\mathbf{C}'\mathbf{C} \qquad \text{(CRE)}$$

Proof. Let $\mathbf{\Pi}$ denote the solution of

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A} + \mathbf{K}(t)\mathbf{B}\mathbf{B}'\mathbf{K}(t) - \mathbf{C}'\mathbf{C} \qquad \text{(CR)}$$

From Theorem 21.1 we have

$$\mathbf{x}_0'\mathbf{\Pi}(0, \mathbf{0}, t)\mathbf{x}_0 = \min_u \int_0^t \mathbf{u}'(\sigma)\mathbf{u}(\sigma) + \mathbf{x}'(\sigma)\mathbf{C}'\mathbf{C}\mathbf{x}(\sigma)\, d\sigma$$

Hence $\mathbf{\Pi}(0, \mathbf{0}, t)$ is monotone increasing, in a matrix sense, as t increases. Moreover, since the realization is minimal, it is controllable and there exists a bounded control which drives $\mathbf{x}$ to zero in any positive interval. Such a control obviously gives a finite value for

$$\int_0^\infty \mathbf{u}'(\sigma)\mathbf{u}(\sigma) + \mathbf{x}'(\sigma)\mathbf{C}'\mathbf{C}\mathbf{x}(\sigma)\, d\sigma$$

and this upper bound is of the form $\mathbf{x}_0'\mathbf{\Pi}_0\mathbf{x}_0$.

For all $\mathbf{x}_0$ the quantity $\mathbf{x}_0'\mathbf{\Pi}(0, \mathbf{0}, t)\mathbf{x}_0$ is thus a monotone increasing function of t and it is bounded from above. Therefore by a well-known theorem in analysis, there is for all $\mathbf{x}_0$ a limit

$$\lim_{t\to\infty} \mathbf{x}_0'\mathbf{\Pi}(0, \mathbf{0}, t)\mathbf{x}_0 = \eta$$

Does there necessarily exist a *matrix* limit

$$\lim_{t\to\infty} \mathbf{\Pi}(0, \mathbf{0}, t) = \mathbf{\Pi}_\infty$$

Yes. To prove this we observe that the scalar

$$[1, 0, \ldots, 0]\mathbf{\Pi}(0, \mathbf{0}, t)[1, 0, \ldots, 0]'$$

has a limit as t approaches infinity and that its value is $\pi_{11}(0, \mathbf{0}, \infty) \overset{\text{def}}{=} \pi_{11\infty}$. Repeating this for a vector with a one in the jth spot shows $\pi_{jj}(0, \mathbf{0}, t)$ has a limit. Hence the diagonal elements of $\mathbf{\Pi}$ approach a limit. Let $\mathbf{x}_0$ now be given by

$$\mathbf{x}_0' = [0, 0, \ldots, 0, \overset{i\text{th}}{1}, \ldots, 0, \overset{j\text{th}}{1}, 0, \ldots, 0]$$

then the limit of

$$\mathbf{x}_0' \mathbf{\Pi}(0, \mathbf{0}, t)\mathbf{x}_0 = \pi_{ii}(0, \mathbf{0}, t) + \pi_{jj}(0, \mathbf{0}, t) + 2\pi_{ij}(0, \mathbf{0}, t)$$

exists and hence π_{ij} $(=\pi_{ij})$ has a limit for all i and j. Since the differential equation (CR) defining $\mathbf{\Pi}$ is time invariant,

$$\lim_{t \to -\infty} \mathbf{\Pi}(t, \mathbf{0}, 0) = \mathbf{\Pi}_\infty$$

The derivative of $\mathbf{\Pi}$ is related to $\mathbf{\Pi}$ by a time invariant differential equation and therefore $\dot{\Pi}$ approaches a constant also; clearly it must be $\mathbf{0}$. Hence Π_∞ satisfies

$$\mathbf{A}'\mathbf{\Pi}_\infty + \mathbf{\Pi}_\infty \mathbf{A} - \mathbf{\Pi}_\infty \mathbf{BB}'\mathbf{\Pi}_\infty = -\mathbf{C}'\mathbf{C}$$

Thus $\mathbf{\Pi}_\infty$ is a solution of equation (CRE). It is positive definite since

$$\lim_{t \to \infty} \mathbf{x}_0' \, \mathbf{\Pi}(0, \mathbf{0}, t)\mathbf{x}_0 = \lim_{t \to \infty} \min_{u} \int_0^t \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}'(t) \, dt$$

and the observability condition guarantees that the integral cannot vanish identically for any nonzero $\mathbf{x}_0$. ▮

The next theorem gives an important property of positive definite solutions of (CRE).

Theorem 2. *Let* [**A**, **B**, **C**] *be a constant miniminal realization. If* $\mathbf{\Pi}_\infty$ *is a positive definite solution of* $\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{KBB}'\mathbf{K} + \mathbf{C}'\mathbf{C} = \mathbf{0}$ *then all solutions of* $\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{BB}'\mathbf{\Pi}_\infty)\mathbf{x}(t)$ *go to zero as t approaches infinity and the integral*

$$\eta_1 = \int_0^\infty e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)'t}\mathbf{C}'\mathbf{C}e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)t} \, dt$$

converges.

Proof. Let $\mathbf{\Pi}_\infty$ be positive definite and let $\mathbf{x}_0$ be an arbitrary initial value of $\mathbf{x}$. For $t > 0$ we have

$$\begin{aligned}
\mathbf{x}_0'\mathbf{\Pi}_\infty \mathbf{x}_0 - \mathbf{x}'(t)\mathbf{\Pi}_\infty \mathbf{x}(t) &= -\int_0^t \dot{\mathbf{x}}'(t)\mathbf{\Pi}_\infty \mathbf{x}(t) + \mathbf{x}'(t)\mathbf{\Pi}_\infty \dot{\mathbf{x}}(t) \, dt \\
&= -\int_0^t \mathbf{x}'(t)(\mathbf{\Pi}_\infty \mathbf{A} + \mathbf{A}'\mathbf{\Pi}_\infty - 2\mathbf{\Pi}_\infty \mathbf{BB}'\mathbf{\Pi}_\infty)\mathbf{x}(t) \, d \\
&= \int_0^t \mathbf{x}'(t)(\mathbf{C}'\mathbf{C} + \mathbf{\Pi}_\infty \mathbf{BB}'\mathbf{\Pi}_\infty)\mathbf{x}(t) \, dt \geqslant 0 \qquad (*)
\end{aligned}$$

As $\mathbf{x}_0$ varies over the set of all vectors of unit length the nonnegative quantity

$$\mathbf{x}_0'\mathbf{\Pi}_\infty \mathbf{x}_0 - \mathbf{x}_0' e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)'t}\mathbf{\Pi}_\infty e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)t}\mathbf{x}_0$$

takes on a minimum value. If that minimum is zero then for the corresponding value of $\mathbf{x}_0$

$$\int_0^t \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t) \, dt + \int_0^t \mathbf{x}'(t)\mathbf{\Pi}_\infty \mathbf{BB}'\mathbf{\Pi}_\infty \mathbf{x}(t) \, dt = 0$$

However if $\mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)$ vanishes identically then the differential equation is $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ and if, in addition, $\mathbf{C}\mathbf{x}(t)$ vanishes identically then using the observability assumption we see that $\mathbf{x}_0$ must be zero. Hence for some $\varepsilon > 0$

$$\mathbf{x}'(t)\mathbf{\Pi}_\infty \mathbf{x}(t) - \mathbf{x}'(t+1)\mathbf{\Pi}_\infty \mathbf{x}(t+1) \geqslant \varepsilon\|\mathbf{x}(t)\|^2$$

This means that

$$\int_0^\infty \mathbf{x}'(t)\mathbf{C}\mathbf{C}'(t)\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{u}(t) \, dt \geqslant \sum_{n=1}^\infty \varepsilon\|\mathbf{x}(n)\|^2$$

Since the integral on the left converges this implies that $\|\mathbf{x}(t)\|$ approaches zero as t approaches infinity. Using the fact that $\mathbf{x}(t)$ approaches zero we see from (*) that

$$\begin{aligned}
\mathbf{x}'(0)\mathbf{\Pi}_\infty \mathbf{x}(0) &= \int_0^\infty \mathbf{x}'(t)[\mathbf{C}'\mathbf{C} + \mathbf{\Pi}_\infty \mathbf{BB}'\mathbf{\Pi}_\infty]\mathbf{x}(t) \, dt \\
&\geqslant \int_0^\infty \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t) \, dt \\
&= \int_0^\infty \mathbf{x}_0' e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)'t}\mathbf{C}'\mathbf{C}e^{(\mathbf{A}-\mathbf{BB}'\mathbf{\Pi}_\infty)t}\mathbf{x}_0 \, dt
\end{aligned}$$

Hence the integral η_1 converges. ▮

Using these two theorems we can proceed to find a complete solution of the Riccati equation.

Suppose $\mathbf{\Pi}_\infty$ satisfies the quadratic equation

$$\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{KBB}'\mathbf{K} + \mathbf{L} = \mathbf{0} \qquad \text{(CRE)}$$

Then rewriting equation (CR) in terms of $\mathbf{\Psi} = \mathbf{K} - \mathbf{\Pi}_\infty$ gives

$$\dot{\mathbf{\Psi}}(t) = -(\mathbf{A}' - \mathbf{\Pi}_\infty\mathbf{BB}')\mathbf{\Psi}(t) - \mathbf{\Psi}(t)(\mathbf{A} - \mathbf{BB}'\mathbf{\Pi}_\infty) + \mathbf{\Psi}(t)\mathbf{BB}'\mathbf{\Psi}(t)$$

This equation is still nonlinear but if it is pre- and post-multiplied by $\mathbf{\Psi}^{-1}$ it becomes linear in $\mathbf{\Psi}^{-1}$ (Use $\frac{d}{dt}\mathbf{\Psi}^{-1} = -\mathbf{\Psi}^{-1}\dot{\mathbf{\Psi}}\mathbf{\Psi}^{-1}$)

$$\frac{d}{dt}\mathbf{\Psi}^{-1}(t) = +(\mathbf{A} - \mathbf{BB}'\mathbf{\Pi}_\infty)\mathbf{\Psi}^{-1}(t) + \mathbf{\Psi}^{-1}(t)(\mathbf{A}' - \mathbf{\Pi}_\infty \mathbf{BB}') - \mathbf{BB}'$$

This is an equation of the type discussed in Section 11. In view of Theorem 2 we can find $\mathbf{Q}_1$ which satisfies

$$(\mathbf{A} - \mathbf{BB}'\boldsymbol{\Pi}_\infty)\mathbf{Q} + \mathbf{Q}(\mathbf{A}' - \boldsymbol{\Pi}_\infty \mathbf{BB}') - \mathbf{BB}' = \mathbf{0}$$

In fact

$$\mathbf{Q}_1 = -\int_0^\infty e^{(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)t}\mathbf{BB}'e^{(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)'t}\,dt$$

In terms of $\mathbf{Q}_1$

$$\boldsymbol{\Psi}^{-1}(t) = \mathbf{Q}_1 + e^{+(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)t}[\boldsymbol{\Psi}^{-1}(0) - \mathbf{Q}_1]e^{+(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)'t}$$

This gives as a solution of the Riccati equation:

$$\boldsymbol{\Pi}(t, \mathbf{K}_0, 0) = \boldsymbol{\Pi}_\infty + \{\mathbf{Q}_1 + e^{(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)t}[(\mathbf{K}_0 - \boldsymbol{\Pi}_\infty)^{-1} - \mathbf{Q}_1]e^{(\mathbf{A}'-\boldsymbol{\Pi}_\infty\mathbf{BB}')t}\}^{-1}$$

To complete the picture on $\boldsymbol{\Pi}$ we include the following result on the uniqueness.

Theorem 3. *If* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{L} = \mathbf{L}'$ *are constant then there is at most one symmetric solution of* $\mathbf{KA} + \mathbf{A}'\mathbf{K} - \mathbf{KBB}'\mathbf{K} + \mathbf{L} = \mathbf{0}$ *having the property that* $\mathbf{A} - \mathbf{BB}'\mathbf{K}$ *has its eigenvalues in* $\operatorname{Re} s < 0$.

Proof. Assume, contrary to what we want to prove, that there are two symmetric solutions, $\mathbf{K}_1$ and $\mathbf{K}_2$ such that $\mathbf{A} - \mathbf{BB}'\mathbf{K}_1$ and $\mathbf{A} - \mathbf{BB}'\mathbf{K}_2$ have their eigenvalue in $\operatorname{Re} s < 0$. All solutions of $\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{BB}'\mathbf{K}_i)\mathbf{x}(t)$ approach zero as t approaches infinity. Moreover, as we have seen in the proof of Theorem 22-2,

$$\int_0^t \mathbf{u}'(\sigma)\mathbf{u}(\sigma) + \mathbf{x}'(\sigma)\mathbf{L}\mathbf{x}(\sigma)\,d\sigma$$

$$= \mathbf{x}'(0)\mathbf{K}_1\mathbf{x}(0) - \mathbf{x}'(t)\mathbf{K}_1\mathbf{x}(t) + \int_0^t \|\mathbf{u}(t) + \mathbf{B}'\mathbf{K}_1 x(t)\|^2\,dt$$

$$= \mathbf{x}'(0)\mathbf{K}_2\mathbf{x}(0) - \mathbf{x}'(t)\mathbf{K}_2\mathbf{x}(t) + \int_0^t \|\mathbf{u}(t) + \mathbf{B}'\mathbf{K}_2 x(t)\|^2\,dt$$

Since $\mathbf{K}_1 \neq \mathbf{K}_2$ and both are symmetric there exists $\mathbf{x}_0$ such that $\mathbf{x}_0'\mathbf{K}_1\mathbf{x}_0 \neq \mathbf{x}_0'\mathbf{K}_2\mathbf{x}_0$. Say $\mathbf{x}_0'\mathbf{K}_1\mathbf{x}_0 \geqslant \mathbf{x}_0'\mathbf{K}_2\mathbf{x}_0$. Then let $\mathbf{u}(t) = -\mathbf{B}'\mathbf{K}_2\mathbf{x}(t)$. Taking the limit as t goes to infinity gives

$$\mathbf{x}_0'\mathbf{K}_2\mathbf{x}_0 = \mathbf{x}_0'\mathbf{K}_1\mathbf{x}_0 + \int_0^\infty \|\mathbf{B}'(\mathbf{K}_1 - \mathbf{K}_2)\mathbf{x}(t)\|^2\,dt$$

which is clearly impossible. Hence we contradict the hypothesis that there exists $\mathbf{K}_1 \neq \mathbf{K}_2$ such that $(\mathbf{A} - \mathbf{BB}'\mathbf{K}_1)$ and $(\mathbf{A} - \mathbf{BB}'\mathbf{K}_2)$ have their eigenvalues in $\operatorname{Re} s < 0$. ▮

Theorem 4. *If* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *is a minimal realization then there is exactly one symmetric positive definite solution of* $\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{KBB}'\mathbf{K} + \mathbf{CC}' = \mathbf{0}$.

Proof. This is an immediate consequence of Theorems 1, 2 and 3.

The use of these theorems permits an analysis of still one more special case, namely minimization of quadratic functionals for linear constant systems on the interval $0 \leqslant t \leqslant \infty$. This problem is of special interest because of the comparatively simple form which the answer takes. The following theorem contains the main results.

Theorem 5. *Let* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *be a constant minimal realization. Let* $\boldsymbol{\Pi}_\infty$ *be the positive definite solution of*

$$\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{KBB}'\mathbf{K} = -\mathbf{C}'\mathbf{C}$$

There exists a control which minimizes

$$\eta = \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t)\,dt$$

for the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{x}(0) = \mathbf{x}_0$$

The minimum value of η *is* $\mathbf{x}_0'\boldsymbol{\Pi}_\infty\mathbf{x}_0$. *The minimizing control in closed loop form is*

$$\mathbf{u}(t) = -\mathbf{B}'\boldsymbol{\Pi}_\infty\mathbf{x}(t)$$

whereas in open loop form it is

$$\mathbf{u}(t) = -\mathbf{B}'\boldsymbol{\Pi}_\infty e^{(\mathbf{A}-\mathbf{BB}'\boldsymbol{\Pi}_\infty)t}\mathbf{x}_0$$

Proof. First observe that if $\|\mathbf{x}(t)\|$ does not approach zero as t approaches infinity then the integral

$$\eta = \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t)\,dt$$

will diverge. This is an immediate consequence of the observability assumption which together with 21-1 implies

$$\min_u \int_t^{t+1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t)\,dt = \mathbf{x}'(t)\boldsymbol{\Pi}(0, \mathbf{0}, 1)\mathbf{x}(t) \geqslant \varepsilon\|\mathbf{x}(t)\|$$

Make the change of variable

$$\mathbf{v}(t) = \mathbf{u}(t) + \mathbf{B}'\boldsymbol{\Pi}_\infty\mathbf{x}(t)$$

so that

$$\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{BB}'\boldsymbol{\Pi}_\infty)\mathbf{x}(t) + \mathbf{B}\mathbf{v}(t)$$

and η becomes

$$\eta = \int_0^\infty [\mathbf{v}(t) - \mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)]'[\mathbf{v}(t) - \mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)] + \mathbf{x}'(t)\mathbf{C}\mathbf{C}'\mathbf{x}(t)dt$$

expanding the integrand, with the use of $\mathbf{\Pi}_\infty \mathbf{A} + \mathbf{A}'\mathbf{\Pi}_\infty - \mathbf{\Pi}_\infty \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty = -\mathbf{C}'\mathbf{C}$, gives

$$\begin{aligned}\eta &= \int_0^\infty \mathbf{v}'(t)\mathbf{v}(t) - [(\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)\mathbf{x}(t) + \mathbf{B}\mathbf{v}(t)]'\mathbf{\Pi}_\infty \mathbf{x}(t) \\ &\quad - \mathbf{x}'(t)\mathbf{\Pi}_\infty[(\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)\mathbf{x}(t) + \mathbf{B}\mathbf{v}(t)]\, dt \\ &= \int_0^\infty \mathbf{v}'(t)\mathbf{v}(t)\, dt - \mathbf{x}'(t)\mathbf{\Pi}_\infty \mathbf{x}(t)\Big]_0^\infty\end{aligned}$$

From Theorem 2 we know that all solutions of $\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)\mathbf{x}(t)$ go to zero as t approaches infinity. Hence the best choice of $\mathbf{v}$ is $\mathbf{0}$. ▮

Example. Consider minimizing

$$\eta = \int_0^T u^2(t) + x^2(t)\, dt$$

for $\dot{x}(t) = u(t)$. In this case the Riccati equation is

$$\dot{k}(t) = k^2(t) - 1$$

We could solve the Riccati equation directly of course, but instead we follow the path outlined in the above discussion. There is a positive definite equilibrium solution $k = 1$ (and also the solution $k = -1$ which is of no interest).

Let $\psi = (k - 1)$, then

$$\dot{\psi}(t) = \psi^2(t) + 2\psi(t)$$

Dividing by ψ^2 and letting $\psi^{-1} = \phi$ gives

$$\dot{\phi}(t) = -(2\phi(t) + 1)$$

Thus $\phi(t) = \alpha e^{-2t} - \frac{1}{2}$ and

$$\begin{aligned}k(t) = \psi(t) + 1 &= \frac{1}{\alpha e^{-2t} - \frac{1}{2}} + 1 = \frac{2}{2\alpha e^{-2t} - 1} + 1 \\ &= \frac{2\alpha e^{-2t} + 1}{2\alpha e^{-2t} - 1}\end{aligned}$$

If $x(0)$ is fixed and $x(T)$ is free, then the minimizing $u(t)$ is $-k(t)x(t)$ where α, in the definition of k, is picked so as to make $k(T)$ zero. The minimum value of the integral in this case is $x^2(0)k(0)$ and

$$k(t) = \frac{1 - e^{2(t-T)}}{1 + e^{2(t-T)}} = \tanh(T - t)$$

Notice that $k(t)$ goes to 1 as $(t - T)$ goes to minus infinity.

There is one final result on linear constant systems which will be of interest in Section 25.

Theorem 6. *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{L} = \mathbf{L}'$ *be constant matrices. Assume there exists* $\mathbf{\Pi}_\infty$, *a negative definite solution of*

$$\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{B}\mathbf{B}'K = -\mathbf{L}$$

such that the eigenvalues of $(\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)$ *lie in the half-plane* Re $s < 0$. *Then there exists a control which minimizes*

$$\eta = \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}\mathbf{x}(t)\, dt$$

for the system $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$; $\mathbf{x}(0) = \mathbf{x}_0$. *The minimum value of* η *is* $\mathbf{x}_0'\mathbf{\Pi}_\infty \mathbf{x}_0$. *The minimizing control in closed-loop form is* $\mathbf{u}(t) = -\mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)$ *whereas in open-loop form it is* $\mathbf{u}(t) = -\mathbf{B}'\mathbf{\Pi}_\infty e^{(\mathbf{A}-\mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)t}\mathbf{x}_0$

Proof. Introduce $\mathbf{v}$ as in the proof of Theorem 5 to get

$$\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{\Pi}_\infty)\mathbf{x}(t) + \mathbf{B}\mathbf{v}(t)$$

Notice that for the altered performance measure η_1 we have

$$\begin{aligned}\eta_1 &\overset{\text{def}}{=} \min_{\mathbf{u}} \int_0^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}\mathbf{x}(t)\, dt + \mathbf{x}'(t_1)\mathbf{\Pi}_\infty \mathbf{x}(t_1) \\ &= \min_{\mathbf{v}} \int_0^{t_1} \mathbf{v}'(t)\mathbf{v}(t)\, dt + \mathbf{x}'(0)\mathbf{\Pi}_\infty \mathbf{x}(0) \\ &= \mathbf{x}'(0)\mathbf{\Pi}_\infty \mathbf{x}(0)\end{aligned}$$

and that this minimum is achieved for $\mathbf{u}(t) = -\mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)$. Moreover, since $\mathbf{\Pi}_\infty < \mathbf{0}$,

$$\eta = \int_0^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}\mathbf{x}(t)\, dt \geqslant \eta_1$$

On the other hand, for the choice $\mathbf{u}(t) = -\mathbf{B}'\mathbf{\Pi}_\infty \mathbf{x}(t)$ we have $\mathbf{x}(t) \to \mathbf{0}$ as $t \to \infty$ so that $\eta = \eta_1$ and hence the choice of $\mathbf{u}$ must be optimal. ▮

Exercises

1. Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ be a minimal realization. Let $\mathbf{\Pi}_\infty$ be the symmetric positive definite solution of $\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{B}\mathbf{B}'\mathbf{K} + \mathbf{C}'\mathbf{C} = \mathbf{0}$. Show that $\mathbf{\Pi}(t, \mathbf{Q}, 0)$ approaches $\mathbf{\Pi}_\infty$ as t approaches $-\infty$ for all $\mathbf{Q} = \mathbf{Q}' \geqslant \mathbf{0}$.

2. Let Π_∞ be a solution of $A'K + KA - KBB'K + M = 0$ such that the eigenvalues of $A - BB'\Pi_\infty$ lie in the half-plane Re $s < 0$. Suppose that N, defined as

$$N = \int_0^\infty e^{(A - BB'\Pi_\infty)t} BB' e^{(A - BB'\Pi_\infty)'t}\, dt$$

is invertible. Then show that $\Pi_\infty - N^{-1}$ is also a solution of

$$A'K + KA - KBB'K + M = 0$$

and that the eigenvalues of $A - BB'\Pi_\infty + BB'N^{-1}$ all lie in the half-plane Re $s > 0$.

3. Deduce from Theorem 1 the fact that every positive definite matrix has a unique symmetric, positive definite square root. (Compare with Exercise 5, Section 12).

4. Let Π_∞ be a positive definite solution of the quadratic matrix equation

$$A'K + KA - KBB'K + C'C = 0$$

If $A = A'$ and $B = C'$ then show that $-\Pi_\infty^{-1}$ also satisfies the same equation. If [A, B, C] is a minimal realization can there be other negative definite solutions?

5. Find α and β such that for

$$\ddot{\chi}(t) + \alpha\dot{\chi}(t) + \beta\chi(t) = 0$$

the integral

$$q = \int_0^\infty \chi^2(t) + [\alpha\dot{\chi}(t) + \beta\chi(t)]^2\, dt$$

is a minimum. Show that the answer does not depend on $\dot{\chi}(0)$ and $\chi(0)$. Explain.

6. Show that the feedback control law for the constant system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

and the penalty functional

$$\eta = \int_0^{t_1} u'(t)u(t) + x'(t)C'Cx(t)\, dt + x'(t_1)Qx(t_1)$$

is of the form $u(t) = Fx(t)$ with F constant if Q satisfies the equation

$$QA + A'Q - QBB'Q + C'C = 0$$

and is positive definite.

7. Let [A, B, C] be a time invariant minimal realization. Find u such that for

$$\dot{x}(t) = Ax(t) + Bu(t); \qquad y(t) = Cx(t)$$

the integral

$$\eta = \int_0^\infty u'(t)u(t)\, dt + \int_1^\infty y'(t)y(t)\, dt$$

is a minimum.

8. The network shown has, at $t = 0$, one volt across the capacitor and no current flowing in the inductor.

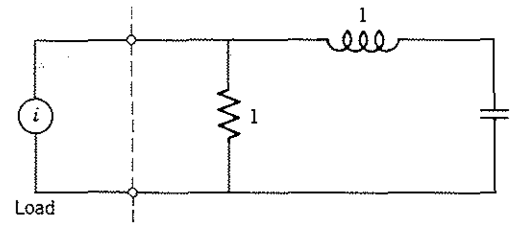


Figure 1

How much of the stored energy can be delivered to a load by using the best possible matching (i.e. choice of $i(t)$)? Infinite time is available to discharge the circuit.

9. The one-dimensional wave equation defined on $0 \leqslant z \leqslant 1$, $0 \leqslant t < \infty$ was discussed in Exercise 5 of Section 6. Proceeding formally, minimize

$$\eta = \int_0^\infty \int_0^1 u^2(z, t) + [\partial x(z, t)/\partial t]^2\, dz\, dt$$

for the equation and boundary conditions given there. Show that

$$u_n(t) = -\alpha_n x_n(t) - \beta_n y_n(t)$$

for a suitable α_n and β_n.

24. CANONICAL EQUATIONS, CONJUGATE POINTS, AND FOCAL POINTS

In the previous sections conditions under which a minimum exists for certain least squares problems were derived together with an explicit formula for the minimizing control and the minimizing trajectory. The constructions given centered around being able to solve a matrix Riccati equation with certain boundary conditions specified at one point in time. In this section we want to study the Riccati equation itself and at the same time consider an alternative point of view which occupies a central position in the further development of this subject.

We begin by establishing an important and somewhat surprising fact. Namely, that the solution of a Riccati equation can be obtained by solving a suitable set of *linear* differential equations.

It is often pointed out in introductory books on differential equations that scalar Riccati equations are related to second order linear equations by a simple transformation. The idea is to relate the equation

$$\ddot{x}(t) + \beta(t)\dot{x}(t) + \gamma(t)x(t) = 0$$

to a first order equation by letting $k(t) = \dot{x}(t)/x(t)$. This makes

$$\dot{k}(t) = \left[\frac{\ddot{x}(t)}{x(t)}\right] - \left[\frac{\dot{x}(t)}{x(t)}\right]^2$$

and so

$$\dot{k}(t) + k^2(t) + \beta(t)k(t) = -\gamma(t)$$

A similar method works in the nonscalar case. Instead of a second order vector equation, however, we start with a pair of first order vector equations.

Lemma. *Let* Φ *be the transition matrix for the set of* $2n$ *linear differential equations*

$$\begin{bmatrix}\dot{\mathbf{x}}(t)\\ \dot{\mathbf{p}}(t)\end{bmatrix} = \begin{bmatrix}\mathbf{A}(t) & -\mathbf{B}(t)\mathbf{B}'(t)\\ -\mathbf{L}(t) & -\mathbf{A}'(t)\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{p}(t)\end{bmatrix}$$

Let $\Phi = [(\Phi_{11}, \Phi_{12}); (\Phi_{21}, \Phi_{22})]$ *be a partitioning of* Φ *into* n *by* n *blocks. If* Π *is defined as*

$$\Pi(t, \mathbf{Q}, t_1) = [\Phi_{22}(t_1, t) - \mathbf{Q}\Phi_{12}(t_1, t)]^{-1}[\mathbf{Q}\Phi_{11}(t_1, t) - \Phi_{21}(t_1, t)]$$

or equivalently as

$$\Pi(t, \mathbf{Q}, t_1) = [\Phi_{21}(t, t_1) + \Phi_{22}(t, t_1)\mathbf{Q}][\Phi_{11}(t, t_1) + \Phi_{12}(t, t_1)\mathbf{Q}]^{-1}$$

then $\Pi(t_1, \mathbf{Q}, t_1) = \mathbf{Q}$ *and*

$$\frac{d}{dt}\Pi(t, \mathbf{Q}, t_1) = -\mathbf{A}'(t)\Pi(t, \mathbf{Q}, t_1) - \Pi(t, \mathbf{Q}, t_1)\mathbf{A}(t) + \Pi(t, \mathbf{Q}, t_1)\mathbf{B}(t)\mathbf{B}'(t)\Pi(t, \mathbf{Q}, t_1) - \mathbf{L}(t)$$

if the indicated inverses exist.

Proof. We will prove only the first form. Previously it has been established (Section 7) that $(d/dt)\Phi_{\mathbf{B}}(t_1, t) = -\Phi_{\mathbf{B}}(t_1, t)\mathbf{B}(t)$. Moreover, $(d/dt)\mathbf{P}^{-1}(t) = -\mathbf{P}^{-1}(t)\dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)$. Using these two facts we will proceed to verify that Π as given here does satisfy the Riccati equation. The boundary condition is clearly correct since $\Phi(t_1, t_1) = \mathbf{I}$. To keep the algebra in hand, define $\mathbf{X}$ and $\mathbf{P}$ by

$$[\mathbf{X}(t), -\mathbf{P}(t)] = [\mathbf{Q}, -I]\begin{bmatrix}\Phi_{11}(t_1, t) & \Phi_{12}(t_1, t)\\ \Phi_{21}(t_1, t) & \Phi_{22}(t_1, t)\end{bmatrix}$$

according to the lemma statement $\Pi(t, \mathbf{Q}, t_1) = \mathbf{P}^{-1}(t)\mathbf{X}(t)$ and hence

$$\frac{d}{dt}\Pi(t, \mathbf{Q}, t_1) = -\mathbf{P}^{-1}(t)\dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)\mathbf{X}(t) + \mathbf{P}^{-1}(t)\dot{\mathbf{X}}(t)$$

Using the properties of Φ we have

$$[\dot{\mathbf{X}}(t), -\dot{\mathbf{P}}(t)] = [\mathbf{X}(t), -\mathbf{P}(t)]\begin{bmatrix}-\mathbf{A}(t) & \mathbf{B}(t)\mathbf{B}'(t)\\ \mathbf{L}(t) & \mathbf{A}'(t)\end{bmatrix}$$

Hence

$$\begin{aligned}\frac{d}{dt}\Pi(t, \mathbf{Q}, t) &= \mathbf{P}^{-1}(t)\mathbf{X}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{P}^{-1}(t)\mathbf{X}(t)\\ &\quad - \mathbf{P}^{-1}(t)\mathbf{P}(t)\mathbf{A}'(t)\mathbf{P}^{-1}(t)\mathbf{P}^{-1}(t)\mathbf{X}(t)\\ &\quad - \mathbf{P}^{-1}(t)\mathbf{X}(t)\mathbf{A}(t) - \mathbf{P}^{-1}(t)\mathbf{P}(t)\mathbf{L}(t)\\ &= \Pi(t, \mathbf{Q}, t_1)\mathbf{B}(t)\mathbf{B}'(t)\Pi(t, \mathbf{Q}, t_1) - \mathbf{A}'(t)\Pi(t, \mathbf{Q}, t_1)\\ &\quad - \Pi(t, \mathbf{Q}, t_1)\mathbf{A}(t) - \mathbf{L}(t)\end{aligned}$$ ▮

Naturally the matrix $\Phi_{22}(t_1, t) - \mathbf{Q}\Phi_{12}(t_1, t)$ which appears in the above expression for Π need not be invertible for all t. Hence Lemma 1 certainly does not imply the existence of a solution of the Riccati equation for all time. However, $\Phi_{22}(t_1, t_1) = \mathbf{I}$ and $\Phi_{12}(t_1, t_1) = \mathbf{0}$ and both are continuous functions of t. Since $\mathbf{I}$ is invertible it is clear that $\Phi_{22}(t_1, t) - \mathbf{Q}\Phi_{12}(t_1, t)$ is also invertible for $|t - t_1|$ sufficiently small regardless of $\mathbf{Q}$. Hence for the first time we have a local existence theorem for the Riccati equation.

Theorem 1. (Local existence theorem for Riccati equations) *If* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{L}$ *are bounded on the interval* $|t - t_0| \leqslant T$ *and if* $\mathbf{K}_0$ *is given, then there exists an* $\varepsilon > 0$ *such that for* $|t - t_0| < \varepsilon$ *a unique solution of*

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)B'(t)\mathbf{K}(t) - \mathbf{L}(t);\ \mathbf{K}(t_0) = \mathbf{K}_0$$

exists on the interval $|t - t_0| < \varepsilon$.

Proof. Existence follows by letting $\mathbf{K}_0 = \mathbf{Q}$ in the above lemma. Uniqueness follows from the fact that in any bounded subset of $R^{n\times n}$ the right side of the Riccati equation satisfies a Lipschitz condition and hence the uniqueness theorem of Exercise 3 of Section 2 applies. ▮

Example. Consider the Riccati equation

$$\dot{k}(t) = k^2(t) + 1$$

The associated linear system equation is the harmonic oscillator equation

$$\begin{bmatrix}\dot{x}(t)\\ \dot{p}(t)\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x(t)\\ p(t)\end{bmatrix}$$

The solution corresponding to $x(0) = 1$ and $p(0) = k_0$ is

$$\begin{bmatrix}x(t)\\ p(t)\end{bmatrix} = \begin{bmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{bmatrix}\begin{bmatrix}1\\ k_0\end{bmatrix}$$

Therefore,

$$\pi(t, k_0, 0) = \frac{k_0 \cos t + \sin t}{\cos t - k_0 \sin t}$$

A set of $2n$ differential equations are said to be in *canonical form* if there exists a scalar function H, called the *Hamiltonian*, such that H depends on $\mathbf{x}$, $\mathbf{p}$ and t and

$$\dot{\mathbf{x}}(t) = \partial H/\partial \mathbf{p}\,|_{\mathbf{x},\mathbf{p},t}$$

$$\dot{\mathbf{p}}(t) = -\partial H/\partial \mathbf{x}\,|_{\mathbf{x},\mathbf{p},t}$$

Obviously, quadratic Hamiltonians give rise to linear equations. For the pair of equations introduced above the appropriate Hamiltonian is

$$H(\mathbf{x}, \mathbf{p}, t) = \tfrac{1}{2}\mathbf{x}'\mathbf{L}(t)\mathbf{x}(t) + \mathbf{p}'(t)\mathbf{A}(t)\mathbf{x}(t) - \tfrac{1}{2}\mathbf{p}'(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{p}(t)$$

In deriving these results from classical variational techniques these equations play a central role. The following theorem indicates one of the reasons.

Theorem 2. *Let* $\mathbf{x}$ *and* $\mathbf{u}$ *be optimal in the sense of Theorem* 21-1, *Theorem* 22-1 *or Theorem* 22-2. *Then there exists an n-vector* $\mathbf{p}$ *such that* $\mathbf{x}$ *and* $\mathbf{p}$ *satisfy the* 2*n canonical equations*

$$\begin{bmatrix}\dot{\mathbf{x}}(t)\\ \dot{\mathbf{p}}(t)\end{bmatrix} = \begin{bmatrix}\mathbf{A}(t) & -\mathbf{B}(t)\mathbf{B}'(t)\\ -\mathbf{L}(t) & -\mathbf{A}'(t)\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{p}(t)\end{bmatrix}; \qquad \begin{bmatrix}\mathbf{x}(t_0)\\ \mathbf{p}(t_0)\end{bmatrix} = \begin{bmatrix}\mathbf{x}_0\\ \mathbf{p}_0\end{bmatrix}$$

with $\mathbf{u}$ *being given by*

$$\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{p}(t)$$

The proof involves a lengthy caculation and is omitted.

In the classical literature on the calculus of variations, the conditions for existence of minima are stated in terms of the nonexistence of conjugate points and focal points. Two points in time t_0 and t_1 are called *conjugate*

points if it is possible to find a nontrivial solution of the canonical equations such that $\mathbf{x}(t_0) = \mathbf{x}(t_1) = \mathbf{0}$. Two points in time t_0 and t_1 are *focal points* if it is possible to find a nontrivial solution of the canonical equations such that $\mathbf{x}(t_0) = \mathbf{p}(t_1) = \mathbf{0}$.

Existence of minima and nonexistence of focal and conjugate points are related as follows. From Theorem 1 of section 21 we know that the free end-point problem with no terminal penalty term has a solution if the associated Riccati equation has a solution $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ which exists on $t_0 \leqslant t \leqslant t_1$. It may be shown that this is the case if and only if no focal points exist on $t_0 \leqslant t \leqslant t_1$. On the other hand, for a fixed end-point problem a sufficient condition for the existence of a minimum is that $\mathbf{\Pi}(t, \mathbf{K}_1, t_1)$ should exist for some symmetric matrix $\mathbf{K}_1$. The precise choice of $\mathbf{K}_1$ is left open and in fact, can be made in such a way as to prolong the existence of $\mathbf{\Pi}(t, \mathbf{K}_1, t_1)$. It may be shown that a choice of $\mathbf{K}$, which yields a solution on $t_0 \leqslant t \leqslant t_1$ is possible if and only if no conjugate points exist on the interval $t_0 \leqslant t \leqslant t_1$.

Example. Consider the minimization of

$$\eta = \int_0^T u^2(t) - x^2(t)\, dt$$

where $\dot{x}(t) = u(t)$ and $x(0) = 0$. The canonical equations are

$$\begin{bmatrix}\dot{x}(t)\\ \dot{p}(t)\end{bmatrix} = \begin{bmatrix}0 & -1\\ +1 & 0\end{bmatrix}\begin{bmatrix}x(t)\\ p(t)\end{bmatrix}$$

so that $x(t) = A \sin t + B \cos t$, $p(t) = -\dot{x}(t)$. Regarded as a fixed end-point problem we see that the interval $0 \leqslant t < \pi$ contains no conjugate points so that for $T < \pi$ and $x(0) = x(T) = 0$ we have

$$\min_u \int_0^T \dot{x}^2(t) - x^2(t)\, dt = 0$$

On the other hand, if this is regarded as a free end-point problem then $x(t) = \sin t$ gives $x(0) = p(\pi/2) = 0$ so that the interval which is free of focal points is only $0 \leqslant t < \pi/2$. (See Figure 1).

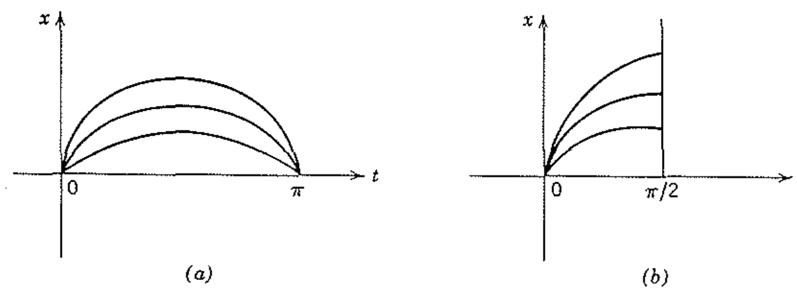


Figure 1. *Conjugate points and focal points for a simple problem.*

This leaves unanswered questions about the effect of terminal cost on existence. One of the easier qualitative results is given in Theorem 3.

Theorem 3. *If* $\mathbf{Q}_1 - \mathbf{Q}_2$ *is nonnegative definite, then for all* $t < t_1$ *the difference* $\mathbf{\Pi}(t, \mathbf{Q}_1, t_1) - \mathbf{\Pi}(t, \mathbf{Q}_2, t_1)$ *is nonnegative definite as long as both* $\mathbf{\Pi}$*'s exist. Hence if the solution passing through* $\mathbf{Q}_2$ *exists on* $t_0 \leqslant t \leqslant t_1$ *then the solution passing through* $\mathbf{Q}_1$ *exists on an interval extending back at least as far as* t_0.

Proof. This is an immediate consequence of Theorem 22.1 which gives $\mathbf{x}'(0)\mathbf{\Pi}(t_0, \mathbf{Q}_1, t_1)\mathbf{x}(0)$ as the minimum value of η. Clearly increasing the terminal penalty increases the total cost. ▮

The question of existence of solutions of the Riccati equation on a whole half-line will now be examined. As is apparent from the example, nontrivial assumptions will be required to keep $\mathbf{\Pi}$ from passing off to infinity at finite values of t. To investigate this question we will assume that $\mathbf{L}(t)$, in addition to being symmetric, is nonnegative definite as well. We also assume that an associated system satisfies a controllability condition.

Theorem 4. *Assume that the controllability Gramian for*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

is positive definite and that $\mathbf{L}(t) = \mathbf{L}'(t)$ *is nonnegative definite for all* t. *If* $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ *is the solution of equation (R) which passes through* $\mathbf{0}$ *at* $t = t_1$, *then given* t_0, $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ *exists on* $t_0 \leqslant t \leqslant t_1$ *regardless of* t_0. *Moreover,*

$$\mathbf{\Pi}(t_0, \mathbf{0}, t_1) \leqslant \mathbf{\Pi}(t_0, \mathbf{0}, t_2) \quad \text{for} \quad t_0 \leqslant t_1 \leqslant t_2(^*) \tag{i}$$

and

$$0 \leqslant \mathbf{\Pi}(t_0, \mathbf{0}, t_2) \leqslant \mathbf{N}(t_0, t_1) \quad \text{for} \quad t_0 \leqslant t_1 \leqslant t_2$$

where

$$\mathbf{N}(t_0, t_1) = [\mathbf{W}(t_0, t_1)]^{-1} + \int_{t_0}^{t_1} \mathbf{X}_1'(\sigma)\mathbf{L}(\sigma)\mathbf{X}_1(\sigma)\, d\sigma$$

with

$$\mathbf{X}_1(\sigma) = \mathbf{\Phi}(\sigma, t_0) - \int_{t_0}^{\sigma} \mathbf{\Phi}(\sigma, \rho)\mathbf{B}(\rho)\mathbf{B}'(\rho)\mathbf{\Phi}'(t_0, \rho)\, d\rho[\mathbf{W}(t_0, t_1)]^{-1}$$

and $\mathbf{\Phi}$ *the transition matrix of* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$.

Proof. From the local existence theorem we know that for $|t - t_1|$ sufficiently small, $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ exists and from Theorem 21-1

* The inequality $\mathbf{A} > \mathbf{B}$ or $\mathbf{A} \geqslant \mathbf{B}$ means, as usual, that $\mathbf{A} - \mathbf{B}$ is positive definite or positive semidefinite.

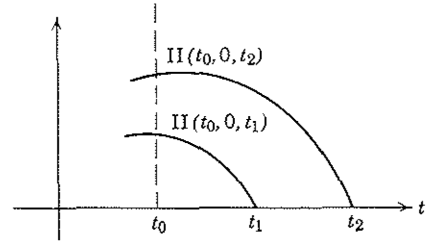


Figure 2. Illustrating inequality (i) of Theorem 4.

$$\mathbf{x}'(t)\mathbf{\Pi}(t, \mathbf{0}, t_1)\mathbf{x}(t) = \min_{\mathbf{u}} \int_t^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt$$

Thus using $\mathbf{L}(t) \geqslant 0$ we see that as long as $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ exists we have

$$\mathbf{0} \leqslant \mathbf{\Pi}(t, \mathbf{0}, t_1)$$

Moreover, if $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ and $\mathbf{\Pi}(t, \mathbf{0}, t_2)$ exist over $t_0 \leqslant t \leqslant t_1$ and $t_0 \leqslant t \leqslant t_2$ then with $\mathbf{x}(t_0) = \mathbf{x}_0$ and $t_2 \geqslant t_1$

$$\begin{aligned} \mathbf{x}_0'\mathbf{\Pi}(t_0, \mathbf{0}, t_1)\mathbf{x}_0 &= \min_{\mathbf{u}} \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt \\ &\leqslant \min_{\mathbf{u}} \int_{t_0}^{t_2} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt \\ &= \mathbf{x}_0'\mathbf{\Pi}(t_0, \mathbf{0}, t_2)\mathbf{x}_0 \end{aligned}$$

Since this holds for all $\mathbf{x}_0$ this validates inequality (i) on the monotone behavior of $\mathbf{\Pi}$.

It remains to show that $\mathbf{\Pi}$ exists over $t_0 \leqslant t \leqslant t_1$ for all t_1. Since the only way $\mathbf{\Pi}$ could fail to exist is for it to pass to infinity for some finite time; we can establish existence by proving boundedness. Consider the control

$$\mathbf{u}_1(t) = \begin{cases} -\mathbf{B}'(t)\mathbf{\Phi}'(t_0, t)[\mathbf{W}(t_0, t_1)]^{-1}\mathbf{x}(t_0) & t_0 \leqslant t \leqslant t_1 \\ \mathbf{0} & t_1 \leqslant t \end{cases}$$

applied at t_0 to the differential equation. As we have seen (Theorem 13.1), this makes $\mathbf{x}(t)$ zero for $t \geqslant t_1$. Hence for this control and resulting response $\mathbf{x}_1(t)$ satisfies

$$\begin{aligned} \int_{t_0}^{\infty} \mathbf{u}_1'(t)\mathbf{u}_1(t) + \mathbf{x}_1'(t)\mathbf{L}(t)\mathbf{x}_1(t)\, dt &\geqslant \min_u \int_{t_0}^{t_2} \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{L}(t)x(t)\, dt \\ &= \mathbf{x}'(t_0)\mathbf{\Pi}(t_0, \mathbf{0}, t_2)\mathbf{x}(t_0) \end{aligned}$$

for all $t_2 > t_1$. An elementary calculation shows that the left side of this inequality is given by $\mathbf{x}'(t_0)\mathbf{N}(t_0, t_1)\mathbf{x}(t_0)$. Since $\mathbf{N}(t_0, t_1)$ is evidently finite, for all t_1 we see that $\mathbf{\Pi}(t, \mathbf{0}, t_1)$ exists on any interval $t_0 \leqslant t \leqslant t_1$. ▮

Exercises

1. Suppose $\mathbf{\Pi}$ satisfies the Riccati equation

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'(t)\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A}(t) + \mathbf{K}(t)\mathbf{B}(t)\mathbf{B}'(t)\mathbf{K}(t) - \mathbf{L}(t)$$

Use the matrix variation of constants formula to show that it satisfies the matrix integral equation

$$\mathbf{\Pi}(t, \mathbf{Q}, t_0) = \mathbf{\Phi}'(t_0, t)\mathbf{Q}\mathbf{\Phi}(t_0, t) + \int_{t_0}^{t} \mathbf{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\mathbf{\Phi}(\sigma, t)\, d\sigma$$

where $\mathbf{\Phi}$ is the transition matrix for

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \tfrac{1}{2}\mathbf{B}(t)\mathbf{B}'(t)\mathbf{\Pi}(t, \mathbf{Q}, t_0)]\mathbf{x}(t)$$

Hence show without any appeal to variational ideas that $\mathbf{\Pi}(t_0, \mathbf{0}, t)$ is monotone increasing with increasing t if $\mathbf{L}$ is nonnegative definite. (J. C. Willems)

2. Given the controllable system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
$$\mathbf{x}(t_0) = \mathbf{x}_0$$

It is desired to determine a matrix $\mathbf{K}$ defined on $t_0 \leqslant t \leqslant t_1$ such that the control

$$\mathbf{u}_0(t) = -\mathbf{B}'(t)\mathbf{K}(t)\mathbf{x}(t)$$

drives the system from $\mathbf{x}_0$ at t_0 to $\mathbf{0}$ at t_1 and minimizes

$$\eta = \int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t)\, dt$$

Find $\mathbf{K}$. What is the relation between $\mathbf{K}$ and the controllability Gramian?

3. Consider the Riccati equation

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A} + \mathbf{K}(t)\mathbf{B}\mathbf{B}'\mathbf{K}(t) - \mathbf{L}(t)$$

with $\mathbf{L}(t) = \mathbf{L}'(t) = \mathbf{L}(t + T)$ and $\mathbf{A}$ and $\mathbf{B}$ constant. Suppose rank $(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B}) = n = \dim \mathbf{x}$. Show that for a suitable value of $\mathbf{Q} = \mathbf{Q}' > \mathbf{0}$ the solution $\mathbf{\Pi}(t, \mathbf{Q}, 0)$ is periodic of period T provided $\mathbf{L}(t)$ is positive definite for all t.

4. Let the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

be controllable and observable, and let the loss functional be

$$J = \int_0^T \mathbf{u}'(t)\mathbf{u}(t) + \mathbf{x}'(t)\mathbf{C}'\mathbf{C}\mathbf{x}(t)\, dt + \mathbf{x}'(T)\mathbf{Q}_1\mathbf{x}(T), \qquad \mathbf{Q}_1 > \mathbf{0}$$

(a) Show that the Riccati equation

$$\dot{\mathbf{K}}(t) = -\mathbf{A}'\mathbf{K}(t) - \mathbf{K}(t)\mathbf{A} + \mathbf{K}(t)\mathbf{B}\mathbf{B}'\mathbf{K}(t) - \mathbf{C}'\mathbf{C}$$
$$\mathbf{K}(T) = \mathbf{Q}_1$$

has a unique solution $\mathbf{K}(t)$ such that $\mathbf{K}(t) > \mathbf{0}$ for all $t \leqslant T$, and hence deduce that $\mathbf{K}^{-1}(t)$ exists for all $t \leqslant T$.

(b) Show that $\mathbf{K}^{-1}$ also satisfies a Riccati equation. Identify the system and cost functional which lead to this second Riccati equation. Is this system controllable and/or observable?

5. Find the control which drives the system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

from $x_1 = 2$, $x_2 = 0$ at $t = 0$ to the circle $x_1^2 + x_2^2 = 1$ at $t = 1$ and minimizes

$$\eta = \int_0^1 u^2(t) + x_1^2(t)\, dt$$

6. Does the function $\mathbf{x}(t) = 0$; $0 \leqslant t \leqslant 2$ minimize

$$\eta = \int_0^2 \dot{x}^2(t) - \gamma x^2(t)\, dt$$

For the boundary condition $x(0) = x(2) = 0$?

7. Consider a unit mass on a spring of unity spring constant. The Lagrangian is $L(x, \dot{x}) = \frac{1}{2}(\dot{x}^2 - x^2)$. If $x(0) = \alpha$ and $x(T) = \beta$ then for what values of T can we assert that the actual motion minimizes the integral of the Lagrangian subject to the fixed end points $x(0) = \alpha$, $x(T) = \beta$? (Compare with *Hamilton's principle* in mechanics.)

8. Show that if $\int_0^{1/2} u^2(\sigma)\, d\sigma \leqslant 1$ then the solution $\pi(t, 0, 0)$ of

$$\dot{k}(t) = k^2(t) + 1 + u(t)$$

is finite at $t = \frac{1}{2}$.
Hint: For $k < 1$ we have

$$\dot{k}(t) \leqslant 2 + u(t)$$

which for $\int_0^{\alpha} u^2(t)\, dt \leqslant \beta$ gives

$$k(\alpha) \leqslant 2\alpha + \frac{\sqrt{\beta}}{\alpha}$$

For $k \geq 1$ we have

$$\dot{k}(t)/k^2(t) \leq 2 + u(t)$$

which for

$$\int_\alpha^{1/2} u^2(\sigma)\,d\sigma \leq 1 - \beta$$

gives

$$k^{-1}(\tfrac{1}{2}) - k^{-1}(\alpha) \leq 2(\tfrac{1}{2} - \alpha) + \frac{\sqrt{1-\beta}}{\frac{1}{2} - \alpha}$$

9. Let $\mathbf{J}$ be a $2n$-by-$2n$ matrix defined in terms of n-by-n identity matrices as

$$\mathbf{J} = \begin{bmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{bmatrix}$$

a $2n$-by-$2n$ matrix is called *symplectic* if $\mathbf{A}'\mathbf{J}\mathbf{A} = \mathbf{J}$. Show that symplectic matrices are invertible and that the product of two symplectic matrices is symplectic. Show that the transition matrix for a canonical system is symplectic.

10. Prove Theorem 2.

25. FREQUENCY RESPONSE INEQUALITIES

The least squares theory of linear constant systems on the interval $0 \leq t \leq \infty$ can be treated entirely in terms of Laplace transformed quantities. We have not done so here because this approach is not well suited for time varying equations or finite intervals. To give the reader some appreciation for the power of this approach in dealing with linear constant problems we discuss here some applications of transform techniques. One of the results (Theorem 2) plays an important role in the treatment of stability of time varying systems. We begin with a simple application of Parseval's equality for sums of exponentials.

Lemma. *Let* $\mathbf{u}$ *be given by* $\mathbf{u}(t) = \mathbf{H}e^{\mathbf{F}t}\mathbf{g}$. *Let* $\mathbf{y}$ *be given by*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t); \qquad \mathbf{x}(0) = \mathbf{0}$$

Assume that the eigenvalues of $\mathbf{A}$ *and* $\mathbf{F}$ *lie in the half-plane* $\operatorname{Re} s < 0$. *If* $\mathbf{R}(s) = \mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ *and* $\alpha^2\mathbf{I} - \mathbf{R}'(-i\omega)\mathbf{R}(i\omega) \geq \mathbf{0}$ *for all real* ω *then*

$$\int_0^\infty \mathbf{y}'(t)\mathbf{y}(t)\,dt \leq \alpha^2 \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t)\,dt$$

Proof. Since $\mathbf{x}(0)$ is zero the transform of $\mathbf{y}$ is $\hat{\mathbf{y}} = \mathbf{R}\hat{\mathbf{u}}$. Using Parseval's relation we have

$$\begin{aligned}\int_0^\infty \mathbf{y}'(t)\mathbf{y}(t)\,dt &= \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} \hat{\mathbf{u}}'(-i\omega)\mathbf{R}'(-i\omega)\mathbf{R}(i\omega)\hat{\mathbf{u}}(i\omega)\,d\omega \\ &\leq \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} \hat{\mathbf{u}}(-i\omega)(\alpha^2\mathbf{I})\hat{\mathbf{u}}(i\omega)\,d\omega \\ &= \alpha^2 \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t)\,dt \ \blacksquare\end{aligned}$$

Using this result it is possible to derive a simple inequality relating the solutions of the quadratic equation

$$\mathbf{KA} + \mathbf{A}'\mathbf{K} - \mathbf{KBB}'\mathbf{K} = -\alpha\mathbf{C}'\mathbf{C}; \quad \alpha \geq 0$$

and the linear equation

$$\mathbf{KA} + \mathbf{A}'\mathbf{K} = -\mathbf{C}'\mathbf{C}$$

First of all notice that if the eigenvalues of $\mathbf{A}$ lie in the half-plane $\operatorname{Re} s < 0$, then positive definite solutions exist for both equations. If we denote the positive definite solution of the quadratic equation by $\mathbf{K}_+$ and if we denote the solution of the linear equation by $\mathbf{K}_0$, then clearly $\mathbf{K}_+ \leq \alpha\mathbf{K}_0$. (One way to see this is to think of the optimal control interpretation of $\mathbf{x}_0'\mathbf{K}_+\mathbf{x}_0$.) Can one bound $\mathbf{K}_+$ from below? The following theorem gives a very useful bound in terms of $\mathbf{K}_0$ and the frequency response.

Theorem 1. *Let* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *be a minimal realization of* $\mathbf{R}$. *Assume that the eigenvalues of* $\mathbf{A}$ *lie in the half-plane* $\operatorname{Re} s < 0$. *If* $\mathbf{K}_+(\alpha)$ *is the positive definite solution of* $\mathbf{KA} + \mathbf{A}'\mathbf{K} - \mathbf{KBB}'\mathbf{K} = -\alpha\mathbf{C}'\mathbf{C}$ *and if* $\mathbf{K}_0$ *is the solution of* $\mathbf{KA} + \mathbf{A}'\mathbf{K} = -\mathbf{C}'\mathbf{C}$, *then for* $\alpha > 0$

$$\alpha\mathbf{K}_0 \geq \mathbf{K}_+(\alpha) \geq \frac{\alpha}{1+\alpha}\mathbf{K}_0$$

provided $\mathbf{I} - \mathbf{R}'(-i\omega)\mathbf{R}(i\omega) \geq 0$.

Proof. We know from Theorem 23.3 that for the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t); \qquad \mathbf{x}(0) = \mathbf{x}_0$$

we have

$$\min_u \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \alpha\mathbf{y}'(t)\mathbf{y}(t)\,dt = \mathbf{x}_0'\mathbf{K}_+(\alpha)\mathbf{x}_0$$

If $\mathbf{u}_0$ and $\mathbf{y}_0$ denote the optimal controls, then

$$\mathbf{y}_0 = \mathbf{C}e^{(\mathbf{A}-\mathbf{BB}'\mathbf{K}_+(\alpha))t}\mathbf{x}_0; \qquad \mathbf{u}_0 = -\mathbf{B}'\mathbf{K}_+(\alpha)e^{(\mathbf{A}-\mathbf{BB}'\mathbf{K}_+(\alpha))t}\mathbf{x}_0$$

Moreover, $\mathbf{y}_0$ can be expressed using transforms as the sum of an initial condition term and the effect of $\hat{\mathbf{u}}_0$, i.e.

$$\hat{\mathbf{y}}_0 = \mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{x}_0 + \mathbf{G}(s)\hat{\mathbf{u}}_0(s) \stackrel{\text{def}}{=} \hat{\mathbf{y}}_1(s) + \hat{\mathbf{y}}_2(s)$$

In terms of this notation

$$\mathbf{x}_0'\mathbf{K}_+(\alpha)\mathbf{x}_0 = \int_0^\infty \alpha[\mathbf{y}_1'(t) + \mathbf{y}_2'(t)][\mathbf{y}_1(t) + \mathbf{y}_2(t)] + \mathbf{u}'(t)\mathbf{u}(t)\,dt$$

Using the preceding lemma we have

$$\int_0^\infty \mathbf{u}'(t)\mathbf{u}(t)\,dt \geqslant \int_0^\infty \mathbf{y}_2'(t)\mathbf{y}_2(t)\,dt$$

Also, from the known relationship between $\mathbf{KA} + \mathbf{A}'\mathbf{K} = -\mathbf{C}'\mathbf{C}$ and quadratic integrals,

$$\mathbf{x}_0'\mathbf{K}_0\mathbf{x}_0 = \int_0^\infty \mathbf{y}_1'(t)\mathbf{y}_1(t)\,dt$$

Denote this last quantity by ρ^2 and let γ^2 be defined by

$$\gamma^2 = \int_0^\infty \mathbf{y}_2'(t)\mathbf{y}_2(t)\,dt$$

Combining these results we have

$$\mathbf{x}_0'\mathbf{K}_+(\alpha)\mathbf{x}_0 \geqslant \alpha\rho^2 - \left|\int_0^\infty 2\alpha y_1'(t)y_2(t)\,dt\right| + (\alpha+1)\gamma^2$$

Now use the Schwartz inequality

$$\left|\int_0^\infty \mathbf{y}_1'(t)\mathbf{y}_2(t)\,dt\right| \leqslant \sqrt{\int_0^\infty \mathbf{y}_1'(t)\mathbf{y}_1(t)\,dt}\sqrt{\int_0^\infty \mathbf{y}_2'(t)\mathbf{y}_2(t)\,dt}$$

to obtain

$$x_0'\mathbf{K}_+(\alpha)\mathbf{x}_0 \geqslant \alpha\rho^2 - 2|\alpha|\rho\gamma + (\alpha+1)\gamma^2$$

Considering this is a function of γ, it has a minimum at $\gamma = |\alpha|\rho/(1+\alpha)$ and the minimum value is $\alpha\rho^2/(1+\alpha)$. Therefore

$$\mathbf{x}'\mathbf{K}_1\mathbf{x}_0 \geqslant \left(\frac{\alpha}{1+\alpha}\right)\mathbf{x}_0'\mathbf{K}_0\mathbf{x}_0 \quad ▮$$

In the above inequalities we have treated α as if it could be negative. This will be capitalized on below.

As a second application of these ideas, we state a theorem on the existence of solutions of least squares problems where the integrand is not positive definite. This problem is more difficult than the corresponding ones with positive semi-definite integrands.

Theorem 2. *Suppose all the eigenvalues of* $\mathbf{A}$ *lie in* $\mathrm{Re}\, s < 0$. *If*

$$\mathbf{I} - \mathbf{R}'(-i\omega)\mathbf{R}(i\omega) \geqslant \mathbf{0}$$

for all real ω *and if* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *is a minimal realization of* $\mathbf{R}$, *then for* $-1 \leqslant \alpha < 0$ *there exists a unique negative definite solution* $\mathbf{K}_+(\alpha)$ *of* $\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{KBB}'\mathbf{K} = -\alpha\mathbf{C}'\mathbf{C}$ *having the property that* $\mathbf{A} - \mathbf{BB}'\mathbf{K}_+(\alpha)$ *has all its eigenvalues in* $\mathrm{Re}\, s \leqslant 0$. *For* $\mathbf{y}$ *and* $\mathbf{u}$ *related by*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

the quantity

$$\eta = \min_u \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \alpha\mathbf{y}'(t)\mathbf{y}(t)\,dt$$

exists and is given by $\mathbf{x}_0'\mathbf{K}_+(\alpha)\mathbf{x}_0$.

Proof. Taking a derivative with respect to α gives

$$[\mathbf{A}' - \mathbf{K}(\alpha)\mathbf{BB}']\,d\mathbf{K}(\alpha) + d\mathbf{K}(\alpha)[\mathbf{A} - \mathbf{BB}'\mathbf{K}(\alpha)] = -d\alpha\mathbf{C}'\mathbf{C}$$

and so for α such that $[\mathbf{A} - \mathbf{BB}'\mathbf{K}(\alpha)]$ has its eigenvalues in $\mathrm{Re}\, s < 0$ we have

$$d\mathbf{K}(\alpha)/d\alpha = -\int_0^\infty e^{[\mathbf{A}-\mathbf{BB}'\mathbf{K}(\alpha)]'t}\mathbf{C}'\mathbf{C}e^{[\mathbf{A}-\mathbf{BB}'\mathbf{K}(\alpha)]t}\,dt$$

Hence starting from $\mathbf{O}$, $\mathbf{K}_+(\alpha)$ is monotone-decreasing, in a matrix sense, with decreasing α. The question is, "How big can $|\alpha|$ get before solutions cease to exist?"

For $|\alpha|$ sufficiently small, the eigenvalues of $\mathbf{A} - \mathbf{BB}'\mathbf{K}_+(\alpha)$ lie in the half plane $\mathrm{Re}\, s < 0$. Hence we have from Theorem 23-6 for $\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) + \mathbf{Bu}(t)$; $\mathbf{y}(t) = \mathbf{Cx}(t)$,

$$\mathbf{x}_0'\mathbf{K}_+(\alpha)\mathbf{x}_0 = \min_u \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) + \alpha\mathbf{y}'(t)\mathbf{y}(t)\,dt$$

However, reasoning exactly as in the proof of Theorem 1, we see that this last integral is bounded from below by

$$K_+(\alpha) \geqslant [\alpha/(1+\alpha)]\mathbf{K}_0; \ \alpha > -1$$

Therefore the solution $\mathbf{K}_+(\alpha)$ valid for $|\alpha|$ small, can be continued up to $\alpha = -1$.

We now need to show that $\lim_{\alpha\to-1}\mathbf{K}_+(\alpha)$ exists and satisfies the given equation. Notice that since $\mathbf{K}_+(\alpha)$ is monotone decreasing with α the limit will exist if we can bound $\mathbf{K}_+(\alpha)$ from below. However, for $\alpha > -1$ we see that $\mathbf{K}(\alpha)$ is bounded from below by (see exercise 5)

$$\mathbf{K}_+(\alpha) \geqslant \left[\int_0^\infty e^{-\mathbf{A}t}\mathbf{BB}'e^{\mathbf{A}'t}\,dt\right]^{-1}$$

Consequently, $\lim_{\alpha \to -1} \mathbf{K}(\alpha)$ does exist and using a continuity argument we see that the limit indeed satisfies the equation. The results contained in the proof of Theorem 23.3 suffices to give uniqueness. ∎

Exercises

1. Suppose that $\mathbf{F}$ is m by m and suppose $\mathbf{I} - \mathbf{F}'(t)\mathbf{F}(t) \geqslant 0$ for all $t \geqslant 0$. If $\mathbf{I} - \mathbf{R}'(i\omega)\mathbf{R}(i\omega) \geqslant \mathbf{0}$ for all real ω and if $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ is a minimal realization of $\mathbf{R}$ then there exists a control $\mathbf{u}$ such that
$$\eta = \int_0^\infty \mathbf{u}'(t)\mathbf{u}(t) - \mathbf{y}'(t)\mathbf{F}'(t)\mathbf{F}(t)\mathbf{y}(t)\,dt$$
is a minimum.
2. Show that if $\mathbf{R}$ is a matrix valued rational function of s which goes to zero at $|s| = \infty$ and if
$$\mathbf{I} - \mathbf{R}'(-s)\mathbf{R}(s)\,|_{s=i\omega} \geqslant \mathbf{0}$$
then there exists a matrix valued rational function $\mathbf{H}$ such that
$$\mathbf{I} - \mathbf{R}'(-s)\mathbf{R}(s) = \mathbf{H}'(-s)\mathbf{H}(s)$$
and the McMillan degree of $\mathbf{H}$ is no larger than that of $\mathbf{R}$.
3. We call a linear constant system $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *passive* if there exists a positive definite matrix $\mathbf{K}$ such that along solutions of
$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$
we have
$$\frac{d}{dt}\mathbf{x}'(t)\mathbf{K}\mathbf{x}(t) \leqslant \mathbf{u}'(t)\mathbf{y}(t)$$
Show that a system is passive if and only if there exists a positive definite solution of a certain matrix equations.
4. Show that if
$$\ddot{y}(t) + 4\dot{y}(t) + y(t) = 0$$
$$\ddot{x}(t) + 4\dot{x}(t) + x(t) = \tfrac{1}{4}u(t)$$
with $x(0) = y(0)$ and $\dot{x}(0) = \dot{y}(0)$ then regardless of the choice of u,
$$\frac{\int_0^\infty x^2(t) + u^2(t)\,dt}{\int_0^\infty y^2(t)\,dt} \geqslant \frac{16}{17}$$

5. Show that in the proof of Theorem 2
$$\mathbf{K}_+(\alpha) \geqslant -\left[\int_0^\infty e^{-\mathbf{A}t}\mathbf{B}\mathbf{B}'e^{-\mathbf{A}'t}\,dt\right]^{-1}$$
for $\alpha > -1$. (Introduce $\mathbf{K}_-(\alpha)$ using Exercise 2, Section 23.)

26. SCALAR CONTROL: SPECTRAL FACTORIZATION

As we have seen, it is always possible to convert a linear nth order scalar equation into a first order vector equation and thus there is no need to provide additional arguments for nth order scalar equations. On the other hand, a number of arbitrary choices must always be made in the process of writing a scalar equation in first order form and these choices frequently obscure the essentials of the problem. Here we discuss an alternative approach which is not so broadly applicable as those discussed in Sections 22 and 23 but which provides a very informative point of view for a certain class of problems. Some results of this section will be used in later developments.

Consider the linear constant scalar equation
$$x^{(n)}(t) + p_{n-1}x^{(n-1)}(t) + \cdots + p_1 x^{(1)}(t) + p_0 x(t) = 0$$
or
$$p(D)x(t) = 0$$
where as usual $D = d/dt$ and $p(D) = D^n + p_{n-1}D^{n-1} + \cdots + p_1 D + p_0$. If $h(D)$ is a second polynomial in D of degree $n - 1$ or less, the integral
$$\eta = \int_{t_0}^{t_1} [h(D)x(t)]^2\,dt$$
can be evaluated by finding a first order vector differential equation representation for the nth order equation, expressing $[h(D)x(t)]^2$ in terms of the state variables, and finally solving an equation of the form $\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{C}$. We are interested in establishing a more direct route.

Everything in this section centers around a result on integrals of the form
$$\eta = \int_a^b \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}\,\phi^{(i)}(t)\phi^{(j)}(t)\,dt \tag{P}$$
where ϕ is an n-times differentiable function of time. In general the value of such an integral will depend not only on the values which ϕ and its derivatives take on at a and b but also on the open interval $a < t < b$. In some cases the dependence on the form of the path disappears; for example
$$\int_a^b 2\dot{\phi}(t)\phi(t)\,dt = \phi^2(b) - \phi^2(a)$$

The integral η defined by equation (P) will be said to be *independent of path* if it can be evaluated in terms of ϕ and its first $n-1$ derivatives at a and b.

A simple characterization of path independent integrals is given by the following lemma.

Lemma 1. *Assume that ϕ is an n times differentiable function of t and that α_{ij} are constants. The integral* (P) *is independent of path if and only if the polynomial*

$$h(s) = \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}[s^i(-s)^j + (-s)^i s^j] \tag{E}$$

vanishes identically.

Proof. If $i+j$ is odd then by using the integration-by-parts formula $(\nu > \mu)$

$$\int_a^b \phi^{(\nu)}(t)\phi^{(\mu)}(t)\,dt = \phi^{(\nu-1)}(t)\phi^{(\mu)}(t)\Big|_a^b - \int_a^b \phi^{(\nu-1)}(t)\phi^{(\mu+1)}(t)\,dt$$

a total of $(|i-j|-1)/2$ times the integral is reduced to one of the form $x^{(p)}(t)x^{(p-1)}(t)$ which can be integrated. If $i+j$ is even the situation is more interesting. Using integration-by-parts a total of $|i-j|/2$ times leaves one with $|i-j|/2$ terms expressed in terms of the end points plus a term $\frac{1}{2}[(-1)^i + (-1)^j]\{x^{([i+j]/2)}\}^2$ inside the integral. Adding all these up gives

$$\eta = \eta_0 + \frac{1}{2}\int_a^b \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}[(-1)^i + (-1)^j]\{\phi^{([i+j]/2)}\}^2\,dt$$

where η_0 is a function of the end points only and the sum is over $i+j$ even. But since

$$\sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}[s^i(-s)^j + (-s)^i s^j] = \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}[(-1)^j + (-1)^i]s^{i+j}$$

We see that the integral term in equation (P) vanishes if and only if equation (E) holds. ▮

It is clear that if the integrability condition holds then η can be expressed as

$$\eta = \int_a^b \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}\,\phi^{(i)}(t)^{(j)}(t)\,dt$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} k_{ij}\,\phi^{(i-1)}(t)\phi^{(j-1)}(t)\Big|_a^b$$

with $k_{ij} = k_{ji}$. That is to say, if η is expressible in terms of the end points it is a quadratic form in ϕ and its first $n-1$ derivatives. Since the times a and b play no role in our applications we introduce the notation

$$\eta(\mathbf{x}) = \int_{t(\mathbf{0})}^{t(\mathbf{x})} \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_{ij}\,\phi^{(i)}(t)\phi^{(j)}\,dt = \mathbf{x}'\mathbf{K}\mathbf{x}$$

to describe a path integral which starts at $\mathbf{0}$ (i.e. $\phi[t(\mathbf{0})] = \phi^{(1)}[t(\mathbf{0})] = \cdots \phi^{(n-1)}[t(\mathbf{0})] = 0$) and ends at a point $\mathbf{x}$ (i.e. $\phi[t(\mathbf{x})] = x_1,\ \phi^{(1)}[t(\mathbf{x})] = x_2 \cdots \phi^{(n-1)}[t(\mathbf{x})] = x_n$). Thus this integral is to be interpreted as a line integral in the state space (see Figure 1).

What we have developed is nothing more than an integral representation for a quadratic form just as is

$$\eta(\mathbf{x}) = \int_0^\infty \mathbf{x}' e^{\mathbf{A}'t}\mathbf{C}e^{\mathbf{A}t}\mathbf{x}\,dt$$

Its importance stems from the fact that under certain conditions derivatives can be calculated simply by removing the integral sign.

With this result in mind, let us return to the problem of evaluating the integral

$$\eta = \int_0^\infty [h(D)x(t)]^2\,dt$$

along solutions of the differential equation

$$p(D)x(t) = 0$$

Suppose that the degree of $p(D)$ exceeds that of $h(D)$. Then this problem could be treated using the results of Section 11 by solving

$$\dot{\mathbf{Q}}(t) = \mathbf{A}'\mathbf{Q}(t) + \mathbf{Q}(t)\mathbf{A} + \mathbf{C}'\mathbf{C}; \qquad \mathbf{Q}(t_1, t_1) = \mathbf{0}$$

With $\mathbf{A}$ and $\mathbf{C}$ taken (for example) from the standard controllable representation of $h(s)/p(s)$.

Let's try a different approach. Suppose it is possible to solve the linear polynomial equation

$$p(s)q(-s) + q(s)p(-s) = h(s)h(-s) \tag{*}$$

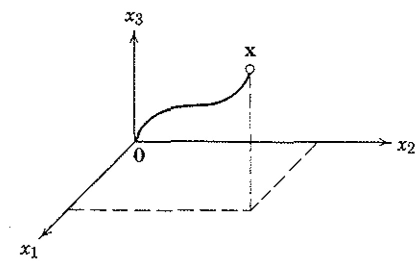


Figure 1. *Illustrating the meaning of* $\eta(\mathbf{x})$.

for $q(s)$. Along solutions of $p(D)x(t) = 0$ we have

$$\int_{t_0}^{t_1} [h(D)x(t)]^2 \, dt = \int_{t_0}^{t_1} [h(D)x(t)]^2 - 2p(D)x(t)q(D)x(t) \, dt$$

but in view of equation (*) and our previous Lemma we see that for some q

$$\int_{t_0}^{t_1} [h(D)x(t)]^2 - 2p(D)x(t)q(D)x(t) \, dt = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} q_{ij} x^{(i)}(t) x^{(j)}(t) \Big|_{t_0}^{t_1}$$

Thus from a knowledge of $q(s)$ we can evaluate the given integral in terms of x and its derivatives at t_0 and t_1.

This focuses attention on the polynomial equation (*). Clearly it is playing a role here equivalent to that played by $\mathbf{A}'\mathbf{Q} + \mathbf{QA} + \mathbf{M} = \mathbf{0}$ in Section 11. Its theory is similar but more elementary.

Theorem 1. *Consider the scalar polynomial equation (*) with $p(s)$ monic and of degree n. Assume that the coefficients of all polynomials are real and that the degree of $h(s)$ is less than n. Then given $p(s)$ and $h(s)$ there exists a unique solution $q(s)$ if all the zeros of $p(s)$ lie in the half-plane* $\operatorname{Re} s < 0$.

Proof. Pick $\mathbf{c}$, $\mathbf{A}$, and $\mathbf{b}$ so that

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{b}u; \qquad y = \mathbf{cx}$$

is the standard controllable representation of $h(s)/p(s)$. Since we have assumed that the zeros of $p(s)$ lie in $\operatorname{Re} s < 0$ Theorem 2 of Section 11 guarantees that we can solve the equation

$$\mathbf{A}'\mathbf{Q} + \mathbf{QA} = -\mathbf{c}'\mathbf{c}$$

If $\mathbf{Q}_0$ is the solution then

$$q(s) = p(s)\mathbf{b}'\mathbf{Q}_0(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$$

satisfies equation (*). To see this add and subtract $\mathbf{Q}_0 s$ from the left side of $\mathbf{A}'\mathbf{Q} + \mathbf{QA} = -\mathbf{c}'\mathbf{c}$ to get

$$-\mathbf{Q}(\mathbf{I}s - \mathbf{A}) - (-\mathbf{I}s - \mathbf{A}')\mathbf{Q} = -\mathbf{c}'\mathbf{c}$$

Now pre- and postmultiply by $\mathbf{b}'(-\mathbf{I}s - \mathbf{A}')^{-1}$ and $(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$ respectively, and remove a minus sign to get

$$\mathbf{b}'(-\mathbf{I}s - \mathbf{A}')^{-1}\mathbf{Qb} + \mathbf{b}'\mathbf{Q}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} = \mathbf{b}'(-\mathbf{I}s - \mathbf{A}')^{-1}\mathbf{cc}'(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$$

If we now multiply by $p(s)p(-s)$ we see that

$$p(s)p(-s)\mathbf{b}'(-\mathbf{I}s - \mathbf{A}')^{-1}\mathbf{Qb} + p(s)p(-s)\mathbf{b}'\mathbf{Q}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} = h(s)h(-s)$$

or

$$p(s)q(-s) + p(-s)q(s) = h(s)h(-s)$$

as claimed.

To prove uniqueness we observe that the equations for the n coefficients of $q(s)$ are linear. We have displayed a solution for all possible choices of $h(s)$. This is an n parameter family. Hence the matrix defining the transformation between the elements of $h(s)$ and those of $q(s)$ must be nonsingular and the solutions unique. ∎

Now let us consider some variational problems. We will immediately treat only the simplest cases—linear, constant, scalar, systems on an infinite time interval.

Starting with

$$p(D)x(t) = u(t); \qquad y(t) = h(D)x(t)$$

find $x(t)$ such that

$$\eta = \int_0^\infty u^2(t) + y^2(t) \, dt = \int_0^\infty [p(D)x(t)]^2 + [h(D)x(t)]^2 \, dt$$

is a minimum. Taking a cue from the proofs in Section 22 we look for some function of the state at $t = 0$ and infinity which we could add to make the integral a perfect square. If this square were to be $[r(D)x(t)]^2$ then we would require that the integrand of

$$\int_0^\infty [p(D)x(t)]^2 + [h(D)x(t)]^2 - [r(D)x(t)]^2 \, dt = \eta - \int_0^\infty [r(D)x(t)]^2 \, dt$$

be a perfect differential. But from Lemma 1 we see that this is the case if

$$p(s)p(-s) + h(s)h(-s) = r(s)r(-s) \tag{SF}$$

Moreover, if $r(s)$ has all its zeros in $\operatorname{Re}[s] < 0$ then an input u which makes

$$r(D)x(t) = 0$$

is surely a good candidate to minimize η.

As an equation in the unknown, r, equation (SF) is nonlinear. The conditions under which there exists a real polynomial $r(s)$ which satisfies it have long been known since they play a wide role in mathematical physics and engineering.

Theorem 2. (Spectral Factorization) *Let $c(s)$ be an even polynomial having real coefficients and being of degree $2n$. If $c(s)$ is nonnegative for* $\operatorname{Re}[s] = 0$ *then there exists a polynomial $r(s)$ which has real coefficients and is of degree n such that*

$$r(s)r(-s) = c(s)$$

Moreover, $r(s)$ can be taken to have all its zeros in the half-plane $\operatorname{Re}[s] \leqslant 0$.

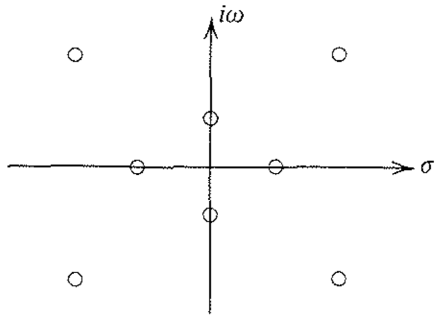


Figure 2. Showing the zero locations of an even polynomial with real coeficients.

Proof. To begin let us observe that since $c(s)$ has real coefficients and since $c(s) = c(-s)$ it follows that if s_i is a zero then so is $-s_i$ and $-\bar{s}_i$ where an overbar denotes complex conjugate. This means that with the possible exception of zeros which are purely real or purely imaginary, all zeros occur in fours. (See Figure 2.) Thus $c(s)$ can be expressed in factored form as

$$c(s) = c_0 \prod_j (s + p_j)(s - p_j) \prod_i [(s^2 - \sigma_i^2)^2 + 2\omega_i^2(s^2 + \sigma_i^2) + \omega_i^4]$$

with the σ_i and ω_i real and non-negative and the p_j purely real or purely imaginary.

If $c(i\omega) \geqslant 0$ for all real ω then clearly any zero of $c(s)$ which lies on the imaginary axis must be of even multiplicity for otherwise $c(s)$ would change sign on the imaginary axis. This means we can assume that all the p_i in the above representation are real and positive, the imaginary axis factors being accounted for by $(s^2 + \omega_i^2)^2$ type factors. We take for $r(s)$

$$r(s) = \sqrt{c_0} \prod_j (s + p_j) \prod_i [(s + \sigma_i)^2 + \omega_i^2]$$

Clearly this meets all the requirements of the theorem and at the same time has all its zeros in the half-plane $\text{Re}[s] \leqslant 0$. ∎

Given an even polynomial $c(s)$ which is nonnegative for $\text{Re}[s] = 0$ we write

$$[c(s)]^+ = r(s)$$

with $r(s)$ given above and call $[c(s)]^+$ the *left half-plane spectral factor* of $c(s)$. We also write

$$[c(s)]^- = r(-s)$$

and call $[c(s)]^-$ the *right half-plane spectral factor of* $c(s)$.

Theorem 3. *Assume that $p(D)$ and $q(D)$ have no common factors and consider the system*

$$p(D)x(t) = u(t); \qquad y(t) = q(D)x(t)$$

Then the control

$$u(t) = -[p(D)p(-D) + q(D)q(-D)]^+ x(t) + p(D)x(t)$$

drives the system from a given initial state at $t = 0$ to the equilibrium point at $t = \infty$. For this choice of u

$$\eta = \int_0^\infty u^2(t) + y^2(t)\, dt$$

$$= \int_{t(0)}^{t(x)} (p(D)x)^2 + (q(D)x)^2 - \{[p(D)p(-D) + q(D)q(-D)]^+ x\}^2\, dt$$

is less than the corresponding value for any other u which drives x to zero in finite or infinite time.

Proof. We will manipulate the expression for η in such a way as to make the minimizing choice of x obvious. Consider adding and subtracting $\{[p(D)p(-D) + q(D)q(-D)]^+ x(t)\}^2$ from η to get

$$\eta = \int_0^\infty \{p(D)x)^2 + (q(D)x(t)\}^2 - \{[p(D)p(-D) + q(D)q(-D)]^+ x(t)\}^2\, dt$$

$$+ \int_0^\infty \{[p(D)p(-D) + q(D)q(-D)]^+ x(t)\}^2\, dt$$

From Lemma 1 on path integrals we see that the first integral is independent of path. The value of x and its derivatives at $t = 0$ are assumed given and we know $x(t)$ and all its derivatives approach zero as t approaches infinity.

Since the end points are fixed at $t = 0$ and at $t = \infty$, we cannot influence the value of that term. Clearly one can do no better than to make

$$[p(D)p(-D) + q(D)q(-D)]^+ x(t) = 0$$

since this will make $x(t)$ go to zero and will also make the second integral in the above expression for η be zero. This will be the case if

$$u(t) = -[p(D)p(-D) + q(D)q(-D)]^+ x(t) + p(D)x(t)$$

and this must be the minimizing choice of $u(t)$. ∎

Finally, notice that if the equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

is in standard controllable form then $\mathbf{x}(t) = [x(t), x^{(1)}(t) \cdots x^{(n-1)}(t)]'$ and so if

$$\mathbf{x}'\mathbf{K}\mathbf{x} = \int_{t(0)}^{t(x)} [p(D)x(t)]^2 + [q(D)x(t)]^2 - [(p(D)p(-D) + q(D)q(-D))^+ x(t)]^2 \, dt$$

then

$$\mathbf{KA} + \mathbf{A}'\mathbf{K} - \mathbf{Kbb}'\mathbf{K} = -\mathbf{c}'\mathbf{c}$$

provided $\mathbf{c}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b} = q(s)/p(s)$.

Exercises

1. Show that the control law which minimizes

$$\eta = \int_0^\infty u^2(t) + x^2(t) \, dt; \qquad x(0) = x_0$$

for the system

$$\dot{x}(t) = \alpha x(t) + 2u(t)$$

is

$$u = \left(-\frac{1}{2}\alpha - \sqrt{\frac{\alpha^2}{4} + 1}\right)x$$

2. Find necessary and sufficient conditions for the existence of

$$\eta = \min_u \int_0^\infty u(t)\, y(t) \, dt$$

for the system

$$p(0)x(t) = u(t); \qquad y(t) = q(0)x(t)$$

3. Show that if $\mathbf{x}(0) = \mathbf{0}$ and all the eigenvalues of $\mathbf{A}$ lie in $\text{Re}[s] < 0$ then for all $T > 0$

$$\int_0^T y^2(t) \, dt \leqslant \max_\omega |\, g(i\omega)|^2 \int_0^T u^2(t) \, dt$$

where

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t); \qquad y(t) = \mathbf{c}\mathbf{x}(t)$$

and

$$g(s) = \mathbf{c}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{b}$$

4. Find $[c(s)]^+$ (i.e. the spectral factor with roots in $\text{Re}\, s < 0$) for the following polynomials.
   (a) $(1 - s^2)$
   (b) $(1 - s^2 + s^4)$
   (c) $(1 + 2\delta s + s^2)(1 - 2\delta s + s^2)$
5. Consider the nth order dynamical system

$$x^{(n)}(t) = u(t)$$

Find $u(t)$ as a function of $x(t)$ and its first $n - 1$ derivative such that

$$\eta = \int_0^\infty u^2(t) + \alpha x^2(t) \, dt$$

is a minimum. Sketch the zeros of $\{[\alpha + s^n][\alpha + (-s)^n]\}^+$ as a function of α.

6. Use integration by parts to show that the values of λ such that the two-point boundary value problem

$$y^{(4)}(t) + \alpha y^{(2)}(t) + \beta y(t) = \lambda y(t); \qquad y(0) = \dot{y}(0) = y(1) = \dot{y}(1) = 0$$

has a nontrivial solution are real.

7. Consider an inhomogeneous, one-dimensional wave equation on $0 \leqslant z \leqslant 1$, $0 \leqslant t < \infty$.

$$\frac{\partial^2 x(z, t)}{\partial t^2} - \frac{\partial^2 x(z, t)}{\partial z^2} = \mathbf{u}(x, t); \qquad x(0, t) = x(1, t) = 0$$

Show that if x is any sufficiently smooth solution then

$$\int_0^1 \left[\frac{\partial x(z, 0)}{\partial z}\right]^2 + \left[\frac{\partial x(z, 0)}{\partial t}\right]^2 dz + \int_0^\infty \int_0^1 \left[\frac{\partial x(z, t)}{\partial t} + u\right]^2 dz\, dt = \int_0^\infty \int_0^1 \left[\frac{\partial x(z, t)}{\partial z}\right]^2 + [u(x, t)]^2 \, dz\, dt$$

hence to minimize the performance measure on the right one should let $u = -\partial x/\partial t$. (Compare with exercise 9, Section 23.)

27. SPECTRAL FACTORIZATION AND $\mathbf{A}'\mathbf{K} + \mathbf{KA} - \mathbf{Kbb}'\mathbf{K} = -\mathbf{c}'\mathbf{c}$

For the scalar input system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

and the functional

$$\eta = \int_0^\infty \mathbf{x}'(t)\mathbf{L}\mathbf{x}(t) + u^2(t) \, dt$$

to minimize we have two methods of possible solution. We can attempt to solve the equation

$$A'K + KA - Kbb'K = -L \tag{CR}$$

or we can attempt to solve the spectral factorization equation

$$p(s)p(-s) + h(s)h(-s) = r(s)r(-s) \tag{SF}$$

We have as yet, however, not discussed a method of passing directly between equations (CR) and (SF) without appealing to the variational problem. The objective of this section is to rectify this situation. We do so by proving a special case of Theorem 23-4 in a new way.

Theorem 1. *Let* $[A, b, c]$ *be a constant minimal realization. Then there exists one or more real symmetric solutions of*

$$A'K + KA - Kbb'K = -cc'$$

exactly one of which is positive definite.

Proof. The rational function m defined by

$$b(-Is - A)^{-1}cc'(Is - A)^{-1} + 1 = m(s)$$

is clearly nonnegative for $s = i\omega$ and hence can be written as

$$b'(-Is - A')^{-1}CC'(Is - A)^{-1}b + 1 = \frac{r(s)r(-s)}{p(s)p(-s)}$$

where $p(s) = \det(Is - A)$. In particular, we can let r be given by

$$r(s) = [m(s)p(s)p(-s)]^{+}$$

Since p is monic and $c'(Is - A)^{-1}b$ goes to zero as $|s|$ approaches infinity we see that r is monic also. Hence $r - p$ is of degree $n - 1$ or less. Since we have assumed $\det(b, Ab, \ldots, A^{n-1}b) \neq 0$ we can without loss of generality assume that (A, b) is in standard controllable form. Obviously there exists an n-vector k such that

$$k'(Is - A)^{-1}b = [r(s) - p(s)]/p(s)$$

Observe that $\det(Is - A + bk') = r(s)$ and hence that $A - bk'$ has all its eigenvalues in the half-plane $\mathrm{Re}[s] \leqslant 0$. However from the equation defining $m(s)$ we see that m has a real part which is bounded from below by 1 for $s = i\omega$ and so m can have no zeros on $\mathrm{Re}[s] = 0$. This means r can have no zeros on $\mathrm{Re}[s] = 0$ and thus $A - bk'$ has its eigenvalues in $\mathrm{Re}[s] < 0$.

Now introduce K_1 as the (unique and positive definite from Theorem 11.3) solution of

$$(A' - kb')K + K(A - bk') = -cc' - kk'$$

Clearly we have

$$A'K_1 + K_1A - K_1bb'K_1 = -CC' - (K_1b - k)'(K_1b - k)'$$

If we can show that $K_1b = k$ then we will have shown that K_1 satisfies (CR). To do this multiply the above equation by -1 and add and subtract $K_1 s$ to get

$$(-Is - A)K_1 + K_1(Is - A) + K_1bb'K_1 = CC' + (K_1b - k)'(K_1b - k)'$$

Now pre and post-multiply this equation by $b'(-Is - A')^{-1}$ and $(Is - A)^{-1}b$ respectively to get

$$\begin{aligned} b'K_1(Is - A)^{-1}b + b'(-Is - A')^{-1}K_1b + b'(-Is - A)^{-1}K_1bb'K_1(Is - A)^{-1}b \\ = b'(-Is - A')^{-1}CC'(Is - A)^{-1}b \\ + b'(-Is - A')^{-1}(K_1b - k)'(K_1b - k)'(Is - A)^{-1}b \end{aligned}$$

Now introduce the notation

$$b'K_1(Is - A)^{-1}b = \frac{n(s)}{p(s)}$$

so that the previous equation becomes

$$\frac{n(s)}{p(s)} + \frac{n(-s)}{p(-s)} + \frac{n(s)n(-s)}{p(s)p(-s)} = \frac{r(s)r(-s)}{p(s)p(-s)} - 1 + \frac{n(s) - [r(s) - p(s)]}{p(s)} + \frac{n(-s) - [r(-s) - p(-s)]}{p(-s)}$$

Now let $n(s) - r(s) + p(s) = \varepsilon(s)$. In terms of $\varepsilon(s)$ and $r(s)$ we have

$$\frac{p(s) + n(s)}{p(s)}\,\frac{p(-s) + n(-s)}{p(-s)} = \frac{r(s)r(-s)}{p(s)p(-s)} + \frac{\varepsilon(s)\varepsilon(-s)}{p(s)p(-s)}$$

or more simply

$$\frac{\varepsilon(s)r(-s)}{p(-s)p(s)} + \frac{\varepsilon(-s)r(s)}{p(s)p(-s)} = 0$$

Clearly $p(s)p(-s)$ and $r(s)r(-s)$ are both nonzero so this last equation implies that

$$\frac{\varepsilon(s)}{r(s)} + \frac{\varepsilon(-s)}{r(-s)} = 0$$

Using the fact that $[r(s)]^{-1}$ is analytic in the half-plane $\mathrm{Re}\, s \geqslant 0$ we see that those two rational functions have no common poles and hence their sum can vanish only if they vanish individually hence $\varepsilon(s) = 0$. Thus

$$(b'K_1 - k')(Is - A)^{-1}b = 0$$

Using the Laurent series expansion of $(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}$ we see that

$$(\mathbf{b}'\mathbf{K}_1-\mathbf{k}')(s^{-1}+\mathbf{A}s^{-2}+\mathbf{A}^2s^{-3}+\cdots+\mathbf{A}^{n-1}s^{-n})\mathbf{b}=0$$

This means that

$$(\mathbf{b}'\mathbf{K}_1-\mathbf{k}')(\mathbf{b},\mathbf{A}\mathbf{b},\ldots,\mathbf{A}^{n-1}\mathbf{b})=0$$

Since the system was assumed to be controllable we see that $\mathbf{b}'\mathbf{K}_1=\mathbf{k}'$ and hence $\mathbf{K}_1$ satisfies equation (CR).

To show that this is the only positive definite solution assume the contrary. Let $\mathbf{K}_2>0$ be a second solution of equation (CR). Then starting with

$$\mathbf{A}'\mathbf{K}_2+\mathbf{K}_2\mathbf{A}-\mathbf{K}_2\mathbf{b}\mathbf{b}'\mathbf{K}_2=-\mathbf{C}\mathbf{C}'$$

we multiply by -1, add and subtract $\mathbf{K}_2 s$, and pre- and post-multiply by $\mathbf{b}'(-\mathbf{I}s-\mathbf{A}')^{-1}$ and $(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}$ respectively to get

$$\mathbf{b}'\mathbf{K}_2(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}+\mathbf{b}'(-\mathbf{I}s-\mathbf{A}')^{-1}\mathbf{K}_2\mathbf{b}+\mathbf{b}'(-\mathbf{I}s-\mathbf{A}')\mathbf{K}_2\mathbf{b}\mathbf{b}'\mathbf{K}_2(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}$$
$$=\mathbf{b}'(-\mathbf{I}s-\mathbf{A}')^{-1}\mathbf{C}\mathbf{C}'(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}$$

Letting $\mathbf{b}'\mathbf{K}_2(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b}=\dfrac{n(s)}{p(s)}$ and reasoning as before shows that

$$1+\frac{n(s)}{p(s)}=\frac{r(s)}{p(s)}$$

as before and so $\mathbf{k}=\mathbf{K}_2\mathbf{b}=\mathbf{K}_1\mathbf{b}$ with the zeros of $\mathbf{A}-\mathbf{b}\mathbf{b}'\mathbf{K}_2$ in $\operatorname{Re} s<0$. Since $\mathbf{K}_2$ satisfies

$$\mathbf{K}_2(\mathbf{A}-\mathbf{b}\mathbf{k}')+(\mathbf{A}'-\mathbf{k}\mathbf{b}')\mathbf{K}_2=-\mathbf{c}\mathbf{c}'-\mathbf{k}\mathbf{k}'$$

and since the equation has a unique solution, we see that $\mathbf{K}_2=\mathbf{K}_1$. ▮

Exercises

1. Given a frequency response function r which has a nonnegative real part for $s=j\omega$, show that it has a minimal realization of the form

$$\dot{\xi}(t)=(\Sigma-\alpha\alpha')\xi(t)+(\beta-d\alpha)u(t);\qquad y(t)=(\beta'+d\alpha)\xi(t)+du(t)$$

provided r has all its poles in the half plane $\operatorname{Re} s>0$ and is bounded at infinity. (Compare with Darlington's form [29].)

2. Show how the techniques of this section can be used to find solutions of the given quadratic matrix equation which are not positive definite or negative definite. (Consider factorizations which are not spectral factorizations).

NOTES AND REFERENCES

20. Minimization of functions in E^n is an extensive subject in its own right. This section simply contains some of the basic facts needed later on. Theorem 2 is related to the idea of a generalized inverse (See e.g. Zadeh and Desoer [91]). It is an easy consequence of the projection theorem (see Loomis and Sternberg [56]) and in fact can be generalized substantially using Hilbert space ideas.
21. The role of the Riccati equation in solving least square probems of this type has long been known. Radon's paper [69] uses it as do standard sources in the calculus of variations (e.g. Gelfand and Fomin [25]). The method of proof given here suffers from the defect associated with all proofs by verification—one is left wondering how to attack the next problem. On the other hand, the results come quickly and with very little machinery. The development is such that there can be no confusion about the question of when a minimum exists. Popov's highly enlightening paper [68] and Gelfand and Fomin [25] describe similar approaches to the problem. The role of the Riccati equation has been emphasized in the control theory literature by Kalman [40].
22. The technique used in the proof of Theorem 1 was apparently first used in the control theory literature by Kalman, Ho, and Narendra [42]. The results themselves can be obtained in many ways. See e.g. Lee and Markus [52] and Athans and Falb [4].
23. Kalman [40] is an excellent source for this problem. Wonham [82] describes a weakening of the hypothesis with respect to controllability and observability. Athans and Falb [4] and Bryson and Ho [18] have a number of detailed examples. The relationship between this approach and Wiener Theory is more or less clear in the scalar case (see Sections 26 and 27) however the general case is decidedly nontrivial. Popov [68] and Anderson [3] have examined some aspects of this question.
24. In more conventional treatments of these problems (Gelfand and Fomin [25]) the canonical equations are derived first and the Riccati equation is obtained from them. The conjugate point conditions are usually introduced before the Riccati equation is on hand rather than afterward as we have done. Results of the type given by Theorem 4 can be found in Kalman [40] and Kleinman [44]. Full specification of the boundary conditions for the canonical equations can be obtained from the so-called transversality conditions [25].
25. These results are closely related to some recent work in stability theory. Theorem 2 is particularly useful in that respect (see Section 33).
26. When applied to scalar problems the general theory tends to render obscure some otherwise simple results. Lemma 1 permits one to obtain

by inspection some results which look quite difficult from other points of view. Special forms of this lemma were used by Wonham and Johnson [83] and Brockett [15]. Special cases of Theorem 3 can be proven using the Wiener-Hopf technique (Newton *et al.* [64]). The spectral factorization theorem could also be derived from the property of the matrix Riccati Equation.

27. These calculations play a basic role in the work of Yacubovich [84–85] and Kalman [37–38]. See Aizerman and Gantmacher [2] for related material.

4

STABILITY

A number of classic papers on theory of feedback systems deal with stability. In fact, stability was probably the first question in what is now called system theory, which was dealt with in a satisfactory way. In the hands of Maxwell, Minorsky and Nyquist, stability questions motivated the introduction of new mathematical tools in engineering and in particular, Nyquist's paper on stability showed in a very precise way, how complex variable methods can be used in the design of man-made systems.

Of course stability theory has been of interest to mathematicians and astronomers for an even longer time and has had a similar stimulating impact on these fields. The specific problem of attempting to prove that the solar system is stable accounted for the introduction of many new methods. Lagrange, Laplace, Liapunov and Poincaré are but a few of the many contributors.

Starting about a decade ago, work relating the approaches of Nyquist and other engineers to the results of Liapunov *et al.*, began to appear. It soon became apparent that a number of interesting results could be obtained in this way and the literature expanded rapidly. In the selection of topics for this chapter we have been completely partial to this work which attempts to use the best of both areas, even so, only the most fundamental questions are considered. None of the interesting results on nonlinear systems is included, although some of these are easy consequences of the basic tools developed.

28. NORMED VECTOR SPACES AND INEQUALITIES

We began Chapter 1 with a discussion of vector spaces. In Section 12 we discussed Linear Transformations and in Section 20 we considered vector spaces with inner products. In this, the last section devoted to Linear Algebra *per se*, we discuss Normed Vector Spaces.

Recall that associated with every vector in an inner product space is a length, $(\langle \mathbf{x}, \mathbf{x}\rangle)^{1/2}$. However in some cases other definitions of length are useful and this leads to a discussion of vector spaces with a norm. A vector space X is said to be a *normed vector space* if corresponding to each element $\mathbf{x}$ in X there is a nonnegative number $|\mathbf{x}|$ which is called the *norm* and which has the properties

(i) $|\mathbf{x}| \geqslant 0$ and $|\mathbf{x}| = 0$ if and only if $\mathbf{x} = \mathbf{0}$ (positive definiteness)
(ii) $|a\mathbf{x}| = |a|\,|\mathbf{x}|$ for all scalars a (scaling identity)
(iii) $|\mathbf{x} + \mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|$ (triangle inequality)

Examples

(i) R^n with $|\mathbf{x}|$ defined as in Section 1,

$$|\mathbf{x}| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$$

(ii) $C^n\,[t_0, t_1]$ with $|\mathbf{f}|$ defined as

$$|\mathbf{f}| = \int_{t_0}^{t_1} \|\mathbf{f}(\sigma)\|\, d\sigma$$

(iii) $R^{n\times m}$ with $|\mathbf{A}|$ defined as

$$|\mathbf{A}| = (\operatorname{tr} \mathbf{A}'\mathbf{A})^{1/2}$$

The verification of properties (i)–(iii) in each case is left to the reader.

In vector analysis the triangle inequality is a basic tool. One useful implication of it is the following lemma.

Lemma 1. *If each of the components of* $\mathbf{x}(\ )$ *are Riemann integrable on* $t_0 \leqslant t \leqslant t_1$ *then*

$$\left|\int_{t_0}^{t} \mathbf{x}(\sigma)\, d\sigma\right| \leqslant \int_{t_0}^{t} |\mathbf{x}(\sigma)|\, d\sigma; \qquad t_0 \leqslant t \leqslant t_1$$

Proof. Replace each integral by the approximating sums and use the triangle inequality to obtain

$$\left|\sum_{i=1}^{v} \mathbf{x}(t_i)\, \Delta_i\right| \leqslant \sum_{i=1}^{v} |\mathbf{x}(t_i)|\, \Delta$$

from which the integral inequality follows immediately. ▮

If L is a linear transformation of a normed vector space X into a normed vector space Y then we call L a *bounded linear transformation* if there exists a constant k such that

$$|L(\mathbf{x})|_Y \leqslant k\,|\mathbf{x}|_X$$

for all $\mathbf{x}$. Otherwise L is said to be *unbounded.* It is easily seen that all linear transformations whose range and domain are finite dimensional vector spaces are necessarily bounded. The greatest lower bound on the set of k's which suffice in this inequality is called the *induced norm* of L. For linear transformations of R^m into R^n this gives a definition for a norm of a matrix. That is, if $\mathbf{A}$ belongs to $R^{n\times m}$ then

$$|\mathbf{A}| = \max_{|\mathbf{x}|=1} |\mathbf{A}\mathbf{x}|$$

We have already introduced the norm $\|\mathbf{x}\|$ for elements of R^n. We reserve the notation $\|\mathbf{A}\|$ for the matrix norm which it induces. The following lemma tells how to calculate this matrix norm.

Lemma 2. *If* $\mathbf{A}$ *is a real* n *by* m *matrix then* $\|\mathbf{A}\| = \|\mathbf{A}'\|$ *and both equal the square root of maximum eigenvalue of* $\mathbf{A}'\mathbf{A}$ *which, in turn, equals the square root of maximum eigenvalue of* $\mathbf{A}\mathbf{A}'$.

Proof. From the definition of $\|\mathbf{A}\|$ we have

$$\begin{aligned}\|\mathbf{A}\| &= \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| \\ &= \max_{\|\mathbf{x}\|=1} (\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x})^{1/2} \\ &= \sqrt{\mu}\end{aligned}$$

where μ is the maximum eigenvalue of $\mathbf{A}'\mathbf{A}$. On the other hand, if λ is a nonzero eigenvalue of $\mathbf{A}'\mathbf{A}$ then it is also an eigenvalue of $\mathbf{A}\mathbf{A}'$; in fact, if $\mathbf{A}'\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ then $\mathbf{A}\mathbf{A}'(\mathbf{A}\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x})$. Hence $\|\mathbf{A}\| = \|\mathbf{A}'\|$. ▮

The basic properties of the induced matrix norm are given in Theorem 1.

Theorem 1. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are real matrices then*

(i) $\|\mathbf{A}\| \geqslant 0$ *and* $\|\mathbf{A}\| = 0$ *if and only if* $\mathbf{A} = \mathbf{0}$
(ii) $\|\alpha\mathbf{A}\| = |\alpha|\,\|\mathbf{A}\|$
(iii) $\|\mathbf{A} + \mathbf{B}\| \leqslant \|\mathbf{A}\| + \|\mathbf{B}\|$
(iv) $\|\mathbf{A}\mathbf{B}\| \leqslant \|\mathbf{A}\| \cdot \|\mathbf{B}\|$
(v) $\|\mathbf{A}'\| = \|\mathbf{A}\|$

provided the dimensions of $\mathbf{A}$ *and* $\mathbf{B}$ *are such that the indicated operations make sense.*

Proof. The first statement is obvious and (ii) follows immediately from the definition of $\|\cdot\|$ for vectors. For (iii) we have from the definition and the triangle inequality

$$\begin{aligned}\|\mathbf{A} + \mathbf{B}\| &= \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}\| \\ &\leqslant \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| + \max_{\|\mathbf{x}\|=1} \|\mathbf{B}\mathbf{x}\| \\ &= \|\mathbf{A}\| + \|\mathbf{B}\|\end{aligned}$$

To prove (iv) we need only the definition of the induced norm.

$$\begin{aligned}\|\mathbf{A}\mathbf{B}\| &= \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{B}\mathbf{x}\| \\ &\leqslant \|\mathbf{A}\| \max_{\|\mathbf{x}\|=1} \|\mathbf{B}\mathbf{x}\| \\ &\leqslant \|\mathbf{A}\| \cdot \|\mathbf{B}\|\end{aligned}$$

The fifth equality follows from the previous lemma. ▮

It is apparent that we have two competing interpretations for $\|\mathbf{A}\|$ when $\mathbf{A}$ is an n by 1 array of real numbers. If we regard $\mathbf{A}$ as a matrix which maps R^1 into R^n then $\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1}\|\mathbf{Ax}\|$ whereas if we regard $\mathbf{A}$ as an n-vector then $\|\mathbf{A}\|$ is just the square root of the sum of the squares of the components. Happily enough, these are numerically equal and hence we can permit a certain amount of carelessness.

A sequence $\mathbf{x}_i$ in a normed vector space is said to be a *Cauchy sequence* if given any $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that for all i and j larger than $N(\varepsilon)$ it follows that $\|\mathbf{x}_i - \mathbf{x}_j\| < \varepsilon$. An element $\mathbf{x}_0$ of X is said to be the *limit of a convergent sequence* $\{\mathbf{x}_i\}$ if given any $\varepsilon > 0$ there exists an integer N such that $\|\mathbf{x}_i - \mathbf{x}_0\| < \varepsilon$ for all $i > N$. If every Cauchy sequence has a limit then the normed vector space is said to be *complete*. Complete normed vector spaces are also called *Banach spaces*. Examples of complete normed vector spaces are:

(i) The space R^n with $|\mathbf{x}|$ defined by

$$|\mathbf{x}| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} \qquad \text{(Euclidean norm)}$$

(ii) The space of continuous functions defined on the interval [0, 1] with

$$|x| = \max_{0 \leqslant t \leqslant 1} |x(t)| \qquad \text{(uniform norm)}$$

Exercises

1. The *spectral radius* of an n by n matrix $\mathbf{A}$ is defined as the maximum of the magnitudes of the eigenvalues of $\mathbf{A}$. Show that any induced matrix norm is greater than or equal to the spectral radius. (Caution: $\mathbf{A}$ may have complex eigenvalues.)
2. Let S denote the set of all complex n-vectors $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ such that $\mathbf{u}'\mathbf{u} + \mathbf{v}'\mathbf{v} = 1$. If $\mathbf{A}$ is an n by n matrix, the subset of the complex plane which $(\mathbf{u}' - i\mathbf{v}')\mathbf{A}(\mathbf{u} + i\mathbf{v})$ sweeps out as $\mathbf{x}$ ranges over S is called the *numerical range* of $\mathbf{A}$. Show that the numerical range is a convex set and that it contains the eigenvalues of $\mathbf{A}$.
3. An n by n matrix is called *normal* if $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}'$. Show that if $\mathbf{A}$ is normal then the numerical range of $\mathbf{A}$ is the smallest convex set which contains the eigenvalues of $\mathbf{A}$.
4. Let S_1 and S_2 be subsets of the complex plane and let S_1/S_2 be the set of all quotients, z/w with z in S_1 and w in S_2. Show that every eigenvalue of $\mathbf{A}^{-1}\mathbf{B}$ belongs to $\sigma(\mathbf{B})/\sigma(\mathbf{A})$ where $\sigma(\mathbf{B})$ and $\sigma(\mathbf{A})$ are the numerical ranges of $\mathbf{B}$ and $\mathbf{A}$ respectively. (Williams)
5. A *Banach Algebra* is a normed vector space V together with a mapping (written as a multiplication) of $V \times V$ into V such that if $\mathbf{A}$ and $\mathbf{B}$ belong to V then $\|\mathbf{AB}\| \leqslant \|\mathbf{A}\| \cdot \|\mathbf{B}\|$. Verify that the space of n by n matrices is a Banach Algebra if $\|\mathbf{A}\|$ is defined as

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

with the vector norm being $\|\mathbf{y}\| = (y_1^2 + y_2^2 + \cdots + y_n^2)^{1/2}$.
6. Show that for $\mathbf{x}$ in R^n we may define a norm as

$$|\mathbf{x}| = \max_{1 \leqslant i \leqslant n} |x_i|$$

What is the matrix norm which this vector norm induces?
7. The space of real n by n matrices can be viewed as an n^2 dimensional inner product space with the inner product being defined as $\langle \mathbf{A}, \mathbf{B}\rangle = \operatorname{tr} \mathbf{A}'\mathbf{B}$. This makes $(\operatorname{tr} \mathbf{A}'\mathbf{A})^{1/2}$ the "length of $\mathbf{A}$." Show that

$$(\operatorname{tr} \mathbf{A}'\mathbf{A})^{1/2} \geqslant \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

8. Verify the following for n-vectors $\mathbf{x}$ and $\mathbf{y}$ and n by n matrices $\mathbf{A}$.

(i) $\|\mathbf{x}\| = \max_{\|\mathbf{y}\|=1} \mathbf{y}'\mathbf{x}$

(ii) $\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\| = \max_{\|\mathbf{y}\|=1} \max_{\|\mathbf{x}\|=1} \mathbf{y}'\mathbf{Ax}$

9. (Contraction Mapping Principle) Let T be a transformation of a Banach space into itself. Suppose that for each $\mathbf{x}$ in the set $B_b = \{\mathbf{x}\colon \|\mathbf{x}\| \leqslant b\}$ we have $\|T(\mathbf{x})\| \leqslant \mathbf{b}$ and suppose that for all $\mathbf{x}_1$ and $\mathbf{x}_2$ in B_b

$$\|T(\mathbf{x}_1) - T(\mathbf{x}_2)\| \leqslant k\,\|\mathbf{x}_1 - \mathbf{x}_2\|; \qquad k < 1$$

Prove that there exists an element $\mathbf{x}^*$ of B_b such that

$$T(\mathbf{x}^*) = \mathbf{x}^*$$

i.e. prove that T leaves one point unchanged. Proceed by successive approximations. Let

$$\mathbf{x}_0 = \mathbf{0}$$
$$\mathbf{x}_{i+1} = T(\mathbf{x}_i)$$

and using the above inequalities show that

$$\|\mathbf{x}_{i+1} - \mathbf{x}_i\| \leqslant k^i b$$

and hence

$$\|\mathbf{x}_{i+m} - \mathbf{x}_i\| \leqslant b(k^i + k^{i+1} \cdots k^{i+m-1})$$

Show that this is a Cauchy sequence and hence has a limit. Prove that the limit is a fixed point. Show that there cannot be two fixed points in B_b. Compare this series of successive approximations with those used in the proof of Theorem 3.1.

29. UNIFORM STABILITY AND EXPONENTIAL STABILITY

We begin our discussion of stability by using some standard inequalities to establish the basic properties of linear differential equations.

The differential equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is said to be *uniformly stable* if there exists a positive constant γ such that for all real t and t_0 in the half-plane $t > t_0$

$$\|\boldsymbol{\Phi}(t, t_0)\| \leqslant \gamma$$

It is clear that uniform stability does not imply that solutions which start near zero will converge to zero. For example, the harmonic oscillator equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is uniformly stable although $\sqrt{x_1^2 + x_2^2}$, the length of $\mathbf{x}$, is constant throughout the motion.

In order to study convergence to zero we introduce a second definition. The equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is said to be *exponentially stable* if there exist positive constants γ and λ such that for all real t and t_0 in the half-plane $t > t_0$,

$$\|\boldsymbol{\Phi}(t, t_0)\| \leqslant \gamma e^{-\lambda(t-t_0)}$$

In Section 8 we defined Liapunov transformations. One of their main uses is in studying stability.

Theorem 1. *If* $\mathbf{Px} = \mathbf{z}$ *defines a Liapunov transformation then* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is uniformly stable (exponentially stable) if and only if*

$$\dot{\mathbf{z}}(t) = [\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]\mathbf{z}(t)$$

is uniformly stable (exponentially stable).

Proof. Suppose that $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is uniformly stable then by definition $\|\boldsymbol{\Phi}_{\mathbf{A}}(t, t_0)\|$ is bounded for $t > t_0$ by a constant N which does not depend on t_0. Since by Theorem 4.3,

$$\boldsymbol{\Phi}_{\mathbf{PAP}^{-1}+\dot{\mathbf{P}}\mathbf{P}^{-1}}(t, t_0) = \mathbf{P}(t)\boldsymbol{\Phi}_{\mathbf{A}}(t, t_0)\mathbf{P}^{-1}(t_0)$$

we have

$$\|\boldsymbol{\Phi}_{\mathbf{PAP}^{-1}+\dot{\mathbf{P}}\mathbf{P}^{-1}}(t, t_0)\| \leqslant \|\mathbf{P}(t)\| \cdot \|\boldsymbol{\Phi}(t, t_0)\| \cdot \|\mathbf{P}^{-1}(t_0)\|$$

Since $\mathbf{P}$ is a Liapunov transformation $\|\mathbf{P}(t)\|$ and $\|\mathbf{P}^{-1}(t)\|$ are bounded by some constant M. As a result the transition matrix for the z-equations is bounded. Since the inverse of a Liapunov transformation is also a Liapunov transformation, it follows that if the z-equation is uniformly stable then the x-equation is also. Exponential stability may be established in the same way. The details are omitted. ∎

Theorem 2. *If* $\mathbf{A}$ *is constant or periodic and bounded, then* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable if and only if for every fixed* t_0, $\|\boldsymbol{\Phi}(t, t_0)\|$ *approaches* 0 *as* t *approaches infinity.*

Proof. The necessity part is obvious. To establish sufficiency we observe that from the results of Section 8 we can always find a Liapunov transformation such that the system is time invariant. From similarity theory (Section 12) we can find a Liapunov transformation to reduce the constant system to one which is in Jordan normal form. For systems in Jordan form, $\boldsymbol{\Phi}$ is a block diagonal matrix with each element being one of the three forms

$$t^k e^{\sigma t} \cos \lambda t,\; t^k e^{\sigma t} \sin \lambda t, \qquad \text{or} \quad t^k e^{\lambda t}$$

since $t^k e^{\sigma t}$ goes to zero if and only if there exist γ and $\lambda > 0$ such that $t^k e^{\sigma t} \leqslant \gamma e^{-\lambda t}$ the theorem is proven. ∎

Such a simple relation between convergence of $\boldsymbol{\Phi}$ and exponential stability generally does not hold. (See Problem 1.) Nonetheless, it turns out that if $\mathbf{A}$ is bounded then exponential stability is, in fact, equivalent to a number of other conditions which at first sight seem less demanding. Notable among these are various conditions on the boundedness of the integrals of $\|\boldsymbol{\Phi}(t, t_0)\|$. Theorem 3 below gives four such conditions. Figure 1 illustrates the two different paths of integration used.

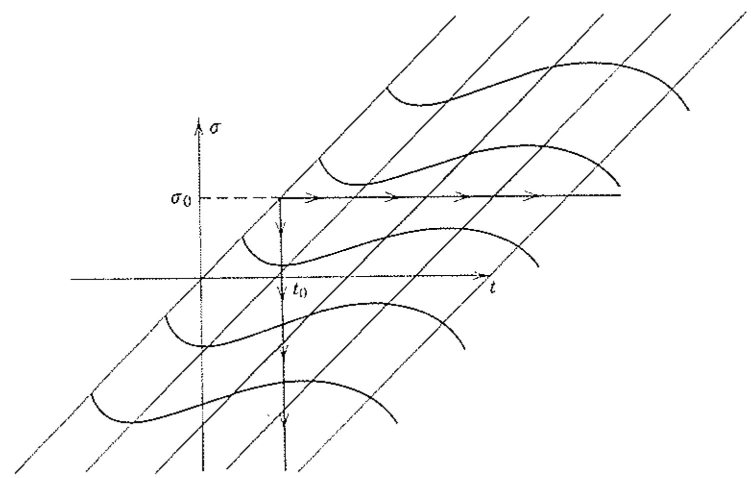


Figure 1. Illustrating $\boldsymbol{\Phi}(t, \sigma)$ *and paths of integration for Theorem 3.*

Theorem 3. *Let* **A** *be bounded on* $(-\infty, \infty)$. *Then any one of the following four statements is equivalent to exponential stability for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$

(i) $\displaystyle\int_{t_0}^{t_1} \|\mathbf{\Phi}(t, t_0)\|^2 \, dt \leqslant M_1$ for all $t_1 \geqslant t_0$

(ii) $\displaystyle\int_{t_0}^{t_1} \|\mathbf{\Phi}(t, t_0)\| \, dt \leqslant M_2$ for all $t_1 \geqslant t_0$

(iii) $\displaystyle\int_{t_0}^{t_1} \|\mathbf{\Phi}(t_1, \sigma)\|^2 \, d\sigma \leqslant M_3$ for all $t_1 \geqslant t_0$

(iv) $\displaystyle\int_{t_0}^{t_1} \|\mathbf{\Phi}(t_1, \sigma)\| \, d\sigma \leqslant M_4$ for all $t_1 \geqslant t_0$

(The constants M_i do not depend on t_0 or t_1.)

Proof. It is easy to see that exponential stability implies (i)–(iv). If the null solution is exponentially stable then $\|\mathbf{\Phi}(t, t_0)\| \leqslant \gamma e^{-\lambda(t-t_0)}$. Using this estimate in the above integrals gives as a suitable choice: $M_1 = M_3 = \gamma^2/2\lambda$ and $M_2 = M_4 = \gamma/\lambda$.

The reverse implications need to be taken one at a time.

(i) Assume there exists a bound M_1. Since **A** is bounded it follows that there exists an α such that

$$\|\dot{\mathbf{\Phi}}(t, t_0)\| = \|\mathbf{A}(t)\mathbf{\Phi}(t, t_0)\| \leqslant \alpha \|\mathbf{\Phi}(t, t_0)\|$$

Hence

$$\begin{aligned}
&\|\mathbf{\Phi}'(t_1, t_0)\mathbf{\Phi}(t_1, t_0) - \mathbf{I}\| \\
&= \left\| \int_{t_0}^{t_1} \dot{\mathbf{\Phi}}'(t, t_0)\mathbf{\Phi}(t, t_0) + \mathbf{\Phi}'(t, t_0)\dot{\mathbf{\Phi}}(t, t_0) \, dt \right\| \\
&\leqslant \int_{t_0}^{t_1} \|\dot{\mathbf{\Phi}}'(t, t_0)\mathbf{\Phi}(t, t_0) + \mathbf{\Phi}'(t, t_0)\dot{\mathbf{\Phi}}(t, t_0)\| \, dt \\
&\leqslant \int_{t_0}^{t} \|\dot{\mathbf{\Phi}}'(t, t_0)\mathbf{\Phi}(t, t_0)\| + \|\mathbf{\Phi}'(t, t_0)\dot{\mathbf{\Phi}}(t, t_0)\| \, dt \\
&\leqslant \int_{t_0}^{t_1} \|\dot{\mathbf{\Phi}}'(t, t_0)\| \cdot \|\mathbf{\Phi}(t, t_0)\| + \|\mathbf{\Phi}'(t, t_0)\| \cdot \|\dot{\mathbf{\Phi}}(t, t_0)\| \, dt \\
&\leqslant 2\alpha \int_{t_0}^{t_1} \|\mathbf{\Phi}(t, t_0)\|^2 \, dt \\
&\leqslant 2\alpha M_1; \qquad t_1 \geqslant t_0
\end{aligned}$$

Using the triangle inequality for matrices gives

$$\begin{aligned}
\|\mathbf{\Phi}'(t_1, t_0)\mathbf{\Phi}(t_1, t_0)\| &\leqslant \|\mathbf{\Phi}'(t_1, t_0)\mathbf{\Phi}(t_1, t_0) - \mathbf{I}\| + \|\mathbf{I}\| \\
&\leqslant 2\alpha M_1 + 1; \qquad t_1 \geqslant t_0
\end{aligned}$$

This implies $\|\mathbf{\Phi}(t_1, t_0)\|$ is itself bounded for $t_1 \geqslant t_0$ by a constant N_1.

To obtain an exponential bound we write $\mathbf{\Phi}(t, t_0) = \mathbf{\Phi}(t, \sigma)\mathbf{\Phi}(\sigma, t_0)$ and observe that

$$\begin{aligned}
\int_{t_0}^{t_1} \|\mathbf{\Phi}(t, t_0)\|^2 \, d\sigma &= \int_{t_0}^{t_1} \|\mathbf{\Phi}(t, \sigma)\mathbf{\Phi}(\sigma, t_0)\|^2 \, d\sigma \\
&\leqslant \int_{t_0}^{t_1} \|\mathbf{\Phi}(t, \sigma)\|^2 \cdot \|\mathbf{\Phi}(\sigma, t_0)\|^2 \, d\sigma \\
&\leqslant N_1 \int_{t_0}^{t_1} \|\mathbf{\Phi}(\sigma, t_0)\|^2 \, d\sigma \\
&\leqslant N_1 M_1
\end{aligned}$$

However, since the integrand on the left does not depend on σ this gives

$$(t - t_0)\|\mathbf{\Phi}(t, t_0)\|^2 \leqslant N_1 M_1, \qquad t \geqslant t_0$$

Thus for $T = 4N_1 M_1$ we have

$$\|\mathbf{\Phi}(t_0 + T, t_0)\| \leqslant \tfrac{1}{2}$$

and in general

$$\|\mathbf{\Phi}(t_0 + nT, t_0)\| \leqslant (\tfrac{1}{2})^n$$

It follows from this that $\|\mathbf{\Phi}(t, t_0)\|$ is exponentially bounded. (See problem 2.)

(ii) The philosophy is the same but we now work with $\mathbf{\Phi}(t_1, t_0)$ instead of $\mathbf{\Phi}'(t_1, t_0)\mathbf{\Phi}(t_1, t_0)$. Use as before, $\|\dot{\mathbf{\Phi}}(t, t_0)\| \leqslant \alpha\|\mathbf{\Phi}(t, t_0)\|$ and generate the string of inequalities

$$\begin{aligned}
\|\mathbf{\Phi}(t_1, t_0) - \mathbf{I}\| &= \left\| \int_{t_0}^{t_1} \dot{\mathbf{\Phi}}(t, t_0) \, dt \right\| \\
&\leqslant \int_{t_0}^{t_1} \|\dot{\mathbf{\Phi}}(t, t_0)\| \, dt \\
&\leqslant \alpha \int_{t_0}^{t_1} \|\mathbf{\Phi}(t, t_0)\| \, dt \\
&\leqslant \alpha M_2; \qquad t_1 \geqslant t_0
\end{aligned}$$

Use the triangle inequality to obtain

$$\|\mathbf{\Phi}(t_1, t_0)\| \leqslant \alpha M_2 + 1$$

Now this lets us establish the inequality

$$\int_{t_0}^{t_1} \|\Phi(t, t_0)\|^2 \, dt \leqslant (\alpha M_2 + 1) \int_{t_0}^{t_1} \|\Phi(t, t_0)\| \, dt$$
$$\leqslant (\alpha M_2 + 1) M_2$$

and so (ii) implies (i) and hence exponential stability.

(iii) Here the argument used in (i) needs to be modified slightly. Since $\mathbf{A}$ is bounded there exists α such that $\|\mathbf{A}(t)\| \leqslant \alpha$ for all t. From Theorem 2 of Section 7 we have

$$\frac{d}{d\sigma} \Phi(t, \sigma) = -\Phi(t, \sigma)\mathbf{A}(\sigma)$$

Hence

$$\left\| \frac{d}{d\sigma} \Phi(t, \sigma) \right\| \leqslant \alpha \|\Phi(t, \sigma)\|$$

However,

$$\|\mathbf{I} - \Phi'(t_1, t_0)\Phi(t_1, t_0)\|$$
$$= \left\| \int_{t_0}^{t_1} \left[\frac{d}{d\sigma} \Phi'(t_1, \sigma)\right] \Phi(t_1, \sigma) + \Phi'(t_1, \sigma) \left[\frac{d}{d\sigma} \Phi(t_1, \sigma)\right] d\sigma \right\|$$

Reasoning as in part (i) this becomes for $t_1 \geqslant t_0$

$$\|\mathbf{I} - \Phi'(t_1, t_0)\Phi(t_1, t_0)\| \leqslant 2\alpha M_3$$

Using the triangle inequality we have

$$\|\Phi'(t_1, t_0)\Phi(t_1, t_0)\| \leqslant \|\Phi'(t_1, t_0)\Phi(t_1, t_0) - \mathbf{I}\| + \|\mathbf{I}\| \leqslant 2\alpha M_3 + 1$$

This shows that $\|\Phi(t, t_0)\|$ is bounded for $t > t_0$ by a constant, N_3, independent of t and t_0.

To get the exponential bound we begin with the observation that for $t > t_0$ we have

$$(t - t_0)\|\Phi(t_1, t_0)\|^2 = \int_{t_0}^{t} \|\Phi(t_1, t_0)\|^2 \, d\sigma$$
$$= \int_{t_0}^{t} \|\Phi(t_1, \sigma)\Phi(\sigma, t_0)\|^2 \, d\sigma$$
$$\leqslant \int_{t_0}^{t} \|\Phi(t_1, \sigma)\|^2 \, \|\Phi(\sigma, t_0)\|^2 \, d\sigma$$
$$\leqslant N_3 \int_{t_0}^{t} \|\Phi(t_1, \sigma)\|^2 \, d\sigma$$
$$\leqslant N_3 M_3$$

Hence

$$\|\Phi(t, t_0)\|^2 \leqslant N_3 M_3 (t - t_0)^{-1}$$

and for $T = 4N_3 M_3$ we have

$$\|\Phi(t_0 + T, t_0)\| \leqslant \tfrac{1}{2}$$

Therefore

$$\|\Phi(t_0 + nT, t_0)\| \leqslant (\tfrac{1}{2})^n$$

and we conclude that Φ is exponentially bounded.

(iv) If a suitable M_4 exists then we can modify the approach used in (iii) in the same way (i) was modified to get (ii). In this way we obtain

$$\|\Phi(t_1, t_0) - \mathbf{I}\| \leqslant \alpha M_4, \qquad t \geqslant t_0$$

And so $\|\Phi(t_1, t_0)\|$ is bounded for $t_1 \geqslant t_0$ by $\alpha M_4 + 1$. However, this gives the inequality

$$\int_{t_0}^{t_1} \|\Phi(t, \sigma)\|^2 \, d\sigma \leqslant (\alpha M_4 + 1) \int_{t_0}^{t_1} \|\Phi(t_1, \sigma)\| \, d\sigma$$
$$\leqslant (\alpha M_4 + 1) M_4$$

Hence (iv) implies (iii) which implies exponential stability. ▮

Exercises

1. Show that all solutions of $\dot{x}(t) = -2t(t+1)^{-2}x(t)$ go to zero as t approaches infinity but that the equation is not exponentially stable.
2. Give a complete argument to show that if $\mathbf{A}$ is bounded on $t_0 \leqslant t \leqslant \infty$ then

$$\|\Phi(t_0 + nT, t_0)\| \leqslant (\tfrac{1}{2})^n$$

implies exponential stability.

3. (Wazewski) Show that the equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable if there exist constants $c \geqslant 0$ and $a > 0$ such that for all positive t_0 and $t > t_0$ we have

$$\int_{t_0}^{t} \lambda_m(\sigma) \, d\sigma \leqslant -a(t - t_0) + c$$

where $\lambda_m(\sigma)$ denotes the largest eigenvalue of $\mathbf{A}(\sigma) + \mathbf{A}'(\sigma)$.
Hint: Show that

$$\frac{d}{dt} \mathbf{x}'(t)\mathbf{x}(t) \leqslant \lambda_m(t)\mathbf{x}'(t)\mathbf{x}(t)$$

4. The second order equation

$$\ddot{x}(t) + 2\dot{x}(t) + f(t)x(t) = 0$$

Can be expressed as

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & -2\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} - f(t)\begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

with $x(t) = x_1(t) - x_2(t)$. Use the results of the previous exercise to show exponential stability if

$$\int_0^t \max[-f(\sigma), f(\sigma) - 4]\, d\sigma$$

is suitably bounded.

5. Show that if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is uniformly stable and $\int_0^\infty \|\mathbf{f}(\sigma)\|\, d\sigma$ exists then all solutions of

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t)$$

are bounded.

6. Show that if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is uniformly stable and $\mathbf{B}$ is such that the integral

$$\mathbf{B}_0 = \int_0^\infty \mathbf{B}(\sigma)\, d\sigma$$

converges, then $\dot{\mathbf{x}}(t) = [\mathbf{A}(t) + \mathbf{B}(t)]\mathbf{x}(t)$ is also uniformly stable.

30. INPUT-OUTPUT STABILITY

In the preceding section we were concerned exclusively with the transition matrix which in turn reflects the behavior of the solution vector $\mathbf{x}$ in the absence of any input. There is a second type of stability which refers to the effects of inputs. It centers around the idea that every bounded input should produce a bounded output if a *system* is to be regarded as being stable. We will say that the system

$$\begin{bmatrix}\dot{\mathbf{x}}(t)\\ \mathbf{y}(t)\end{bmatrix} = \begin{bmatrix}\mathbf{A}(t) & \mathbf{B}(t)\\ \mathbf{C}(t) & \mathbf{D}(t)\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{u}(t)\end{bmatrix}$$

is *BIBO stable* if for all t_0 and for zero initial conditions at $t = t_0$, every bounded input defined on $[t_0, \infty)$ gives rise to a bounded response on $[t_0, \infty)$. We will say that this system is *uniformly BIBO stable* if there exists constant k (independent of t_0) such that for all t_0 the statements

$$\mathbf{x}(t_0) = \mathbf{0}$$
$$\|\mathbf{u}(t)\| \leqslant 1; \qquad t \geqslant t_0$$

imply $\|\mathbf{y}(t)\| \leqslant k$ for $t \geqslant t_0$. We want to treat Uniform BIBO Stability in some detail in this section. In particular, our interest is in establishing circumstances under which it is equivalent to exponential stability. The following lemma will be a basic tool.

Lemma 1. *Let* $\mathbf{y}$ *and* $\mathbf{u}$ *be related by*

$$\mathbf{y}(t) = \int_{t_0}^t \mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\, d\sigma$$

There exists a constant k such that for $\mathbf{u}$ *constrained by*

$$\|\mathbf{u}(t)\| \begin{cases} = 0; & t \leqslant t_0 \\ \leqslant 1; & t > t_0 \end{cases}$$

it follows that

$$\|\mathbf{y}(t)\| \leqslant k; \qquad \text{all } t > t_0$$

if and only if there exists a constant N such that

$$N \geqslant \int_{t_0}^t \|\mathbf{T}(t, \sigma)\|\, d\sigma; \qquad \text{all } t > t_0$$

Proof. Sufficiency. Assume that the bound on the integral holds, then if $\|\mathbf{u}(t)\| \leqslant 1$ the following string of inequalities is valid.

$$\begin{aligned}\|\mathbf{y}(t)\| &= \left\|\int_{t_0}^t \mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\, d\sigma\right\| \\ &\leqslant \int_{t_0}^t \|\mathbf{T}(t, \sigma)\mathbf{u}(\sigma)\|\, d\sigma \\ &\leqslant \int_{t_0}^t \|\mathbf{T}(t, \sigma)\| \cdot \|\mathbf{u}(\sigma)\|\, d\sigma \\ &\leqslant \int_{t_0}^t \|\mathbf{T}(t, \sigma)\|\, d\sigma \\ &\leqslant N; \qquad t \geqslant t_0\end{aligned}$$

Necessity. If for some t_0 it follows that for every finite k there exists a time t_k such that

$$\int_{t_0}^{t_k} \|\mathbf{T}(t_k, \sigma)\|\, d \geqslant k$$

then there can be no bound on $\|\mathbf{y}(t)\|$. To prove this use $\|\mathbf{y}\| = \max \mathbf{v}'\mathbf{y}$ with $\|\mathbf{v}\| = 1$. For a given t_k

$$\begin{aligned}\|\mathbf{y}_0(t_k)\| &= \int_{t_0}^{t_k} \max_{\|\mathbf{v}(t_k)\| = 1}\ \max_{\|\mathbf{u}(\sigma)\| = 1} \mathbf{v}'(t_k)\mathbf{T}(t_k, \sigma)\mathbf{u}(\sigma)\, d\sigma \\ &= \int_{t_0}^{t_k} \|\mathbf{T}(t_k, \sigma)\|\, d\sigma\end{aligned}$$

Hence $\|y_0(t_k)\|$ cannot be bounded if no bound exists on the integral of the norm of **T**. ▮

As an immediate consequence of this lemma we have the following theorem relating exponential stability and input-output stability.

Theorem 1. *Let* **A** *be bounded on* $(-\infty, \infty)$. *The system*

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

is uniformly BIBO stable if and only if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable.*

Proof. By the previous lemma and the variation of constants formula we see that this system is BIBO stable if and only if there exists a constant N such that for all t

$$\int_{-\infty}^{t} \|\mathbf{\Phi}(t, \sigma)\| d\sigma \leqslant N$$

However, from part (iv) of Theorem 3 of Section 29 we see that this holds if and only if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable. ▮

Corollary. *Let A be bounded on $(-\infty, \infty)$. Assume that the system*

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

is uniformly BIBO stable. Then if $\|\mathbf{u}(t)\|$ is bounded on (t_0, ∞) and if $\|\mathbf{u}(t)\|$ approaches zero as t approaches infinity then $\|\mathbf{x}(t)\|$ also approaches zero as t approaches infinity.

Proof. The system is BIBO stable and hence it is also exponentially stable. Since $\|\mathbf{u}(t)\|$ goes to zero as t goes to infinity, it follows that given any μ greater than zero there exists a time t_1 such that for $t \geqslant t_1$ $\|\mathbf{u}(t)\| \leqslant \mu$. From the bound derived above it is clear that the change in the response **x** due to truncating u at t_1 is bounded by $N\mu$. On the other hand, given $\rho > 0$ there exists a time t_2 such that for $t > t_1 + t_2$, the response **x** is bounded by $\rho\|\mathbf{x}(t_1)\| \geqslant \|\mathbf{x}(t)\|$; all $t > t_1 + t_2$. Hence given any number $\varepsilon > 0$ and any input satisfying the hypothesis there exists a time $t_1 + t_2$ such that for $t > t_1 + t_2$ the length of $\mathbf{x}(t)$ is less than ε. ▮

The following theorem includes Theorem 1 as a special case.

Theorem 2. *Let* **A** *and* **B** *be bounded on* $(-\infty, \infty)$. *Assume there exist positive constants ε and δ such that* $\mathbf{W}(t_0, t_0 + \delta) \geqslant \varepsilon\mathbf{I}$; *both ε and δ are independent of t_0. Then the system*

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} \begin{bmatrix} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

is uniformly BIBO stable if and only if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable.*

Proof. Sufficiency. Clearly boundedness of **B** and exponential stability of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ imply that there exists **N** such that

$$\int_{-\infty}^{t} \|\mathbf{\Phi}(t, \sigma)\mathbf{B}\alpha\| \, d\sigma \leqslant N$$

Using Lemma 1 this gives uniform BIBO stability.

Necessity. Assume such an ε and δ are given. Write **I** as

$$\mathbf{I} = \int_{\sigma-\delta}^{\sigma} \mathbf{\Phi}(\sigma, \rho)\mathbf{B}(\rho)\mathbf{B}'(\rho)\mathbf{\Phi}'(\sigma, \rho) \, d\rho \mathbf{W}^{-1}(\sigma - \delta, \sigma)$$

The hypothesis insures that $\mathbf{W}^{-1}(\sigma - \delta, \sigma)$ is bounded independent of σ. Since $\mathbf{B}(\rho)$ is bounded, and $\mathbf{\Phi}(\sigma, \rho)$ is bounded for $|\sigma - \rho| \leqslant \delta$, we see there exists a constant M such that

$$\|\mathbf{B}'(\rho)\mathbf{\Phi}'(\sigma, \rho)\mathbf{W}^{-1}(\sigma - \delta, \sigma)\| \leqslant M$$

Premultiply the above equation by $\mathbf{\Phi}(t, \sigma)$ and use this bound to get

$$\|\mathbf{\Phi}(t, \sigma)\| \leqslant M \int_{\sigma-\delta}^{\sigma} \|\mathbf{\Phi}(t, \rho)\mathbf{B}(\rho)\| \, d\rho$$

Now observe that

$$\int_{t-n\delta}^{t} \|\mathbf{\Phi}(t, \rho)\mathbf{B}(\rho)\| \, d\rho = \int_{t-\delta}^{t} \|\mathbf{\Phi}(t, \rho)\mathbf{B}(\rho)\| \, d\rho + \int_{t-2\delta}^{t-\delta} \|\mathbf{\Phi}(t, \rho)\mathbf{B}(\rho)\| \, d\rho + \cdots$$

$$\geqslant [\|\mathbf{\Phi}(t, t)\| + \|\mathbf{\Phi}(t, t-\delta)\| + \cdots + \|\mathbf{\Phi}(t, t - n\delta + \delta)\|]M^{-1}$$

So we see that

$$\|\mathbf{\Phi}(t, t)\| + \|\mathbf{\Phi}(t, t-\delta)\| + \|\mathbf{\Phi}(t, t-2\delta)\| + \cdots \leqslant nM$$

Since $\|\mathbf{\Phi}(t, \sigma)\| \leqslant e^{\alpha\delta}$ for $|t - \sigma| \leqslant \delta$ this insures that there exists $\rho > 0$ such that

$$\int_{-\infty}^{t} \|\mathbf{\Phi}(t, \sigma)\| \, d\sigma \leqslant \rho e^{\alpha\delta} nM \ ▮$$

Finally, to conclude this section, we give a result which uses both controllability and observability in the hypothesis.

Theorem 3. *Let* **A**, **B** *and* **C** *be bounded on* $(-\infty, \infty)$. *Assume that there exist positive constants ε_1 and δ_1 such that* $\mathbf{W}(t_0, t_0 + \delta_1) \geqslant \varepsilon_1 \mathbf{I}$ *and assume that there exists positive constants ε_2 and δ_2 such that*

$$\mathbf{M}(t_0, t_0 + \delta_2) \geqslant \varepsilon_2 \mathbf{I}$$

Then the system

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{C}(t) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix}$$

is uniformly BIBO stable if and only if $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable.*

Proof. The fact that exponential stability implies BIBO stability is easily established. Its proof is omitted.

Suppose uniform BIBO stability does not imply exponential stability. Then by the previous theorem it must not imply uniform BIBO stability of the system

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} \qquad (*)$$

Indeed, if there is no bound on the state there can be none on $\|\mathbf{y}\|$ for setting $\mathbf{u}$ equal to zero on $t_0 \leqslant t \leqslant t_0 + \delta$ gives

$$\int_{t_0}^{t_0+\delta} \|\mathbf{y}(t)\| \, dt = \mathbf{x}'(t_0)\mathbf{M}(t_0, t_0 + \delta)\mathbf{x}(t_0)$$

so

$$\max_{t_0 \leqslant t \leqslant t_0 + \delta} \|\mathbf{y}(t)\| \geqslant \delta^{-1}\mathbf{x}'(t_0)\mathbf{M}(t_0, t_0 + \delta)\mathbf{x}(t_0)$$
$$\geqslant \delta^{-1}\|\mathbf{x}(t_0)\| \, \lambda_{\min}(Mt_0, t_0 + \delta)$$

Hence if $\|\mathbf{x}(t)\|$ is not bounded for all bounded u, $\|\mathbf{y}\|$ cannot be either. However, by the previous theorem, uniform BIBO stability of (*) implies exponential stability of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$. ▮

Now that basic theorems have been proven, let's return to the relationship between BIBO stability and uniform BIBO stability. More specifically, can it happen that the system is BIBO stable but not uniformly BIBO stable? For the class of systems under consideration here the answer is no. However, the proof of this is very tedious if the argument is constructed from first principles (Kaplan). Otherwise, the relatively deep "principle of uniform boundedness" can be invoked to establish equivalence easily (Desoer and Thomasian). Having remarked that equivalence can be shown we are content to drop the subject.

Exercises

1. Show that for the system

$$x^{(n)}(t) + p_{n-1}(t)x^{(n-1)}(t) + \cdots + p_1(t)x^{(1)}(t) + p_0(t)x(t) = u(t)$$
$$y(t) = x(t)$$

uniform BIBO stability implies $x(t)$ goes to zero at an exponential rate (i.e. $|x(t)| \leqslant \gamma e^{-\lambda t}$) for $u \equiv 0$, provided the coefficients p_i are bounded.

2. Is the system

$$\dot{\mathbf{x}}(t) = x(t) + e^{2t}u$$
$$y(t) = e^{-2t}x(t)$$

BIBO stable? Is the homogeneous equation exponentially stable?

31. LIAPUNOV'S DIRECT METHOD

Determining whether or not all solutions of a particular linear differential equation remain bounded or go to zero as t approaches infinity can be quite difficult. Even two dimensional equations such as the Mathieu equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -a + 2q\cos 2t & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

must ordinarily be studied numerically in order to determine if particular values of parameters give rise to transition matrices which are bounded. On the other hand, it is possible to generate some useful sufficient conditions which, if satisfied, guarantee that all solutions will be bounded or even approach zero. This section is devoted to the study of one method of generating such sufficient conditions.

In order to determine the stability properties of solutions of differential equations, we will introduce certain scalar functions of $\mathbf{x}$ and t and study their evolution in time. The basic idea has its origin in classical dynamics where stability criteria involving the scalar *energy* are quite useful. Liapunov developed this method of investigating stability further in his memoir [55] and although the subject has evolved considerably since, we refer to this approach using his name.

Let $v(\mathbf{x}, t)$ be a quadratic form in $\mathbf{x}$ such that v is once differentiable with respect to t, and v and $(\partial v/\partial t)$ are bounded according to $v(\mathbf{x}, t) \leqslant k\,\|\mathbf{x}\|^2$ and $\|(\partial v(\mathbf{x}, t)/\partial t)\| \leqslant k_1\|\mathbf{x}\|^2$. Suppose G is a subset of the state space. We will say that v is a *quadratic Liapunov function for the differential equation* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *on the set* G if $\dot{v}$, the time rate of change of v along the solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ as computed by

$$\dot{v} = \sum_{i=1}^{m} \frac{\partial v}{\partial x_i}\dot{x}_i + \frac{\partial v}{\partial t}$$

satisfies $\dot{v}(\mathbf{x}, t) \leqslant w(\mathbf{x}) \leqslant 0$ for all $\mathbf{x}$ in G and all t.

Lets now make this notation more explicit. Since v is a quadratic form it can be written as

$$v(\mathbf{x}, t) = \mathbf{x}'\mathbf{Q}(t)\mathbf{x}$$

Differentiating this gives

$$\dot{v}(\mathbf{x}, t) = \dot{\mathbf{x}}'\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{x}'(t)\mathbf{Q}(t)\dot{\mathbf{x}}(t) + \mathbf{x}'(t)\dot{\mathbf{Q}}(t)\mathbf{x}(t)$$
$$= \mathbf{x}'(t)[\mathbf{A}'(t)\mathbf{Q}(t) + \mathbf{Q}(t)\mathbf{A}(t) + \dot{\mathbf{Q}}(t)]\mathbf{x}(t)$$

If $w(\mathbf{x})$ is to be an upper bound on $\dot{v}(\mathbf{x}, t)$ then we can always let $w(\mathbf{x})$ be a (time independent) quadratic form $w(\mathbf{x}) = \mathbf{x}'\mathbf{W}\mathbf{x}$.

Thus, a quadratic Liapunov function on a set G is a quadratic form whose derivative, $\dot{v}$, as computed along solutions of the given equation, is nonpositive. Our definition does not restrict the sign of v although it restricts the sign of $\dot{v}$. This is in contrast with a great deal of the literature on Liapunov's method.

Throughout this section we will give a fixed meaning to the letter E in accordance with

$$E = \{\mathbf{x}: w(\mathbf{x}) = 0; \mathbf{x} \in G\}$$

If G is a closed, bounded subset of the state space and if $v(\mathbf{x}, t)$ is a Liapunov function on G for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)x(t)$ then since $v(\mathbf{x}, t)$ is bounded by $k\|\mathbf{x}\|^2$ it is bounded on G. Moreover, v is monotone decreasing and since it is bounded on G, one might expect that either $\mathbf{x}$ leaves G or else $\dot{v}(\mathbf{x}, t)$ approaches zero. A result of this type is given by the following very useful theorem.

Theorem 1. *If v is a quadratic Liapunov function for $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}(t)$ on a closed bounded set G and if $\|\mathbf{A}(t)\|$ is bounded on $-\infty < t < \infty$ then any solution of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ which remains in G for all $t > t_0$ approaches E.*

Proof. Consider a solution $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ which remains in G for $t \geq t_0$. Since $\mathbf{A}$ is bounded and $\mathbf{x}$ is bounded, it follows that $\dot{\mathbf{x}}$ is bounded. Moreover

$$v(\mathbf{x}, t)\Big|_{t_0, \mathbf{x}_0}^{t_1, \mathbf{x}_1} = \int_{t_0}^{t_1} \dot{v}(\mathbf{x}, t)\, dt$$

Hence if $v(x, t)$ is bounded on $t_0 \leq t < \infty$ it follows that the integral

$$v_0 = \int_{t_0}^{\infty} \dot{v}(\mathbf{x}, t)\, dt$$

must converge. Since $\dot{v}(\mathbf{x}, t)$ is less than $\mathbf{x}'\mathbf{W}\mathbf{x}$ it follows that

$$\int_{t_0}^{\infty} \mathbf{x}_0' \mathbf{\Phi}'(t, t_0)\mathbf{W}\mathbf{\Phi}(t, t_0)\mathbf{x}_0\, dt < \infty$$

However, because $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ and its derivative $\mathbf{A}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ are bounded it follows that

$$\lim_{t\to\infty} \mathbf{x}_0' \mathbf{\Phi}'(t, t_0)\mathbf{W}\mathbf{\Phi}(t, t_0)\mathbf{x}_0 = 0$$

and this implies $\mathbf{x}(t)$ approaches E. ▌

Example. (LaSalle) To illustrate the use of Theorem 1 we consider the second order equation

$$\ddot{x}(t) + a(t)\dot{x}(t) + x(t) = 0$$

It can be written as

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -a(t)x_2(t) - x_1(t)$$

If a is positive, then $v(\mathbf{x}) = x_1^2 + x_2^2$ is a Liapunov function on any set G in the state space because $\dot{v}(x, t) = -a(t)x_2^2(t)$.

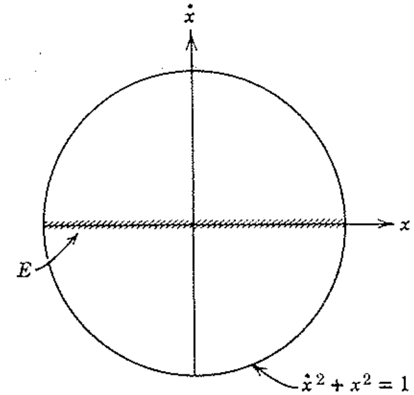


Figure 1. Applying Theorem 1 to $\ddot{x}(t) + a(t)\,\dot{x}(t) + x(t) = 0$.

If in addition we assume that a is bounded and $a(t) \geq \varepsilon > 0$ for all t. Then we can apply the above theorem and conclude that any solution which stays in a compact set approaches $E = \{\mathbf{x}: x_2 = 0\}$. In this case a simple *ad hoc* argument shows that all solutions are bounded. For example, $v(\mathbf{x}) = \|\mathbf{x}\|^2$ is monotone nonincreasing. Thus $\|\mathbf{x}(t)\| \leq \|\mathbf{x}(t_0)\|$, and hence $\mathbf{x}$ is bounded. Our conclusion is that all solutions are bounded and that $x_2(t)$ approaches zero as t approaches infinity. If we write the equations of motion as $\ddot{x}(t) + \varepsilon\dot{x}(t) + \mathbf{x}(t) = (a(t) + \varepsilon)\dot{x}(t)$ and use the corollary of Theorem 30-1 we see that $\|\mathbf{x}\|$ goes to zero.

Theorem 2. *If there exists a quadratic Liapunov function for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ such that v is negative for some $\mathbf{x}$ and t and $\dot{v}(\mathbf{x}, t) \leq \mathbf{x}'\mathbf{W}\mathbf{x}$ with $\mathbf{W}$ negative definite, then $\dot{\mathbf{x}}(t) = \mathbf{A}(t)x(t)$ is not stable.*

Proof. From Theorem 1 every solution which remains bounded approaches E. However since $\mathbf{W}$ is negative definite, $E = \{0\}$. Thus all solutions which are bounded approach zero. However if $v(\mathbf{x}_0, t_0)$ is negative $v(\mathbf{x}, t)$ remains negative and, in fact, $v(\mathbf{x}, t)$ remains less than $v(\mathbf{x}_0, t_0)$. Since $v(\mathbf{0}, t) \equiv 0$, $\mathbf{\Phi}(t, t_0)\mathbf{x}_0$ cannot approach zero and hence must be unbounded. ▌

This same type of an argument can be used with a slightly weaker hypothesis to conclude the absence of exponential stability.

Theorem 3. *If there exists a quadratic Liapunov function for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *on the whole state space such that* $v(\mathbf{x}, t)$ *is negative for some* $\mathbf{x}$ *and* t, *then* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is not exponentially stable.*

Proof. Since v is a Liapunov function on the whole state space it follows that v is nonincreasing as a function of time. Hence if it is negative for some $\mathbf{x}_0$ and t_0 the solution starting at $\mathbf{x}_0$ at time t_0 cannot go to zero because $v(\mathbf{0}, t) = 0$. ▮

In order to qualify as a Liapunov function it is necessary that $\mathbf{x}'(t)\mathbf{Q}(t)\mathbf{x}(t)$ be bounded from above by $k\,\|\mathbf{x}\|^2$. Hence $e^t x^2$ can never be a Liapunov function according to our definition. In order to prove stability or exponential stability, however, we need a lower bound on v of the form

$$v(\mathbf{x}, t) \geqslant \varepsilon\,\|\mathbf{x}\|^2$$

as well.

Theorem 4. *If* $v(\mathbf{x}, t)$ *is a quadratic Liapunov function for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *on the whole state space and if for some constant* $\varepsilon > 0$,

$$v(\mathbf{x}, t) \geqslant \varepsilon\,\|\mathbf{x}\|^2$$

then $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is uniformly stable.*

Proof. Since v is a Liapunov function its derivative is nonpositive. Hence for all $\mathbf{x}_0$ and all $t \geqslant t_0$ we have

$$v[\boldsymbol{\Phi}(t, t_0)\mathbf{x}_0, t] \leqslant v(\mathbf{x}_0, t_0)$$

However, since $k\,\|\mathbf{x}\|^2 \geqslant v(\mathbf{x}, t) \geqslant \varepsilon\,\|\mathbf{x}\|^2$ we see that if $\|\mathbf{x}_0\| = 1$ then $v(\mathbf{x}_0, t_0) \geqslant \varepsilon$ and therefore $\|\mathbf{x}(t)\|^2 \leqslant k/\varepsilon$ for all $t \geqslant t_0$. This establishes uniform stability. ▮

Our final result in this direction is a theorem on exponential stability.

Theorem 5. *If* $v(\mathbf{x}, t)$ *is a quadratic Liapunov function for* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *on the whole state space and if for some constant* $\varepsilon > 0$

$$v(\mathbf{x}, t) \geqslant \varepsilon\,\|\mathbf{x}\|^2; \qquad \text{all } t \geqslant t_0$$

then $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable if for some constant* $\mu > 0$, *and all* $t \geqslant t_0$

$$\dot{v}(\mathbf{x}, t) \leqslant -\mu\,\|\mathbf{x}\|^2$$

Proof. Since $v(\mathbf{x}, t)$ is a Liapunov function it is bounded from above by $k\,\|\mathbf{x}\|^2$. Hence we have

$$\frac{d}{dt}v(\mathbf{x}, t) < -\mu\,\|\mathbf{x}\|^2 < -(\mu/k)v(\mathbf{x}, t)$$

Hence $v(\mathbf{x}, t)$ is bounded from above by $e^{-(\mu/k)(t-t_0)}v(\mathbf{x}_0, t_0)$. Since $v(x, t) \geqslant \varepsilon\,\|\mathbf{x}\|$ this means

$$\|\mathbf{x}(t)\| < \varepsilon^{-1}e^{-(\mu/k)(t-t_0)}v(\mathbf{x}_0, t_0)$$

which establishes exponential stability. ▮

The question naturally arises as to whether or not there actually exists a Liapunov function which establishes stability if in fact, stability exists, for otherwise a search for a Liapunov function might be in vain. The following result shows that if exponential stability is on hand, then there always exists a Liapunov function meeting the requirements of Theorem 5.

Theorem 6. *Let* $\mathbf{A}$ *be bounded and suppose that the equation* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable. Let* $\mathbf{L}(\ )$ *be a bounded symmetric matrix. Then the integral*

$$\mathbf{Q}(t) = \int_t^{\infty} \boldsymbol{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\boldsymbol{\Phi}(\sigma, t)\,d\sigma$$

converges for all real t *and the derivative of* $\mathbf{x}'\mathbf{Q}(t)\mathbf{x}$ *along solutions of* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is* $-\mathbf{x}'\mathbf{L}(t)\mathbf{x}$. *Moreover, if* $\mathbf{L}(t) \geqslant \varepsilon\mathbf{I}$ *for all* t *then there exists constants* η_1 *and* η_2 *such that for all* t

$$\mathbf{I}\eta_1 \geqslant \mathbf{Q}(t) \geqslant \mathbf{I}\eta_2 > \mathbf{0}$$

Proof. Let $m\mathbf{I}$ be an upper bound on $\mathbf{L}$. Since $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable there exists γ and λ, both positive, such that

$$\|\boldsymbol{\Phi}(\sigma, t)\| \leqslant \gamma e^{-\lambda(\sigma-t)}; \qquad \sigma \geqslant t$$

Using these two estimates we see that $\mathbf{Q}$ is bounded according to

$$\begin{aligned}\mathbf{Q}(t) &\leqslant \int_t^{\infty} m\gamma^2 e^{-2\lambda(\sigma-t)}\mathbf{I}\,d\sigma \\ &= (m/2\lambda)\gamma^2\mathbf{I}\end{aligned}$$

This gives the upper bound on $\mathbf{Q}$.

To compute the derivative observe that

$$\begin{aligned}\frac{d}{dt}\mathbf{x}'\mathbf{Q}(t)\mathbf{x} &= \mathbf{x}'[\mathbf{A}'(t)\mathbf{Q}(t) + \mathbf{Q}(t)\mathbf{A}(t) + \dot{\mathbf{Q}}(t)]\mathbf{x}(t) \\ &= \mathbf{x}'\int_t^{\infty} \mathbf{A}'(t)\boldsymbol{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\boldsymbol{\Phi}(\sigma, t) + \boldsymbol{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\boldsymbol{\Phi}(\sigma, t)\mathbf{A}(t)\,d\sigma\,\mathbf{x} \\ &\quad + \mathbf{x}'\frac{d}{dt}\int_t^{\infty} \boldsymbol{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\boldsymbol{\Phi}(\sigma, t)\,d\sigma\mathbf{x} \\ &= \mathbf{x}'\int_t^{\infty} \frac{d}{dt}[\boldsymbol{\Phi}'(\sigma, t)\mathbf{L}(\sigma)\boldsymbol{\Phi}(\sigma, t)]\,d\sigma\mathbf{x} \\ &= -\mathbf{x}'\mathbf{L}(t)\mathbf{x}\end{aligned}$$

In going from the second line to the third we have used the differential equation for Φ and in going to the final expression we have used the fact that $\Phi(\sigma, t)$ goes to zero as σ approaches infinity and also the fact that $\Phi(t, t) = \mathbf{I}$.

To show that $x'Q(t)x$ is bounded from below as indicated provided $\mathbf{L}(t) \geqslant \varepsilon\mathbf{I} > 0$, observe that if a is an upper bound on $\|\mathbf{A}(t)\|$ valid for all t then

$$\left\| \frac{d}{dt} \Phi(t, t_0)\mathbf{x}_0 \right\| \leqslant a \|\mathbf{x}_0\|$$

and hence for $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$

$$\|\dot{\mathbf{x}}\| \leqslant a \|\mathbf{x}\|$$

Integrating this inequality gives

$$\|\Phi(\sigma, t)\mathbf{x}\| \geqslant e^{-a|\sigma - t|} \|\mathbf{x}\|$$

Using this on $\mathbf{x}'\mathbf{Q}(t)\mathbf{x}$ gives

$$\begin{aligned} \mathbf{x}'\mathbf{Q}(t)\mathbf{x} &= \int_t^\infty \mathbf{x}'\Phi'(\sigma, t)\mathbf{L}(\sigma)\Phi(\sigma, t)\mathbf{x}\, d\sigma \\ &\geqslant \|\mathbf{x}\|^2 \varepsilon \int_t^\infty e^{-2a(\sigma - t)}\, d\sigma \\ &\geqslant \|\mathbf{x}\|^2 \varepsilon/2a \quad \blacksquare \end{aligned}$$

Notice that the special case where $\mathbf{A}$ is a constant has been treated in Section 11 where the equivalence between solving

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{L}$$

and the evaluation of

$$\mathbf{Q} = \int_0^\infty e^{\mathbf{A}'t}\mathbf{L}e^{\mathbf{A}t}\, dt$$

was established.

Exercises

1. Find q_{12} and q_{22} such that the quadratic form

$$v(\mathbf{x}) = [x_1, x_2]\begin{bmatrix} 1 & q_{12} \\ q_{12} & q_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is positive definite and $\dot{v}$ as computed along the solutions of

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\eta(t) & -2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is nonpositive for $0 \leqslant \eta(t) \leqslant 4$.

2. If $\mathbf{z} = \mathbf{P}\mathbf{x}$ is a Liapunov transformation and if $\mathbf{x}'\mathbf{Q}(t)\mathbf{x}$ is a Liapunov function for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ which establishes exponential stability, then show that $\mathbf{z}'\mathbf{P}'(t)\mathbf{Q}(t)\mathbf{P}(t)\mathbf{z}$ is a Liapunov function which establishes the exponential stability of $\dot{\mathbf{z}}(t) = [\mathbf{P}(t)\mathbf{A}(t)\mathbf{P}^{-1}(t) + \dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t)]\mathbf{z}(t)$. Use this result to find a Liapunov function for $\dot{\mathbf{x}}(t) = e^{\mathbf{F}t}\mathbf{A}e^{-\mathbf{F}t}\mathbf{x}(t)$ assuming $e^{\mathbf{F}t}$ is periodic and $\mathbf{A} + \mathbf{F}$ has its eigenvalues in Re $s < 0$.
3. Compute $\dot{v}$ for $v = e^{3t}x^2$ and $\dot{x} = -x$. Compute v for $\dot{v} = e^{-3t}x^2$ and $\dot{x} = x$. Explain why theorem 1 is not violated.

32. SOME PERTURBATIONAL RESULTS

Among the most basic results in stability theory are those which assert that the behavior of the solutions of a particular vector equation is about the same as that of a "perturbed" version. A simple result of this type will serve to illustrate the idea.

Theorem 1. *Assume that the equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable. Then there exists an ε greater than zero such that if $\|\mathbf{B}(t)\| < \varepsilon$ for all t, then the null solution of $\dot{\mathbf{x}}(t) = [\mathbf{A}(t) + \mathbf{B}(t)]\mathbf{x}(t)$ is also exponentially stable.*

Proof. In the proof of Theorem 6 of Section 31, let $\mathbf{L} = \mathbf{I}$.

Then the resulting $\mathbf{Q}$ is positive definite and along solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}(t)\mathbf{x}(t)$ the derivative of $\mathbf{x}'\mathbf{Q}\mathbf{x}$ is

$$\begin{aligned} \frac{d}{dt}\mathbf{x}'\mathbf{Q}\mathbf{x} &= \mathbf{x}'[\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + \dot{\mathbf{Q}} + \mathbf{Q}\mathbf{B} + \mathbf{B}'\mathbf{Q}]\mathbf{x} \\ &= -\mathbf{x}'\mathbf{x} + \mathbf{x}'[\mathbf{Q}\mathbf{B} + \mathbf{B}'\mathbf{Q}]\mathbf{x} \end{aligned}$$

Given $p > 0$ there exists $\varepsilon > 0$ such that if $\|\mathbf{B}(t)\| < \varepsilon$ for all t, then

$$\|\mathbf{Q}\mathbf{B}(t) + \mathbf{B}'(t)\mathbf{Q}\| \leqslant \rho < 1$$

Hence, $\mathbf{x}'\{-\mathbf{I} + \mathbf{Q}\mathbf{B} + \mathbf{B}'\mathbf{Q}\}\mathbf{x}$ is negative definite and $\mathbf{x}'\mathbf{Q}\mathbf{x}$ is a quadratic Liapunov function which is positive definite. The set E is just the origin. Using Theorem 31-1 we see that $\mathbf{x}(t)$ approaches zero as t approaches infinity. $\blacksquare$

This result is only the simplest and most basic of many in the same vein which require the magnitude of the perturbation to be small. This type of estimate is not effective in treating perturbations which are large but slowly varying. For example, if we have $\ddot{x}(t) + 2\dot{x}(t) + (5 + 4\cos\omega t)x(t) = 0$ and express it as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -5 - 4\cos\omega t & -2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

then the preceding theorem cannot be used regardless of the value of ω, although for ω sufficiently small, intuition suggests that the null solution should be asymptotically stable. The following theorem shows this to be the case.

Theorem 2. *Consider a time varying matrix* $\mathbf{A}$. *Assume that there exists a bound b such that* $\|\mathbf{A}(t)\| < b$ *for all t and assume the eigenvalues of* $\mathbf{A}$ *lie in the half-plane* Re $s \leqslant \mu < 0$ *for all t. Then there exists an* $\varepsilon > 0$ *such that if* $\|\dot{\mathbf{A}}(t)\| \leqslant \varepsilon$ *for all t the null solution of* $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ *is exponentially stable.*

Proof. Let $\mathbf{Q}$ be the solution of

$$\mathbf{QA} + \mathbf{A}'\mathbf{Q} = -\mathbf{I}$$

Since $\mathbf{A}$ depends on t, $\mathbf{Q}$ will also. Since $\mathbf{A}$ is bounded and since the eigenvalues of $\mathbf{A}$ lie in Re $s \leqslant \mu < 0$, $\mathbf{Q}$ is also bounded. (See Problem 8.) Moreover, $\dot{\mathbf{Q}}\mathbf{A} + \mathbf{A}'\dot{\mathbf{Q}} = -\mathbf{Q}\dot{\mathbf{A}} - \dot{\mathbf{A}}'\mathbf{Q}$ and so,

$$\dot{\mathbf{Q}}(t) = \int_0^\infty e^{\mathbf{A}'(t)\sigma}[\mathbf{Q}(t)\dot{\mathbf{A}}(t) + \dot{\mathbf{A}}'(t)\mathbf{Q}(t)]e^{\mathbf{A}(t)\sigma}\, d\sigma$$

Clearly $\|\dot{\mathbf{Q}}(t)\|$ can be made less than one for all t by making $\|\dot{\mathbf{A}}\|$ suitably small. To complete the proof we observe that the time rate of change of $\mathbf{x}'\mathbf{Q}\mathbf{x}$ along solutions of the given equation is

$$\frac{d}{dt}\mathbf{x}'\mathbf{Q}\mathbf{x}\Big|_t = \mathbf{x}'(\mathbf{A}'\mathbf{Q} + \mathbf{QA} + \dot{\mathbf{Q}})\mathbf{x}\Big|_t$$

$$= -\mathbf{x}'(t)\mathbf{x}(t) + \mathbf{x}'(t)\dot{\mathbf{Q}}(t)\mathbf{x}(t)$$

which is less than zero if $\|\dot{\mathbf{Q}}\|$, hence if $\|\dot{\mathbf{A}}\|$, is small. ▮

To complement the above result on slowly varying coefficients we now establish a result which requires the coefficients to be rapidly varying and to have a suitable average value.

Theorem 3. *Let* $\mathbf{A}$ *depend continuously on t and assume it is periodic of period T. Suppose that*

$$\overline{\mathbf{A}} = \frac{1}{T}\int_0^T \mathbf{A}(\sigma)\, d\sigma$$

has all its eigenvalues in the half plane Re $s < 0$. *Then there exists a number* α_0 *such that the null solution of*

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\alpha t)\mathbf{x}(t)$$

is exponentially stable for all α *larger than* α_0.

Proof. From the Peano-Baker series we know that the transition matrix for this equation is given by

$$\mathbf{\Phi}(t, t_0) = \mathbf{I} + \int_0^t \mathbf{A}(\alpha\sigma)\, d\sigma + \frac{1}{2!}\int_0^t \mathbf{A}(\alpha\sigma_1)\int_0^{\sigma_1}\mathbf{A}(\alpha\sigma_2)\, d\sigma_2\, d\sigma_1 + \cdots$$

now

$$\mathbf{\Phi}(t + T/\alpha, t_0) = \mathbf{I} + \int_{t_0}^{t_0+T/\alpha}\mathbf{A}(\alpha\sigma)\, d\sigma + \frac{1}{2!}\int_{t_0}^{t_0+t/\alpha}\int_{t_0}^{\sigma_1}\mathbf{A}(\alpha\sigma_1)\,\mathbf{A}(\alpha\sigma_2)d\sigma_2\, d\sigma_1$$

$$= \mathbf{I} + \frac{1}{\alpha}\overline{\mathbf{A}} + \text{terms of order } \left(\frac{1}{\alpha}\right)^2$$

Letting $1/\alpha = \varepsilon$ we see that at $\varepsilon = 0$ the eigenvalues are on the unit disk but

$$\frac{d}{d\varepsilon}\mathbf{\Phi}(t_0 + \varepsilon T, t_0) = \overline{\mathbf{A}}$$

hence for ε sufficiently small the eigenvalues of $\mathbf{\Phi}$ lie inside the unit disk. ▮

None of the three preceding theorems is sufficiently sharp to permit an analysis of the equation $\ddot{x}(t) + \varepsilon(\sin \omega t)\dot{x}(t) + x(t) = 0$, regardless of how small $|\varepsilon|$ is required to be. As an example of the more delicate perturbation theorems applicable to problems of this type consider the following theorem of Liapunov.

Theorem 4. *Let* $\mathbf{A}$ *be a constant matrix with* $\mathbf{A} = -\mathbf{A}'$ *and assume that the eigenvalues of* $\mathbf{A}$ *are distinct* mod $2\pi i/T$. *Let* $\mathbf{B}(t) = \mathbf{B}'(t) = -\mathbf{B}(-t)$ *for all t, and assume* $\mathbf{B}$ *is periodic of period T. Then there exists an* $\varepsilon_0 > 0$ *such that if* $|\varepsilon| < \varepsilon_0$ *all solutions of*

$$\dot{\mathbf{x}}(t) = [\mathbf{A} + \varepsilon\mathbf{B}(t)]\mathbf{x}(t)$$

remain bounded.

Proof. We begin by showing that if λ is an eigenvalue of $\mathbf{\Phi}(t_0 + T, t_0)$ then λ^{-1} is also. Since the eigenvalues of a nonsingular matrix $\mathbf{M}$ are the reciprocals of those of $\mathbf{M}^{-1}$, this is equivalent to showing that if λ is an eigenvalue of $\mathbf{\Phi}(t_0 + T, t_0)$ it is also an eigenvalue of $\mathbf{\Phi}^{-1}(t_0 + T, t_0) = \mathbf{\Phi}(t_0, t_0 + T)$. Observe that if $\dot{x}(t) = \mathbf{A}\mathbf{x}(t) + \varepsilon\mathbf{B}(t)\mathbf{x}(t)$ then,

$$\frac{d}{dt}\mathbf{x}(-t) = -\mathbf{A}\mathbf{x}(-t) - \varepsilon\mathbf{B}(-t)\mathbf{x}(-t)$$

$$= \mathbf{A}'\mathbf{x}(-t) + \varepsilon\mathbf{B}(t)\mathbf{x}(-t)$$

Hence $\mathbf{x}$ satisfies the adjoint differential equation when it is regarded as a function of $-t$. This means $\mathbf{\Phi}(t_0, t_0 + T) = \mathbf{\Phi}'(t_0 + T, t_0)$. Using the fact

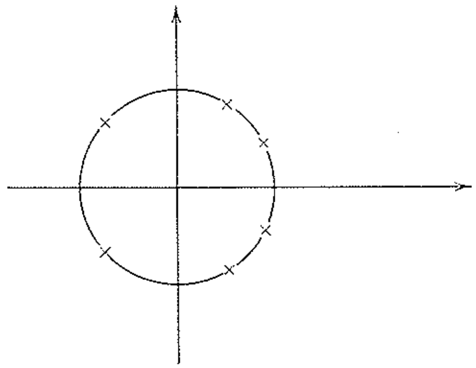


Figure 1. The unit disk and the eigenvalues of e^{AT}.

that the eigenvalues of any matrix are identical to those of its transpose we see that $\Phi(t_0, t_0 + T)$ and $\Phi^{-1}(t_0, t_0 + T)$ have the same eigenvalues.

Now at $\varepsilon = 0$ the eigenvalues of $\Phi(t_0 + T, t_0)$ are just those of $e^{\mathbf{A}t}$ and since $\mathbf{A} = -\mathbf{A}'$ these all lie on the unit disk. Moreover, since the eigenvalues of $\mathbf{A}$ are distinct and remain so when $n2\pi i/T$ is subtracted from them we see that the eigenvalues of $e^{\mathbf{A}t}$ are distinct. From the Peano-Baker series we know that $\Phi(t_0 + T, t_0)$ depends continuously on ε.

Now use the fact that the eigenvalues of a matrix depend continuously on its entries. At $\varepsilon = 0$ the eigenvalues of $\Phi(t_0 + T, t_0)$ are distributed as shown by the crosses in the figure. They move continuously as ε is increased but since they must occur in reciprocal pairs they must lie on the circle until two eigenvalues come together. Thus using Theorem 10-3 we see that the null solution is stable for $|\varepsilon|$ small. ▮

Exercises

1. If Φ is a symplectic matrix (Exercise 9, Section 24) show that its eigenvalues occur in reciprocal pairs. Use Theorem 4 to obtain a stability theorem for periodic canonical systems.
2. Construct a Liapunov function for the equation

$$\ddot{x}(t) + a\dot{x}(t) + bx(t) = 0$$

which has a negative definite derivative provided a and b are positive.

3. Show that all the eigenvalues of $\mathbf{A}$ have real parts less than $-\sigma$ if and only if for every symmetric positive definite matrix $\mathbf{C}$ there is a unique symmetric positive definite $\mathbf{Q}$ such that

$$\mathbf{A}'\mathbf{Q} + \mathbf{Q}\mathbf{A} + 2\sigma\mathbf{Q} = -\mathbf{C}$$

4. Let $\mathbf{A}$ be periodic of period T and assume that $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable. Show that there exists a positive definite matrix $\mathbf{Q}$ which is also periodic of period T such that $\mathbf{x}'\mathbf{Q}\mathbf{x}$ is a Liapunov function for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ on the whole state space.
5. Let $\mathbf{A}$ depend on t in a continuous way and assume $\mathbf{A}' = \mathbf{A}$. Let $\lambda_m(t)$ be the maximum eigenvalue of $\mathbf{A}(t)$. Show that if the integral satisfies

$$\int_{t_0}^{t} \lambda_m(t)\, dt \leqslant -\varepsilon(t - t_0)$$

then the equation $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ is exponentially stable. (See Exercise 3, Section 29).

6. Use the previous result to show that if $\mathbf{R}$ is a symmetric frequency response having a Laurent expansion $R(s) = L_0 s^{-1} + L_1 s^{-2} + \cdots$ and if the Hankel matrices

$$\mathbf{H}_n = \begin{bmatrix} \mathbf{L}_0 & \mathbf{L}_1 & \mathbf{L}_2 & \cdots & \mathbf{L}_r \\ \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 & \cdots & \mathbf{L}_{r+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{L}_2 & \mathbf{L}_{r+1} & \mathbf{L}_{r+2} & \cdots & \mathbf{L}_{2r} \end{bmatrix}; \qquad n = 0, 1, 2, \ldots$$

are all positive definite or positive semidefinite, then the equation $\dot{\mathbf{x}}(t) = \mathbf{A} - \mathbf{B}\mathbf{K}(t)\mathbf{C}\mathbf{x}(t)$ is stable if $\mathbf{K} = \mathbf{K}'$ and the average value of the maximum eigenvalue of $\mathbf{A} - \mathbf{B}\mathbf{K}(t)\mathbf{C}$ has a negative real part.

7. Consider the controllable system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$

with $\mathbf{A}$ constant and $\mathbf{B}$ constant. Let $\mathbf{W}(0, t_1)$ be given by

$$\mathbf{W}(0, t_1) = \int_0^{t_1} e^{-\mathbf{A}\sigma}\mathbf{B}\mathbf{B}'e^{-\mathbf{A}'\sigma}\, d\sigma$$

Show that if t_1 is positive all solutions of the linear system

$$\dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{B}\mathbf{B}'\mathbf{W}^{-1}(0, t_1))\mathbf{x}(t)$$

are bounded on $[0, \infty)$. (Canales)

8. If $\mathbf{A}$ is an n by n matrix whose entries are less than M in magnitude and whose eigenvalues lie in the half-plane $\text{Re}\, s \leqslant \varepsilon < 0$, find a bound on $\|\mathbf{A}^{-1}\|$.
9. Consider the second order equation of the Mathieu type

$$\ddot{x}(t) + [1 + \varepsilon \cos \omega t]x(t) = 0$$

For what values of ω is it impossible to use Theorem 4 to establish stability for $|\varepsilon|$ suitably small?

10. Show that all solutions of the second order equation

$$\ddot{x}(t) + [1 + f(t)]x(t) = 0$$

are bounded if $f(t) = -f(-t) = f(t+1)$ for all t and

$$\eta = \int_0^1 f^2(\sigma)\, d\sigma$$

is sufficiently small.

11. Show that if $p(t) = -p(-t) = p(t+1) \not\equiv 0$ then the equation

$$\ddot{x}(t) + p(t)x(t) = 0$$

is uniformly stable if

$$\int_0^1 p^2(t)\, dt < 4$$

(Liapunov).

33. FREQUENCY DOMAIN STABILITY CRITERIA

One of the main themes of system theory is the fruitful interplay between stationary weighting patterns and their associated frequency response functions. We have seen that although the frequency response function is the Laplace transform of the associated weighting pattern, it is not necessary to view it as such or, in fact, to know anything about Laplace transforms to make effective use of it. Moreover, it is an observed fact that in many problems the introduction of a frequency response function at the right moment results in a dramatic simplification. An early and still very important use of frequency response characterization to investigate the stability of dynamical systems was developed by Nyquist in connection with feedback amplifier design. His results, which apply only to a special class of problems, will be discussed in the following section together with some generalizations. Our objectives here are to prove a basic result which is in many ways broader in scope and to establish some associated techniques.

We begin with a lemma on equations of a special type.

Lemma 1. *If* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *is a minimal realization of a frequency response* $\mathbf{R}$ *and if* $\mathbf{F}$ *is a bounded, matrix valued, function of time, then for the system*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{B}\mathbf{F}(t)\mathbf{C}\mathbf{x}(t); \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \qquad \text{(MF)}$$

it follows that $\|\mathbf{x}(t)\|$ *approaches zero as* t *approaches infinity provided* $\|\mathbf{y}(t)\|$ *approaches zero as* t *approaches infinity.*

Proof. For the system (MF) we know (Theorem 24-4) that for fixed $\mathbf{x}(t)$,

$$\int_t^{t+\sigma} \mathbf{u}'(\rho)\mathbf{u}(\rho) + \mathbf{y}'(\rho)\mathbf{y}(\rho)\, d\rho \geqslant \mathbf{x}'(t)\mathbf{K}(\sigma)\mathbf{x}(t)$$

with $\mathbf{K}(\sigma)$ being nonnegative definite for all $\sigma > 0$. In view of the observability assumption the integral cannot vanish for $\mathbf{x}(t) \neq 0$ so $\mathbf{K}$ must be positive definite. To apply this to the problem at hand, let $\mathbf{u}(t) = -\mathbf{F}(t)\mathbf{C}\mathbf{y}(t)$ to get

$$\int_t^{t+\sigma} \mathbf{y}'(\rho)[\mathbf{I} + \mathbf{C}'\mathbf{F}'(\rho)\mathbf{F}(\rho)\mathbf{C}]\mathbf{y}(\rho)\, d\rho \geqslant \mathbf{x}'(t)\mathbf{K}(\sigma)\mathbf{x}(t)$$

Since $\mathbf{F}$ is bounded we see that if $\|\mathbf{y}(t)\|$ approaches zero, then $\|\mathbf{x}(t)\|$ must also. ▮

Theorem 1. *Let* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *be a minimal realization of the frequency response function* $\mathbf{R}$ *and assume that the eigenvalues of* $\mathbf{A}$ *lie in the half-plane* $\operatorname{Re} s < 0$. *Suppose that for all real* ω

$$\mathbf{I} - \mathbf{R}'(-i\omega)\mathbf{R}(i\omega) \geqslant \mathbf{0}$$

Then if $\mathbf{I} - \mathbf{F}'(t)\mathbf{F}(t) \geqslant \varepsilon\mathbf{I} > \mathbf{0}$ *for all* $t > 0$, *it follows that all solutions of*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{B}\mathbf{F}(t)\mathbf{C}\mathbf{x}(t)$$

are bounded and approach zero as t *approaches infinity.*

Proof. From Theorem 25-2 it follows that there exists a negative definite solution of equation

$$\mathbf{A}'\mathbf{K} + \mathbf{K}\mathbf{A} - \mathbf{K}\mathbf{B}\mathbf{B}'\mathbf{K} = \mathbf{C}'\mathbf{C}$$

Denote this solution by $-\mathbf{K}_0$. Then along solutions of the given differential equation

$$\begin{aligned}\frac{d}{dt}\mathbf{x}'\mathbf{K}_0\mathbf{x} &= \mathbf{x}'[-\mathbf{K}_0\mathbf{B}\mathbf{B}'\mathbf{K}_0 - \mathbf{C}'\mathbf{C} - (\mathbf{K}_0\mathbf{B}\mathbf{F}\mathbf{C}) - (\mathbf{K}_0\mathbf{B}\mathbf{F}\mathbf{C})']\mathbf{x}\\ &= -\mathbf{y}_1'\mathbf{y}_1 - \mathbf{y}_2'\mathbf{y}_2 - \mathbf{y}_1'\mathbf{F}\mathbf{y}_2 - \mathbf{y}_2'\mathbf{F}'\mathbf{y}_1\\ &= -(\mathbf{y}_1 + \mathbf{F}\mathbf{y}_2)'(\mathbf{y}_1 + \mathbf{F}\mathbf{y}_2) - \mathbf{y}_2'(\mathbf{I} - \mathbf{F}'\mathbf{F})\mathbf{y}_2\end{aligned}$$

where $\mathbf{y}_1 = \mathbf{B}'\mathbf{K}_0\mathbf{x}$ and $\mathbf{y}_2 = \mathbf{C}'\mathbf{x}$. This last expression is clearly negative semidefinite as a function of $\mathbf{x}$, so we see from Theorem 31-1, that all solutions approach the set $\{\mathbf{x}: \mathbf{y}_2 = \mathbf{0}\}$. Hence by Lemma 1 it follows that $\mathbf{x}$ goes to zero. ▮

Example. Consider the scalar equation

$$\ddot{x}(t) + (\alpha(t) + \delta)\dot{x}(t) + (\beta(t) + 1)x(t) = 0$$

If we make the identifications,

$$\mathbf{R}(s) = \begin{bmatrix} \dfrac{1}{s^2 + \delta s + 1} \\[2ex] \dfrac{s}{s^2 + \delta s + 1} \end{bmatrix}; \qquad \mathbf{F}(t) = [\beta(t), \alpha(t)]$$

then we see that all solutions are bounded and approach zero as t approaches infinity provided

$$\sup_t[\beta^2(t) + \alpha^2(t)] \cdot \sup_\omega \left| \frac{1 + \omega^2}{(1 - \omega^2)^2 + \delta\omega^2} \right| < 1$$

There are two simple auxiliary ideas which greatly extend the usefulness of this result. Let $\mathbf{F}_0$ be a constant matrix having the same dimensions as $\mathbf{F}$. Then one can either study equation (MF) or else study the equivalent system

$$\dot{x}(t) = (\mathbf{A} - \mathbf{BF}_0\mathbf{C})\mathbf{x}(t) - \mathbf{B}(\mathbf{F}(t) - \mathbf{F}_0)\mathbf{Cx}(t)$$

This leads to the following corollary.

Corollary 1. *Let* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *be a minimal realization of the frequency response function* $\mathbf{R}$. *Assume that there exists a constant matrix* $\mathbf{F}_0$ *such that the eigenvalues of* $\mathbf{A} - \mathbf{BF}_0\mathbf{C}$ *lie in the half-plane* Re $s < 0$. *Let*

$$\mathbf{R}_0(s) = \mathbf{C}(\mathbf{I}s - \mathbf{A} + \mathbf{BF}_0\mathbf{C})^{-1}\mathbf{B}$$

Suppose that for all real ω

$$\mathbf{I} - \mathbf{R}_0'(-i\omega)\mathbf{R}_0(i\omega) \geqslant \mathbf{0}$$

Then if for all t

$$\mathbf{I} - [\mathbf{F}(t) - \mathbf{F}_0]'[\mathbf{F}(t) - \mathbf{F}_0] \geqslant \varepsilon\mathbf{I} > \mathbf{0}$$

it follows that all solutions of

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) - \mathbf{BF}(t)\mathbf{Cx}(t)$$

are bounded and approach zero as t *approaches infinity.*

This result essentially permits one to apply Theorem 1 to any equation in the same feedback equivalence class.

Theorem 1 as it stands, cannot be improved by letting $\mathbf{x}(t) = \mathbf{Pz}(t)$ and studying the $\mathbf{z}$ equation. It is coordinate free in that sense. However, in another sense it is not. If $\mathbf{P}_1\mathbf{Q}_1 = \mathbf{I}$ and $\mathbf{P}_2\mathbf{Q}_2 = \mathbf{I}$, then it is enough that the hypothesis be satisfied for $\mathbf{R}_1 = \mathbf{Q}_2\mathbf{C}(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{BP}_1$, and $\mathbf{F}(t) = \mathbf{Q}_1\mathbf{F}(t)\mathbf{P}_2$.

Corollary 2. *Let* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ *be a minimal realization of the frequency response function* $\mathbf{R}$ *and assume that the eigenvalues of* $\mathbf{A}$ *lie in* Re $s < 0$. *Assume that there exist constant matrices* $\mathbf{P}_1$, $\mathbf{Q}_1$, $\mathbf{P}_2$, *and* $\mathbf{Q}_2$ *such that* $\mathbf{P}_1\mathbf{Q}_1 = \mathbf{I}$ *and* $\mathbf{P}_2\mathbf{Q}_2 = \mathbf{I}$. *Let* $\mathbf{R}_1(s) = \mathbf{Q}_2\mathbf{R}(s)\mathbf{P}_1$ *and let* $\mathbf{F}_1(t) = \mathbf{Q}_1\mathbf{F}(t)\mathbf{P}_2$. *Suppose that for all real* ω

$$\mathbf{I} - \mathbf{R}_1'(-i\omega)\mathbf{R}_1(i\omega) \geqslant \mathbf{0}$$

and suppose that for all real t

$$\mathbf{I} - \mathbf{P}_2'\mathbf{F}'(t)\mathbf{Q}_1'\mathbf{Q}_1\mathbf{F}(t)\mathbf{P}_2 \geqslant \varepsilon\mathbf{I} > \mathbf{0}$$

Then it follows that all solutions of

$$\dot{\mathbf{x}}(t) = \mathbf{Ax}(t) - \mathbf{BF}(t)\mathbf{Cx}(t)$$

go to zero as t *approaches infinity.*

Exercises

1. Apply Theorem 1 to the equation

$$\ddot{x}(t) + 2\dot{x}(t) + (2 + \alpha \sin t)x(t) = 0$$

in the form

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -(2 + \alpha \sin t) & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

How large can α be if the hypothesis is to be fulfilled? Make a change of coordinates according to

$$\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Apply Theorem 1 again. Can ω be chosen in such a way as to allow for a larger value of α than was found before?

2. Evaluate $\mathbf{C}'(\mathbf{I}s - \mathbf{A} + \mathbf{BF}_0\mathbf{C})^{-1}\mathbf{B}$ in terms of $\mathbf{R}(s) = \mathbf{C}'(\mathbf{I}s - \mathbf{A})^{-1}\mathbf{B}$ and $\mathbf{F}_0$. (See Section 19.)
3. Let $\mathbf{R}$ and $\mathbf{\Omega}$ be defined as shown:

$$\mathbf{R}(s) = \begin{bmatrix} r(s) & 0 \\ 0 & r(s) \end{bmatrix}; \qquad \mathbf{\Omega} = \begin{bmatrix} 0 & \omega_0 \\ -\omega_0 & 0 \end{bmatrix}$$

Show that the system

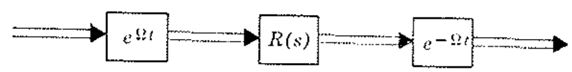


Figure 1

has a stationary weighting pattern. Call it $\mathbf{R}_1(s)$. Then show that

$$\mathbf{I} - \alpha\mathbf{R}_1'(-s)\mathbf{R}_1(s)|_{s=i\omega} \geqslant \mathbf{0}; \qquad \text{all } \omega$$

if

$$1 - \alpha r(-s)r(s)|_{s=i\omega} \geqslant 0; \qquad \text{all } \omega$$

and conversely.

34. NYQUIST CRITERION: POSITIVE REAL FUNCTIONS

When specialized to the case where **R** and **F** are scalars, the results discussed in Section 33 become much more concrete and easier to apply. In particular, comparisons between the necessary and sufficient conditions available for stability of time invariant linear differential equations and the sufficient conditions for stability of time varying linear differential equations are easier to make. The actual comparisons are postponed until the next section. Here we set the background.

Let r be a rational function of a complex variable $s = \sigma + i\omega$. The locus of points

$$\Gamma(r) = \{u + iv \colon u = \operatorname{Re} r(i\omega),\ v = \operatorname{Im} r(i\omega);\ -\infty \leqslant \omega \leqslant \infty\}$$

is called the *Nyquist Locus of r*. That is, the Nyquist Locus is the image of the line $\operatorname{Re}[s] = 0$ under the mapping r. If $\Gamma(r)$ is bounded we will say that the Nyquist Locus encircles a point $u_0 + iv_0$ ρ times if $u_0 + iv_0$ is not on the Nyquist Locus and $2\pi\rho$ is the net increase in the argument of $r(i\omega) - u_0 - iv_0$ as ω increases from $-\infty$ to $+\infty$. We regard clockwise rotations as the direction of increasing argument. Negative encirclements are counterclockwise encirclements.

We will frequently assume that the Nyquist Locus is bounded. This will be the case if and only if the degree of the denominator of r equals or exceeds that of the numerator and r has no poles on the line $\operatorname{Re}[s] = 0$.

The first result we will give relating to the Nyquist Locus is the following version of a feedback stability theorem generally attributed to Nyquist. Mathematically, the result is best viewed as an application of the principle of the argument.

Theorem 1. *Suppose r has a bounded Nyquist Locus. If r has ν poles in the half-plane $\operatorname{Re}[s] > 0$ then $r/(1 + kr)$ has $\rho + \nu$ poles in the half-plane $\operatorname{Re}[s] > 0$ if the point $-1/k + i0$ is not on the Nyquist Locus and $\Gamma(r)$ encircles $-(1/k + i0)$ ρ times in the clockwise sense.*

Proof. If r' denotes the derivative of r then for any simple closed curve C an integration around this curve in a clockwise sense gives

$$\frac{1}{2\pi i}\int_C r'(s)/r(s)\, ds = \mu - \nu$$

where μ is the number of zeros of r inside C and ν is the number of poles of r in the same region, repeated zeros and poles being counted according to their multiplicity. To see this, observe that the singularities of the integrand inside C are the poles and zeros of r inside C. Write r as

$$r(s) = (s + s_i)^{m_i} r_i(s)$$

and notice that the residue of r'/r at every pole of r of multiplicity m_i is $-m_i$ and that the residue at every zero of r of multiplicity m is m_i. That is

$$\frac{r'(s)}{r(s)} = \frac{m_i}{(s + s_i)} + \frac{r_i(s)}{r(s)}$$

with r_i/r analytic near $-s_i$. The desired result then follows from applying the residue theorem to all such expansions with $-s_i$ inside C.

Moreover, by direct integration we see that

$$\frac{1}{2\pi i}\int r'(s)/r(s)\, ds = \frac{1}{2\pi i}\ln r$$

Thus, if we integrate around a closed curve C we get no change in the magnitude of $\ln r$ but the argument of r will change by a multiple of 2π equal to the number of times the image of C encircles the origin in the r plane. Thus,

$$\mu - \nu = \frac{1}{2\pi i}\int_\Gamma r'(s)/r(s) ds = \begin{cases}\text{number of times the image of } C(r)\\ \text{encircles the origin in the } r\text{-plane}\end{cases}$$

Consider the frequency response of the closed loop system $r/(1 + kr)$. Clearly it has zeros where r has zeros and poles where $1 + kr$ has zeros. Let C be a contour consisting of the imaginary axis and notice that the Nyquist locus of $1 + kr$ encircles the origin if and only if the Nyquist locus of r encircles the $-1/k$ point. Using the integral expression for $\mu - \nu$ we obtain the desired result. ▌

A second result which we require and which uses the Nyquist Locus, concerns a special class of rational functions which we now define. If $q(s)$ and $p(s)$ are polynomials without common factors we will say that $q(s)/p(s)$ is *positive real* if

(i) $\operatorname{Re} q(i\omega)/p(i\omega) \geqslant 0$ for all real ω.
(ii) $q(s) + p(s)$ has all its zeros in $\operatorname{Re}[s] < 0$.

We want to give a necessary and sufficient condition for $(\alpha r + 1)/(\beta r + 1)$ to be positive real in terms of the Nyquist Locus of r. Naturally r itself is positive real if $\Gamma(r)$ lies in the closed right half-plane and the zeros of $1 + r$ lie in the open left half-plane. In order to study positive realness of the bilinear transformation of r we introduce the following notation.

Let α and β be real numbers with $\alpha < \beta$. If $\alpha\beta > 0$ the symbol $D(\alpha, \beta)$ stands for a disk in the complex plane defined by

$$D(\alpha, \beta) = \left\{u + iv\colon \left[u + \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\right]^2 + v^2 < \frac{1}{4}\left|\frac{1}{\alpha} - \frac{1}{\beta}\right|^2\right\}$$

If $\alpha\beta < 0$ then $\alpha < 0 < \beta$ and $D(\alpha, \beta)$ is the complement of a disk.

$$D(\alpha, \beta) = \left\{u + iv: \left[u + \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\right]^2 + v^2 > \frac{1}{4}\left|\frac{1}{\alpha} + \frac{1}{\beta}\right|\right\}$$

If $\alpha = 0$ then $\beta > 0$ and $D(\alpha, \beta)$ is a half plane

$$D(0, \beta) = \left\{u + iv: u < -\frac{1}{\beta}\right\}$$

and if $\alpha < 0$, $\beta = 0$, $D(\alpha, \beta)$ is the half plane

$$D(\alpha, \beta) = \left\{u + iv: u > -\frac{1}{\alpha}\right\}$$

For all other cases $D(\alpha, \beta)$ is left undefined. (See Figure 1)

In the statement of Theorem 2 below, properties of $(\alpha r + 1)/(\beta r + 1)$ are related to the Nyquist Locus of r. Here and elsewhere, when reference is made to encirclements of the D region described above by the Nyquist locus, it should be understood that this is possible only if $\alpha\beta > 0$.

Theorem 2. *Suppose the Nyquist Locus of r is bounded and suppose r has v poles in the half-plane* $\text{Re}[s] > 0$. *Let z be given by*

$$z = \frac{r + 1/\alpha}{r + 1/\beta}$$

(a) *If* $\Gamma(r)$ *does not intersect* $D(\alpha, \beta)$ *and encircles it v times in the counterclockwise sense then z is positive real.*

(b) *If* $\Gamma(r)$ *does not intersect* $D(\alpha, \beta)$ *and encircles it fewer than v times in the counterclockwise sense then z is not positive real.*

Proof. Write r as the ratio of two polynomials q/p which do not have common factors. Then

$$z = \frac{q + p/\alpha}{q + p/\beta}$$

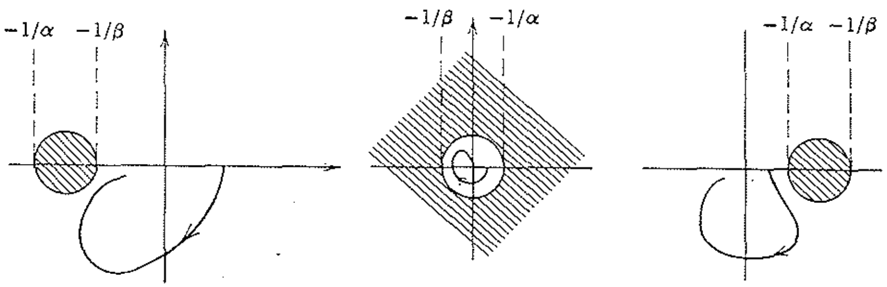


Figure 1. Defining $D(\alpha, \beta)$.

and if $\alpha \neq \beta$ then the polynomials $q_0 = q + p/\alpha$ and $p_0 = q + p/\beta$ are without common factors. Indeed, an inversion of the bilinear map defining z gives (assuming $\alpha \neq \beta$)

$$\frac{q}{p} = \frac{q_0 - p_0}{\beta q_0 - \alpha p_0}$$

If q_0 and p_0 have a common factor then this is surely a common factor of q and p, contrary to the hypothesis.

The mapping

$$z = \frac{r + 1/\alpha}{r + 1/\beta}$$

takes the exterior of the disk $D(\alpha, \beta)$ in the r-plane into the half-plane $\text{Re}[z] \geqslant 0$. Hence, if $\Gamma(r)$ does not intersect $D(\alpha, \beta)$

$$0 \leqslant \text{Re}[z(j\omega)] = \text{Re}\,\frac{r(j\omega) + 1/\alpha}{r(j\omega) + 1/\beta}$$
$$= \frac{p_0(j\omega)q_0(-j\omega) + p_0(-j\omega)q_0(j\omega)}{2p_0(j\omega)^2}$$

We conclude that if the Nyquist locus $\Gamma(r)$ does not intersect the disk $D(\alpha, \beta)$ then for all real ω

$$p_0(j\omega)q_0(-j\omega) + p_0(-j\omega)q_0(j\omega) \geqslant 0$$

We have shown that our assumptions insure that q_0 and p_0 do not have common factors and that z is nonnegative on the line $\text{Re}[s] \equiv 0$. It remains to show only that $p_0 + q_0$ has, or does not have, all its zeros in $\text{Re}[s] < 0$ according to whether (a) or (b) is satisfied. The polynomial $p_0 + q_0$ takes the form

$$p_0 + q_0 = 2\left[q + \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)p\right]$$

The hypothesis prevents $\Gamma(r)$ from passing through the point $r = -\frac{1}{2}(1/\alpha + 1/\beta)$ (i.e., the center of the disk). Thus $p_0 + q_0$ has no zeros on $\text{Re}[s] = 0$. According to Nyquist's theorem, $p_0 + q_0$ has $v + \rho$ zeros in the half-plane $\text{Re}[s] > 0$ where ρ denotes the number of times $\Gamma(r)$ encircles the point $r = -\frac{1}{2}(1/\alpha + 1/\beta)$ in the clockwise sense. Thus $p_0 + q_0$ has $v + \rho$ zeros in the half-plane $\text{Re}[s] > 0$. When (a) is satisfied z is positive real and when (b) is satisfied z is not positive real. ▮

Exercises

1. Suppose that q and p do not have common factors. Show that q/p is positive real if (i) $p(iw)q(-iw) + p(-iw)q(iw) \geqslant 0$; (ii) all poles of p on

$\mathrm{Re}[s] = 0$ are simple, and (iii) the residues of q/p at these poles are real and positive.

2. Show that if r is given by

$$r(s) = \frac{s + a}{s^2 + bs + 1}$$

then r is positive real if and only if $b \geqslant a \geqslant 0$.

3. Show that

$$z(s) = \frac{s^2 + q_1 s + q_0}{s^2 + p_1 s + p_0}$$

is positive real if and only if q_0, q_1, p_0, p_1 are all positive and

$$(q_0 + p_0 - q_1 p_1)^2 \leqslant 4 q_0 p_0$$

4. Suppose that q and p do not have common factors and suppose further that $q(iw)/p(iw)$ has a positive real part for some w. Then q/p is positive real if and only if

$$p^2(D)x(t) + f q^2(D)x(t) = 0$$

is uniformly stable for all constant f in the interval $0 < f < \infty$.

5. If r_1 and r_2 are positive real, show that $r_1 + r_2$ and $(r_1)^{-1} + (r_2)^{-1}$ are also positive real.

6. Consider the polynomial with real coefficients

$$p(s) = as^5 + bs^4 + cs^3 + ds^2 + es + f;$$

Use the Nyquist's Criterion to show that it has its zeros in the half-plane $\mathrm{Re}\, s < 0$ if and only if the polynomial

$$q(s) = b^2 s^4 + (bc - ad)s^3 + bds^2 + (be - af)s + b$$

does. Also show that the Hurwitz determinants for $p(s)$ are positive if and only if those of $r(s)$ are positive.

35. THE CIRCLE CRITERION

When specialized to the case where **F** and **G** are scalars, Theorem 1 of Section 33 states that if the Nyquist Locus lies inside the circle $\{z: |z| \leqslant 1\}$ and if f is less than one in magnitude for all t, then all solutions are bounded. It was this form of the basic frequency response theorem that Bongiorno [10] established using the additional assumption that f is periodic. Some manipulation of this special case yields the more appealing statement that all solutions are bounded if the zeros of r lie in $\mathrm{Re}\, s < 0$ and the Nyquist locus of r does not encircle the disk $D(\alpha, \beta)$ where $\alpha < f(t) < \beta$. This is directly comparable with the stability part of the Nyquist Criterion.

To get an actual generalization of the Nyquist Criterion, however, is another matter. Nyquist's theorem also gives instability criteria and nothing done thus far gives any information on instability. This change, in effect, necessitates a whole new approach to the problem. The techniques to be used here are based on a scalar representation of the equations of motion which we now introduce.

Consider the system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t); \qquad y = \mathbf{c}\mathbf{x}(t) \tag{S}$$

and consider the related differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{b}f(t)\mathbf{c}\mathbf{x}(t) \tag{F}$$

If the system (S) is controllable, then there exists a nonsingular matrix **P** which puts the system in standard controllable form. For (S) in standard controllable form the differential equation (F) takes the form

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_{n-1}(t) \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -p_0 & -p_1 & -p_0 & \cdots & -p_{n-1} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + f(t) \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \\ -q_0 & -q_1 & -q_2 & \cdots & -q_{n-1} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix}$$

and in particular, the variable x_1 satisfies the nth order differential equation

$$p(D)x_1(t) + f(t)q(D)x_1(t) = 0$$

Thus (F) can always be reduced to this form if the system (S) is controllable. For notational convenience we let $x_1 = x$ and write the basic equation as

$$p(D)x(t) + f(t)q(D)x(t) = 0 \tag{SF}$$

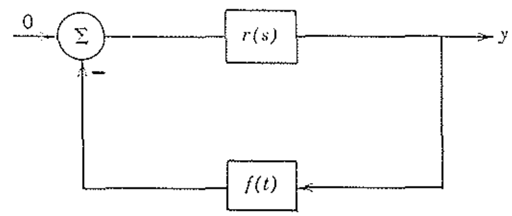


Figure 1. The system configuration for the circle criterion.

Theorem 1. (Circle Theorem) *Let* $[\mathbf{A}, \mathbf{b}, \mathbf{c}]$ *be a minimal realization of the frequency response function* r. *Assume that* ν *eigenvalues of* $\mathbf{A}$ *lie in the half-plane* $\operatorname{Re} s > 0$ *and that no eigenvalues of* $\mathbf{A}$ *lie on the line* $\operatorname{Re} s = 0$. *If* f *is a piecewise continuous function bounded according to* $\alpha + \varepsilon \leqslant f(t) \leqslant \beta - \varepsilon$, $\varepsilon > 0$ *then*

(i) *all solutions of* $\dot{\mathbf{x}}(t) = [\mathbf{A} - \mathbf{b}f(t)\mathbf{c}]\mathbf{x}(t)$ *are bounded and go to zero as* t *approaches infinity provided the Nyquist locus of* r *does not intersect the disk* $D(\alpha, \beta)$ *and encircles it exactly* ν *times in the counterclockwise sense.*

(ii) *at least one solution of* $\dot{\mathbf{x}}(t) = [\mathbf{A} - \mathbf{b}f(t)\mathbf{c}]\mathbf{x}(t)$ *does not remain bounded as* t *approaches infinity provided the Nyquist Locus of* r *does not intersect the disk* $D[\alpha, \beta]$ *and encircles it fewer than* ν *times in the counterclockwise sense.*

Proof. Since the Nyquist locus of r does not intersect the disk $D(\alpha, \beta)$ we know from Theorem 2 of Section 34 that

$$z = \frac{r + 1/\alpha}{r + 1/\beta} = \frac{q + p/\alpha}{q + p/\beta}$$

is nonnegative on the line $\operatorname{Re}[s] \equiv 0$ and in addition, that z is positive real if (i) is satisfied and that z is not positive real if (ii) is satisfied. Since $z(i\omega) \geqslant 0$ for all real ω we see that

$$\operatorname{Re}[\alpha q(i\omega) + p(i\omega)][\beta q(-i\omega) + p(i\omega)] \geqslant 0$$

for all real ω. Thus the definition

$$r(D) = \{\operatorname{Ev}[\alpha q(D) + p(D)][\beta q(-D) + p(-D)]\}^{-}$$

involving spectral factorization makes sense.

Consider the definition (see Section 26)

$$v(\mathbf{x}) = \int_{t(0)}^{t(\mathbf{x})} [\alpha q(D) + p(D)]x(t)[\beta q(D) + p(D)]x(t) - [r(D)x(t)]^2\, dt$$

where $\mathbf{x} = (x; x^{(1)}; \ldots, x^{(n-1)})'$ with n being the degree of p. We will show that $v(\mathbf{x})$ is a Liapunov function for the equation in question.

First observe that the integral in the definition of $v(\mathbf{x})$ is independent of path (see Lemma 1 of Section 26) and hence unambiguously defines a function of $\mathbf{x}$. From our definition of a Liapunov function we need only show that along solutions of equation (SF) $\dot{v}(\mathbf{x}, t) \leqslant w(\mathbf{x}) \leqslant 0$. If we rewrite equation (SF) as

$$[p(D) + \alpha q(D)]x(t) + (f(t) - \alpha)q(D)x(t) = 0$$

and then multiply by $[(\beta q(D) + p(D)]x(t)$ we see that along solutions of

equation

$$v[\mathbf{x}(t_2)] - v[\mathbf{x}(t_1)] = \int_{t_1}^{t_2} -[\beta - f(t)][f(t) - \alpha][q(D)x(t)]^2 - [r(D)x(t)]^2\, dt$$

so

$$\dot{v}(\mathbf{x}, t) = -[\beta - f(t)][f(t) - \alpha][q(D)x(t)]^2 - [r(D)x(t)]^2$$

This is clearly nonpositive if $\alpha \leqslant f(t) \leqslant \beta$ and so $v(\mathbf{x})$ is a Liapunov function. It should be pointed out that since $r(D)$ and $q(D)$ may be of degree n, the right side of the equation for $\dot{v}$, as it is written may depend on $x, x^{(1)} \cdots x^{(n-1)}$ *and* $x^{(n)}$ hence the differential equation (SF) itself must be used to re-express this in terms of $\mathbf{x} = (x; x^{(1)}; \ldots; x^{(n-1)})$. From problem 8 (this section) we know that the Liapunov function $v(\mathbf{x})$ is positive definite if z is positive real and neither positive definite nor positive semidefinite if z is not positive real.

We now complete the two parts of the proof separately.

(i) We have a positive definite time independent quadratic form such that

$$\dot{v}(\mathbf{x}, t) \leqslant -\varepsilon^2(q(D)x)^2$$

This implies via Theorem 31-1 that $(n(D)x)^2 \to 0$. By lemma 33.1 we see $\mathbf{x} \to 0$.

(ii) We have $v(\mathbf{x}) < 0$ for some $\mathbf{x}$ and $v(\mathbf{0}) = 0$. Using Theorem 31-1 we see that every solution which remains bounded approaches $q(D)x = 0$ which means by lemma 33.1 that $\mathbf{x} \to \mathbf{0}$. However this is impossible if v is initially negative. Hence those solutions for which $v(0)$ is negative go to infinity as t approaches infinity. ▮

Example. Consider the second order equation

$$\ddot{x}(t) + 2\dot{x}(t) + x(t) + f(t)x(t) = 0$$

The Nyquist locus and some appropriate D regions are shown in Figure 2. A quick calculation shows that this equation is stable if for some $\alpha > 0$

$$\alpha^2 \leqslant f(t) + 1 \leqslant (\alpha + 2)^2$$

and thus the null solution of

$$\ddot{x}(t) + 2\dot{x}(t) + g(t)x(t) = 0$$

is stable for

$$\alpha^2 \leqslant g(t) \leqslant (\alpha + 2)^2$$

Using the instability part of the theorem we see that the null solution is unstable if $f(t) < -1$ for all time.

Unfortunately the circle criterion usually does not give the least restrictive conditions possible on $f(t)$. Since this is an important point we will use the second order equation of this example to illustrate.

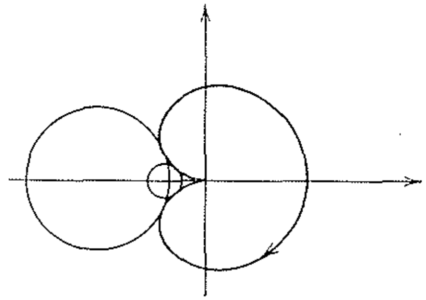


Figure 2. The Nyquist Locus for $1/(s^2+2s+1)$ *and some circles appropriate for applying the circle criterion.*

Theorem 2. *All solutions of the equation in the above example are bounded if* $0 < f(t) < v$, v *being the solution of*

$$2 = \sqrt{v}\exp\frac{\cos^{-1}(1/\sqrt{v})}{\sqrt{v-1}};\ v > 1 \tag{T}$$

Proof. The plan of the proof is to construct a closed curve in the state space which the trajectories always remain parallel to or else cross from the outside inward.

Consider the solutions in the phase plane. The slope of the trajectory at any point is

$$m = (d\dot{x})/(dx) = -2 - f(t)(x/\dot{x})$$

Now consider the trajectory in the first quadrant of the phase plane if one lets $f(t) = 0$ and starts at $\dot{x} = 1$, $x = 0$. This is a straight line which ends at $\dot{x} = 0$, $x = \frac{1}{2}$. In the fourth quadrant consider the trajectory obtained by setting f equal to a constant whose value is such that if the trajectory starts at $\dot{x} = 0$, $x = \frac{1}{2}$ it crosses the line $x = 0$ at $\dot{x} = -1$. (See Figure 3.)

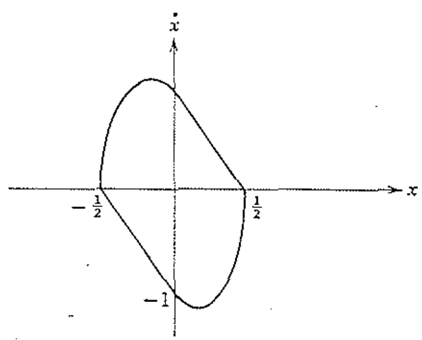


Figure 3. Showing some trajectories of $\ddot{x}(t)+2\dot{x}(t)+f(t)x(t)=0$.

An elementary calculation using the transition matrix for two dimensional equations shows that it takes $(0 \leqslant \tan^{-1}x < \pi)$

$$t = \frac{1}{\sqrt{v-1}}\tan^{-1} - \sqrt{v-1} \tag{T}$$

units of time to move from $\dot{x} = 0$ to $x = 0$ along solutions of

$$\ddot{x}(t) + 2\dot{x}(t) + vx(t) = 0$$

Imposing the above boundary conditions then we see that v must be chosen so as to satisfy equation (T). We see from the expression for the slope of the trajectories that for any $f(t)$ such that $0 \leqslant f(t) \leqslant v$, the trajectories which start inside this contour will remain inside for all time. ▮

It may be shown that v is about 11.6 compared with an upper limit of 4 predicted by the circle criterion.

In order to conclude stability using the circle criteria it is necessary that $[\mathbf{A} - \mathbf{bc}'f(t)]$ have all its eigenvalues in the half-plane $\text{Re}[s] < 0$ for all permissible values of $f(t)$. On the other hand, the results of Theorem 2 clearly indicate that this condition is not sufficient since it is not sufficient for this special case. In order to be able to use the circle criterion to establish instability, it is necessary to have at least one eigenvalue of $[\mathbf{A} - \mathbf{bc}'f(t)]$ in $\text{Re}[s] > 0$ for all permissible values of $f(t)$. Thus far we have no evidence to indicate whether or not this condition in itself implies instability. The following example shows that like stability, instability cannot be concluded on such weak evidence as eigenvalue locations.

Example. (Brockett and Lee) Consider the differential equation

$$\ddot{x}(t) - 2\dot{x}(t) + 2x(t) + f(t)[3\dot{x}(t) - 4x(t)] = 0$$

with $f(t)$ constrained for all t by

$$0 \leqslant f(t) \leqslant 1$$

Now the zeros of $(D^2 - 2D + 2) + f(t)(3D - 4)$ are sketched in Figure 4. As f varies between zero and 1 at least one zero stays in the half-plane $\text{Re}[s] \geqslant 1$. We will show that in spite of this there is a time function f which makes the equation exponentially stable.

Let $f(t)$ vary periodically between the limits 0 and 1 according to the relationship

$$f(t) = \begin{cases} 0 & \text{for } 0 \leqslant t - nT < T_1 \quad n = 0, 1, 2, \ldots \\ 1 & \text{for } T_1 \leqslant t - nT < T \quad n = 0, 1, 2, \ldots \end{cases}$$

where $T_1 = \tan^{-1} 3$, $T_2 = \pi$, $T = T_1 + T_2$ (see Figure 4).

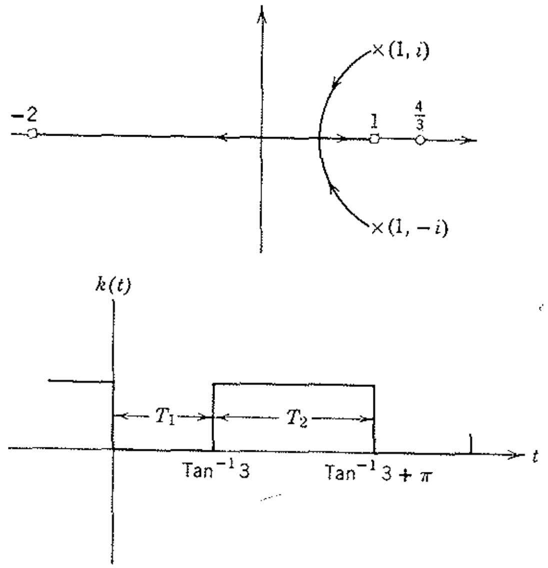


Figure 4. Illustrating the example.

For $f(t)$ given above, the general form of the solution is

$$x(t) = p_1(t)e^{\lambda_1 t} + p_2(t)^{\lambda_2 t}$$

where $p_1(t)$ and $p_2(t)$ are bounded periodic functions dependent on the initial state having period T, and the λ_i are the so-called characteristic exponents. A straightforward but lengthy calculation shows that the quantities $e^{\lambda i} = z_i$ satisfy

$$z_i^2 + 2[e^{T_1 - 2T_2} \cos T_1]z_i + e^{2T_1 - T_2} = 0$$

Use of the quadratic formula shows that $|z_1| < 1$ and $|z_2| < 1$. Thus regardless of the initial state, $\|x(t)\|^2$ approaches zero at an exponential rate.

Accordingly, the example equation is not uniformly stable for any fixed gain satisfying $0 \leqslant f \leqslant 1$ but is uniformly stable for at least one time-varying gain satisfying the same constraints.

Exercises

1. Let f be periodic and let $[\mathbf{A}, \mathbf{b}, \mathbf{c}]$ be a minimal realization of r. Assume that the circle criterion permits one to conclude that the null solution of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{b}f(t)\mathbf{c}\mathbf{x}(t)$ is exponentially stable for $\alpha \leqslant f(t) \leqslant \beta$. Show that if $\alpha \leqslant f(t) \leqslant \beta$ then all solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - \mathbf{b}f(t)\mathbf{c}\mathbf{x}(t) + \mathbf{b}$ approach a periodic solution and show that the average value of this solution is less than $\mathbf{c}\mathbf{A}^{-1}\mathbf{b}$. (Wolaver)

2. The second order equation

$$\ddot{x}(t) + \eta(t)\dot{x}(t) + \gamma(t)\dot{x} = 0$$

can be converted to a second order equation containing only one time variation in various ways. Let $y(t) = \rho(t)x(t)$ and verify that

$$\ddot{y}(t) + [-2\dot{\rho}(t)/\rho(t) + \eta(t)]\dot{y}(t) + [\gamma(t) - \eta(t)\dot{\rho}(t)/\rho(t) + 2\dot{\rho}^2(t)/\rho(t) - \ddot{\rho}(t)/\rho(t)]y(t) = 0$$

Suppose, for example, that η is periodic and $\bar{\eta}$ is its average value. Show that if ρ satisfies

$$2\dot{\rho}(t) = [\bar{\eta} + \eta(t)]\rho(t)$$

then y satisfies an equation of the form $\ddot{y}(t) + \bar{\eta}\dot{y}(t) + v(t)y(t) = 0$ and that x is bounded if and only if y is.

3. The Damped Mathieu equation is

$$\ddot{x} + \delta\dot{x} + (a - 2b \cos 2t)x = 0$$

Use the circle criterion to obtain an estimate for the range of a, b and δ for which this equation is uniformly stable.

4. Consider a piecewise constant system of the form

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}; \qquad \begin{aligned} &\mathbf{A}(t+\omega) = \mathbf{A}(t) \\ &\mathbf{A}_1 \text{ for } 0 \leqslant t \leqslant \alpha_1\omega \\ &\mathbf{A}_2 \text{ for } \omega\alpha_1 \leqslant t < (\alpha_1 + \alpha_2)\omega \\ &\vdots \\ &\mathbf{A}_n \text{ for } (\alpha_1 + \alpha_2 \cdots + \alpha_{n-1})\omega \leqslant t < (\alpha_1 + \alpha_2 \cdots + \alpha_n) = \omega \end{aligned}$$

Show that this system is stable as $\omega \to 0$ if $\sum_{i=1}^{n} \alpha_i \mathbf{A}_i$ has all its eigenvalues in the left half-plane. (Desoer-Page)

5. Let $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ be a minimal realization. Use the Gronwall-Bellman inequality (exercise 9, section 2) to show that the linear equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) - f(t)\mathbf{b}\mathbf{c}\mathbf{x}(t)$$

is exponentially stable if $|f(t)| < \alpha$ and

$$|\mathbf{c}e^{\mathbf{A}t}\mathbf{b}| \leqslant (1/\alpha\lambda)e^{-\lambda t}$$

Show that the circle criterion gives a better result. (J. L. Willems)

6. Use of the technique of the proof of the circle criterion to show that null solution of

$$p(D)x(t) + f[q(D)x(t)] = 0$$

is asymptotically stable in the large if $f(\sigma)/\sigma \geqslant \varepsilon > 0$ for all real σ and

$(1+\alpha s)q(s)/p(s)$ is positive real for some $\alpha > 0$. *Hint:* Show that

$$v(\mathbf{x}) = \int_{t(0)}^{t(x)} (1+\alpha D)q(D)x\,p(D)x - \{\mathrm{Ev}[(1+\alpha D)q(D)p(-D)]^{+}x\}^2\,dt$$

$$+ \alpha \int_{t(0)}^{t(x)} f[q(D)x(t)]Dq(D)x\,dt$$

is a Liapunov function. (This is a version of Popov's Theorem)

7. Consider a scalar nonlinear equation of the form

$$p(D)x(t) + f[q(D)x(t)] = u(t)$$

with $u(t)$ periodic of period T. Suppose that f is differentiable and that $\alpha \leqslant f'(x) \leqslant \beta$. Moreover, suppose that

$$p(D)x(t) + k(t)q(D)x(t) = 0$$

is exponentially stable for any $k(\ )$ such that $\alpha \leqslant k(t) \leqslant \beta$ for all t. Show that all solutions of the original nonlinear equation approach the same periodic solution and that the least period of this solution is T or less. Hence conclude that *jump phenomena* and *subharmonic responses* are impossible in this type of system.

8. Show that $v(\mathbf{x})$ as given in the proof of Theorem 1 is positive definite if z is positive real and is not positive definite otherwise. *Hint:* If z is positive real evaluate the integral along a solution of

$$[\alpha q(D) + p(D) + \beta q(D) + p(D)x(t)] = 0$$

If z is not positive real assume v is positive definite and use Liapunov Theory to get a contradiction.

NOTES AND REFERENCES

28. Kolmogorov and Fomin's two small volumes [45] are excellent sources for material on normed linear spaces. Varga [75] and Marcus and Minc [57] are both good references on vector and matrix inequalities.

29. The relationship between exponential stability and integrability of $\|\Phi(t,\sigma)\|^i$ is considered in Kaplan [43] and Kalman and Bertram [41]. Original work was done by Perron around 1930. We have tried to be systematic in stating side by side 4 equivalent conditions on the integral of the norm of Φ. However it is clear that integrability of the first and second powers of the norm are not all that might be of interest.

30. There is considerable confusion in the literature about the distinction between BIBO stability and what we have called uniform BIBO stability. The equivalence of these two ideas for linear systems has been correctly established in several places and in several different settings (Kaplan [43], Desoer and Thomasian [21], Youla [88]) but very often the distinction has been erroneously overlooked. The basic trick used to establish Theorem 2 is due to Silverman and Anderson [73].

31. Good surveys of the basic stability definitions can be found in the nice monographs of Hahn [30] and LaSalle and Lefschetz [51]. The survey of Kalman and Bertram [41] is very readable and contains a number of interesting applications. Our basic approach is taken from LaSalle [49] who also emphasizes the importance of Yoshizawa's paper (Yoshizawa [86]). One merit of this approach is that it allows one to give a single satisfactory definition of what one means by a Liapunov function. By using Yoshizawa's theorem we can come to conclusions about time-varying equations using a Liapunov function which is only negative semi-definite. This was apparently not widely recognized before the appearance of Yoshizawa's paper.

32. Theorem 1 is completely standard and can be found e.g. in Coddington and Levinson [20]. Theorem 2 is contained in Rosenbrock [70]. Erugin [22] refers to Theorem 3 as originating with Demidovich.

33. These results are easy consequence of the main theorem of Popov [68] (also Anderson [3]). They have also been obtained by others. The type of proof given here has the merit of producing a Liapunov function without requiring a knowledge of matrix spectral factorization.

34. Nyquist's result [65] is widely used in feedback system design. Together with the root locus technique [43], it has been for many years the basis of design for control systems and electronic amplifiers. The concept of a positive real function originated in engineering circles with Brune who discovered that a positive real function could always be realized as the impedance of a linear passive electrical network [29]. These functions are also important in mathematics and have been studied by Bochner and Mathias. It may seem strange to put these two ideas together here, but in fact it is quite difficult to separate them as is evident, for example, in Hurwitz's paper.

35. Many people have worked on the stability criterion referred to here as the circle criterion. References [9, 10, 48, 61, 62, 71, 72, 92] in addition to others are devoted to the subject. Reference [12] contains some historical remarks. The instability part of this theorem was discovered later [16]. Desoer, Narendra, Popov, Sandberg, J. C. Willems, J. L. Willems, Yacubovich, and Zames have done a great deal of interesting work on this type of problem including generalizations far beyond the result given here.

REFERENCES

1. R. Abraham, *Linear and Multilinear Algebra*, W. A. Benjamin, Inc., New York, 1966.
2. M. A. Aizerman and R. F. Gantmacher, *Absolute Stability of Regulator Systems*, San Francisco, Holden-Day, 1964. (Translation of the 1963 Russian Edition).
3. B. D. O. Anderson, "A System Theory Criterion for Positive Real Matrices," *SIAM J. on Control*, Vol. 5, No. 2, May 1967, 171–182.
4. M. Athans and P. Falb, *Optimal Control*, New York, McGraw-Hill, 1966.
5. E. F. Beckenbach and R. Bellman, *Inequalities*, Berlin, Springer-Verlag, 1961.
6. R. Bellman, *Introduction to Matrix Analysis*, New York, McGraw-Hill, 1960.
7. R. Bellman and R. Kalaba (Editors), *Mathematical Trends in Control Theory*, New York, Dover, 1964.
8. G. Birkhoff and G. Rota, *Ordinary Differential Equations*, Boston, Ginn and Company, 1962.
9. J. J. Bongiorno, Jr., "Real frequency stability criteria for linear time-varying systems," *Proc. IEEE*, vol. 52, July 1964, 832–841.
10. J. J. Bongiorno, Jr., and D. Graham, "An extension of the Nyquist Barkhausen stability criterion to linear lumped-parameter systems with time-varying elements," *IEEE Trans. on Automatic Control* (Correspondence), vol. AC-8, April 1963, 166–170.
11. T. F. Bridgeland, "Stability of Linear Signal Transmissions Systems," *SIAM Review*, Vol. 5, No. 1, 1963, 7–32.
12. R. W. Brockett, "The Status of Stability Theory for Deterministic Systems," *IEEE Trans. on Automatic Control*, vol. AC-11, July, 1966, 596–606.
13. R. W. Brockett, "Variational Methods for the Stability of Periodic Equations," *Differential Equations and Dynamical Systems*, (Jack Hale and J. P. LaSalle, Eds.) Academic Press, 1966.
14. R. W. Brockett, "Poles, Zeros, and Feedback: State Space Interpretation," *IEEE Trans. on Automatic Control*, Vol. AC-10, No. 2, April 1965, 129–135.

15. R. W. Brockett, "On the stability of nonlinear feedback systems," *IEEE Trans. on Applications and Industry*, vol. 83, November 1964, 443–449.
16. R. W. Brockett and H. B. Lee, "Frequency Domain Instability Criteria for Time-Varying and Nonlinear Systems," *Proc. IEEE*, vol. 55, No. 5, 1967.
17. R. W. Brockett and M. D. Mesarovic, "The reproducibility of multivariable systems," *Math Analysis and Applications*, Vol. 10, 1965.
18. A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Blaisdell, Waltham, Mass., 1969.
19. L. Cesari, *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*, New York, Academic Press, 1963 (Second Edition).
20. E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw Hill, 1955.
21. C. A. Desoer and A. J. Thomasian, "A Note on Zero-State Stability of Linear Systems," *Proceedings of the First Allerton Conference on Circuit and System Theory*, Urbana, Illinois, University of Illinois, 1963.
22. N. P. Erugin, *Linear Systems of Ordinary Differential Equations*, Academic Press, New York, 1966. (Translation of the 1963 Russian original.)
23. W. Fulks, *Advanced Calculus*, John Wiley, New York, 1961.
24. F. R. Gantmacher, *The Theory of Matrices*, Vols. I and II, New York, Chelsea Publishing Company, 1959.
25. I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice Hall, Englewood Cliffs, N. J., 1963.
26. E. G. Gilbert, "Controllability and Observability in Multivariable Control Systems," *SIAM Journal for Control*, Vol. 1, No. 2, 1963, 128–151.
27. W. H. Greub, *Linear Algebra* (Third Edition), Springer-Verlag, New York, 1967.
28. M. Gruber and J. L. Willems, "On a generalization of the Circle Criterion," *Proceedings of the Fourth Allerton Conference on Circuits and System Theory*, University of Illinois, Urbana, Illinois, 1966.
29. E. A. Guillemin, *Synthesis of Passive Networks*, New York, Wiley, 1957.
30. W. Hahn, *Theory and Application of Liapunov's Direct Method*, Prentice Hall, 1963. (Translation of 1959 German edition.)
31. J. Hale, *Oscillation in Nonlinear Systems*, New York, McGraw-Hill, 1963.
32. P. R. Halmos, *Finite Dimensional Vector Spaces*, Princeton, New Jersey, Van Nostrand, 1958.
33. Ch. Hermite, "Sur le nombre des racines d'une équation algébrique comprises entre des limites données," *J. Reine Angew. Math.*, Vol. 52, 1850, 39–51.

34. B. L. Ho and R. E. Kalman, "Effective Construction of Linear State-Variable Models from Input-Output Data," Regelungstechnik, vol. 12, pp. 545–548, 1966.
35. R. E. Kalman, "On the Structural Properties of Linear Constant, Multivariable Systems," Proceedings of the 1966 IFAC Conference, Butterworths, London.
36. R. E. Kalman, "Canonical Structure of Linear Dynamical Systems," *Proc. National Acad. Sci. U.S.*, vol. 48, 596–600, 1962.
37. R. E. Kalman, "When is a linear control system optimal?" *Trans. ASME* (*J. Basic Engrg.*). pt. D., March 1964, 51–60.
38. R. E. Kalman, "Liapunov Functions for the Problem of Lurie in Automatic Controls," *Proc. Nat. Acad. of Sci., U.S.A.*, Vol. 49, No. 3, 1963, 201–205.
39. R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," *SIAM Journal for Control*, Vol. 1, No. 2, 1963, 152–192.
40. R. E. Kalman, "Contributions to the Theory of Optimal Control," *Boletin de la Sociedad Matematica Mexicana*, 1960.
41. R. E. Kalman and J. E. Bertram, "Control System Design Via the 'Second Method' of Liapunov, Part I, Continuous-Time Systems; Part II, Discrete Time Systems," *ASME Journal of Basic Engineering*, Vol. 82, 1960, 371–393, 394–400.
42. R. E. Kalman, Y. C. Ho, and K. S. Narendra, "Controllability of Linear Dynamical Systems," *Contributions to the Theory of Differential Equations*, Vol. 1, Interscience, 1963.
43. W. Kaplan, *Operational Methods for Linear Systems*, Reading, Mass., Addison-Wesley, 1962.
44. D. L. Kleinman, "On the linear regulator problem and the matrix Riccati equation," M.I.T. Electronic Systems Lab., Cambridge, Mass., Rept. ESL-R-271, June, 1966.
45. A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis*, vols. I and II, Graylock Press, Rochester, N.Y., 1957.
46. N. N. Krasovskii, *Stability of Motion*, Stanford, California, Stanford University Press, 1963. (Translation of the 1954 Russian edition.)
47. E. Kreindler and P. E. Sarachick, "On the concepts of controllability and observability of linear systems," *IEEE Trans. on Automatic Control*, Vol. AC-9, April 1964, 129–136.
48. J. Kudrewicz, "Stability of nonlinear systems with feedback," Automation and Remote Control, Vol. 25, August 1964 (English translation pp. 1027–1037).
49. J. P. LaSalle, "An Invariance Principle in the Theory of Stability," *Differential Equations and Dynamical Systems*, J. Hale and J. P. LaSalle, Eds., Academic Press, 1967.

50. LaSalle, J. P., "The Time-optimal Control Problem," in *Contributions to Differential Equations*, vol. V, 1–24, Princeton University Press, Princeton, N.J., 1960.
51. J. P. LaSalle and S. Lefschetz, *Stability by Liapunov's Direct Method*, New York, Academic, 1961.
52. E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, John Wiley, New York, 1967.
53. S. Lefschetz, *Differential Equations: Geometric Theory*, New York, Interscience, 1957.
54. S. H. Lehnigk, *Stability Theorems for Linear Motions*, Prentice Hall, Englewood Cliffs, N.J., 1966.
55. A. M. Liapunov, "Probleme general de la stabilite du mouvement," *Ann. Fac. Sci. Toulouse*, Vol. 9, 203–474, (1907) (French Translation of Russian paper published in 1893 in *Comm. Soc. Math. Kharkow*) French translation also reprinted as Vol. 17 in *Ann. of Math Studies*, Princeton, New Jersey, 1949.
56. L. H. Loomis and Shlomo Sternberg, *Advanced Calculus*, Addison-Wesley, Reading, Mass., 1968.
57. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Inc., Boston, 1964.
58. B. McMillan, "Introduction to Formal Realizability Theory," *Bell System Technical Journal*, Vol. 31, 1952, 217–279, March, 591–600, May.
59. K. Meyer, "On the existence of Lyapunov functions for the problem of Lure," *SIAM J. on Control*, vol. 3, No. 3, 1966, 373–383.
60. B. S. Morgan, "The synthesis of linear multivariable feedback," Ph. D. dissertation, University of Michigan, Ann Arbor, 1963.
61. K. S. Narendra and R. M. Goldwyn, "A geometrical criterion for the stability of certain nonlinear, nonautonomous systems," *IEEE Trans. on Circuit Theory* (Correspondence), Vol. CT-11, September 1964, 406–408.
62. B. N. Naumov and Y. Z. Tsypkin, "A frequency criterion for absolute process stability in nonlinear automatic control systems," *Automation and Remote Control*, Vol. 25, June 1964. (English translation, 765–778.)
63. Nemytskii and V. V. Stepanov, *Qualitative Theory of Differential Equations*, Princeton, New Jersey, Princeton University Press, 1960. (Translation of an earlier Russian edition with some additions.)
64. G. C. Newton, Jr., L. A. Gould and J. F. Kaiser, *Analytic Design of Linear Feedback Controls*, New York, John Wiley, 1957.
65. H. Nyquist, "The Regeneration Theory," *Bell Systems Technical Journal*, Vol. 11, 1932, 126–147.
66. M. C. Pease, *Methods of Matrix Algebra*, Academic Press, New York, 1965.
67. L. S. Pontryagin, "Optimal Control Processes," *Uspekhi Mat. Nauk USSR*, Vol. 45, p.3.

68. V. M. Popov, "Hyperstability and Optimality of Automatic Systems with Several Control Functions," *Rev. Roum. Sci.-Electrotechn. et Energ.*, Vol. 9, 1964, 629–690.
69. J. Radon, "Zum Problem von Lagrange," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universitat*. Vol. 6, 1928, 273–299.
70. H. H. Rosenbrock, "The Stability of Linear Time-dependent Control Systems," *J. of Electronics and Control*, Vol. 15, No. 1, p. 73, July, 1963.
71. I. W. Sandberg, "A frequency domain condition for the stability of systems containing a single time-varying nonlinear element," *Bell Systems Tech. J.*, Vol. 43, 1964, 1601–1638.
72. I. W. Sandberg, "On the stability of solutions of linear differential equations with periodic coefficients," *J. SIAM*, Vol. 12, No. 2, 1964, 487–496.
73. L. M. Silverman and B. D. O. Anderson, "Controllability, Observability, and Stability of Linear Systems," *SIAM J. Control*, Vol. 6, No. 1, 1968.
74. P. Slepian and L. Weinberg, "Positive Real Matrices," *Journal of Mathematics and Mechanics*, Vol. 9, No. 1, 1960, 71–83.
75. R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Inc., New Jersey, 1962.
76. L. Weinberg, *Network Analysis and Synthesis*, McGraw-Hill, 1962.
77. L. Weiss and R. E. Kalman, "Contributions to Linear System Theory," *International Journal for Engineering Science*, Vol. 3, pp. 141–171.
78. J. C. Willems, "On the Asymptotic Stability of the Null Solution of Linear Differential Equations with Periodic Coefficients," *IEEE Trans. on Automatic Control*, Vol. AC-13, No. 1, Feb. 1968, pp. 65-72.
79. J. C. Willems, *The Analysis of Feedback Systems*, M.I.T. Press, Cambridge, Mass., 1970.
80. J. L. Willems, *Stabiliteitsonderziek van systemen met een niet-lineair tijdsafhankelijk element door middel van de directe methode van Liapunov, met teopassingen in de Sterkstroomtechnoek*. Doctoral Thesis, University of Ghent, 1967.
81. M. Wonham, "On Pole Assignment in Multi-Input Controllable Linear Systems," *IEEE Trans. on Automatic Control*, Vol. AC-12, No. 6, Dec., 1967, 660–665.
82. M. Wonham, "On a Matrix Riccati Equation of Stochastic Control," *SIAM J. Control*, Vol. 6, No. 4, 1968, 681–697.
83. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," *ASME Journal of Basic Engineering*, March, 1964, 107–115.
84. V. A. Yacubovich, "Absolute Stability of Nonlinear Control Systems in Critical Cases," Parts I and II, *Automation and Remote Control*, Vol. 24, Nos. 3 and 6, Oct. 1963 and Jan. 1964, 273–282, 655–668.

85. V. A. Yacubovich, "The solution of certain matrix inequalities in automatic control theory," *Dokl. Akad. Nauk USSR*, vol. 143, 1962, 1304–1307.
86. T. Yoshizawa, "Asymptotic Behavior of Solutions of a System of Differential Equations," *Contributions to Differential Equations*, Vol. 1, No. 3, 1963, 371–387.
87. D. C. Youla, "The Synthesis of Linear Dynamical Systems from Prescribed Weighting Patterns," *J. SIAM Appl. Math.*, Vol. 14, No. 3, 1966.
88. D. C. Youla, "On the Stability of Linear Systems," *IEEE Trans. on Circuit Theory*, Vol. CT-10, No. 2, 1963, 276–279.
89. D. C. Youla, "On the Factorization of Rational Matrices," *IRE Trans. on Information Theory*, Vol. IT-7, July 1961, 172–189.
90. D. C. Youla and Plinio Tissi, "*n*-port Synthesis via Reactance Extraction—Part I," *IEEE International Convention Record*.
91. L. Zadeh and C. A. Desoer, *Linear System Theory*, New York, McGraw-Hill, 1963.
92. G. Zames, "On the input-output stability of time-varying nonlinear feedback systems—Parts I and II," *IEEE Trans. on Automatic Control*, Vol, AC-11, April 1966, 228–238, and July 1966, 465–476.
93. P. Zeiger, "Ho's Algorithm, Commutative Diagrams, and the Uniqueness of Minimal Linear Systems," *Information and Control*, Vol. 11, pp. 71–79, 1967.

GLOSSARY OF NOTATION

A. Spaces, Norms, and Transformations

R^n	Cartesian n-space
E^n	Euclidian n-space (R^n + inner product)
$C^m(t_0, t_1)$	Set of continuous maps of $[t_0, t_1]$ into R^m
$C_*^m(t_0, t_1)$	$C^m(t_0, t_1)$ + inner product
$R^{n \times m}$	The space of all real n by m matrices
$E^{n \times m}$	$R^{n \times m}$ + inner product
$\|\mathbf{x}\|$	$(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$
$\|\mathbf{A}\|$	Max of $\|\mathbf{Ax}\|$ for $\|\mathbf{x}\| = 1$
$\mathbf{I}(\mathbf{I}_n)$	Identity matrix (of dimension n)
$\mathbf{0}(\mathbf{0}_n)$	Zero matrix (of dimension n)
$(\mathbf{A}, \mathbf{B})$	Column partition of matrices
$(\mathbf{A}; \mathbf{B})$	Row partition of matrices
λ_i ; s_i	Eigenvalues
$\mathbf{0} > 0$ $(\mathbf{0} \geqslant 0)$	$\mathbf{0}$ is symmetric and positive definite (nonnegative definite)
L	Linear transformation
L^*	Adjoint transformation

B. Differential Equations

$\mathbf{\Phi}(t, t_0)$	Transition matrix
$\mathbf{\Phi}_{\mathbf{A}}(t, t_0)$	Transition matrix for $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$
$\mathbf{x}$, $\mathbf{z}$	State vectors (n-dimensional)
$\mathbf{p}$	Adjoint equation variable (n-dimensional)
$e^{\mathbf{A}t}$	Matrix exponential
$\mathbf{P}^{-1}(t)e^{\mathbf{R}(t-t_0)}\mathbf{P}(t_0)$	Floquet–Liapunov representation of $\mathbf{\Phi}$

C. Linear Systems

$\mathbf{u}$	Input variable (m-dimensional)
$\mathbf{y}$	Output variable (q-dimensional)
$\mathbf{x}$	State variable (n-dimensional)
$\mathbf{W}(t_0, t_1)$	Controllability Gramian
$\mathbf{W}_T$	$(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})(\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{n-1}\mathbf{B})'$
$\mathbf{M}(t_0, t_1)$	Observability Gramian
$\mathbf{M}_T$	$(\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1})' (\mathbf{C}; \mathbf{CA}; \ldots; \mathbf{CA}^{n-1})$
$\mathbf{T}(t, \sigma)$	Weighting pattern
$[\mathbf{A}, \mathbf{B}, \mathbf{C}]$, $[\mathbf{F}, \mathbf{G}, \mathbf{H}]$	Realizations
s	Laplace transform variable
$\hat{\mathbf{u}}$, $\hat{\mathbf{x}}$, etc.	Laplace transforms
$\mathbf{R}$	Frequency response
$p(s)$	Least common multiple of the denominator of $\hat{R}(s)$
$\mathbf{H}_r$	Hankel matrix

D. Least Squares

$\mathbf{W}(t_0, t_1)$	Controllability Gramian
$\mathbf{K}$	Riccati equation variable
$\mathbf{X}$, $\mathbf{P}$	Hamiltonian equation variables
$\mathbf{\Pi}(t, \mathbf{K}_0, t_0)$	Riccati equation solution
$\mathbf{\Pi}_\infty$	Positive definite solution of $\mathbf{A'K} + \mathbf{KA} - \mathbf{KBB'K} = -\mathbf{L}$
$\mathbf{R}$	Frequency response
$[p(s)]^+$	Left half-plane spectral factor
$Evh(s)$	Even part of $h(s) = \frac{1}{2}[h(s) + h(-s)]$
η	Penalty functional
$\mathbf{x}'(t_1)\mathbf{Q}\mathbf{x}(t_1)$	Terminal penalty
$\int_{t_0}^{t_1} \mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t)\, dt$	Trajectory penalty
$\int_{t_0}^{t_1} \mathbf{u}'(t)\mathbf{u}(t)\, dt$	Control penalty

E. Stability

$\mathbf{F}(t)$	Feedback gain
$\mathbf{Q}$	Solution of $\mathbf{A'K} + \mathbf{KA} - \mathbf{KBB'K} - \mathbf{L} = 0$

$v(\mathbf{x})$	Liapunov function
$\mathbf{R}$	Frequency response
λ_i	Eigenvalues
Γ	Nyquist locus

INDEX