

Tratamento Numérico
de Equações Diferenciais
MAP 5725

Alexandre Megiorin Roma Rudimar Luiz N6s

16 de fevereiro de 2012

Prefácio

Aqui vem o prefácio.

Sumário

Lista de Figuras	vii
Lista de Tabelas	x
Lista de Algoritmos	xii
Lista de Símbolos	xiii
1 Introdução	1
1.1 O Problema de Cauchy	3
1.1.1 Exercícios	7
1.2 Discretização do Problema de Cauchy	8
1.2.1 Exercícios	14
1.3 Suplemento teórico	15
1.3.1 Teorema do Valor Médio	15
1.3.2 Conjunto convexo	15
1.3.3 Problema de valor inicial bem posto	16
1.3.4 Teorema de Picard	17
1.3.5 Teorema Fundamental do Cálculo	17
1.3.6 Série de Taylor de uma função de uma variável	17
1.3.7 Série de Taylor de uma função de duas variáveis	19
1.4 Nota biográfica	21
1.5 Exercícios resolvidos	23
2 Métodos de passo único	29
2.1 Erro local de discretização	30
2.2 Consistência e ordem de consistência	31
2.2.1 Consistência	31
2.2.2 Ordem de consistência	32
2.2.3 Exercícios	33
2.3 Erro global de discretização	33
2.4 Convergência	34
2.5 Expansão do erro global de discretização	37
2.5.1 Estimativa do erro global de discretização	38
2.5.2 Estimativa da ordem de convergência	39
2.5.3 Depuração do código computacional	39
2.5.4 Exercícios	43
2.6 Suplemento teórico	43

2.6.1	Teorema de Weierstrass	43
2.6.2	Teorema do Confronto	43
2.7	Nota biográfica	44
2.8	Exercícios resolvidos	46
3	Métodos de passo único de altas ordens	51
3.1	Métodos da Série de Taylor	51
3.2	Métodos de Runge-Kutta Explícitos	52
3.2.1	Caracterização dos Métodos de Runge-Kutta	52
3.2.2	Métodos de Runge-Kutta explícitos de R -estágios	54
3.2.3	Métodos de Runge-Kutta de ordens superiores	55
3.2.4	Exercícios	56
3.3	Controle do passo de integração	57
3.3.1	Exercícios	57
3.4	Suplemento teórico	59
3.4.1	Regra da cadeia	59
3.5	Nota biográfica	59
3.6	Exercícios resolvidos	61
4	Estabilidade dos métodos de passo único	75
4.1	Estabilidade absoluta	75
4.1.1	Exercícios	78
4.2	Suplemento teórico	79
4.2.1	Instabilidade inerente	79
4.3	Exercícios resolvidos	81
5	Métodos de passo múltiplo lineares	85
5.1	Definição	85
5.2	Dedução	86
5.3	Erro Local de Discretização	87
5.3.1	Expansão em Série de Taylor	88
5.4	Consistência	89
5.4.1	Ordem de consistência	90
5.5	Convergência	91
5.5.1	Condições necessárias à convergência	91
5.5.2	Exercícios	93
5.6	Exercícios resolvidos	95
6	Estabilidade dos métodos de passo múltiplo	103
6.1	Equações de diferenças lineares	103
6.1.1	Exercícios	107
6.2	Divergência	108
6.2.1	Exercícios	110
6.3	Polinômios característicos	112
6.4	Zero-Estabilidade	114
6.4.1	Exercícios	116

6.5	Condições iniciais	118
6.6	Erro global de discretização	118
6.7	Estabilidade absoluta	120
6.7.1	Exercícios	121
6.8	Limitante para o erro local de discretização	122
6.8.1	Exercícios	125
6.9	Limitante para o erro global de discretização	125
6.9.1	Exercícios	127
6.10	Suplemento teórico	127
6.10.1	Teorema do Valor Médio para Integrais	127
6.11	Exercícios resolvidos	128
7	Métodos Preditores-Corretores	143
7.1	Definição	145
7.1.1	Exercícios	145
7.2	Erro local de discretização	148
7.3	Dispositivo de Milne	151
7.4	Estabilidade absoluta	152
7.4.1	Exercícios	156
7.5	Controle do passo de integração	156
7.5.1	Exercícios	156
7.6	Suplemento teórico	158
7.6.1	Teorema do Ponto Fixo	158
7.7	Exercícios resolvidos	159
8	Equações diferenciais rígidas	163
8.1	Definição	163
8.2	Métodos numéricos para equações rígidas	165
8.2.1	Exercícios	166
8.3	Exercícios resolvidos	167
A	Exercícios complementares	169
B	Exercícios computacionais	173
C	Algoritmos	177
D	O Teorema de Picard	179
D.1	Teoremas auxiliares	179
D.2	O Teorema de Picard	179
	Bibliografia	184
	Índice Remissivo	185

Lista de Figuras

1.1	Interpretação geométrica do Método de Euler.	10
2.1	Interpretação geométrica do erro local de discretização.	30
3.1	Interpretação geométrica do Método de Euler Modificado.	71
3.2	Interpretação geométrica do Método de Euler Aprimorado.	72
5.1	Comparação das soluções exata e numérica do Problema de Cauchy (5.17), onde a solução numérica foi obtida através do método (5.16) com (a) $h = 0,1$ e (b) $h = 0,05$	96
5.2	Comparação das soluções exata e numérica do Problema de Cauchy (5.17), onde a solução numérica foi obtida através do método (5.16) com $h = 0,025$	97
5.3	Comportamento do erro na solução numérica do Problema de Cauchy (5.17) obtida através do método (5.16) com refinamento do passo de integração h	98
6.1	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,1$	129
6.2	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,05$	130
6.3	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,025$	131
6.4	Comparação das soluções exata e numérica do problema de valor inicial (6.71), onde a solução numérica foi obtida através do método (6.70) com $h = 0,1$	132
6.5	Comparação das soluções exata e numérica do problema de valor inicial (6.71), onde a solução numérica foi obtida através do método (6.70) com (a) $h = 0,05$ e (b) $h = 0,025$	133

- 6.6 Comparação das soluções exata e numérica do problema de valor inicial (6.74), onde a solução numérica foi obtida através do Método de Runge-Kutta Clássico com $h = 0,2$ e $h = 0,4$ 135

Lista de Tabelas

2.1	Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.35) no instante $t = 1$	41
2.2	Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.36) no instante $t \approx 1,292695719373$ - Butcher [5].	41
2.3	Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.37) no instante $t = 1$ - Butcher [5].	42
4.1	Erros globais de discretização produzidos pelo Método de Euler na solução numérica de (4.1) para dois passos de integração.	76
4.2	Intervalos de estabilidade absoluta para os Métodos de Runge-Kutta de ordem R com R estágios.	78
5.1	Pontos interpolados no Método de Simpson.	86
6.1	Erro global de discretização na solução do problema de valor inicial (6.15) com o método (6.14) e $h = 0,01$	110
6.2	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,1$	129
6.3	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,05$	130
6.4	Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,025$	131
6.5	Comparação das soluções exata e numérica do problema de valor inicial (6.74), onde a solução numérica foi obtida através do Método de Runge-Kutta Clássico com $h = 0,2$ e $h = 0,4$	134
7.1	Características dos Métodos de Adams-Bashforth.	144
7.2	Características dos Métodos de Adams-Moulton.	144
7.3	Polinômios característicos que definem o Método Preditor-Corretor.	158
7.4	Polinômios característicos que definem o Método Preditor-Corretor.	159

8.1	Erro global de discretização do problema de valor inicial (8.3) em $t = \frac{3}{2}$, com $h = 0,1$ e os Métodos de Euler, Runge-Kutta Clássico e Adams-Bashforth-Moulton de quarta ordem - Burden [7].	164
8.2	Solução numérica do sistema rígido (8.4)-(8.5) pelo Método de Runge-Kutta Clássico com $h = 0,05$ - Burden [7].	165
8.3	Solução numérica do sistema rígido (8.4)-(8.5) pelo Método de Runge-Kutta Clássico com $h = 0,1$ - Burden [7].	165
8.4	Erro global de discretização cometido em $t = 1$ na solução do problema de valor inicial (8.10), onde a solução numérica foi obtida através dos Métodos de Runge-Kutta Clássico e do Trapézio com $h = 0,2$ - Burden [7].	167
8.5	Erro global de discretização cometido em $t = 1$ na solução do problema de valor inicial (8.10), onde a solução numérica foi obtida através dos Métodos de Runge-Kutta Clássico e do Trapézio com $h = 0,25$ - Burden [7].	168

Lista de Algoritmos

3.3.1 Controle do passo de integração no Método de Euler Modificado. . .	58
C.0.1 Algoritmo para o Método de Euler.	177
C.0.2 Algoritmo para o Método de Euler Aprimorado.	178

Lista de Símbolos

$\frac{\partial f}{\partial x}$	Derivada parcial de f em relação a x
$\frac{\partial f}{\partial y}$	Derivada parcial de f em relação a y
$y(t_{k+1})$	Solução exata do Problema de Cauchy no instante $k + 1$
y_{k+1}	Aproximação de $y(t_{k+1})$
α	Erro local de discretização
e	Erro global de discretização
C_{p+1}	Constante do erro local de discretização principal

Capítulo 1

Introdução

Descreve-se nesta seção um modelo matemático [4, 8] que pode ser empregado em medicina para diagnosticar o Diabetes (*Diabetes Melito*). O corpo humano utiliza a *glicose*, um tipo de açúcar, como principal fonte de energia. O diabetes é uma desordem pancreática caracterizada pela concentração persistente e crescente de glicose no sangue (hiperglicemia) e na urina (glicosúria) causada pela deficiência de um hormônio produzido pelo pâncreas, a *insulina*.

Os diabetes mais comuns são os do Tipo I, Tipo II e o gestacional. O diabetes do Tipo I, embora possa surgir em qualquer idade, aparece geralmente antes dos 30 anos, inclusive em crianças e adolescentes, e depende de insulina para o seu controle (insulinodependente). Algumas causas possíveis para o seu aparecimento são infecção virótica, dieta inadequada e desordens no sistema imunológico. Mais frequente, o diabetes do Tipo II não depende de insulina para o seu controle e aparece na maturidade ou na velhice, em geral após os 40 anos de idade, tendo como causas a obesidade, a dieta inadequada e a gravidez. O diabetes gestacional afeta sobretudo mulheres gestantes com mais de 30 anos de idade que já tiveram filhos pesando acima de 4kg ao nascer, ou são obesas, ou adquiriram muito peso durante a gestação.

Os sintomas do diabetes são sede excessiva (polidipsia), exacerbada produção urinária (poliúria), desidratação, cansaço e perda de peso. Também podem ocorrer a perda de visão, da sensibilidade nas mãos e pés, câibras e constipação. A hiperglicemia severa e persistente pode levar ao coma. O controle da glicose no sangue é feito por intermédio da administração de drogas ou insulina e pode reduzir a suscetibilidade a algumas complicações associadas a doença tais como a retinopatia.

Pode-se diagnosticar o diabetes empregando o *Teste de Tolerância a Glicose* (GTT). Neste, o indivíduo em jejum ingere uma dose elevada de glicose e, por um período de três a cinco horas, colhem-se amostras sucessivas de sangue que são utilizadas para tabelar a variação da concentração de glicose no sangue ao longo do tempo.

Seja \tilde{g}_0 a concentração de glicose no sangue do indivíduo ao chegar ao laboratório. Se $\tilde{g}(t)$ for a concentração de glicose medida no instante $t \geq t_0$, define-se

$$g(t) = \tilde{g}(t) - \tilde{g}_0 \quad (1.1)$$

como sendo a variação da concentração de glicose no sangue com relação a concentração de chegada, onde t_0 é o instante em que o indivíduo terminou de ingerir completamente a dose de glicose para realizar o exame.

Um modelo matemático [4] que relaciona as variações das concentrações de glicose e de insulina, $g(t)$ e $h(t)$ respectivamente, com relação a seus níveis de chegada é dado pelo sistema de equações

$$\frac{d}{dt}g(t) = -k_1g(t) - k_2h(t), \quad (1.2)$$

$$\frac{d}{dt}h(t) = +k_3g(t) - k_4h(t), \quad (1.3)$$

onde os parâmetros $k_i \geq 0$, $1 \leq i \leq 4$, são determinados de forma a se obter qualitativa e quantitativamente os valores das funções $g(t)$ e $h(t)$ ao longo de um intervalo de tempo (no caso, de 3 a 5 horas). A exemplo de $g(t)$, define-se a variação de concentração de insulina $h(t)$ como

$$h(t) = \tilde{h}(t) - \tilde{h}_0, \quad (1.4)$$

onde \tilde{h}_0 é a concentração desse hormônio no sangue do indivíduo ao chegar ao laboratório. Após uma noite em jejum, assume-se que os níveis de glicose e de insulina, \tilde{g}_0 e \tilde{h}_0 respectivamente, são níveis que expressam o equilíbrio do sistema metabólico.

Quando um aumento na concentração desvia a glicose de seu nível de equilíbrio, parte dela é absorvida pelos tecidos e parte armazenada no fígado, forçando assim um decréscimo em sua concentração no sangue. Similarmente, um aumento na concentração de insulina também acarreta um decréscimo nos níveis da glicose sanguínea pois o hormônio facilita seu armazenamento nos tecidos e no fígado. Este processo é descrito pela equação (1.2). Por outro lado, um aumento na concentração de glicose faz com que mais insulina seja secretada, aumentando a concentração desta na corrente sanguínea. O organismo, ao identificar este aumento na concentração hormonal, tende a fazê-la baixar novamente. Este mecanismo de autorregulação é descrito pela equação (1.3).

O modelo matemático (1.2) e (1.3) é constituído de um sistema de *equações diferenciais ordinárias* (EDOs) lineares de primeira ordem não homogêneas, as quais necessitam de *condições iniciais* $g(t_0) = g_0$ e $h(t_0) = h_0$ para garantir a unicidade

de solução (problema de valor inicial).

Como não existe um procedimento prático simples para medir a concentração de hormônios, elimina-se a função $h(t)$ do modelo matemático. Derivando-se (1.2) e utilizando-se (1.3), chega-se à equação diferencial ordinária linear de segunda ordem homogênea

$$\frac{d^2}{dt^2}g(t) + 2\alpha\frac{d}{dt}g(t) + \omega^2g(t) = 0, \quad (1.5)$$

onde $\alpha = \frac{k_1 + k_4}{2}$ e $\omega^2 = k_1k_4 + k_2k_3$ são parâmetros a serem determinados com o auxílio das medidas efetuadas no teste GTT. Neste caso, as condições iniciais passam a ser dadas por $g(t_0) = g_0$ e $\frac{d}{dt}g(t_0) = \dot{g}_0$, onde $g(t_0)$ e $\frac{d}{dt}g(t_0)$ são, respectivamente, a concentração de glicose no início do teste e a “velocidade” com que esta concentração tende a se modificar inicialmente.

O diagnóstico é dado observando-se o período $T = \frac{2\pi}{\omega}$. Se T for menor que 4 horas, o teste indica normalidade.

Finalmente, se o comportamento da função (1.4) for conhecido ao longo do tempo, uma simplificação do modelo (1.2) e (1.3) é dada por

$$\begin{cases} \dot{y}(t) = \frac{d}{dt}y(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad (1.6)$$

onde $y(t) = g(t)$, $f(t, y(t)) = -k_1y(t) - k_2h(t)$ e $t \in [0, \xi]$, sendo ξ o tempo de duração do teste GTT.

1.1 O Problema de Cauchy

Um *problema de valor inicial (p.v.i.)* definido por uma equação diferencial ordinária de primeira ordem com condição inicial, como (1.6), é chamado *Problema de Cauchy* [1, 5, 10, 18].

O Problema de Cauchy

$$\begin{cases} \dot{y}(t) = \frac{d}{dt}y(t) = f(t, y(t)), & t \in [a, b] \\ y(t_0) = y(a) = y_0 \end{cases} \quad (1.7)$$

Como uma equação diferencial ordinária de ordem superior a um pode, através de uma substituição adequada, ser escrita como um sistema de equações diferenciais ordinárias de primeira ordem, (1.7) também representa um sistema de equações diferenciais. Neste caso, tem-se

$$\dot{y}(t) = [\dot{y}_1(t) \quad \dot{y}_2(t) \quad \cdots \quad \dot{y}_m(t)]^T,$$

$$f(t, y(t)) = \begin{bmatrix} f_1(t, y(t)) \\ f_2(t, y(t)) \\ \vdots \\ f_m(t, y(t)) \end{bmatrix} = \begin{bmatrix} f_1(t, y_1(t), y_2(t), \dots, y_m(t)) \\ f_2(t, y_1(t), y_2(t), \dots, y_m(t)) \\ \vdots \\ f_m(t, y_1(t), y_2(t), \dots, y_m(t)) \end{bmatrix}$$

e

$$\begin{cases} y_1(t_0) = y_{10} \\ y_2(t_0) = y_{20} \\ \vdots \\ y_m(t_0) = y_{m0} \end{cases}.$$

Da teoria de equações diferenciais [2, 15, 18], sabe-se que para uma equação diferencial ordinária de ordem m ou para um sistema de m equações diferenciais ordinárias de primeira ordem são necessárias exatamente m condições iniciais para garantir unicidade de solução. Para outras equações diferenciais, a unicidade de solução é determinada não somente pelas condições iniciais, como também pelas condições de contorno. Neste caso, a solução deve satisfazer a restrições nos extremos do intervalo de integração.

Exemplo 1.1. *Solucione o p.v.i.*

$$\begin{cases} \frac{d}{dt}y(t) = 3ty(t), & t > 0 \\ y(0) = 1 \end{cases}. \quad (1.8)$$

Em (1.8), tem-se uma equação diferencial ordinária linear, de primeira ordem, homogênea.

$$\begin{aligned} \frac{d}{dt}y(t) &= 3ty(t) \\ \frac{1}{y(t)} \frac{d}{dt}y(t) &= 3t \\ \int \frac{1}{y(t)} \frac{d}{dt}y(t) dt &= \int 3t dt \\ \ln|y(t)| &= \frac{3}{2}t^2 + C_1 \\ y(t) &= Ce^{\frac{3}{2}t^2} \end{aligned} \quad (1.9)$$

Usando a condição inicial em (1.9), tem-se que

$$y(0) = C \Rightarrow C = 1.$$

Logo,

$$\boxed{y(t) = e^{\frac{3}{2}t^2}}. \quad (1.10)$$

Exemplo 1.2. Solucione o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = 2y(t) + e^{3t}, & t > 0 \\ y(0) = 1 \end{cases}. \quad (1.11)$$

Em (1.11), tem-se uma equação diferencial ordinária linear, de primeira ordem, não homogênea.

Fator integrante: $e^{\int -2dt} = e^{-2t}$

$$\begin{aligned} e^{-2t} \left[\frac{d}{dt}y(t) - 2y(t) \right] &= e^{-2t}e^{3t} \\ e^{-2t} \frac{d}{dt}y(t) - 2e^{-2t}y(t) &= e^t \\ \frac{d}{dt} [e^{-2t}y(t)] &= e^t \\ \int \frac{d}{dt} [e^{-2t}y(t)] dt &= \int e^t dt \\ e^{-2t}y(t) &= e^t + C \\ y(t) &= e^{3t} + Ce^{2t} \end{aligned} \quad (1.12)$$

Usando a condição inicial em (1.12), tem-se que

$$y(0) = 1 + C \Rightarrow C = 0.$$

Logo,

$$\boxed{y(t) = e^{3t}}. \quad (1.13)$$

Exemplo 1.3. Reescreva o p.v.i.

$$\begin{cases} y''(t) + 2y'(t) - 3y(t) = 0, & t > 0 \\ y(0) = k_1 \\ y'(0) = k_2 \end{cases} \quad (1.14)$$

como um sistema de duas equações diferenciais ordinárias de primeira ordem.

Convenções:

$$y(t) = y_1(t); \quad (1.15)$$

$$y'(t) = y'_1(t) = y_2(t); \quad (1.16)$$

$$y''(t) = y'_2(t). \quad (1.17)$$

Substituindo-se (1.15), (1.16) e (1.17) em (1.14), tem-se que:

$$\begin{cases} y'_1(t) = y_2(t) \\ y'_2(t) = -2y_2(t) + 3y_1(t) \\ y_1(0) = k_1 \\ y_2(0) = k_2 \end{cases} .$$

O Problema de Cauchy (1.7) admite uma única solução $y \in C^1([a, b])$ se $f(t, y)$ for contínua e de Lipschitz.

Teorema 1.1 (Condição de existência e unicidade de solução do Problema de Cauchy). *Seja o problema de valor inicial*

$$\begin{cases} \frac{d}{dt}y(t) = f(t, y(t)) \\ y(a) = y_0 \end{cases}, \quad (1.18)$$

onde $f : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ é contínua na primeira variável e satisfaz a Condição de Lipschitz na segunda variável. Então, existe uma única solução para o p.v.i. (1.18).

A demonstração do teorema de existência e unicidade de solução do Problema de Cauchy pode ser encontrada em Butcher [5].

Definição 1.1 (Condição de Lipschitz). *Uma função $f(t, y)$ satisfaz a Condição de Lipschitz na variável y , $(t, y) \in \Omega \subset \mathbb{R} \times \mathbb{R}^n$, se e somente se existir uma constante $L > 0$ tal que*

$$\|f(t, y_1) - f(t, y_2)\| \leq L\|y_1 - y_2\| \quad (1.19)$$

para quaisquer par de pontos (t, y_1) e (t, y_2) em Ω .

Na Definição 1.1, L é a constante de Lipschitz e pode-se mostrar a desigualdade de Lipschitz (1.19) através do Teorema do Valor Médio [9].

Exemplo 1.4 (Condição de Lipschitz). *Sejam*

$$\begin{aligned} D &= \{(t, y); 1 \leq t \leq 2, -3 \leq y \leq 4\} \text{ e} \\ f(t, y) &= t|y|. \end{aligned}$$

Assim:

$$\begin{aligned} \|f(t, y_1) - f(t, y_2)\| &= \|t|y_1| - t|y_2|\| \\ &= |t| \||y_1| - |y_2|\| \\ &\leq 2\|y_1 - y_2\|. \end{aligned}$$

Portanto, $f(t, y)$ satisfaz a Condição de Lipschitz com constante $L = 2$.

Quando $f(t, y(t))$ é contínua, o Problema de Cauchy (1.7) pode ser colocado em uma *forma integral*, ou seja,

$$y(t) - y(t_0) = \int_{t_0}^t f(s, y(s)) ds. \quad (1.20)$$

A forma (1.20) é obtida aplicando-se o Teorema Fundamental do Cálculo a (1.7).

De agora em diante, o Problema de Cauchy (1.7) será solucionado numericamente e admitir-se-á que o mesmo tem uma única solução.

1.1.1 Exercícios

Exercício 1.1. *Verifique que a solução (1.10) satisfaz de fato o p.v.i. (1.8).*

Exercício 1.2. *Verifique que a solução (1.13) satisfaz de fato o p.v.i. (1.11).*

Exercício 1.3. *Obtenha (1.5) a partir de (1.2) e (1.3).*

Exercício 1.4. *Reescreva o modelo (1.5) como um sistema de equações diferenciais ordinárias de primeira ordem.*

Sugestão: Use $y_1(t) = g(t)$ e $y_2(t) = \frac{d}{dt}g(t)$.

1.2 Discretização do Problema de Cauchy

Com as técnicas matemáticas conhecidas, pode-se determinar a solução exata de um pequeno número de equações diferenciais. Para a esmagadora maioria é possível apenas calcular soluções aproximadas. Contudo, estas não podem ser determinadas em todos os pontos do domínio. Isto é fácil de entender quando se pensa nas restrições impostas pelo uso do computador, tais como aritmética de ponto flutuante (nem todos os números reais podem ser representados) e limitações de tempo e memória.

Ao solucionar numericamente o Problema de Cauchy (1.7), objetiva-se determinar uma aproximação da solução exata em um *conjunto discreto finito de pontos* definido por

$$a = t_0, t_1 = t_0 + \Delta t_1, t_2 = t_1 + \Delta t_2, \dots, b = t_n = t_{n-1} + \Delta t_n,$$

onde cada Δt_k , $1 \leq k \leq n$, denota o *passo de integração* necessário para ir de t_{k-1} a t_k . Adota-se um passo de integração uniforme, isto é,

$$\Delta t_1 = \Delta t_2 = \dots = \Delta t_n = \Delta t = h = \frac{b-a}{n},$$

$$t_k = a + k\Delta t = a + kh,$$

onde n é o número de partições uniformes efetuadas no intervalo $[a, b]$. Uma vez obtida a solução aproximada nos pontos t_k , pode-se empregar as técnicas estudadas em um curso introdutório de Cálculo Numérico [7, 13, 14, 17] (interpolação polinomial e aproximação de funções por splines ou pelo método dos mínimos quadrados) para obter uma aproximação contínua válida em todo o intervalo de estudo.

Para ilustrar uma das possíveis abordagens empregadas na discretização do Problema de Cauchy, considere-se sua forma diferencial (1.7). Supondo que a solução $y(t)$ seja única e que tenha pelo menos m derivadas contínuas, pode-se expandí-la em uma Série de Taylor

$$y(t_k + h) = y(t_k) + hy^{(1)}(t_k) + \frac{h^2}{2!}y^{(2)}(t_k) + \dots + \frac{h^m}{m!}y^{(m)}(\xi), \quad (1.21)$$

para algum ponto ξ , $t_k < \xi < t_{k+1}$. Lembrando que $y(t_k + h) = y(t_{k+1})$, reescreve-se (1.21) como

$$\frac{y(t_{k+1}) - y(t_k)}{h} = y^{(1)}(t_k) + \frac{h}{2!}y^{(2)}(t_k) + \dots + \frac{h^{m-1}}{m!}y^{(m)}(\xi). \quad (1.22)$$

Quando h é suficientemente pequeno tem-se que

$$\frac{y(t_{k+1}) - y(t_k)}{h} \approx y^{(1)}(t_k) = f(t_k, y(t_k)). \quad (1.23)$$

De (1.23) deriva-se o Método de Euler, uma proposta simples de discretização do Problema de Cauchy.

Método de Euler

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, h) \end{cases} \quad (1.24)$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$ e $\Phi(t_k, y_k, h) = f(t_k, y_k)$.

Exemplo 1.5. *Discretize o Problema de Cauchy*

$$\begin{cases} \frac{d}{dt}y(t) = e^{2t}y(t), \quad t \in [0, 1] \\ y(0) = 1 \end{cases}, \quad (1.25)$$

cuja solução exata é

$$y(t) = e^{\frac{e^{2t}-1}{2}},$$

usando o Método de Euler.

$$\begin{cases} y(0) = 1 \\ y_{k+1} = y_k + h e^{2t_k} y_k \\ = (1 + h e^{2t_k}) y_k \end{cases}$$

$$0 \leq k \leq n-1$$

$$k=0 \Rightarrow y_1 = (1 + h e^{2t_0}) y_0$$

$$k=1 \Rightarrow y_2 = (1 + h e^{2t_1}) y_1$$

$$k=2 \Rightarrow y_3 = (1 + h e^{2t_2}) y_2$$

$$\vdots$$

$$y_{k+1} = (1 + h e^{2t_k}) y_k$$

Exemplo 1.6. *Calcule as soluções numérica e exata do p.v.i. (1.25) em $t = 1$. No cálculo da solução numérica, empregue a discretização do Exemplo 1.5 com $h = \frac{1}{2}$.*

$$k=0 \Rightarrow y_1 = \left(1 + \frac{1}{2}e^0\right)(1) = \frac{3}{2}$$

$$k=1 \Rightarrow y_2 = \left(1 + \frac{1}{2}e^1\right)\left(\frac{3}{2}\right) = \frac{3}{4}(e+2) \approx 3,539$$

$$y(1) = e^{\frac{e^2-1}{2}} \approx 24,399$$

No Método de Euler, calcula-se a aproximação y_{k+1} da solução no instante t_{k+1} conhecendo-se apenas o valor da aproximação no instante de tempo t_k imediatamente anterior. Métodos numéricos com esta propriedade são denominados de *métodos de passo único* ou *métodos de um passo* (traduções livres dos termos *single step methods* e *one step methods*, respectivamente). Além disso, o Método de Euler é um *método explícito*, isto é, não é necessário solucionar qualquer tipo de equação algébrica: a aproximação y_{k+1} pode ser calculada diretamente pela soma das parcelas do lado direito de (1.24).

Geometricamente, o Método de Euler (1.24) fornece a aproximação y_{k+1} por meio da reta tangente à curva que define a solução exata. A equação da reta tangente ao gráfico da solução exata no ponto (t_0, y_0) é dada por

$$r(t) - y_0 = (t - t_0)y'(t_0) = (t - t_0)f(t_0, y_0).$$

Quando $t = t_0 + h = t_1$, tem-se que $r(t_0 + h) = y_0 + h f(t_0, y_0) = y_1$. A Figura 1.1 ilustra a interpretação geométrica.

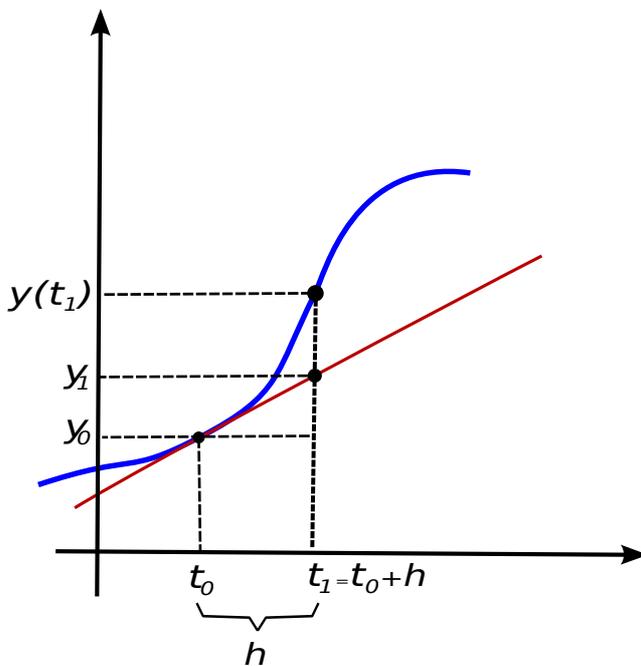


Figura 1.1: Interpretação geométrica do Método de Euler.

Outros três métodos que se enquadram na classe dos métodos de passo único são o Método de Euler Implícito, o Método de Euler Aprimorado e o Método do Trapézio.

Método de Euler Implícito

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_{k+1}, h) \end{cases} \quad (1.26)$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$ e

$$\Phi(t_k, y_{k+1}, h) = f(t_k + h, y_{k+1}).$$

Método de Euler Aprimorado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, h) \end{cases} \quad (1.27)$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$, $\Phi(t_k, y_k, h) = \frac{1}{2}(\kappa_1 + \kappa_2)$
e

$$\begin{cases} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f(t_k + h, y_k + h \kappa_1) \end{cases} .$$

Método do Trapézio

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, y_{k+1}, h) \end{cases} \quad (1.28)$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$ e

$$\Phi(t_k, y_k, y_{k+1}, h) = \frac{1}{2} (f(t_k, y_k) + f(t_k + h, y_{k+1})).$$

O Método de Euler Aprimorado (1.27), assim como o Método de Euler (1.24), é um método explícito. Já nos Métodos de Euler Implícito (1.26) e do Trapézio (1.28), para avançar a solução no tempo é preciso solucionar uma equação algébrica cuja incógnita é y_{k+1} . Por exemplo, se $f(t, y) = \sin(y + \pi t)$, reescreve-se o Método do Trapézio (1.28) como

$$y_{k+1} - \frac{h}{2} \sin(y_{k+1} + \pi t_{k+1}) = y_k + \frac{h}{2} \sin(y_k + \pi t_k). \quad (1.29)$$

Em (1.29), para cada instante de tempo é necessário empregar um método como o de Newton [7, 13, 14, 17] para determinar a raiz y_{k+1} . Um método como o Método de Euler Implícito ou o Método do Trapézio é chamado *método implícito*.

Em uma outra abordagem, empregam-se na discretização do Problema de Cauchy técnicas de quadratura numérica ao invés de expansões em Série de Taylor. Essa abordagem é uma das formas naturais de se obter *métodos de passo múltiplo* (*multistep methods*). Nestes, para determinar y_{k+1} , deve-se conhecer aproximações da solução em p instantes anteriores $y_k, y_{k-1}, y_{k-2}, \dots, y_{k-p+1}$. São por este motivo conhecidos como *métodos de p passos*. Como exemplo, são apresentados os Método de Adams-Bashforth e de Adams-Moulton de quatro passos.

Método de Adams-Bashforth

$$\begin{cases} y_0 = y(t_0), & y_p \text{ pré-determinado}, \quad 1 \leq p \leq 3 \\ y_{k+1} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}) \end{cases} \quad (1.30)$$

com $3 \leq k \leq n-1$, $h = \frac{b-a}{n}$, $f_{k-m} = f(t_{k-m}, y_{k-m})$ e $0 \leq m \leq 3$.

Método de Adams-Moulton

$$\begin{cases} y_0 = y(t_0), & y_p \text{ pré-determinado}, \quad 1 \leq p \leq 3 \\ y_{k+1} = y_k + \frac{h}{720}(251f_{k+1} + 646f_k - 264f_{k-1} + 106f_{k-2} - 19f_{k-3}) \end{cases} \quad (1.31)$$

com $3 \leq k \leq n-1$, $h = \frac{b-a}{n}$, $f_{k-m} = f(t_{k-m}, y_{k-m})$ e $-1 \leq m \leq 3$.

O método (1.30) é apenas um dos muitos métodos pertencentes a classe dos métodos de Adams-Bashforth. Todos eles são explícitos e envolvem um número maior ou menor de passos no tempo. O mesmo se pode dizer do método (1.31), com a diferença de que é um método implícito.

Nos métodos de passo múltiplo são necessárias as aproximações da solução em alguns dos instantes iniciais. Por exemplo, para empregar (1.30) tem que se conhecer as aproximações y_1, y_2 e y_3 , além da condição inicial y_0 . Estas aproximações podem ser determinadas através do emprego de um método de passo único.

O Método de Simpson, outro método de passo múltiplo, pode ser obtido através da forma integral (1.20) do Problema de Cauchy. Considerando o intervalo de tempo

$t_{k-1} \leq t \leq t_{k+1}$, chega-se a

$$y(t_{k+1}) - y(t_{k-1}) = \int_{t_{k-1}}^{t_{k+1}} \frac{d}{ds} y(s) ds = \int_{t_{k-1}}^{t_{k+1}} f(s, y(s)) ds. \quad (1.32)$$

Empregando-se quadratura numérica [7, 13, 14, 17], aproxima-se a integral em (1.32) pelo Método de Simpson. Dessa maneira, integra-se o polinômio de segundo grau $p_2(s)$ que interpola $f(s, y(s))$ nos pontos s_{n-1} , s_n e s_{n+1} . O polinômio $p_2(s)$ é escrito na forma de Lagrange como

$$p_2(s) = f(t_{k+1}, y(t_{k+1}))L_{k+1}(s) + f(t_k, y(t_k))L_k(s) + f(t_{k-1}, y(t_{k-1}))L_{k-1}(s), \quad (1.33)$$

onde $L_j(s)$ são os polinômios de Lagrange

$$L_j(s) = \prod_{i=k-1}^{k+1} \frac{s - t_i}{t_j - t_i}, \quad i \neq j, \quad (1.34)$$

para $j = k - 1$, k e $k + 1$. Substituindo-se (1.34) em (1.33) e admitindo-se um passo de integração constante no tempo, chega-se a

$$\begin{aligned} p_2(s) &= f(t_{k+1}, y(t_{k+1})) \frac{(s - t_{k-1})(s - t_k)}{2h^2} - f(t_k, y(t_k)) \frac{(s - t_{k-1})(s - t_{k+1})}{h^2} \\ &+ f(t_{k-1}, y(t_{k-1})) \frac{(s - t_k)(s - t_{k+1})}{2h^2}, \end{aligned} \quad (1.35)$$

o qual é integrado no intervalo $[t_{k-1}, t_{k+1}]$. Como resultado, tem-se o Método de Simpson

$$y(t_{k+1}) - y(t_{k-1}) \approx \frac{h}{3} [f(t_{k+1}, y(t_{k+1})) + 4f(t_k, y(t_k)) + f(t_{k-1}, y(t_{k-1}))], \quad (1.36)$$

um método de dois passos implícito.

Método de Simpson

$$\begin{cases} y_0 = y(t_0), & y_1 \text{ pré-determinado} \\ y_{k+1} = y_{k-1} + \frac{h}{3}(f_{k+1} + 4f_k + f_{k-1}) \end{cases} \quad (1.37)$$

com $1 \leq k \leq n - 1$, $h = \frac{b-a}{n}$, $f_{k-m} = f(t_{k-m}, y_{k-m})$ e $-1 \leq m \leq 1$.

Além da expansão em Série de Taylor e da quadratura numérica, outras abordagens são possíveis no desenvolvimento de métodos para solucionar numericamente equações diferenciais ordinárias, como por exemplo a extrapolação [14] e a interpolação por splines [10].

Exemplo 1.7. *Discretize o Problema de Cauchy*

$$\begin{cases} \frac{d}{dt}y(t) = \sin(y) \\ y(t_0) = y_0 \end{cases}$$

usando:

1. o Método de Euler Aprimorado (1.27);

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + \frac{h}{2} [\sin(y_k) + \sin(y_k + h \sin(y_k))] \end{cases}$$

2. o Método do Trapézio (1.28).

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + \frac{h}{2} (\sin(y_k) + \sin(y_{k+1})) \end{cases}$$

1.2.1 Exercícios

Exercício 1.5. *Discretize o modelo simplificado (1.6) para o diabetes usando o Método de Euler Aprimorado. Escreva um programa em linguagem C [6, 16] que forneça a solução aproximada g_{k+1} dados t_0 , t_{final} , n e $g(t_0)$. Prepare um arquivo de saída do tipo*

t_0	g_0
t_1	g_1
\vdots	
t_n	g_n

que será utilizado para traçar o gráfico da solução aproximada.

Exercício 1.6. *Discretize o modelo simplificado do diabetes (1.6) utilizando o Método de Adams-Bashforth (1.30).*

Exercício 1.7. *Empregue o Método do Trapézio*

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, y_{k+1}, h) \end{cases}$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$, e

$$\Phi(t_k, y_k, y_{k+1}, h) = \frac{1}{2} (f(t_k, y_k) + f(t_k + h, y_{k+1}))$$

para discretizar o modelo biológico presa-predador definido por

$$\begin{cases} \dot{x} = 1,2x - x^2 - \frac{xy}{x+0,2} \\ \dot{y} = \frac{1,5xy}{x+0,2} - y \end{cases} .$$

1.3 Suplemento teórico

1.3.1 Teorema do Valor Médio

Teorema 1.2. *Seja $f(t)$ uma função que satisfaz as seguintes hipóteses:*

1. $f(t)$ é contínua no intervalo $[a, b]$;
2. $f(t)$ é derivável no intervalo (a, b) .

Então, existe um número $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

ou, de maneira equivalente,

$$f(b) - f(a) = f'(c)(b - a).$$

1.3.2 Conjunto convexo

Definição 1.2. *Um conjunto $D \subset \mathbb{R}^2$ é convexo se para quaisquer pontos (t_1, y_1) e (t_2, y_2) pertencentes a D , o ponto*

$$((1 - \lambda)t_1 + \lambda t_2, (1 - \lambda)y_1 + \lambda y_2)$$

também pertence a D para todo $\lambda \in [0, 1]$.

Teorema 1.3. *Se $f(t, y)$ é definida em um conjunto convexo $D \subset \mathbb{R}^2$ e existe uma constante $L > 0$ tal que*

$$\left| \frac{\partial}{\partial y} f(t, y) \right| \leq L \quad \forall (t, y) \in D,$$

então $f(t, y)$ satisfaz a Condição de Lipschitz em D na variável y com constante de Lipschitz L .

1.3.3 Problema de valor inicial bem posto

Definição 1.3. *O problema de valor inicial*

$$\begin{cases} \frac{d}{dt}y(t) = f(t, y(t)), & t \in [a, b] \\ y(a) = \alpha \end{cases} \quad (1.38)$$

é dito bem posto se:

1. Existe uma única solução $y(t)$;
2. $\forall \epsilon > 0 \exists \kappa(\epsilon) > 0$ com a propriedade de que, sempre que $|\epsilon_0| < \epsilon$ e $\delta(t)$ for contínua com $|\delta(t)| < \epsilon$ em $[a, b]$, a solução única $z(t)$ para o problema

$$\begin{cases} \frac{d}{dt}z(t) = f(t, z(t)) + \delta(t), & t \in [a, b] \\ z(a) = \alpha + \epsilon_0 \end{cases} \quad (1.39)$$

existe com

$$|z(t) - y(t)| < \kappa(\epsilon)\epsilon, \quad \forall t \in [a, b].$$

O problema de valor inicial (1.39) é chamado *problema perturbado* associado ao problema de valor inicial original (1.38).

Teorema 1.4. *Seja $D = \{(t, y); a \leq t \leq b \text{ e } y \in]-\infty, \infty[\}$. Se $f(t, y)$ é uma função contínua que satisfaz a Condição de Lipschitz na variável y em D , então o problema de valor inicial*

$$\begin{cases} \frac{d}{dt}y(t) = f(t, y(t)) \\ y(a) = \alpha \end{cases}$$

é bem posto.

1.3.4 Teorema de Picard

Teorema 1.5. *Seja $f(t, y)$ contínua e de Lipschitz em $\Omega = I_a \times B_b$, onde $I_a = \{t; |t - t_0| \leq a\}$ e $B_b = \{y; |y - y_0| \leq b\}$. Se $\|f(t, y)\| \leq M$ em Ω , então existe uma única solução de*

$$\begin{cases} \dot{y}(t) = \frac{d}{dt}y(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases}$$

em I_α , onde $\alpha = \min\{a, \frac{b}{M}\}$.

A demonstração do Teorema de Picard pode ser encontrada em Doering [2], Sotomayor [18] ou no Apêndice D.

1.3.5 Teorema Fundamental do Cálculo

Teorema 1.6. *Se $f(x)$ é contínua em $[a, b]$, então*

$$\int_a^b f(x)dx = F(b) - F(a),$$

onde $F(x)$ é qualquer primitiva de $f(x)$, isto é, $F(x)$ é uma função tal que

$$F'(x) = f(x).$$

1.3.6 Série de Taylor de uma função de uma variável

Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função de classe C^{n+1} . A Série de Taylor de $f(t)$ com centro em $t = t_0$ ($t - t_0 = h$) é definida como

$$f(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(t_0)}{k!} (t - t_0)^k, \quad |t - t_0| < R, \quad (1.40)$$

onde R é o raio de convergência da série.

Truncando-se a série (1.40) em n , pode-se escrever a função $f(t)$ como

$$f(t) = p_n(t) + R_n(t). \quad (1.41)$$

Em (1.41), $p_n(t)$ é o polinômio de Taylor de grau n , enquanto $R_n(t)$ é o resto da expansão, ou seja,

$$p_n(t) = \sum_{k=0}^n \frac{f^{(k)}(t_0)}{k!} (t - t_0)^k \text{ e } R_n(t) = f(t) - p_n(t) = \frac{f^{(n+1)}(\varepsilon)}{(n+1)!} (t - t_0)^{n+1},$$

com $\varepsilon \in (t_0, t)$.

Considerando-se $n = 0$ em (1.41), tem-se que

$$f(t) = f(t_0) + R_0(t). \quad (1.42)$$

Como

$$R_0(t) = f'(\varepsilon)(t - t_0),$$

(1.42) é reescrita como

$$f(t) - f(t_0) = f'(\varepsilon)(t - t_0), \quad (1.43)$$

exatamente o que afirma o Teorema do Valor Médio.

A Série de Taylor tem muitas aplicações, como por exemplo, no desenvolvimento de técnicas de solução numérica de equações diferenciais ordinárias e equações diferenciais parciais.

Exemplo 1.8. Calcule o valor de uma função $f(t)$ de classe $C^\infty((-\infty, \infty))$ em $t + h$.

Para tanto, basta expandir $f(t)$ em Série de Taylor e truncá-la em um n determinado.

$$1. \quad f(t+h) = \underbrace{f(t) + hf^{(1)}(t)}_{p_1(t)} + \underbrace{\frac{h^2}{2}f^{(2)}(\varepsilon)}_{R_1(t)}$$

$$2. \quad f(t+h) = \underbrace{f(t) + hf^{(1)}(t) + \frac{h^2}{2}f^{(2)}(t)}_{p_2(t)} + \underbrace{\frac{h^3}{6}f^{(3)}(\varepsilon)}_{R_2(t)}$$

$$3. \quad f(t+h) = \underbrace{f(t) + hf^{(1)}(t) + \frac{h^2}{2}f^{(2)}(t) + \frac{h^3}{6}f^{(3)}(t)}_{p_3(t)} + \underbrace{\frac{h^4}{24}f^{(4)}(\varepsilon)}_{R_3(t)}$$

4.

$$f(t+h) = \underbrace{f(t) + hf^{(1)}(t) + \frac{h^2}{2}f^{(2)}(t) + \dots + \frac{h^{10}}{10!}f^{(10)}(t)}_{p_{10}(t)} + \underbrace{\frac{h^{11}}{10!}f^{(11)}(\varepsilon)}_{R_{10}(t)}$$

5.

$$f(t+h) = \underbrace{f(t) + hf^{(1)}(t) + \frac{h^2}{2}f^{(2)}(t) + \dots + \frac{h^{100}}{100!}f^{(100)}(t)}_{p_{100}(t)} + \underbrace{\frac{h^{101}}{100!}f^{(101)}(\varepsilon)}_{R_{100}(t)}$$

1.3.7 Série de Taylor de uma função de duas variáveis

Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função de classe C^{n+1} . Pode-se expandir $f(t, y)$ em uma Série de Taylor com centro em $(t, y) = (t_0, y_0)$ ($t - t_0 = h$ e $y - y_0 = k$) em duas etapas. Primeiramente, expande-se na direção t e, em seguida, na direção y .

Expandindo-se na direção t :

$$\begin{aligned} f(t_0 + h, y_0 + k) &= \underbrace{f(t_0, y_0 + k)}_I + \\ &+ \underbrace{h f_t(t_0, y_0 + k)}_{II} + \\ &+ \frac{h^2}{2!} \underbrace{f_{tt}(t_0, y_0 + k)}_{III} + \\ &+ \frac{h^3}{3!} \underbrace{f_{ttt}(\varepsilon_1, y_0 + k)}_{IV}. \end{aligned} \quad (1.44)$$

Expandindo-se cada termo (I, II, III e IV) na direção y :

$$\begin{aligned} f(t_0, y_0 + k) &= f(t_0, y_0) + k f_y(t_0, y_0) + \frac{k^2}{2!} f_{yy}(t_0, y_0) + \\ &+ \frac{k^3}{3!} f_{yyy}(t_0, \varepsilon_2); \\ f_t(t_0, y_0 + k) &= f_t(t_0, y_0) + k f_{ty}(t_0, y_0) + \frac{k^2}{2!} f_{tyy}(t_0, y_0) + \\ &+ \frac{k^3}{3!} f_{tyyy}(t_0, \varepsilon_3); \\ f_{tt}(t_0, y_0 + k) &= f_{tt}(t_0, y_0) + k f_{tty}(t_0, y_0) + \frac{k^2}{2!} f_{ttyy}(t_0, y_0) + \\ &+ \frac{k^3}{3!} f_{ttyyy}(t_0, \varepsilon_4); \end{aligned}$$

$$\begin{aligned}
f_{ttt}(\varepsilon_1, y_0 + k) &= f_{ttt}(\varepsilon_1, y_0) + kf_{ttty}(\varepsilon_1, y_0) + \frac{k^2}{2!}f_{tttyy}(\varepsilon_1, y_0) + \\
&+ \frac{k^3}{3!}f_{tttyyy}(\varepsilon_1, \varepsilon_5).
\end{aligned}$$

Substituindo-se a expansão de cada termo (I,II,III e IV) em (1.44):

$$\begin{aligned}
f(t_0 + h, y_0 + k) &= f(t_0, y_0) + kf_y(t_0, y_0) + \\
&+ \frac{k^2}{2!}f_{yy}(t_0, y_0) + \frac{k^3}{3!}f_{yyy}(t_0, \varepsilon_2) + \\
&+ h \left[f_t(t_0, y_0) + kf_{ty}(t_0, y_0) + \frac{k^2}{2!}f_{tyy}(t_0, y_0) + \frac{k^3}{3!}f_{tyyy}(t_0, \varepsilon_3) \right] + \\
&+ \frac{h^2}{2!} \left[f_{tt}(t_0, y_0) + kf_{tty}(t_0, y_0) + \frac{k^2}{2!}f_{ttyy}(t_0, y_0) + \frac{k^3}{3!}f_{ttyyy}(t_0, \varepsilon_4) \right] + \\
&+ \frac{h^3}{3!} \left[f_{ttt}(\varepsilon_1, y_0) + kf_{ttty}(\varepsilon_1, y_0) + \frac{k^2}{2!}f_{tttyy}(\varepsilon_1, y_0) + \frac{k^3}{3!}f_{tttyyy}(\varepsilon_1, \varepsilon_5) \right];
\end{aligned}$$

$$\begin{aligned}
f(t_0 + h, y_0 + k) &= f(t_0, y_0) + hf_t(t_0, y_0) + kf_y(t_0, y_0) + \\
&+ \frac{h^2}{2!}f_{tt}(t_0, y_0) + \frac{2hk}{2!}f_{ty}(t_0, y_0) + \frac{k^2}{2!}f_{yy}(t_0, y_0) + \\
&+ \frac{h^3}{3!}f_{ttt}(t_0, y_0) + \frac{3h^2k}{3!}f_{tty}(t_0, y_0) + \frac{3hk^2}{3!}f_{tyy}(t_0, y_0) + \frac{k^3}{3!}f_{yyy}(t_0, y_0) + \\
&+ \frac{h^4}{4!}f_{tttt}(t_0, y_0) + \frac{4h^3k}{4!}f_{itty}(t_0, y_0) + \frac{6h^2k^2}{4!}f_{ttyy}(t_0, y_0) + \\
&+ \frac{4hk^3}{4!}f_{tyyy}(t_0, y_0) + \frac{k^4}{4!}f_{yyyy}(t_0, y_0) + \\
&+ R(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5).
\end{aligned}$$

Uma generalização para a expansão em Série de Taylor em torno do ponto (t_0, y_0) é dada por

$$f(t, y) = \sum_{n=0}^{\infty} \left[\frac{1}{n!} \sum_{i=0}^n C_i^n h^i k^{n-i} \frac{\partial^n}{\partial t^i \partial y^{n-i}} f(t_0, y_0) \right],$$

com $n = 0, 1, 2, \dots$, $C_i^n = \frac{n!}{i!(n-i)!}$, $t - t_0 = h = \Delta t$ e $y - y_0 = k = \Delta y$.

1.4 Nota biográfica



Augustin Louis Cauchy¹ (1789-1857)

Nascido em Paris em 21 de Agosto de 1789, Cauchy teve uma infância conturbada por eventos ligados à Revolução Francesa. Buscando um lugar mais seguro para viver, seu pai mudou-se para Arcueil quando Cauchy tinha quatro anos. Entretanto, a escassez de alimentos o forçou a retornar a Paris pouco tempo depois.

Pierre Simon Laplace e Joseph-Louis Lagrange eram amigos da família e visitavam com frequência sua casa. Ao que parece, Lagrange se interessou pela educação matemática de Cauchy, aconselhando seu pai a inscrevê-lo primeiro na École Centrale du Panthéon (1802), onde Cauchy estudou por dois anos os idiomas clássicos. Em 1804, Cauchy teve aulas de Matemática e, em 1805, fez o exame de admissão para a École Polytechnique, tendo sido examinado por Jean-Baptiste Biot.

Na École Polytechnique, Cauchy teve André-Marie Ampère como instrutor de análise. Graduou-se em 1807 e foi estudar engenharia na École des Ponts et Chaussées. Aluno excepcional, como trabalho prático de seu curso, Cauchy foi incumbido de auxiliar no projeto da construção do Canal de Ourcq, trabalhando sob a orientação de Pierre Girard.

Seu primeiro emprego foi nas instalações portuárias de Cherbourg, a serviço da frota de Napoleão (1810). Apesar do trabalho ser extenuante, Cauchy dedicou-se simultaneamente à leitura dos livros *Mécanique Céleste*, de Laplace, e *Théorie des Fonctions*, de Lagrange, partindo em seguida para a pesquisa em matemática. Encorajado por Adrien-Marie Legendre, escreveu e publicou trabalhos na área de geometria em 1811 e 1812.

Cauchy adoece e sob licença médica volta a Paris, onde submeteu um trabalho sobre funções simétricas no jornal da École Polytechnique em fins de 1812 (publicado em 1815). Com o término de sua licença médica, Cauchy retorna a contragosto para seu trabalho em Cherbourg, o qual ele abandona pouco tempo depois. Retorna a Paris onde tenta, sem sucesso, encontrar uma posição acadêmica. Consegue, entretanto, trabalhar como engenheiro no projeto do Canal de Ourcq junto com Pierre

¹J.J. O'Connor e E.F. Robertson, *MacTutor History of Mathematics*
<http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Cauchy.html>
 Carl B. Boyer, *História da Matemática* [3]

Girard.

Ingressa em 1815 na Académie de Sciences após ter resolvido uma das conjecturas de Pierre de Fermat. Foi apenas em 1817 que ele consegue uma posição no Collège de France quando Biot se afasta para participar de uma expedição à Escócia. Cauchy foi o primeiro a estudar rigorosamente a convergência de séries infinitas e a definir o conceito de integral. Escreveu o texto *Cours d'analyse* em 1821 para os estudantes da École Polytechnique, o qual continha os teoremas básicos do cálculo com demonstrações rigorosas. Iniciou estudos sobre o cálculo de resíduos em 1826, tendo introduzido o conceito de função de uma variável complexa em 1829.

Cauchy tinha uma personalidade difícil e não se dava particularmente bem com seus colegas cientistas. Com frequência trazia sua fé religiosa, o catolicismo, para o ambiente de trabalho. Em uma ocasião, tomou o partido dos jesuítas colocando-se contra a Académie de Sciences.

Em 1830, Cauchy pede licença e se afasta de seu trabalho viajando para Turin e depois Praga, onde passa a tutorar o neto de Charles X. Retorna a Paris em 1838 e, devido a alterações políticas, ele reassume sua posição na Académie de Sciences mas não suas posições de professor, uma vez que o regime político vigente exigiu dele um juramento de fidelidade, o qual Cauchy se recusou a fazer.

Sua produtividade caiu nos anos subsequentes, embora ele tenha desenvolvido importantes trabalhos em equações diferenciais e em física matemática. Foi apenas em 1848, com uma outra mudança política, que Cauchy reassumiu suas posições acadêmicas.

Os últimos anos da vida de Cauchy foram marcados por disputas acadêmicas com Joseph Liouville e científicas com Jean-Marie Duhamel, o que o isolaram ainda mais junto aos colegas e o amarguraram até o fim.

Cauchy faleceu em Sceaux, próximo de Paris, em 23 de maio de 1857, deixando um legado científico vasto com mais de 780 trabalhos em revistas e uma obra publicada em 27 volumes. Foi pioneiro no estudo de análise e teoria da permutação de grupos. Desenvolveu pesquisas em convergência de séries infinitas, equações diferenciais, determinantes, probabilidade, física matemática e análise complexa, entre outras áreas.

1.5 Exercícios resolvidos

1. Mostre que o Método de Euler

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{cases}$$

com

$$\Phi(t_k, y_k, h) = f(t_k, y_k) \quad \text{e} \quad h = \frac{b-a}{n}$$

é um método linear explícito e de passo único.

Solução

$$y_{k+1} = y_k + h\Phi(t_k, y_k) \Rightarrow y_{k+1} = y_k + hf_k \Rightarrow \underbrace{y_{k+1} - y_k - hf_k}_{F(y_k, y_{k+1}, f_k)} = 0$$

$$F(y_k, y_{k+1}, f_k) = y_{k+1} - y_k - hf_k$$

O operador F acima define o Método de Euler. Para mostrar que o método é linear, basta mostrar que o operador F é linear. Assim sendo, deve-se verificar que

$$F(y_k + \alpha\tilde{y}_k, y_{k+1} + \alpha\tilde{y}_{k+1}, f_k + \alpha\tilde{f}_k) = F(y_k, y_{k+1}, f_k) + \alpha F(\tilde{y}_k, \tilde{y}_{k+1}, \tilde{f}_k).$$

$\begin{aligned} F(y_k + \alpha\tilde{y}_k, y_{k+1} + \alpha\tilde{y}_{k+1}, f_k + \alpha\tilde{f}_k) &= y_{k+1} + \alpha\tilde{y}_{k+1} + y_k + \\ &+ \alpha\tilde{y}_k + f_k + \alpha\tilde{f}_k \\ &= y_{k+1} - y_k - hf_k + \\ &+ \alpha\tilde{y}_{k+1} - \alpha\tilde{y}_k - \alpha\tilde{f}_k \\ &= y_{k+1} - y_k - hf_k + \\ &+ \alpha(\tilde{y}_{k+1} - \tilde{y}_k - \tilde{f}_k) \\ &= F(y_k, y_{k+1}, f_k) + \\ &+ \alpha F(\tilde{y}_k, \tilde{y}_{k+1}, \tilde{f}_k). \end{aligned}$

Portanto, o Método de Euler é linear.

Com relação ao Método de Euler, tem-se ainda que:

- O método é explícito porque para calcular y_{k+1} utilizam-se apenas as informações da k -ésima iteração, no caso y_k , h e $f(t_k, y_k)$;

- É de passo único pois para calcular a aproximação no instante t_{k+1} são necessárias somente as informações do instante imediatamente anterior t_k .

2. Seja

$$\ddot{x}(t) + 0,12\dot{x}(t) + 2x(t) = 0 \quad (1.45)$$

com condições iniciais $x(0) = 1$ e $\dot{x}(0) = 0$.

- Reescreva a EDO de segunda ordem como um sistema de EDOs de primeira ordem;
- Discretize o sistema obtido no item (a) usando o Método de Euler.

Solução

- Convenções adotadas:

$$y_1(t) = x(t) \Rightarrow \dot{y}_1(t) = \dot{x}(t), \quad (1.46)$$

$$y_2(t) = \dot{y}_1(t) = \dot{x}(t) \Rightarrow \dot{y}_2(t) = \ddot{x}(t). \quad (1.47)$$

Substituindo-se (1.46) e (1.47) em (1.45), obtém-se:

$$\dot{y}_2(t) + 0,12y_2(t) + 2y_1(t) = 0 \Rightarrow \dot{y}_2(t) = -0,12y_2(t) - 2y_1(t). \quad (1.48)$$

Assim, tem-se o sistema de EDOs a seguir.

$$\boxed{\begin{cases} \dot{y}_1(t) = y_2(t) \\ \dot{y}_2(t) = -0,12y_2(t) - 2y_1(t) \end{cases}} \quad (1.49)$$

Como

$$x(0) = 1 \Rightarrow y_1(0) = 1 \text{ e } \dot{x}(0) = 0 \Rightarrow y_2(0) = 0, \quad (1.50)$$

os dados iniciais do sistema (1.49) são

$$\boxed{\begin{cases} y_1(0) = 1 \\ y_2(0) = 0 \end{cases}} \quad (1.51)$$

(b) Discretização do sistema de EDOs pelo Método de Euler.

$$\begin{bmatrix} y_1^{k+1} \\ y_2^{k+1} \end{bmatrix} = \begin{bmatrix} y_1^k \\ y_2^k \end{bmatrix} + h \begin{bmatrix} y_2^k \\ -0, 12y_2^k - 2y_1^k \end{bmatrix}$$

ou

$$\begin{bmatrix} y_{1,k+1} \\ y_{2,k+1} \end{bmatrix} = \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix} + h \begin{bmatrix} y_{2,k} \\ -0, 12y_{2,k} - 2y_{1,k} \end{bmatrix}$$

3. Obtenha o Método do Trapézio

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, y_{k+1}, h), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{cases}$$

$$\Phi(t_k, y_k, y_{k+1}, h) = \frac{1}{2} (f(t_k, y_k) + f(t_{k+1}, y_{k+1}))$$

e o Método de Simpson

$$\begin{cases} y_0 = y(t_0) \\ y_{k+2} = y_k + \frac{h}{3}(f_{k+2} + 4f_{k+1} + f_k), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{cases}$$

$$f_{k+i} = f(t_{k+i}, y_{k+i}), \quad i = 0, 1, 2$$

utilizando a quadratura numérica.

Solução

(a) Método do Trapézio

$$\frac{d}{dt}y(t) = f(t, y)$$

$$\int_{t_k}^{t_{k+1}} \frac{d}{ds}y(s)ds = \int_{t_k}^{t_{k+1}} f(s, y)ds \Rightarrow [y(s)]_{t_k}^{t_{k+1}} = \int_{t_k}^{t_{k+1}} f(s, y)ds$$

$$y(t_{k+1}) - y(t_k) = \int_{t_k}^{t_{k+1}} f(s, y)ds$$

$$y_{k+1} - y_k = \int_{t_k}^{t_{k+1}} f(s, y)ds$$

$$\boxed{y_{k+1} = y_k + \int_{t_k}^{t_{k+1}} f(s, y)ds} \quad (1.52)$$

Inicialmente, aproxima-se $f(s, y)$ pelo polinômio interpolador de Lagrange de primeiro grau nos pontos t_k e t_{k+1} . Depois disso, integra-se esse polinômio no intervalo $[t_k, t_{k+1}]$.

Pontos de interpolação

t_k	t_{k+1}
$f(t_k, y(t_k))$	$f(t_{k+1}, y(t_{k+1}))$

Para simplificar, denota-se f_k por $f(t_k, y(t_k))$.

$$\begin{aligned} p_1(s) &= f_k L_k(s) + f_{k+1} L_{k+1}(s) \\ &= f_k \frac{s - t_{k+1}}{t_k - t_{k+1}} + f_{k+1} \frac{s - t_k}{t_{k+1} - t_k} \end{aligned}$$

$$p_1(s) = -\frac{f_k}{h}(s - t_{k+1}) + \frac{f_{k+1}}{h}(s - t_k)$$

$$\begin{aligned} \int_{t_k}^{t_{k+1}} p_1(s) ds &= -\frac{f_k}{h} \int_{t_k}^{t_{k+1}} (s - t_{k+1}) ds + \\ &+ \frac{f_{k+1}}{h} \int_{t_k}^{t_{k+1}} (s - t_k) ds \end{aligned}$$

$$\begin{aligned} \int_{t_k}^{t_{k+1}} p_1(s) ds &= -\frac{f_k}{h} \left[\frac{s^2 - 2st_{k+1}}{2} \right]_{t_k}^{t_{k+1}} + \\ &+ \frac{f_{k+1}}{h} \left[\frac{s^2 - 2st_k}{2} \right]_{t_k}^{t_{k+1}} \end{aligned}$$

$$\begin{aligned} \int_{t_k}^{t_{k+1}} p_1(s) ds &= \frac{f_k}{h} \left[\frac{t_{k+1}^2 - 2t_k t_{k+1} + t_k^2}{2} \right] + \\ &+ \frac{f_{k+1}}{h} \left[\frac{t_{k+1}^2 - 2t_k t_{k+1} + t_k^2}{2} \right] \end{aligned}$$

$$\int_{t_k}^{t_{k+1}} p_1(s) ds = \frac{f_k}{2h}(t_{k+1} - t_k)^2 + \frac{f_{k+1}}{2h}(t_{k+1} - t_k)^2$$

$$\int_{t_k}^{t_{k+1}} p_1(s) ds = \frac{f_k}{2h}h^2 + \frac{f_{k+1}}{2h}h^2$$

$$\int_{t_k}^{t_{k+1}} p_1(s) ds = h \left[\frac{1}{2}(f_k + f_{k+1}) \right]$$

Substituindo-se o resultado acima (a integral de $p_1(s)$ no intervalo $[t_k, t_{k+1}]$) em (1.52), tem-se que:

$$\begin{aligned}
 y_{k+1} &= y_k + h \left[\frac{1}{2} (f_k + f_{k+1}) \right] \\
 &= y_k + h \underbrace{\left[\frac{1}{2} (f(t_k, y_k) + f(t_{k+1}, y_{k+1})) \right]}_{\Phi(t_k, y_k, y_{k+1}, h)} \\
 &= y_k + h\Phi(t_k, y_k, y_{k+1}, h).
 \end{aligned}$$

Portanto, o método do Trapézio é definido por

$$y_{k+1} = y_k + h\Phi(t_k, y_k, y_{k+1}, h)$$

onde

$$\Phi(t_k, y_k, y_{k+1}, h) = \frac{1}{2} (f(t_k, y_k) + f(t_{k+1}, y_{k+1})).$$

(b) Método de Simpson

$$\begin{aligned}
 \frac{d}{dt}y(t) &= f(t, y) \\
 \int_{t_k}^{t_{k+2}} \frac{d}{ds}y(s)ds &= \int_{t_k}^{t_{k+2}} f(s, y)ds \Rightarrow [y(s)]_{t_k}^{t_{k+2}} = \int_{t_k}^{t_{k+2}} f(s, y)ds \\
 y(t_{k+2}) - y(t_k) &= \int_{t_k}^{t_{k+2}} f(s, y)ds \\
 y_{k+2} - y_k &= \int_{t_k}^{t_{k+2}} f(s, y)ds \\
 \boxed{y_{k+2} = y_k + \int_{t_k}^{t_{k+2}} f(s, y)ds} & \quad (1.53)
 \end{aligned}$$

Aproxima-se $f(s, y)$ em (1.53) pelo polinômio interpolador de Lagrange de grau 2 nos pontos t_k , t_{k+1} e t_{k+2} , sendo estes igualmente espaçados. Feito isso, integra-se no intervalo $[t_k, t_{k+2}]$ o polinômio obtido.

Pontos de interpolação

t_k	t_{k+1}	t_{k+2}
f_k	f_{k+1}	f_{k+2}

$$\begin{aligned}
p_2(s) &= f_k L_k(s) + f_{k+1} L_{k+1}(s) + f_{k+2} L_{k+2}(s) \\
&= f_k \frac{(s - t_{k+1})(s - t_{k+2})}{(t_k - t_{k+1})(t_k - t_{k+2})} + \\
&\quad + f_{k+1} \frac{(s - t_k)(s - t_{k+2})}{(t_{k+1} - t_k)(t_{k+1} - t_{k+2})} + \\
&\quad + f_{k+2} \frac{(s - t_k)(s - t_{k+1})}{(t_{k+2} - t_k)(t_{k+2} - t_{k+1})} \\
&= \frac{f_k}{2h^2} (s - t_{k+1})(s - t_{k+2}) + \\
&\quad - \frac{f_{k+1}}{h^2} (s - t_k)(s - t_{k+2}) + \\
&\quad + \frac{f_{k+2}}{2h^2} (s - t_k)(s - t_{k+1})
\end{aligned}$$

$$\begin{aligned}
\int_{t_k}^{t_{k+2}} p_2(s) ds &= \frac{f_k}{2h^2} \int_{t_k}^{t_{k+2}} (s - t_{k+1})(s - t_{k+2}) ds + \\
&\quad - \frac{f_{k+1}}{h^2} \int_{t_k}^{t_{k+2}} (s - t_k)(s - t_{k+2}) ds + \\
&\quad + \frac{f_{k+2}}{2h^2} \int_{t_k}^{t_{k+2}} (s - t_k)(s - t_{k+1}) ds \quad (1.54)
\end{aligned}$$

$$\begin{aligned}
\int_{t_k}^{t_{k+2}} p_2(s) ds &= \frac{f_k}{2h^2} \frac{2h^3}{3} + \frac{f_{k+1}}{h^2} \frac{4h^3}{3} + \frac{f_{k+2}}{2h^2} \frac{2h^3}{3} \quad (1.55) \\
&= \frac{f_k h}{3} + \frac{4f_{k+1} h}{3} + \frac{f_{k+2} h}{3}
\end{aligned}$$

$$= \frac{h}{3} (f_k + 4f_{k+1} + f_{k+2}) \quad (1.56)$$

Observação: A passagem de (1.54) para (1.55) não é trivial. Ela exige considerável manipulação algébrica.

Substituindo-se (1.56) em (1.53), obtém-se o Método de Simpson

$$y_{k+2} = y_k + \frac{h}{3} (f_k + 4f_{k+1} + f_{k+2})$$

Capítulo 2

Métodos de passo único

De modo geral, um *método de passo único* (ou *método de um passo*) para o Problema de Cauchy (1.7) tem a forma abaixo.

Método de passo único

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, t_{k+1}, y_k, y_{k+1}, h) \end{cases} \quad (2.1)$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$ e $h = \frac{b-a}{n}$.

Se em (2.1) a função Φ não depende de nenhuma informação no instante t_{k+1} , então (2.1) é um método de passo único *explícito*. Caso contrário, diz-se que o método de passo único (2.1) é *implícito*. Como exemplo de métodos de passo único explícito e implícito, respectivamente, citam-se os métodos de Euler (1.24) e do Trapézio (1.28).

Embora as definições e resultados teóricos a serem apresentados a seguir também sejam válidos para métodos implícitos, por simplicidade, eles serão colocados apenas no contexto de métodos explícitos. Para tanto, considere-se um método de passo único explícito reescrito como

$$\begin{cases} y_0 = y(t_0) \\ \frac{y_{k+1} - y_k}{h} - \Phi(t_k, y_k, h) = 0 \end{cases} \quad (2.2)$$

2.1 Erro local de discretização

Definição 2.1 (Erro local de discretização). *Dado o método numérico de passo único, o erro local de discretização é definido por*

$$\alpha_k = \frac{y(t_{k+1}) - y(t_k)}{h} - \Phi(t_k, y(t_k), h). \quad (2.3)$$

Observe que na definição do erro local de discretização (2.3), utiliza-se a solução exata $y(t)$ do Problema de Cauchy e não uma aproximação como é feito em (2.2). Multiplicando-se (2.3) pelo passo de integração h chega-se a

$$h \alpha_k = d_k = y(t_{k+1}) - [y(t_k) + h \Phi(t_k, y(t_k), h)] = y(t_{k+1}) - y_{k+1}. \quad (2.4)$$

Em (2.4), a diferença entre a solução exata $y(t_{k+1})$ e a solução aproximada y_{k+1} pode ser interpretada como sendo o erro produzido em uma única aplicação do método numérico. O caráter local é conferido ao se partir da solução exata no instante anterior, o que garante que nenhum outro tipo de erro tenha sido cometido anteriormente. Na Figura (2.1), pode-se visualizar a interpretação geométrica do erro local de discretização $d_k = h\alpha_k = y(t_{k+1}) - y_{k+1}$.

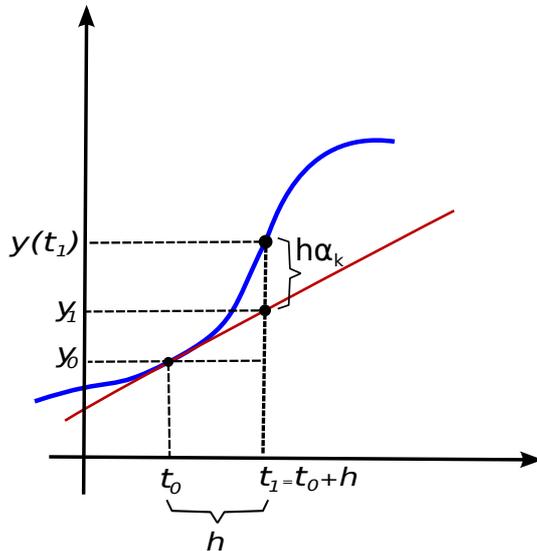


Figura 2.1: Interpretação geométrica do erro local de discretização.

2.2 Consistência e ordem de consistência

2.2.1 Consistência

Supondo a continuidade de Φ como função do passo de integração h e lembrando que $y(t)$ é a solução do Problema de Cauchy (1.7), o limite do erro local de discretização (2.3) quando o passo de integração tende a zero é

$$\begin{aligned}\lim_{h \rightarrow 0} \alpha_k &= \lim_{h \rightarrow 0} \left[\frac{y(t_k + h) - y(t_k)}{h} - \Phi(t_k, y(t_k), h) \right] \\ &= y'(t) - \Phi(t, y(t), 0) \\ &= f(t, y(t)) - \Phi(t, y(t), 0).\end{aligned}$$

O limite anterior deve ser entendido como tendo $hk = t - t_0$ fixado. É razoável esperar/impor que $\lim_{h \rightarrow 0} \alpha_k = 0$, o que implica $\Phi(t, y, 0) = f(t, y)$ desde que $\Phi(t_k, y(t_k), h)$ seja contínua.

Definição 2.2 (Consistência). *Um método de passo único é consistente com a equação diferencial ordinária (ou simplesmente consistente) se e somente se*

$$\Phi(t, y, 0) = f(t, y)$$

ou, equivalentemente, se e somente se

$$\lim_{h \rightarrow 0} \alpha_k = 0. \quad (2.5)$$

Exemplo 2.1. *Verifique que o Método de Euler Aprimorado é consistente.*

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + \frac{h}{2} [f(t_k, y_k) + f(t_{k+1}, y_k + hf(t_k, y_k))] \end{cases}$$

$$\begin{aligned}\Phi(t, y, h) &= \frac{1}{2} [f(t, y) + f(t + h, y + hf(t, y))] \\ \Phi(t, y, 0) &= \frac{1}{2} [f(t, y) + f(t, y)] = f(t, y)\end{aligned} \quad (2.6)$$

Note-se que, por hipótese, $f(t, y)$ é contínua e Lipschitziana. Assim, $\Phi(t, y, h)$ é contínua no terceiro argumento pois é dada por uma soma de funções contínuas (a segunda parcela do lado direito de (2.6) é uma função composta de função contínua, portanto contínua).

2.2.2 Ordem de consistência

Definição 2.3 (Ordem de consistência). *Se existirem constantes positivas C , h_0 e q , independentes do tamanho do passo de integração h e do subíndice k , com $0 < h \leq h_0$, tais que o erro local de discretização satisfaça*

$$\max_k |\alpha_k| \leq Ch^q, \quad (2.7)$$

então o método numérico tem ordem de consistência q .

A desigualdade (2.7) expressa o quão rapidamente o erro local de discretização vai a zero à medida que h diminui.

Exemplo 2.2. *Determine a ordem de consistência, bem como a constante C , do Método de Euler*

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + hf(t_k, y_k) \end{cases} .$$

O erro local de discretização para o Método de Euler é dado por

$$\alpha_k = \frac{y(t_{k+1}) - y(t_k)}{h} - f(t_k, y(t_k)). \quad (2.8)$$

Truncando-se no segundo termo a expansão em Série de Taylor de $y(t_{k+1})$ com centro em $t = t_k$, obtém-se

$$y(t_{k+1}) = y(t_k) + (t_{k+1} - t_k)y'(t_k) + \frac{(t_{k+1} - t_k)^2}{2!}y''(\xi), \quad (2.9)$$

com $\xi \in (t_k, t_{k+1})$.

Como $h = t_{k+1} - t_k$, a substituição de (2.9) em (2.8) resulta em

$$\begin{aligned} \alpha_k &= \frac{y(t_k) + hy'(t_k) + \frac{h^2}{2!}y''(\xi) - y(t_k)}{h} - f(t_k, y(t_k)) \\ &= y'(t_k) + \frac{h}{2!}y''(\xi) - f(t_k, y(t_k)) \\ &= \frac{h}{2}y''(\xi). \end{aligned} \quad (2.10)$$

De (2.10), segue que

$$|\alpha_k| = h \left| \frac{1}{2} y''(\xi) \right| \Rightarrow \max_k |\alpha_k| \leq h \max_{\xi \in I} \frac{|y''(\xi)|}{2}. \quad (2.11)$$

Em (2.11), supõe-se que $y''(t)$ seja contínua, o que garante, pelo Teorema de Weierstrass, a existência de mínimo e máximo absolutos no intervalo.

De (2.11) e da definição (2.7), conclui-se que o Método de Euler tem ordem de consistência 1 (um) com constante $C = \max_{\xi \in I} \frac{|y''(\xi)|}{2}$.

Notação: Se existem constantes positivas C , h_0 e q , independentes do tamanho do passo de integração h e do subíndice k , tais que, para $0 < h \leq h_0$, $\|\alpha_k\| \leq Ch^q$, então, por simplicidade, escreve-se

$$\alpha_k = O(h^q)$$

(lê-se: "o erro de discretização local no ponto t_k tem ordem q ").

2.2.3 Exercícios

Exercício 2.1. Verifique que os Métodos de Euler Implícito e do Trapézio são consistentes.

Exercício 2.2. Determine a ordem de consistência do Método de Euler Aprimorado.

2.3 Erro global de discretização

Definição 2.4 (Erro global de discretização). O erro global de discretização no instante $t = t_k$ é dado por

$$e(t_k, h) = y(t_k) - y_k = e_k \quad (2.12)$$

e representa o erro total acumulado cometido até o k -ésimo passo de integração.

No que segue, uma vez mais será considerado $t - t_0 = hk$ fixado. É importante salientar que y_k não é calculado a partir de valores exatos no instante de tempo anterior e sim a partir de valores obtidos pela aplicação do método numérico (em todos os instantes de tempo anteriores).

2.4 Convergência

Definição 2.5 (Convergência). *Um método numérico é convergente em $t = t_k$ se e somente se*

$$\lim_{h \rightarrow 0} e(t, h) = 0, \quad (2.13)$$

isto é, se e somente se o erro global de discretização tende a zero quando h tende a zero ou, equivalentemente, quando k tende a infinito sendo $h = (t - t_0)/k$. O método numérico é convergente se for convergente para todo t no intervalo de estudo (para qualquer Problema de Cauchy).

Para determinar quais condições são suficientes para um método de passo único explícito convergir, analisa-se o comportamento do erro global de discretização. Para tanto, considere-se a expressão proveniente do produto do erro local de discretização pelo passo de integração h

$$y(t_{k+1}) = y(t_k) + h\Phi(t_k, y(t_k), h) + h\alpha_k, \quad (2.14)$$

associada a um método de passo único explícito geral

$$y_{k+1} = y_k + h\Phi(t_k, y_k, h). \quad (2.15)$$

Efetuada-se a diferença entre (2.14) e (2.15), tem-se

$$y(t_{k+1}) - y_{k+1} = y(t_k) - y_k + h[\Phi(t_k, y(t_k), h) - \Phi(t_k, y_k, h)] + h\alpha_k,$$

isto é, a evolução do erro global de discretização é dada por

$$e_{k+1} = e_k + h[\Phi(t_k, y(t_k), h) - \Phi(t_k, y_k, h)] + h\alpha_k. \quad (2.16)$$

Supondo que a função $\Phi(t, y, h)$ satisfaz a condição de Lipschitz na variável y , ou seja,

$$\|\Phi(t, y_1, h) - \Phi(t, y_2, h)\| \leq L\|y_1 - y_2\|,$$

para quaisquer t e h e para uma constante positiva L , e que Φ seja contínua para todas as variáveis, tem-se para (2.16) que

$$\begin{aligned} \|e_{k+1}\| &\leq \|e_k\| + h\|\Phi(t_k, y(t_k), h) - \Phi(t_k, y_k, h)\| + h\|\alpha_k\| \\ &\leq \|e_k\| + hL\|y(t_k) - y_k\| + h\|\alpha_k\| \\ &\leq \|e_k\| + hL\|e_k\| + h\|\alpha_k\|, \end{aligned}$$

isto é,

$$\|e_{k+1}\| \leq (1 + hL)\|e_k\| + h\|\alpha_k\|, \quad \text{para todo } k. \quad (2.17)$$

Supondo-se que o erro local de discretização α_k seja limitado,

$$\max_k \|\alpha_k\| \leq d, \quad (2.18)$$

é possível reescrever (2.17) como

$$\|e_{k+1}\| \leq (1 + hL)\|e_k\| + hd, \quad 0 \leq k \leq n-1. \quad (2.19)$$

Note-se que, da arbitrariedade do subíndice k em (2.19), tem-se

$$\begin{aligned} \|e_1\| &\leq (1 + hL)\|e_0\| + hd \\ \|e_2\| &\leq (1 + hL)\|e_1\| + hd \\ \|e_3\| &\leq (1 + hL)\|e_2\| + hd \\ &\vdots \\ \|e_{k+1}\| &\leq (1 + hL)\|e_k\| + hd. \end{aligned} \quad (2.20)$$

Por substituição recursiva das desigualdades de (2.20), obtêm-se

$$\begin{aligned} \|e_1\| &\leq (1 + hL)\|e_0\| + hd \\ \|e_2\| &\leq (1 + hL)^2\|e_0\| + [(1 + hL) + 1]hd \\ \|e_3\| &\leq (1 + hL)^3\|e_0\| + [(1 + hL)^2 + (1 + hL) + 1]hd \\ &\vdots \\ \|e_k\| &\leq (1 + hL)^k\|e_0\| + \\ &\quad + \left[(1 + hL)^{k-1} + \dots + (1 + hL)^2 + (1 + hL) + 1 \right]hd. \end{aligned} \quad (2.21)$$

A segunda das parcelas na desigualdade (2.21) é a soma dos k termos de uma progressão geométrica de termo inicial 1 (um) e razão $(1 + hL)$ ¹. Logo,

$$\|e_k\| \leq (1 + hL)^k\|e_0\| + \frac{(1 + hL)^k - 1}{L}d. \quad (2.22)$$

Pela convexidade da função exponencial, pode-se mostrar que $e^t \geq (1 + t)$ e, conseqüentemente,

$$(e^{hL})^k \geq (1 + hL)^k. \quad (2.23)$$

Empregando-se a desigualdade (2.23) em (2.22), constata-se que

$$\|e_k\| \leq e^{khL}\|e_0\| + \frac{e^{khL} - 1}{L}d, \quad (2.24)$$

¹A soma S_n dos n termos de uma progressão geométrica é dada por $S_n = \frac{a_1(q^n - 1)}{q - 1}$, onde a_1 é o primeiro termo e q é a razão da progressão.

onde d é a constante de limitação do erro local de discretização.

Teorema 2.1 (Delimitação do erro global de discretização). *Seja um método de passo único explícito definido por*

$$\begin{cases} y_0 &= y(t_0) \\ y_{k+1} &= y_k + h \Phi(t_k, y_k, h) \end{cases},$$

onde $\Phi(t, y, h)$ é uma função contínua em seus argumentos e satisfaz a condição de Lipschitz para a variável y , isto é, existe uma constante $L > 0$ tal que

$$\|\Phi(t, y_1, h) - \Phi(t, y_2, h)\| \leq L \|y_1 - y_2\|.$$

Além disso, se o erro local de discretização for limitado, ou seja,

$$\max_k \|\alpha_k\| \leq d,$$

então o erro global de discretização satisfaz a delimitação

$$\|e_k\| \leq e^{khL} \|e_0\| + \frac{e^{khL} - 1}{L} d.$$

Se $\|e_0\| = 0$ e se o método numérico for consistente de ordem q , isto é,

$$\max_k \|\alpha_k\| \leq C h^q,$$

tem-se que

$$0 \leq \|e_k\| \leq \frac{e^{khL} - 1}{L} C h^q. \quad (2.25)$$

Em (2.25), o lado direito da desigualdade tende a zero quando h tende a zero. Dessa forma, pelo *Teorema do Confronto*, o limite do valor absoluto do erro global de discretização tende a zero quando o passo de integração tende a zero. Portanto, um método de passo único consistente é convergente (caso a função Φ seja contínua em seus argumentos e Lipschitziana no segundo argumento). É possível mostrar a volta: sob as mesmas condições de regularidade da função Φ , um método de passo único explícito convergente é consistente. Note-se que o teorema de convergência abaixo pode ser estendido a métodos de passo único implícitos.

Teorema 2.2 (Convergência para métodos de passo único). *Um método de passo único explícito*

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h) \end{cases},$$

onde Φ é Lipschitziana em y e contínua em seus argumentos, é convergente se e somente se for consistente.

Exemplo 2.3. *O Método de Euler (1.24) é consistente com ordem de consistência um. Nesse método, $\Phi(t, y, h) = f(t, y)$. Como $f(t, y)$ é contínua e de Lipschitz (Problema de Cauchy), a função $\Phi(t, y, h)$ também é contínua e de Lipschitz. Portanto, o Método de Euler é convergente.*

2.5 Expansão do erro global de discretização

O comportamento exibido na prática por um determinado método numérico depende fortemente da regularidade da função $f(t, y)$ que define o Problema de Cauchy. À luz de um resultado teórico bastante importante (Stoer [14]) é possível se averiguar computacionalmente com que ordem o método numérico empregado converge.

Teorema 2.3 (Expansão assintótica do erro global de discretização). *Seja uma função $f(t, y)$ com $N + 2$ derivadas parciais com relação a y , contínuas e limitadas na faixa*

$$\{(t, y); a = t_0 \leq t \leq t_f = b, y \in \mathbb{R}^n\}.$$

Além disso, seja $\eta(t, h)$ a solução numérica obtida através de um método de passo único

$$\eta_{k+1} = \eta_k + h\Phi(t_k, \eta_k, h)$$

de ordem p para $y(t)$, determinada com o passo de integração h , onde $y(t)$ é a solução do Problema de Cauchy

$$\begin{cases} \frac{d}{dt}y(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases}.$$

Nessas condições, a solução numérica $\eta(t, h)$ admite expansão em potências de h da forma

$$\eta(t, h) = y(t) + h^p e_p(t) + h^{p+1} e_{p+1}(t) + \cdots + h^N e_N(t) + h^{N+1} E_{N+1}(t, h) \quad (2.26)$$

com $e_j(t_0) = 0$, $j = p, p+1, \dots$, válida para todo $t \in [a, b]$ e para todo $h = \frac{t-t_0}{n}$, $n = 1, 2, \dots$

Note-se que as funções $e_j(t)$ são independentes de h e o resto $E_{N+1}(t, h)$ é limitado para t fixado e para todo $h = \frac{t-t_0}{n}$, $n = 1, 2, \dots$. A expansão em série assintótica (2.26) pode ser escrita em função do erro global de discretização no instante t como

$$-e(t, h) = \eta(t, h) - y(t) = \sum_{j=p}^N h^j e_j(t) + h^{N+1} E_{N+1}(t, h). \quad (2.27)$$

O resultado teórico (2.27) pode ser utilizado na prática para estimar o erro global de discretização e a ordem do método, sendo também extremamente útil no processo de depuração do código computacional.

2.5.1 Estimativa do erro global de discretização

Para um dado passo de integração h , obtenha-se $\eta(t, h)$. Para o mesmo instante de tempo t , obtenha-se $\eta(t, \frac{h}{2})$. Para h suficientemente pequeno, em primeira aproximação, tem-se de (2.27) que

$$-e(t, h) = \eta(t, h) - y(t) \approx e_p(t) h^p, \quad (2.28)$$

$$-e(t, \frac{h}{2}) = \eta\left(t, \frac{h}{2}\right) - y(t) \approx e_p(t) \left(\frac{h}{2}\right)^p. \quad (2.29)$$

Efetuada-se a diferença entre (2.28) e (2.29), chega-se a

$$\eta(t, h) - \eta\left(t, \frac{h}{2}\right) \approx e_p(t) \left[h^p - \left(\frac{h}{2}\right)^p \right] = e_p(t) \left(\frac{h}{2}\right)^p (2^p - 1)$$

e, portanto, a

$$e_p(t) \left(\frac{h}{2}\right)^p \approx \frac{\eta(t, h) - \eta(t, \frac{h}{2})}{2^p - 1}. \quad (2.30)$$

Substituindo-se (2.30) em (2.29), para $h > 0$ suficientemente pequeno, obtém-se

$$e(t, \frac{h}{2}) \approx -\frac{\eta(t, h) - \eta(t, \frac{h}{2})}{2^p - 1}, \quad (2.31)$$

uma estimativa do erro global de discretização em t , tendo sido calculadas as aproximações $\eta(t, h)$ e $\eta(t, \frac{h}{2})$ e sendo conhecida a ordem do método empregado. Embora as considerações anteriores tenham sido apresentadas supondo-se que a ordem

teórica do método seja p , na prática verifica-se uma ordem \bar{p} que, dependendo da regularidade da função $f(t, y)$, pode exibir um comportamento distinto da ordem p .

2.5.2 Estimativa da ordem de convergência

A expansão (2.26) fornece um meio útil para se estimar a ordem p (ordem \bar{p}) com a qual o método numérico em uso converge para a solução de um determinado Problema de Cauchy.

Para tanto, obtenha-se $\eta(t, 2h)$, $\eta(t, h)$ e $\eta(t, \frac{h}{2})$, as aproximações numéricas da solução exata no instante t empregando passos de integração $2h$, h e $\frac{h}{2}$, respectivamente, para um passo de integração $h > 0$ suficientemente pequeno. O valor absoluto do quociente entre as diferenças $\eta(t, 2h) - \eta(t, h)$ e $\eta(t, h) - \eta(t, \frac{h}{2})$ fornece, em primeira aproximação,

$$\left| \frac{\eta(t, 2h) - \eta(t, h)}{\eta(t, h) - \eta(t, \frac{h}{2})} \right| \approx \left| \frac{e_{\bar{p}}(t) (2^{\bar{p}} - 1) h^{\bar{p}}}{e_{\bar{p}}(t) (1 - 2^{-\bar{p}}) h^{\bar{p}}} \right| = 2^{\bar{p}}. \quad (2.32)$$

Calculando o logaritmo de base 2 de (2.32)

$$\log_2 \left(\left| \frac{\eta(t, 2h) - \eta(t, h)}{\eta(t, h) - \eta(t, \frac{h}{2})} \right| \right) \approx \log_2 (2^{\bar{p}}) = \bar{p},$$

tem-se uma estimativa \bar{p} para a ordem p do método.

Este procedimento deve ser executado para várias triplas de passos sucessivamente menores $(2h, h, \frac{h}{2})$, $(h, \frac{h}{2}, \frac{h}{4})$, \dots , obtendo-se assim uma sequência de aproximações $\bar{p}_1, \bar{p}_2, \dots$, que converge para a ordem \bar{p} do método.

2.5.3 Depuração do código computacional

Nas seções anteriores, as estimativas apresentadas assumem que o método numérico já tenha passado por um processo conhecido como *verificação / validação* e que, portanto, ele esteja implementado corretamente e funcionando perfeitamente. Durante a programação, entretanto, é necessário o uso de estratégias de depuração para remover eventuais erros de lógica e/ou de coeficientes ou parâmetros que tenham sido introduzidos inadvertidamente. Para isto, empregam-se a aproximação do erro global de discretização (2.28) e um Problema de Cauchy com solução exata conhecida. Tal estratégia é denominada *validação por solução manufaturada*.

Para verificar se o código computacional está correto, escolhe-se um Problema de Cauchy com solução suave (isto é, suficientemente diferenciável). Em (2.28), o erro global de discretização é conhecido uma vez que se tem à mão a solução exata

$y(t)$ do Problema de Cauchy. Deseja-se então determinar a ordem p do método. Para tanto, determina-se $\eta(t, h)$ e $\eta(t, \frac{h}{2})$, as aproximações numéricas da solução exata do problema no instante t , empregando-se passos de integração $h > 0$ e $\frac{h}{2}$, respectivamente. O valor absoluto do quociente entre (2.28) e (2.29) fornece

$$\left| \frac{\eta(t, h) - y(t)}{\eta(t, \frac{h}{2}) - y(t)} \right| \approx \left| \frac{e_{p^*}(t)h^{p^*}}{e_{p^*}(t) (\frac{h}{2})^{p^*}} \right| = 2^{p^*}, \quad (2.33)$$

ou seja,

$$p^* \approx \log_2 \left(\left| \frac{\eta(t, h) - y(t)}{\eta(t, \frac{h}{2}) - y(t)} \right| \right). \quad (2.34)$$

Quando esta estratégia é executada para passos de integração progressivamente menores $h, \frac{h}{2}, \frac{h}{4}, \frac{h}{8}, \dots$, obtém-se pelo uso sucessivo de (2.34) uma sequência $p_1^*, p_2^*, p_3^*, p_4^*, \dots$, que converge à ordem p que a teoria prevê para o método numérico, desde que este tenha sido implementado corretamente e aplicado a um Problema de Cauchy com garantias de solução única suficientemente diferenciável.

O procedimento de verificação por solução manufaturada não pode, por motivo algum, ser negligenciado. Para o bom programador, ele precede o uso regular do método no problema que se deseja solucionar numericamente. Por cautela, o procedimento de verificação deve ser efetuado para várias soluções manufaturadas com complexidade variada.

Exemplo 2.4. *Comprove numericamente a ordem de convergência do Método de Euler aplicado à solução do problema de valor inicial*

$$\begin{cases} \frac{d}{dt}y(t) &= -20y(t), \quad t \in [0, 1] \\ y(0) &= 1 \end{cases}, \quad (2.35)$$

cuja solução exata é $y(t) = e^{-20t}$ (verifique!).

Na Tabela (2.1), apresenta-se a razão entre dois erros globais para passos de integração progressivamente menores (razão de refinamento dois) em $t = 1$. Observa-se que a ordem estimada, localizada na última coluna, tende a um, a ordem de convergência prevista pela teoria para o Método de Euler.

Exemplo 2.5. *Verifique a ordem de convergência numérica do Método de Euler aplicado à solução do problema de valor inicial*

$$\begin{cases} \frac{d}{dt}y(t) &= -y(t) \tan(t) - \frac{1}{\cos(t)} \\ y(0) &= 1 \end{cases} \quad (2.36)$$

$h = n^{-1}$	$ e(t, h) $	$\left \frac{e(t, h)}{e\left(t, \frac{h}{2}\right)} \right $	$\log_2 \left \frac{e(t, h)}{e\left(t, \frac{h}{2}\right)} \right $
$2,000000 \times 10^{-1}$	2,430000E+02		
$1,000000 \times 10^{-1}$	9,999999E-01	2,430000E+02	7,924813E+00
$5,000000 \times 10^{-2}$	2,061154E-09	4,851652E+08	2,885390E+01
$2,500000 \times 10^{-2}$	2,060244E-09	1,000441E+00	6,367371E-04
$1,250000 \times 10^{-2}$	1,960019E-09	1,051135E+00	7,194787E-02
$6,250000 \times 10^{-3}$	1,534787E-09	1,277063E+00	3,528297E-01
$3,125000 \times 10^{-3}$	9,876378E-10	1,553998E+00	6,359843E-01
$1,562500 \times 10^{-3}$	5,632015E-10	1,753614E+00	8,103308E-01
$7,812500 \times 10^{-4}$	3,010566E-10	1,870749E+00	9,036163E-01
$3,906250 \times 10^{-4}$	1,556782E-10	1,933839E+00	9,514678E-01
$1,953125 \times 10^{-4}$	7,916378E-11	1,966533E+00	9,756547E-01
$9,765625 \times 10^{-5}$	3,991780E-11	1,983170E+00	9,878083E-01

Tabela 2.1: Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.35) no instante $t = 1$.

em $t \approx 1,292695719373$ (raiz da equação $e^t \cos(t) = 1$), sabendo que a solução exata de (2.36) é $y(t) = \cos(t) - \sin(t)$ (comprove!).

$h = n^{-1}$	$ e(t, h) $	$\left \frac{e(t, h)}{e\left(t, \frac{h}{2}\right)} \right $	$\log_2 \left \frac{e(t, h)}{e\left(t, \frac{h}{2}\right)} \right $
$5,000000 \times 10^{-2}$	1,130400E-03		
$2,500000 \times 10^{-2}$	2,561790E-04	4,412540E+00	2,141609E+00
$1,250000 \times 10^{-2}$	6,115026E-05	4,189336E+00	2,066722E+00
$6,250000 \times 10^{-3}$	1,494962E-05	4,090422E+00	2,032250E+00
$3,125000 \times 10^{-3}$	3,696597E-06	4,044157E+00	2,015839E+00
$1,562500 \times 10^{-3}$	9,191362E-07	4,021816E+00	2,007847E+00
$7,812500 \times 10^{-4}$	2,291629E-07	4,010842E+00	2,003905E+00
$3,906250 \times 10^{-4}$	5,721340E-08	4,005406E+00	2,001949E+00
$1,953125 \times 10^{-4}$	1,429410E-08	4,002588E+00	2,000933E+00
$9,765625 \times 10^{-5}$	3,573300E-09	4,000251E+00	2,000091E+00

Tabela 2.2: Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.36) no instante $t \approx 1,292695719373$ - Butcher [5].

Na Tabela (2.2), apresenta-se a razão entre dois erros globais para passos

de integração progressivamente menores (razão de refinamento dois) em $t \approx 1,292695719373$. Observa-se que a ordem estimada, localizada na última coluna, tende a dois, diferente da ordem de convergência um prevista pela teoria para o Método de Euler. A ordem mais elevada é justificada pelo cancelamento do termo mais significativo na expansão assintótica do erro global de discretização para $t \approx 1,292695719373$, raiz da equação $e^t \cos(t) = 1$ [5]. Já em $t = \frac{\pi}{4}$, verifica-se ordem estimada um, a ordem de convergência prevista pela teoria para o Método de Euler.

Exemplo 2.6. Verifique a ordem de convergência numérica do Método de Euler aplicado à solução do problema de valor inicial

$$\begin{cases} \frac{d}{dt}y(t) &= -\frac{ty(t)}{1-t^2} \\ y(0) &= 1 \end{cases} \quad (2.37)$$

em $t = 1$, sabendo que a solução exata de (2.37) é $y(t) = \sqrt{1-t^2}$ (comprove!).

$h = n^{-1}$	$ e(t, h) $	$\left \frac{e(t, h)}{e(t, \frac{h}{2})} \right $	$\log_2 \left \frac{e(t, h)}{e(t, \frac{h}{2})} \right $
$1,2500000000 \times 10^{-1}$	3,012018700E-01		
$6,2500000000 \times 10^{-2}$	2,072697687E-01	1,453188E+00	5,392210E-01
$3,1250000000 \times 10^{-2}$	1,441738248E-01	1,437638E+00	5,237004E-01
$1,5625000000 \times 10^{-2}$	1,009724646E-01	1,427853E+00	5,138473E-01
$7,8125000000 \times 10^{-3}$	7,100787890E-02	1,421990E+00	5,079109E-01
$3,9062500000 \times 10^{-3}$	5,005564440E-02	1,418579E+00	5,044464E-01
$1,9531250000 \times 10^{-3}$	3,533418900E-02	1,416635E+00	5,024680E-01
$9,7656250000 \times 10^{-4}$	2,496156840E-02	1,415544E+00	5,013562E-01
$4,8828125000 \times 10^{-4}$	1,764145320E-02	1,414938E+00	5,007392E-01
$2,4414062500 \times 10^{-4}$	1,247093200E-02	1,414606E+00	5,004001E-01
$1,2207031250 \times 10^{-4}$	8,816964600E-03	1,414425E+00	5,002153E-01
$6,1035156250 \times 10^{-5}$	6,234037200E-03	1,414327E+00	5,001153E-01
$3,0517578125 \times 10^{-5}$	4,407942200E-03	1,414274E+00	5,000615E-01

Tabela 2.3: Verificação da ordem de convergência do Método de Euler aplicado ao problema de valor inicial (2.37) no instante $t = 1$ - Butcher [5].

Na Tabela (2.3), apresenta-se a razão entre dois erros globais para passos de integração progressivamente menores (razão de refinamento dois) em $t = 1$. Observa-se que a ordem estimada, localizada na última coluna, tende a meio, diferente da ordem de convergência um prevista pela teoria para o Método de Euler. A justificativa para a perda de ordem é que a Condição de Lipschitz não é mantida quando $t = 1$ e $y = 0$ [5].

2.5.4 Exercícios

Exercício 2.3. *Mostre que um método de passo único explícito convergente é consistente.*

Exercício 2.4. *Estime numericamente a ordem de convergência do Método de Euler Aprimorado (1.27) com a estratégia de solução manufaturada empregando o problema modelo (2.35) no instante final $t = 1$.*

Exercício 2.5. *Estime numericamente a ordem de convergência do Método de Euler (1.24) com a estratégia de solução manufaturada empregando o problema modelo (2.36) no instante final $t = \frac{\pi}{4}$.*

Exercício 2.6. *Estime numericamente a ordem de convergência do Método de Euler (1.24) com a estratégia de solução manufaturada empregando o problema modelo (2.37) no instante final $t = \frac{1}{2}$.*

2.6 Suplemento teórico

2.6.1 Teorema de Weierstrass

Teorema 2.4. *Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função contínua no conjunto A fechado e limitado. Então, existem pontos de máximo e mínimo absoluto de f em A , isto é, existem $x_0, x_1 \in A$ tais que*

$$f(x_0) \leq f(x) \leq f(x_1)$$

para todo $x \in A$.

2.6.2 Teorema do Confronto

Teorema 2.5. *Se $f(t) \leq g(t) \leq h(t)$ quando t está próximo de a (exceto possivelmente em a) e*

$$\lim_{t \rightarrow a} f(t) = \lim_{t \rightarrow a} h(t) = L,$$

então

$$\lim_{t \rightarrow a} g(t) = L.$$

2.7 Nota biográfica



Leonhard Euler (1707-1783)

Euler nasceu em Basel, Suíça, em 15 de abril de 1707. A vontade do pai de Euler era que este seguisse sua carreira na igreja. Assim, ele foi para a Universidade de Basel estudar para o ministério. Entretanto, geometria logo se tornou seu assunto preferido e ele obteve o consentimento do pai para mudar para matemática após a persuasão de Johann Bernoulli, que se tornou seu professor.

Euler entrou para a Academia de Ciências de São Petesburgo dois anos após sua fundação por Catarina I. Foi professor na Academia em 1730 e de matemática em 1733. Em 1741, a convite de Frederico, o Grande, foi integrado à Academia de Ciências de Berlin, onde permaneceu por 25 anos. Durante este período escreveu mais de 200 artigos, três livros sobre análise matemática e uma publicação científica popular: *Letters to a Princess of Germany*.

Em 1766 retornou à Rússia e ficou completamente cego depois de uma cirurgia de catarata. Euler já havia perdido a visão do olho direito aos 31 anos.

Foi após 1765 que Euler produziu aproximadamente metade de seu trabalho. Mesmo após sua morte, a Academia de São Petesburgo continuou a publicar seus trabalhos inéditos por mais de 50 anos. Sua obra completa contém 886 livros e artigos. É de sua autoria a notação $f(x)$ (1734), i para a raiz quadrada de -1 (1777), π para o número pi e Σ para somatório (1755).

Euler trabalhou com Christian Goldbach em teoria dos números e com Alexis Claude de Clairot em teoria lunar, introduziu as funções beta e gama (Integrais

Eulerianas) e fez contribuições em mecânica do contínuo, problema dos três corpos, elasticidade, acústica, teoria da onda de luz, hidráulica e música. Com seu trabalho *Theory of the Motions of Rigid Bodies* (1765), lançou os fundamentos da mecânica analítica. Euler faleceu em São Petesburgo, Rússia, em 18 de setembro de 1783.

2.8 Exercícios resolvidos

1. Considere o Método do Trapézio.

- Calcule o erro local de discretização do método;
- Verifique que o método é consistente;
- Determine um delimitante superior do erro global de discretização.

Solução

(a)

$$\alpha_k = \frac{y(t_{k+1}) - y(t_k)}{h} - \Phi(t_k, y(t_k), y(t_{k+1}), h)$$

$$\alpha_k = \frac{y(t_{k+1}) - y(t_k)}{h} - \frac{f(t_k, y(t_k)) + f(t_{k+1}, y(t_{k+1}))}{2}$$

$$2h \alpha_k = 2[y(t_{k+1}) - y(t_k)] - h[f(t_k, y(t_k)) + f(t_{k+1}, y(t_{k+1}))]$$

$$2h \alpha_k = 2 \left[y(t_k) + h y'(t_k) + \frac{h^2}{2!} y''(t_k) + \frac{h^3}{3!} y'''(\varepsilon_1) - y(t_k) \right] +$$

$$- h [y'(t_k) + y'(t_{k+1})]$$

$$2h \alpha_k = 2 \left[h y'(t_k) + \frac{h^2}{2!} y''(t_k) + \frac{h^3}{3!} y'''(\varepsilon_1) \right] +$$

$$- h \left[y'(t_k) + y'(t_k) + h y''(t_k) + \frac{h^2}{2!} y'''(\varepsilon_2) \right]$$

$$2h \alpha_k = 2h y'(t_k) + h^2 y''(t_k) + \frac{h^3}{3} y'''(\varepsilon_1) +$$

$$- 2h y'(t_k) - h^2 y''(t_k) - \frac{h^3}{2} y'''(\varepsilon_2)$$

$$2h \alpha_k = \frac{h^3}{3} y'''(\varepsilon_1) - \frac{h^3}{2} y'''(\varepsilon_2)$$

$$\alpha_k = \frac{h^2}{6} y'''(\varepsilon_1) - \frac{h^2}{4} y'''(\varepsilon_2)$$

$$\alpha_k = h^2 \left[\frac{y'''(\varepsilon_1)}{6} - \frac{y'''(\varepsilon_2)}{4} \right]$$

Portanto, o erro local de discretização é dado por

$$\alpha_k = h^2 \left[\frac{y'''(\varepsilon_1)}{6} - \frac{y'''(\varepsilon_2)}{4} \right], \quad \varepsilon_1, \varepsilon_2 \in]t_k, t_{k+1}[.$$

(b)

$$\lim_{h \rightarrow 0} \alpha_k = \lim_{h \rightarrow 0} h^2 \left[\frac{y'''(\varepsilon_1)}{6} - \frac{y'''(\varepsilon_2)}{4} \right] = 0.$$

Logo, o Método do Trapézio é consistente.

(c)

$$\begin{aligned} y(t_{k+1}) &= y(t_k) + h \Phi(t_k, y(t_k), y(t_{k+1}), h) + h \alpha_k & A \\ y_{k+1} &= y_k + h \Phi(t_k, y_k, y_{k+1}, h) & B \end{aligned}$$

Fazendo-se A-B:

$$\begin{aligned} y(t_{k+1}) - y_{k+1} &= y(t_k) - y_k + \\ &+ h [\Phi(t_k, y(t_k), y(t_{k+1}), h) - \Phi(t_k, y_k, y_{k+1}, h)] + \\ &+ h \alpha_k; \end{aligned}$$

$$\begin{aligned} \|e_{k+1}\| &= \|e_k + \\ &+ h \left[\frac{f(t_k, y(t_k)) + f(t_{k+1}, y(t_{k+1}))}{2} - \frac{f(t_k, y_k) + f(t_{k+1}, y_{k+1})}{2} \right] + \\ &+ h \alpha_k\|; \end{aligned}$$

$$\begin{aligned} \|e_{k+1}\| &= \|e_k + \\ &+ \frac{h}{2} [f(t_k, y(t_k)) - f(t_k, y_k) + f(t_{k+1}, y(t_{k+1})) - f(t_{k+1}, y_{k+1})] + \\ &+ h \alpha_k\|; \end{aligned}$$

$$\begin{aligned} \|e_{k+1}\| &\leq \|e_k\| + \frac{h}{2} \|f(t_k, y(t_k)) - f(t_k, y_k)\| + \\ &+ \frac{h}{2} \|f(t_{k+1}, y(t_{k+1})) - f(t_{k+1}, y_{k+1})\| + \\ &+ h \|\alpha_k\|; \end{aligned}$$

$$\begin{aligned} \|e_{k+1}\| &\leq \|e_k\| + \frac{L_1 h}{2} \|y(t_k) - y_k\| + \\ &+ \frac{L_2 h}{2} \|y(t_{k+1}) - y_{k+1}\| + h \|\alpha_k\|. \end{aligned}$$

Sendo $L = \max\{L_1, L_2\}$ e $d = \max_k \|\alpha_k\|$:

$$\begin{aligned}
\|e_{k+1}\| &\leq \|e_k\| + \frac{Lh}{2}\|e_k\| + \frac{Lh}{2}\|e_{k+1}\| + hd; \\
\left(1 - \frac{Lh}{2}\right)\|e_{k+1}\| &\leq \left(1 + \frac{Lh}{2}\right)\|e_k\| + hd; \\
\|e_{k+1}\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)\|e_k\| + \frac{hd}{1 - \frac{Lh}{2}}; \\
\|e_{n+1}\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)\|e_n\| + \frac{hd}{1 - \frac{Lh}{2}}. \tag{2.38}
\end{aligned}$$

Para $n = 0$ em (2.38):

$$\|e_1\| \leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)\|e_0\| + \frac{hd}{1 - \frac{Lh}{2}}.$$

Para $n = 1$ em (2.38):

$$\begin{aligned}
\|e_2\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)\|e_1\| + \frac{hd}{1 - \frac{Lh}{2}}; \\
\|e_2\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right) \left[\left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)\|e_0\| + \frac{hd}{1 - \frac{Lh}{2}} \right] + \frac{hd}{1 - \frac{Lh}{2}}; \\
\|e_2\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^2 \|e_0\| + \left(1 + \frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right) \frac{hd}{1 - \frac{Lh}{2}}.
\end{aligned}$$

Para $n = 2$ em (2.38):

$$\begin{aligned}
\|e_3\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^3 \|e_0\| + \\
&\quad + \left[1 + \frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}} + \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^2 \right] \frac{hd}{1 - \frac{Lh}{2}}.
\end{aligned}$$

Para $n = k - 1$ em (2.38):

$$\begin{aligned}
\|e_k\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^k \|e_0\| + \\
&+ \left[1 + \frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}} + \dots + \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^{k-1}\right] \frac{hd}{1 - \frac{Lh}{2}}; \\
\|e_k\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^k \|e_0\| + \\
&+ \left[\frac{\left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^k - 1}{\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}} - 1}\right] \frac{hd}{1 - \frac{Lh}{2}}; \\
\|e_k\| &\leq \left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^k \|e_0\| + \\
&+ \left[\left(\frac{1 + \frac{Lh}{2}}{1 - \frac{Lh}{2}}\right)^k - 1\right] \frac{d}{L}. \tag{2.39}
\end{aligned}$$

Utilizando a convexidade da função exponencial em (2.39), obtém-se a desigualdade

$$\|e_k\| \leq e^{kLh} \|e_0\| + \frac{e^{kLh} - 1}{L} d,$$

a qual fornece um limitante para o erro global de discretização.

Capítulo 3

Métodos de passo único de altas ordens

Os Métodos da Série de Taylor exemplificam uma maneira de se obter métodos numéricos de ordem mais elevada. Felizmente, há outras formas para se atingir este propósito uma vez que tais métodos são geralmente difíceis de aplicar devido ao cálculo das derivadas da função $f(t, y)$. Os Métodos de Runge-Kutta constituem uma forma alternativa e mais prática de determinar métodos com ordem mais elevada.

3.1 Métodos da Série de Taylor

Expandindo-se $y(t_k + h)$ em Série de Taylor com centro em $t = t_k$, obtém-se

$$\begin{aligned} y(t_k + h) &= y(t_{k+1}) = y(t_k) + h y'(t_k) + \frac{h^2}{2!} y''(t_k) + \\ &+ \cdots + \frac{h^{q+1}}{(q+1)!} y^{(q+1)}(\xi), \quad \xi \in (t_k, t_k + h). \end{aligned} \quad (3.1)$$

Sabendo-se que $y'(t) = f(t, y(t))$ (Problema de Cauchy), pode-se reescrever (3.1) como

$$\begin{aligned} y(t_k + h) &= y(t_{k+1}) = y(t_k) + h f(t_k, y(t_k)) + \frac{h^2}{2!} \frac{d}{dt} f(t_k, y(t_k)) + \\ &+ \cdots + \frac{h^{q+1}}{(q+1)!} \frac{d^q}{dt^q} f(\xi, y(\xi)). \end{aligned} \quad (3.2)$$

Notação:

$$\begin{aligned}
 y^{(1)}(t) &= f(t, y(t)) \\
 y^{(2)}(t) &= \frac{d}{dt}f(t, y(t)) = \frac{\partial f}{\partial t} \frac{dt}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \\
 &= \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) [f](t, y(t)) \\
 y^{(3)}(t) &= \frac{d^2}{dt^2}f(t, y(t)) = \left(\frac{\partial^2 f}{\partial t^2} + 2f \frac{\partial^2 f}{\partial t \partial y} + f^2 \frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial t} \frac{\partial f}{\partial y} + f \left(\frac{\partial f}{\partial y} \right)^2 \right) \\
 &= \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right)^2 [f](t, y(t)) \\
 &\vdots \\
 y^{(j)}(t) &= \frac{d^{j-1}}{dt^{j-1}}f(t, y(t)) = \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right)^{j-1} [f](t, y(t))
 \end{aligned}$$

A expansão (3.2) sugere o método numérico

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, h) \end{cases}, \quad (3.3)$$

onde

$$\Phi(t, y, h) = f(t, y) + \frac{h}{2!} Df(t, y) + \frac{h^2}{3!} D^2f(t, y) + \dots + \frac{h^{q-1}}{q!} D^{q-1}f(t, y),$$

com $D = \frac{\partial}{\partial t} + f \frac{\partial}{\partial y}$.

O Método da Série de Taylor (3.3) com desenvolvimento até os termos de ordem q tem ordem q (erro local de discretização $O(h^q)$).

3.2 Métodos de Runge-Kutta Explícitos

3.2.1 Caracterização dos Métodos de Runge-Kutta

Um método pertence à classe dos Métodos de Runge-Kutta explícitos se

1. for um método de passo único explícito;
2. substituir as derivadas de $f(t, y)$ por cálculos de $f(t, y)$;
3. concordar com o Método da Série de Taylor (3.3) até o termo de q -ésima ordem.

Exemplo 3.1. O Método de Euler Modificado é um Método de Runge-Kutta.

Método de Euler Modificado (ou Método do Ponto Médio)

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h), \end{cases} \quad (3.4)$$

com $t_{k+1} = t_k + h$, $h = \frac{b-a}{n}$, $0 \leq k \leq n-1$ e

$$\Phi(t_k, y_k, h) = f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}f(t_k, y_k)\right).$$

Por inspeção, observa-se que o método (3.4) é de passo único e explícito. Além disso, não existem cálculos de derivadas de $f(t, y)$. Basta então verificar se (3.4) concorda com o Método da Série de Taylor (3.3) para alguma ordem q .

Desenvolvendo-se $f(t + \Delta t, y + \Delta y)$ em Série de Taylor em torno de (t_k, y_k) , onde $\Delta t = \frac{h}{2}$ e $\Delta y = \frac{h}{2}f(t, y)$, tem-se

$$\begin{aligned} f(t_k + \Delta t, y_k + \Delta y) &= f(t_k, y_k) + \left[\frac{\partial}{\partial t}f(t_k, y_k) \quad \frac{\partial}{\partial y}f(t_k, y_k) \right] \begin{bmatrix} \Delta t \\ \Delta y \end{bmatrix} + \\ &+ \frac{1}{2!} \begin{bmatrix} \Delta t & \Delta y \end{bmatrix} \begin{bmatrix} \frac{\partial^2}{\partial t^2}f(t_k, y_k) & \frac{\partial^2}{\partial t \partial y}f(t_k, y_k) \\ \frac{\partial^2}{\partial t \partial y}f(t_k, y_k) & \frac{\partial^2}{\partial y^2}f(t_k, y_k) \end{bmatrix} \begin{bmatrix} \Delta t \\ \Delta y \end{bmatrix} + \\ &+ O(\Delta t^3, \Delta y^3), \end{aligned}$$

ou ainda,

$$\begin{aligned} \Phi(t_k, y_k, h) &= f_k + \frac{h}{2} \left(\frac{\partial f_k}{\partial t} + f_k \frac{\partial f_k}{\partial y} \right) + \frac{1}{2!} \frac{h^2}{4} \left(\frac{\partial^2 f_k}{\partial t^2} + 2f_k \frac{\partial^2 f_k}{\partial t \partial y} + f_k^2 \frac{\partial^2 f_k}{\partial y^2} \right) + \\ &+ O(h^3), \end{aligned} \quad (3.5)$$

sendo $f_k = f(t_k, y_k)$.

Substituindo-se (3.5) em (3.4), obtém-se

$$\begin{aligned} y_{k+1} &= y_k + hf_k + \frac{h^2}{2!} \left(\frac{\partial f_k}{\partial t} + f_k \frac{\partial f_k}{\partial y} \right) + \frac{1}{2!} \frac{h^3}{4} \left(\frac{\partial^2 f_k}{\partial t^2} + 2f_k \frac{\partial^2 f_k}{\partial t \partial y} + f_k^2 \frac{\partial^2 f_k}{\partial y^2} \right) + \\ &+ O(h^4), \end{aligned} \quad (3.6)$$

cujos termos concordam com os do Método da Série de Taylor (3.3) até a segunda ordem (compare!). O Método de Euler Modificado é um Método de Runge-Kutta de ordem 2.

3.2.2 Métodos de Runge-Kutta explícitos de R -estágios

Nos métodos de Runge-Kutta (ou RK), o número de estágios diz respeito ao número de vezes que $f(t, y)$ deve ser calculada. Os métodos RK explícitos com R -estágios têm a forma

$$\begin{cases} y_0 &= y(t_0) \\ y_{k+1} &= y_k + h\Phi(t_k, y_k, h), \end{cases}$$

onde

$$\Phi(t, y, h) = \sum_{r=1}^R c_r \kappa_r, \quad (3.7)$$

com

$$\begin{aligned} \kappa_1(t, y) &= f(t, y) \\ \kappa_2(t, y) &= f(t + ha_2, y + hb_{21}\kappa_1) \\ \kappa_3(t, y) &= f(t + ha_3, y + hb_{31}\kappa_1 + hb_{32}\kappa_2) \\ &\vdots \\ \kappa_r(t, y) &= f\left(t + ha_r, y + h \sum_{s=1}^{r-1} b_{rs}\kappa_s\right), \quad 2 \leq r \leq R, \end{aligned} \quad (3.8)$$

e os parâmetros a_r , b_{rs} e c_r satisfazem as relações

$$\sum_{r=1}^R c_r = 1, \quad a_r = \sum_{s=1}^{r-1} b_{rs}, \quad 2 \leq r \leq R. \quad (3.9)$$

Para um Método RK de dois estágios ($R = 2$), de (3.7)-(3.9) tem-se

$$\begin{aligned} \Phi(t_k, y_k, h) &= c_1 \kappa_1 + c_2 \kappa_2 = c_1 f(t_k, y_k) + \\ &+ c_2 f(t_k + ha_2, y_k + hb_{21}f(t_k, y_k)), \end{aligned} \quad (3.10)$$

com $a_2 = b_{21}$.

Exemplo 3.2. *Tem-se no Método de Euler Modificado (3.4)*

$$\Phi(t, y, h) = f\left(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)\right). \quad (3.11)$$

Comparando-se (3.11) com (3.10), constata-se que $c_1 = 0$ e $c_2 = 1$ (logo $c_1 + c_2 = 1$) e que $a_2 = b_{21} = \frac{1}{2}$. Assim, o Método de Euler Modificado é um Método de Runge-Kutta de segunda ordem com 2 estágios.

Exemplo 3.3. *Tem-se no Método de Euler Aprimorado (1.27),*

$$\Phi(t, y, h) = \frac{1}{2}f(t, y) + \frac{1}{2}f(t + h, y + hf(t_k, y_k)). \quad (3.12)$$

Comparando (3.12) com (3.10), constata-se que $c_1 = \frac{1}{2}$ e $c_2 = \frac{1}{2}$ (logo $c_1 + c_2 = 1$) e que $a_2 = b_{21} = 1$. Assim, o Método de Euler Aprimorado é um Método de Runge-Kutta de segunda ordem com 2 estágios.

É interessante observar que o Método de Euler é um Método de Runge-Kutta de primeira ordem com um estágio e que existem infinitos métodos de dois estágios que terão segunda ordem se satisfizerem as relações (3.9).

3.2.3 Métodos de Runge-Kutta de ordens superiores

Método de Runge-Kutta de Terceira Ordem com Três Estágios (RK33)

$$\Phi(t, y, h) = \frac{1}{6}(\kappa_1 + 4\kappa_2 + \kappa_3), \quad (3.13)$$

com

$$\begin{cases} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{h}{2}, y + \frac{h}{2}\kappa_1\right) \\ \kappa_3 = f(t + h, y - h\kappa_1 + 2h\kappa_2) \end{cases} .$$

Método de Runge-Kutta de Quarta Ordem com Quatro Estágios (RK44)

$$\Phi(t, y, h) = \frac{1}{6}(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4), \quad (3.14)$$

com

$$\begin{cases} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{h}{2}, y + \frac{h}{2}\kappa_1\right) \\ \kappa_3 = f\left(t + \frac{h}{2}, y + \frac{h}{2}\kappa_2\right) \\ \kappa_4 = f(t + h, y + h\kappa_3) \end{cases} .$$

Método de Runge-Kutta de Quarta Ordem com Cinco Estágios (RK45)

$$\Phi(t, y, h) = \frac{25}{216}\kappa_1 + \frac{1408}{2565}\kappa_3 + \frac{2197}{4104}\kappa_4 - \frac{1}{5}\kappa_5, \quad (3.15)$$

com

$$\left\{ \begin{array}{l} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{h}{4}, y + \frac{h}{4}\kappa_1\right) \\ \kappa_3 = f\left(t + \frac{3h}{8}, y + \frac{3h}{32}\kappa_1 + \frac{9h}{32}\kappa_2\right) \\ \kappa_4 = f\left(t + \frac{12h}{13}, y + \frac{1932h}{2197}\kappa_1 - \frac{7200h}{2197}\kappa_2 + \frac{7296h}{2197}\kappa_3\right) \\ \kappa_5 = f\left(t + h, y + \frac{439h}{216}\kappa_1 - 8h\kappa_2 + \frac{3680h}{513}\kappa_3 - \frac{845h}{4104}\kappa_4\right) \end{array} \right.$$

Método de Runge-Kutta de Quinta Ordem com Seis Estágios (RK56)

$$\Phi(t, y, h) = \frac{16}{135}\kappa_1 + \frac{6656}{12825}\kappa_3 + \frac{28561}{56430}\kappa_4 - \frac{9}{50}\kappa_5 + \frac{2}{55}\kappa_6, \quad (3.16)$$

com

$$\left\{ \begin{array}{l} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{h}{4}, y + \frac{h}{4}\kappa_1\right) \\ \kappa_3 = f\left(t + \frac{3h}{8}, y + \frac{3h}{32}\kappa_1 + \frac{9h}{32}\kappa_2\right) \\ \kappa_4 = f\left(t + \frac{12h}{13}, y + \frac{1932h}{2197}\kappa_1 - \frac{7200h}{2197}\kappa_2 + \frac{7296h}{2197}\kappa_3\right) \\ \kappa_5 = f\left(t + h, y + \frac{439h}{216}\kappa_1 - 8h\kappa_2 + \frac{3680h}{513}\kappa_3 - \frac{845h}{4104}\kappa_4\right) \\ \kappa_6 = f\left(t + \frac{h}{2}, y - \frac{8h}{27}\kappa_1 + 2h\kappa_2 - \frac{3544h}{2565}\kappa_3 + \frac{1859h}{4104}\kappa_4 - \frac{11h}{40}\kappa_5\right) \end{array} \right.$$

3.2.4 Exercícios**Exercício 3.1.** *Proponha um Método RK de primeira ordem e dois estágios.*

Exercício 3.2. *Mostre que o Método de Euler Aprimorado (1.27) concorda com o Método da Série de Taylor até o termo de segunda ordem.*

3.3 Controle do passo de integração

Nos métodos de Euler Modificado (3.4) e Runge-Kutta RK33 (3.13), os erros locais de discretização são definidos, respectivamente, por

$$\alpha_k^{EM} = \frac{y(t_{k+1}) - y(t_k)}{h} - \kappa_2 \quad (3.17)$$

e

$$\alpha_k^{RK33} = \frac{y(t_{k+1}) - y(t_k)}{h} - \frac{1}{6}(\kappa_1 + 4\kappa_2 + \kappa_3). \quad (3.18)$$

Efetuada-se a diferença entre (3.17) e (3.18), chega-se a

$$\alpha_k^{EM} - \alpha_k^{RK33} = -\kappa_2 + \frac{1}{6}(\kappa_1 + 4\kappa_2 + \kappa_3) \quad (3.19)$$

$$= \frac{1}{6}(\kappa_1 - 2\kappa_2 + \kappa_3). \quad (3.20)$$

Como $\alpha_k^{RK33} = O(h^3)$, pode-se reescrever (3.19) na forma

$$\alpha_k^{EM} = \frac{1}{6}(\kappa_1 - 2\kappa_2 + \kappa_3) + O(h^3),$$

ou ainda na forma

$$\alpha_k^{EM} \approx \frac{1}{6}(\kappa_1 - 2\kappa_2 + \kappa_3). \quad (3.21)$$

Empregando-se (3.21) como uma estimativa para o erro local de discretização do Método de Euler Modificado, pode-se propor a estratégia descrita no Algoritmo (3.3.1) para controlar o passo de integração no Método de Euler Modificado.

A combinação dos Métodos RK45 (3.15) e RK56 (3.16) é conhecida como *Método de Runge-Kutta-Fehlberg* [7]. Essa combinação é comumente empregada na prática para controlar automaticamente o passo de integração.

3.3.1 Exercícios

Exercício 3.3. *Escreva um algoritmo que empregue o Método de Euler com controle automático do passo de integração baseado nos erros locais de discretização dos Métodos de Euler e de Euler Modificado.*

Exercício 3.4. *Escreva um algoritmo que empregue o Método de Runge-Kutta 45 com controle automático do passo de integração baseado nos erros locais de*

1. Dados (t_k, y_k) , $h > 0$ e $\epsilon > 0$, calcular

$$\begin{aligned}\kappa_1 &\leftarrow f(t_k, y_k) \\ \kappa_2 &\leftarrow f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_1\right) \\ \kappa_3 &\leftarrow f\left(t_k + h, y_k - h\kappa_1 + 2h\kappa_2\right)\end{aligned}$$

2. Estimar o erro local de discretização

$$\alpha_k^{EM} \leftarrow \frac{1}{6}(\kappa_1 - 2\kappa_2 + \kappa_3)$$

Se $|\alpha_k^{EM}| > \epsilon$ então reduzir o passo de integração

$$h \leftarrow \beta h, \quad 0 < \beta < 1$$

retornar a 1

senão aumentar o passo de integração

$$h \leftarrow \delta h, \quad \delta > 1$$

Fim se

3. $y_k \leftarrow y_k + h\Phi_{EM}$

$$t_k \leftarrow \min\{t_k + h, t_{final}\}$$

4. Se $0 < t_k < t_{final}$ então retornar a 1

senão parar

Fim se

algoritmo 3.3.1: Controle do passo de integração no Método de Euler Modificado.

discretização dos Métodos de Runge-Kutta 45 e de Runge-Kutta 56 (Runge-Kutta-Felberg). Verifique que a estimativa para o erro local de discretização é dada por

$$\alpha_k^{RKF} \approx \frac{1}{360}\kappa_1 - \frac{128}{4275}\kappa_3 - \frac{2197}{75240}\kappa_4 + \frac{1}{50}\kappa_5 + \frac{2}{55}\kappa_6.$$

(Sugestão: veja o algoritmo proposto por Burden [7].)

3.4 Suplemento teórico

3.4.1 Regra da cadeia

Teorema 3.1. Se $z = f(x, y)$, $x = g(t)$ e $y = h(t)$, então

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.$$

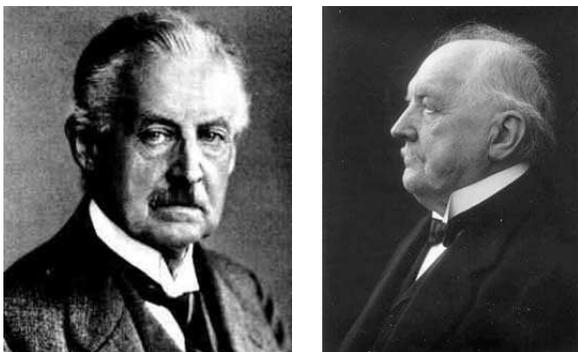
Se $z = f(x, y)$, $x = g(u, v)$ e $y = h(u, v)$, então

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}$$

e

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}.$$

3.5 Nota biográfica



Carl David Tolmé Runge (1856-1927), à esquerda, e Wilhelm Martin Kutta (1867-1944), à direita.

Carl David Tolmé Runge (1856-1927)

Runge nasceu em Bremen, Alemanha, em 30 de agosto de 1856. Aos 19 anos, após deixar a escola, Runge, juntamente com sua mãe, passou seis meses visitando centros culturais na Itália. Quando retornou, entrou para a Universidade de Munique para estudar literatura. Entretanto, seis semanas após seu ingresso, mudou para matemática e física. Foi aluno de Max Planck, de quem se tornou amigo. Quando retornou a Berlin, voltou seus estudos para matemática pura. Sua dissertação de

doutorado foi em geometria diferencial. Mais tarde, influenciado por Kronecker, trabalhou num procedimento para solução de equações algébricas nas quais as raízes se expressam como séries infinitas de funções racionais dos coeficientes. Após sua visita a Mittag-Laffler em Estocolmo (1884), publicou um grande número de artigos na *Acta Mathematica*.

Runge obteve uma cadeira em Hanôver em 1886, onde permaneceu por 18 anos. Desenvolveu grande trabalho experimental e publicou uma grande quantidade de resultados. Em 1904, Klein persuadiu Göttingen a oferecer a Runge uma cadeira de matemática aplicada, na qual permaneceu até 1925. Runge faleceu em Göttingen, Alemanha, em 3 de janeiro de 1927.

Wilhelm Martin Kutta (1867-1944)

Kutta nasceu em Pitschen (agora Byczna), Polônia, em 3 de novembro de 1867. Estudou em Breslau de 1885 a 1890 e depois foi para Munique, onde estudou de 1891 a 1894. Em Munique foi assistente de von Dyck e, de 1898 a 1899, esteve na Universidade de Cambridge, Inglaterra. Kutta obteve postos em Munique, Jena e Aachen. Tornou-se professor em Stuttgart em 1911, onde dedicou-se com total força e paixão ao ensino de matemática para engenheiros, lá permanecendo até 1935.

Kutta era um homem muito culto, conhecendo em profundidade música, arte e história da literatura. Somente por acaso descobriu-se que ele também falava árabe. É conhecido principalmente por suas contribuições em resolução numérica de equações diferenciais ordinárias, os métodos de Runge-Kutta (1902), e pelas célebres fórmulas de aplicações conformes, as transformações de Kutta-Zhukovskii (1910), com as quais, anos mais tarde, gerações de engenheiros viriam a calcular perfis de aerofólios. Kutta faleceu em Fürstfeldbruck, Alemanha, em 25 de dezembro de 1944. Seu túmulo na Igreja de Santa Madalena, abandonado, transformou-se com o tempo em uma trilha.

3.6 Exercícios resolvidos

1. Determine os parâmetros a_2 , c_1 e c_2 do Método de Runge-Kutta com 2 estágios de modo que este tenha ordem máxima. Mostre que a ordem não pode exceder dois e deduza os Métodos de Euler Aprimorado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{cases}$$

$$\Phi(t_k, y_k, h) = \frac{\kappa_1 + \kappa_2}{2} \quad \text{e} \quad \begin{cases} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f(t_k + h, y_k + h\kappa_1) \end{cases}$$

e de Euler Modificado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{cases}$$

$$\Phi(t_k, y_k, h) = \kappa_2 \quad \text{e} \quad \begin{cases} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_1) \end{cases}$$

ambos com $h = \frac{b-a}{n}$.

Solução

Método da Série de Taylor de ordem 3:

$$\begin{cases} y^{(1)}(t) = f(t, y) \\ y^{(2)}(t) = f_t(t, y) + f_y(t, y)f(t, y) \\ y^{(3)}(t) = f_{tt}(t, y) + 2f_{ty}(t, y)f(t, y) + f_{yy}(t, y)f^2(t, y) + \\ + f_t(t, y)f_y(t, y) + f_y^2(t, y)f(t, y) \end{cases} ;$$

$$y_{k+1} = y_k + hy^{(1)}(t_k) + \frac{h^2}{2!}y^{(2)}(t_k) + \frac{h^3}{3!}y^{(3)}(t_k) + O(h^4);$$

$$\begin{aligned} y_{k+1} &= y_k + hf(t_k, y_k) + \frac{h^2}{2!} [f_t(t_k, y_k) + f_y(t_k, y_k)f(t_k, y_k)] + \\ &+ \frac{h^3}{3!} [f_{tt}(t_k, y_k) + 2f_{ty}(t_k, y_k)f(t_k, y_k) + f_{yy}(t_k, y_k)f^2(t_k, y_k)] + \\ &+ \frac{h^3}{3!} [f_t(t_k, y_k)f_y(t_k, y_k) + f_y^2(t_k, y_k)f(t_k, y_k)] + O(h^4). \quad (3.22) \end{aligned}$$

Método de Runge-Kutta de 2 estágios:

$$\begin{aligned} y_{k+1} &= y_k + h(c_1\kappa_1 + c_2\kappa_2) \\ \kappa_1 &= f(t_k, y_k) \\ \kappa_2 &= f(t_k + a_2h, y_k + hb_{21}\kappa_1). \end{aligned}$$

Como $a_2 = b_{21}$:

$$\begin{aligned} y_{k+1} &= y_k + h(c_1\kappa_1 + c_2\kappa_2) \\ \kappa_1 &= f(t_k, y_k) \\ \kappa_2 &= f(t_k + a_2h, y_k + ha_2\kappa_1). \end{aligned} \tag{3.23}$$

Expandindo κ_2 em série de Taylor com centro em (t_k, y_k) até ordem 2, obtém-se

$$\begin{aligned} \kappa_2 &= f(t_k, y_k) + a_2h[f_t(t_k, y_k) + \kappa_1f_y(t_k, y_k)] + \\ &+ \frac{a_2^2h^2}{2!}[f_{tt}(t_k, y_k) + 2\kappa_1f_{ty}(t_k, y_k) + \kappa_1^2f_{yy}(t_k, y_k)]. \end{aligned} \tag{3.24}$$

Substituindo (3.24) em (3.23), tem-se que:

$$\begin{aligned} y_{k+1} &= y_k + hc_1f(t_k, y_k) + hc_2f(t_k, y_k) + \\ &+ h^2a_2c_2[f_t(t_k, y_k) + f(t_k, y_k)f_y(t_k, y_k)] + \\ &+ \frac{h^3a_2^2c_2}{2!}[f_{tt}(t_k, y_k) + 2f(t_k, y_k)f_{ty}(t_k, y_k) + f^2(t_k, y_k)f_{yy}(t_k, y_k)]; \\ y_{k+1} &= y_k + (c_1 + c_2)hf(t_k, y_k) + \\ &+ a_2c_2h^2[f_t(t_k, y_k) + f(t_k, y_k)f_y(t_k, y_k)] + \\ &+ \frac{a_2^2c_2h^3}{2!}[f_{tt}(t_k, y_k) + 2f(t_k, y_k)f_{ty}(t_k, y_k) + f^2(t_k, y_k)f_{yy}(t_k, y_k)]. \end{aligned} \tag{3.25}$$

Comparando (3.22) e (3.25), conclui-se que

$$\boxed{\begin{cases} c_1 + c_2 = 1 \\ a_2c_2 = \frac{1}{2!} \end{cases}}.$$

Assim, se $c_1 = c_2 = \frac{1}{2}$, então $a_2 = 1$ e tem-se o *Método de Euler Aprimorado*; se $c_1 = 0$ e $c_2 = 1$, então $a_2 = \frac{1}{2}$ e tem-se o *Método de Euler Modificado*.

Note-se que não é possível estabelecer concordância entre os termos de terceira ordem de (3.22) e (3.25). Em (3.25) há três termos com o fator h^3 , enquanto que em (3.22) há cinco termos com esse fator. Portanto, a ordem não pode exceder dois.

2. Deduza o método de Runge-Kutta de terceira ordem no qual $c_2 = c_3$ e $a_2 = a_3$.

Solução

$$\begin{aligned} y_{k+1} &= y_k + h \sum_{r=1}^3 c_r \kappa_r \\ &= y_k + h (c_1 \kappa_1 + c_2 \kappa_2 + c_3 \kappa_3). \end{aligned}$$

Como $c_2 = c_3$, tem-se que

$$y_{k+1} = y_k + h (c_1 \kappa_1 + c_2 \kappa_2 + c_2 \kappa_3), \quad (3.26)$$

sendo

$$\begin{aligned} \kappa_1 &= f(t, y), \\ \kappa_2 &= f(t + ha_2, y + hb_{21}\kappa_1), \\ \kappa_3 &= f(t + ha_3, y + hb_{31}\kappa_1 + hb_{32}\kappa_2). \end{aligned}$$

De $a_r = \sum_{s=1}^{r-1} b_{rs}$ e $a_2 = a_3$, obtém-se $a_2 = b_{21}$, $a_3 = b_{31} + b_{32}$ e $a_2 = b_{31} + b_{32}$, o que implica que $b_{31} = a_2 - b_{32}$.

Assim:

$$\begin{aligned} \kappa_1 &= f(t, y); \\ \kappa_2 &= f(t + ha_2, y + ha_2\kappa_1); \\ \kappa_3 &= f(t + ha_2, y + h(a_2 - b_{32})\kappa_1 + hb_{32}\kappa_2). \end{aligned}$$

Supondo que $f(t, y)$ tem derivadas parciais contínuas até a ordem necessária, expande-se κ_2 e κ_3 em torno do ponto (t, y) até o termo de segunda ordem. Durante a expansão, todos os termos de terceira ordem são desprezados porque na substituição em (3.26) esses termos têm ordem quatro.

- Expansão de κ_2 :

$$\begin{aligned}
\kappa_2 &= f + ha_2(f_t + \kappa_1 f_y) + \frac{h^2 a_2^2}{2}(f_{tt} + 2\kappa_1 f_{ty} + \kappa_1^2 f_{yy}); \\
\kappa_2 &= f + ha_2(f_t + f f_y) + \frac{h^2 a_2^2}{2}(f_{tt} + 2f f_{ty} + f^2 f_{yy}); \\
\kappa_2 &= f + ha_2 F + \frac{h^2 a_2^2}{2} G, \tag{3.27}
\end{aligned}$$

onde $F = f_t + f f_y$ e $G = f_{tt} + 2f f_{ty} + f^2 f_{yy}$.

- Expansão de κ_3 :

$$\begin{aligned}
\kappa_3 &= f + h \{a_2 f_t + [(a_2 - b_{32})\kappa_1 + b_{32}\kappa_2] f_y\} + \\
&+ \frac{h^2}{2} \{a_2^2 f_{tt} + 2a_2 [(a_2 - b_{32})\kappa_1 + b_{32}\kappa_2] f_{ty} + [(a_2 - b_{32})\kappa_1 + b_{32}\kappa_2]^2 f_{yy}\}; \\
\kappa_3 &= f + h \underbrace{\{a_2 f_t + [(a_2 - b_{32})f + b_{32}\kappa_2] f_y\}}_{(A)} + \\
&+ \frac{h^2}{2} \underbrace{\{a_2^2 f_{tt} + 2a_2 [(a_2 - b_{32})f + b_{32}\kappa_2] f_{ty} + [(a_2 - b_{32})f + b_{32}\kappa_2]^2 f_{yy}\}}_{(B)}. \tag{3.28}
\end{aligned}$$

– Desenvolvendo-se (A):

$$\begin{aligned}
A &= a_2 f_t + [(a_2 - b_{32})f + b_{32}\kappa_2] f_y \\
&= a_2 f_t + a_2 f f_y - b_{32} f f_y + b_{32} f_y \left(f + ha_2 F + \frac{h^2 a_2^2}{2} G \right) \\
&= a_2 f_t + a_2 f f_y - b_{32} f f_y + b_{32} f f_y + ha_2 b_{32} F f_y + \frac{h^2 a_2^2}{2} b_{32} G f_y \\
&= a_2 (f_t + f f_y) - b_{32} f f_y + b_{32} f f_y + ha_2 b_{32} F f_y + \frac{h^2 a_2^2}{2} b_{32} G f_y \\
&= a_2 F + ha_2 b_{32} F f_y + \frac{h^2 a_2^2 b_{32}}{2} G f_y.
\end{aligned}$$

$$A = a_2 F + ha_2 b_{32} F f_y + \frac{h^2 a_2^2 b_{32}}{2} G f_y \tag{3.29}$$

– Desenvolvendo-se (B):

$$\begin{aligned}
 B &= a_2^2 f_{tt} + 2a_2 [(a_2 - b_{32})f + b_{32}\kappa_2] f_{ty} + \\
 &+ [(a_2 - b_{32})f + b_{32}\kappa_2]^2 f_{yy} \\
 &= a_2^2 f_{tt} + 2a_2^2 f f_{ty} - 2a_2 b_{32} f f_{ty} + \underbrace{2a_2 b_{32} f_{ty} \kappa_2}_{B_1} + \\
 &+ \underbrace{[(a_2 - b_{32})f + b_{32}\kappa_2]^2 f_{yy}}_{B_2}. \quad (3.30)
 \end{aligned}$$

Substituindo (3.27) em B_1 e B_2 , chega-se a:

$$B_1 = 2a_2 b_{32} f f_{ty} + 2ha_2^2 b_{32} F f_{ty};$$

$$B_2 = a_2^2 f^2 f_{yy} + 2ha_2^2 b_{32} F f f_{yy}.$$

Em B_1 e B_2 , descartaram-se termos com fator h^p , $p > 1$. Substituindo B_1 e B_2 em (3.30), obtém-se

$$B = a_2^2 G + 2ha_2^2 b_{32} F f_{ty} + 2ha_2^2 b_{32} F f f_{yy}. \quad (3.31)$$

Substituindo (3.29) e (3.31) em (3.28), tem-se que

$$\begin{aligned}
 \kappa_3 &= f + h \left(a_2 F + ha_2 b_{32} F f_y + \frac{h^2 a_2^2 b_{32}}{2} G f_y \right) \\
 &+ \frac{h^2}{2} (a_2^2 G + 2ha_2^2 b_{32} F f_{ty} + 2ha_2^2 b_{32} F f f_{yy}) \\
 \kappa_3 &= f + ha_2 F + \frac{h^2}{2} (a_2^2 G + 2a_2 b_{32} F f_y). \quad (3.32)
 \end{aligned}$$

Em (3.32), descartou-se o termo com fator h^3 .

Substituindo κ_1 , κ_2 e κ_3 em (3.26), tem-se que

$$\begin{aligned}
 y_{k+1} &= y_k + h \left[c_1 f + c_2 \left(f + ha_2 F + \frac{h^2 a_2^2}{2} G \right) \right] + \\
 &+ hc_2 \left[f + ha_2 F + \frac{h^2}{2} (a_2^2 G + 2a_2 b_{32} F f_y) \right] \\
 y_{k+1} &= y_k + h(c_1 + 2c_2)f + 2h^2 a_2 c_2 F + h^3 (a_2^2 c_2 + a_2 b_{32} c_2 F f_y). \quad (3.33)
 \end{aligned}$$

O Método da Série de Taylor de terceira ordem é dado por

$$y_{k+1} = y_k + hf + \frac{h^2}{2} F + \frac{h^3}{6} (G + F f_y). \quad (3.34)$$

Para que (3.33) coincida com (3.34), deve-se ter

$$\left\{ \begin{array}{l} c_1 + 2c_2 = 1 \\ 2a_2c_2 = \frac{1}{2} \\ a_2^2c_2 = \frac{1}{6} \\ a_2^2b_{32}c_2 = \frac{1}{6} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} a_2 = \frac{2}{3} \\ b_{32} = \frac{2}{3} \\ c_1 = \frac{1}{4} \\ c_2 = \frac{3}{8} \end{array} \right. .$$

Com os valores anteriores, tem-se que $a_3 = \frac{2}{3}$, $c_3 = \frac{3}{8}$ e $b_{31} = 0$. Assim, o Método de Runge-Kutta de terceira ordem, com as condições impostas, é dado por

$$y_{k+1} = y_k + \frac{h}{4} \left(\kappa_1 + \frac{3}{2}\kappa_2 + \frac{3}{2}\kappa_3 \right) \left\{ \begin{array}{l} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{2h}{3}, y + \frac{2h}{3}\kappa_1\right) \\ \kappa_3 = f\left(t + \frac{2h}{3}, y + \frac{2h}{3}\kappa_2\right) \end{array} \right. .$$

3. Mostre que o Método de Runge-Kutta Clássico (RK44)

$$\left\{ \begin{array}{l} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h), \quad t_{k+1} = t_k + h, \quad 0 \leq k \leq n-1 \end{array} \right. ,$$

$$\Phi(t_k, y_k, h) = \frac{1}{6}(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4), \quad \left\{ \begin{array}{l} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_1\right) \\ \kappa_3 = f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_2\right) \\ \kappa_4 = f\left(t_k + h, y_k + h\kappa_3\right) \end{array} \right.$$

e $h = \frac{(b-a)}{n}$, aplicado à equação diferencial $\dot{y} = py$, p constante, fornece

$$\frac{y_{k+1}}{y_k} = e^{ph} + O(ph)^5.$$

Solução

Calculando κ_1 , κ_2 , κ_3 e κ_4 para a equação $\dot{y} = py$, tem-se que

$$\begin{aligned}
\kappa_1 &= f(t_k, y_k) = py_k, \\
\kappa_2 &= f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_1\right) = p\left(y_k + \frac{h}{2}py_k\right) = \left(p + \frac{p^2h}{2}\right)y_k, \\
\kappa_3 &= f\left(t_k + \frac{h}{2}, y_k + \frac{h}{2}\kappa_2\right) = p\left[y_k + \frac{h}{2}\left(p + \frac{p^2h}{2}\right)y_k\right] \\
&= \left(p + \frac{p^2h}{2} + \frac{p^3h^2}{4}\right)y_k, \\
\kappa_4 &= f\left(t_k + h, y_k + h\kappa_3\right) = p\left[y_k + h\left(p + \frac{p^2h}{2} + \frac{p^3h^2}{4}\right)y_k\right] \\
&= \left(p + p^2h + \frac{p^3h^2}{2} + \frac{p^4h^3}{4}\right)y_k.
\end{aligned}$$

Aplicando os resultados acima no Método de Runge-Kutta de quarta ordem, obtém-se

$$\begin{aligned}
y_{k+1} &= y_k + \frac{h}{6}(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4), \\
y_{k+1} &= y_k + \\
&+ \frac{h}{6}\left[py_k + 2\left(p + \frac{p^2h}{2}\right)y_k + 2\left(p + \frac{p^2h}{2} + \frac{p^3h^2}{4}\right)y_k + \left(p + p^2h + \frac{p^3h^2}{2} + \frac{p^4h^3}{4}\right)y_k\right], \\
y_{k+1} &= \left(1 + \frac{ph}{6} + \frac{ph}{3} + \frac{p^2h^2}{6} + \frac{ph}{3} + \frac{p^2h^2}{6} + \frac{p^3h^3}{12} + \frac{ph}{6} + \frac{p^2h^2}{6} + \frac{p^3h^3}{12} + \frac{p^4h^4}{24}\right)y_k, \\
y_{k+1} &= \left(1 + ph + \frac{p^2h^2}{2} + \frac{p^3h^3}{6} + \frac{p^4h^4}{24}\right)y_k, \\
\frac{y_{k+1}}{y_k} &= 1 + ph + \frac{p^2h^2}{2!} + \frac{p^3h^3}{3!} + \frac{p^4h^4}{4!}. \tag{3.35}
\end{aligned}$$

Como

$$e^{ph} = \sum_{n=0}^{\infty} \frac{(ph)^n}{n!} = 1 + ph + \frac{p^2h^2}{2!} + \frac{p^3h^3}{3!} + \frac{p^4h^4}{4!} + O((ph)^5),$$

pode-se reescrever (3.35) como

$$\boxed{\frac{y_{k+1}}{y_k} = e^{ph} + O(p^5h^5).$$

4. Considere o problema de valor inicial

$$\begin{cases} \dot{y} = e^{2t}, & 0 \leq t \leq 1 \\ y(0) = \frac{1}{2} \end{cases}.$$

(a) Estime h para que o erro local de truncamento para o Método de Euler seja menor que 10^{-4} .

- (b) A estimativa para h obtida no item (a) pode ser usada como aproximação inicial do passo de integração para o Método de Runge-Kutta-Fehlberg? Justifique.

Solução

- (a) O Método de Euler

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + hf(t_k, y_k) \end{cases}$$

tem erro local de discretização dado por

$$\alpha_k = \frac{y(t_{k+1}) - y(t_k)}{h} - f(t_k, y(t_k)). \quad (3.36)$$

Truncando-se no segundo termo a expansão em série de Taylor de $y(t_{k+1})$ com centro em $t = t_k$, obtém-se

$$y(t_{k+1}) = y(t_k) + (t_{k+1} - t_k)y'(t_k) + \frac{(t_{k+1} - t_k)^2}{2!}y''(\xi), \quad (3.37)$$

com $\xi \in (t_k, t_{k+1})$.

Denotando-se $h = t_{k+1} - t_k$ e substituindo-se (3.37) em (3.36), tem-se que

$$\begin{aligned} \alpha_k &= \frac{y(t_k) + hy'(t_k) + \frac{h^2}{2!}y''(\xi) - y(t_k)}{h} - f(t_k, y(t_k)) \\ &= y'(t_k) + \frac{h}{2!}y''(\xi) - f(t_k, y(t_k)) \\ &= \frac{h}{2}y''(\xi). \end{aligned} \quad (3.38)$$

Assim,

$$\begin{aligned} |\alpha_k| &= \frac{h}{2} |y''(\xi)|, \\ \max_k |\alpha_k| &\leq \frac{h}{2} \max_{\xi \in I} |y''(\xi)|, \\ \frac{h}{2} \max_{\xi \in I} |y''(\xi)| &< 10^{-4}, \\ h &< \frac{2 \cdot 10^{-4}}{\max_{\xi \in I} |y''(\xi)|}. \end{aligned} \quad (3.39)$$

Ainda,

$$\begin{aligned} I &= [0, 1], \\ y(t) &= \frac{e^{2t}}{2}, \\ y'(t) &= e^{2t}, \\ y''(t) &= 2e^{2t}. \end{aligned} \quad (3.40)$$

Como a função (3.40) é estritamente crescente no intervalo $[0, 1]$, assumindo o valor máximo $2e^2$ em $t = 1$, pode-se reescrever (3.39) como

$$h < \frac{2 \cdot 10^{-4}}{2e^2},$$

$$\boxed{h < \frac{10^{-4}}{e^2} \approx 1,35335 \times 10^{-5}.} \quad (3.41)$$

(b) Uma aproximação para o erro local de discretização do Método de Runge-Kutta-Felhberg é dada por

$$\alpha_k^{RKF} \approx \frac{1}{360}\kappa_1 - \frac{128}{4275}\kappa_3 - \frac{2197}{75240}\kappa_4 + \frac{1}{50}\kappa_5 + \frac{2}{55}\kappa_6. \quad (3.42)$$

Calculando (3.42) com $f(t, y(t)) = e^{2t}$, $t_0 = 0$ e $h = \frac{10^{-4}}{e^2}$, obtém-se

$$\kappa_1 = (t_0, y_0) = e^{2t_0} = e^0 = 1,$$

$$\begin{aligned} \kappa_2 &= f\left(t_0 + \frac{h}{4}, y_0 + \frac{h}{4}\kappa_1\right) \\ &= e^{2(t_0 + \frac{h}{4})} = e^{\frac{h}{2}} \approx 1,000006767, \end{aligned}$$

$$\begin{aligned} \kappa_3 &= f\left(t_0 + \frac{3h}{8}, y_0 + \frac{3h}{32}\kappa_1 + \frac{9h}{32}\kappa_2\right) \\ &= e^{2(t_0 + \frac{3h}{8})} = e^{\frac{3h}{4}} \approx 1,00001015, \end{aligned}$$

$$\begin{aligned} \kappa_4 &= f\left(t_0 + \frac{12h}{13}, y_0 + \frac{1932h}{2197}\kappa_1 - \frac{7200h}{2197}\kappa_2 + \frac{7296h}{2197}\kappa_3\right) \\ &= e^{2(t_0 + \frac{12h}{13})} = e^{\frac{24h}{13}} \approx 1,000024985, \end{aligned}$$

$$\begin{aligned} \kappa_5 &= f\left(t_0 + h, y_0 + \frac{439h}{216}\kappa_1 - 8h\kappa_2 + \frac{3680h}{513}\kappa_3 - \frac{845h}{4104}\kappa_4\right) \\ &= e^{2(t_0 + h)} = e^{2h} \approx 1,000027067, \end{aligned}$$

$$\begin{aligned}\kappa_6 &= f\left(t_0 + \frac{h}{2}, y_0 - \frac{8h}{27}\kappa_1 + 2h\kappa_2 - \frac{3544h}{2565}\kappa_3 + \frac{1859h}{4104}\kappa_4 - \frac{11h}{40}\kappa_5\right) \\ &= e^{2(t_0 + \frac{h}{2})} = e^h \approx 1,000013534,\end{aligned}$$

$$\alpha_0^{RKF} \approx 1,96784 \times 10^{-11}. \quad (3.43)$$

Como no Método de Runge-Kutta Fehlberg tem-se que

$$|\alpha_0^{RKF}| = 1,96784 \times 10^{-11} < \epsilon = 10^{-4}, \quad (3.44)$$

pode-se usar $\frac{10^{-4}}{e^2}$ como valor inicial para o passo de integração h . O mesmo se verifica para $h < \frac{10^{-4}}{e^2}$.

5. Interprete geometricamente o Método de Euler Modificado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h); \quad t_{k+1} = t_k + h; \quad 0 \leq k \leq n-1 \end{cases}$$

$$\Phi(t, y, h) = \kappa_2 \quad \begin{cases} \kappa_1 = f(t, y) \\ \kappa_2 = f\left(t + \frac{h}{2}, y + \frac{h}{2}\kappa_1\right) \end{cases}$$

e o Método de Euler Aprimorado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h); \quad t_{k+1} = t_k + h; \quad 0 \leq k \leq n-1 \end{cases}$$

$$\Phi(t, y, h) = \frac{\kappa_1 + \kappa_2}{2} \quad \begin{cases} \kappa_1 = f(t, y) \\ \kappa_2 = f(t + h, y + h\kappa_1) \end{cases}$$

ambos com $h = \frac{b-a}{n}$.

Solução

Em ambos os casos, considere-se o p.v.i. $\dot{y} = f(t, y)$ com $y(t_0) = y_0$ e solução $y(t)$. Ressalta-se que a situação ideal das ilustrações é quando $(t_k, y_k) = (t_0, y_0)$.

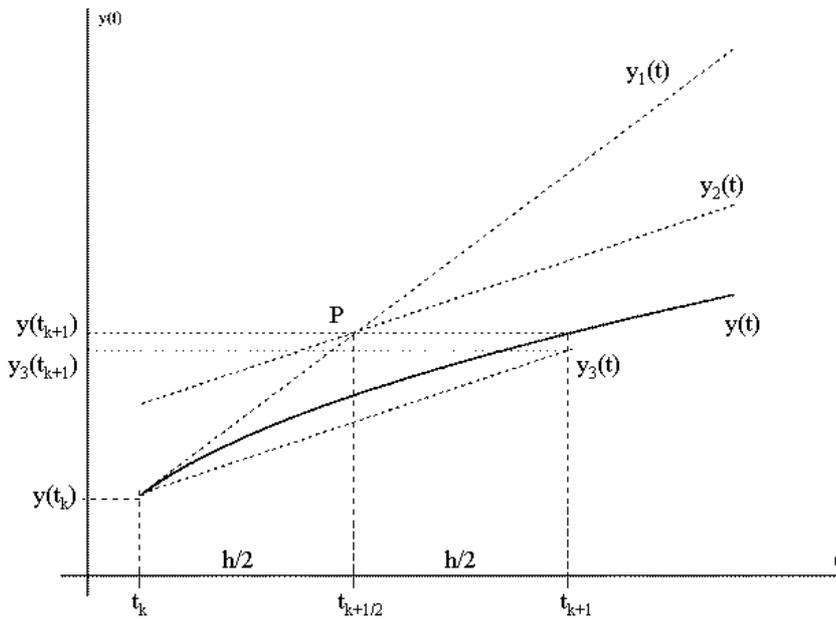


Figura 3.1: Interpretação geométrica do Método de Euler Modificado.

(a) **Método de Euler Modificado**

Na Figura (3.1), $y(t)$ é o gráfico da solução analítica. Seja $y_1(t)$ a reta que passa pelo ponto (t_k, y_k) e que tem inclinação $f(t_k, y_k)$. Com isso, tem-se

$$y_1(t) = y_k + (t - t_k)f(t_k, y_k). \quad (3.45)$$

As coordenadas do ponto P para $t = t_{k+\frac{1}{2}}$ em $y_1(t)$ são dadas por $(t_{k+\frac{1}{2}}, y_k + \frac{h}{2}f(t_k, y_k))$.

Seja agora $y_2(t)$ a reta que passa pelo ponto P e tem inclinação $f(t_{k+\frac{1}{2}}, y_k + \frac{h}{2}f(t_k, y_k))$. Logo,

$$y_2(t) = \left[y_k + \frac{h}{2}f(t_k, y_k) \right] + (t - t_{k+\frac{1}{2}})f(t_{k+\frac{1}{2}}, y_k + \frac{h}{2}f(t_k, y_k)). \quad (3.46)$$

Seja agora $y_3(t)$ a reta que passa por (t_k, y_k) e tem por inclinação a inclinação da reta $y_2(t)$. Assim,

$$y_3(t) = y_k + (t - t_k)f\left(t_{k+\frac{1}{2}}, y_k + \frac{h}{2}f(t_k, y_k)\right). \quad (3.47)$$

Calculando $y_3(t_{k+1})$, obtém-se uma aproximação para o valor que a solução $y(t)$ assume no ponto t_{k+1} . Isso define o Método de Euler Modificado, dado por

$$y_{k+1} = y_k + hf\left(t_{k+\frac{1}{2}}, y_k + \frac{h}{2}f(t_k, y_k)\right).$$

(b) **Método de Euler Aprimorado**

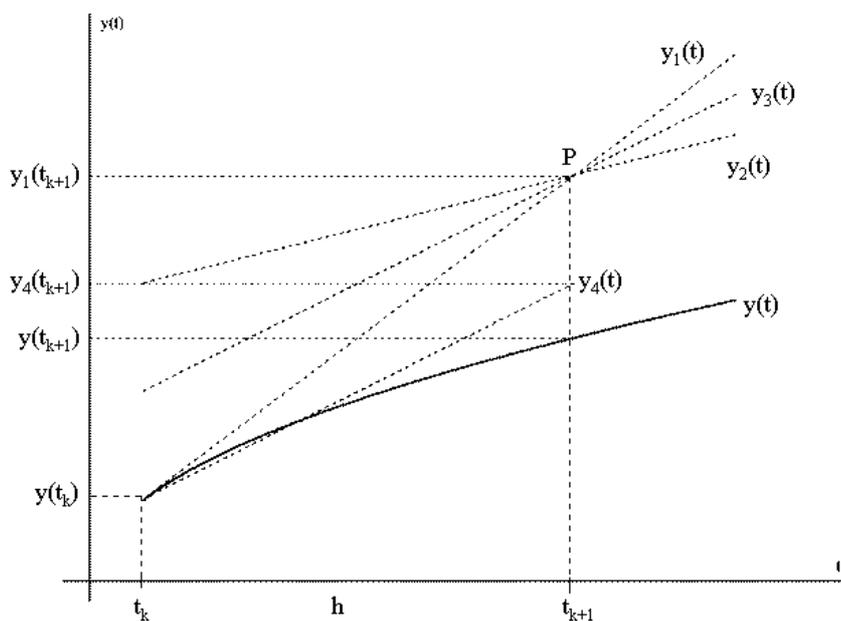


Figura 3.2: Interpretação geométrica do Método de Euler Aprimorado.

Na Figura (3.2), $y(t)$ é o gráfico da solução analítica. Seja $y_1(t)$ a reta que passa pelo ponto (t_k, y_k) e que tem inclinação $f(t_k, y_k)$. Dessa forma, tem-se

$$y_1(t) = y_k + (t - t_k)f(t_k, y_k). \quad (3.48)$$

As coordenadas do ponto P para $t = t_{k+1}$ em $y_1(t)$ são dadas por $(t_{k+1}, y_k + hf(t_k, y_k))$.

Seja agora $y_2(t)$ a reta que passa pelo ponto P e tem inclinação $f(t_{k+1}, y_k + hf(t_k, y_k))$. Portanto,

$$y_2(t) = [y_k + hf(t_k, y_k)] + (t - t_{k+1})f(t_{k+1}, y_k + hf(t_k, y_k)). \quad (3.49)$$

Observe-se que $y_1(t_{k+1})$ é a aproximação gerada pelo Método de Euler.

Seja agora $y_3(t)$ a reta que passa por P e tem por inclinação a média das inclinações das retas $y_1(t)$ e $y_2(t)$. Logo,

$$\begin{aligned} y_3(t) &= [y_k + hf(t_k, y_k)] + \\ &+ (t - t_{k+1}) \frac{f(t_k, y_k) + f(t_{k+1}, y_k + hf(t_k, y_k))}{2}. \end{aligned} \quad (3.50)$$

Considere-se agora a reta $y_4(t)$ que passa pelo ponto (t_k, y_k) e tem a mesma inclinação de $y_3(t)$. Assim,

$$y_4(t) = y_k + (t - t_k) \frac{f(t_k, y_k) + f(t_{k+1}, y_k + hf(t_k, y_k))}{2}. \quad (3.51)$$

Calculando $y_4(t_{k+1})$, obtém-se uma aproximação para o valor que a solução $y(t)$ assume no ponto t_{k+1} . Observe-se pela Figura (3.2) que essa aproximação é mais precisa que a aproximação gerada pelo Método de Euler. Isso define o Método de Euler Aprimorado, dado por

$$y_{k+1} = y_k + \frac{h}{2} [f(t_k, y_k) + f(t_k + h, y_k + hf(t_k, y_k))].$$

Capítulo 4

Estabilidade dos métodos de passo único

Seja o Problema de Cauchy

$$\begin{cases} \frac{d}{dt}y(t) = -10y(t), & t \in [2, 6] \\ y(2) = 1000 \end{cases}, \quad (4.1)$$

cuja solução exata é $y(t) = 1000e^{-10t+20}$ (verifique!).

A obtenção de uma solução numérica para (4.1) demanda, na prática, a escolha de um passo de integração $h > 0$ além, é claro, de um método numérico. Por exemplo, considerem-se as aproximações obtidas com o Método de Euler (1.24)

$$\begin{aligned} y_{k+1} &= y_k + hf(t_k, y_k) \\ &= y_k + h(-10y_k) \\ &= (1 - 10h)y_k, \end{aligned}$$

com passos de integração $h = 0,125$ e $h = 0,5$. A Tabela (4.1) mostra o erro global de discretização para cada um dos passos de integração.

Da análise da Tabela (4.1), constata-se que o Método de Euler produz um erro global de discretização aceitável para $h = 0,125$, porém inadmissível para $h = 0,5$. Qual é então o motivo para o comportamento da solução numérica em ambos os casos? A escolha apropriada do passo de integração h está associada ao conceito de *estabilidade absoluta* (h fixo).

4.1 Estabilidade absoluta

Para compreender a origem dos problemas de estabilidade que se pode ter na escolha de um passo de integração h para um determinado método numérico, analisa-

		h = 0,125	h = 0,5
<i>k</i>	<i>t_k</i>	<i>y(t_k)</i>	$ y(t_k) - y_k $
0	2,0	1,000E+03	0,000E+00
1	2,5	6,738E+00	2,832E+00
2	3,0	4,540E-02	3,014E-02
3	3,5	3,059E-04	2,463E-04
4	4,0	2,061E-06	1,828E-06
5	4,5	1,389E-08	1,298E-08
6	5,0	9,358E-11	9,003E-11
7	5,5	6,305E-13	6,166E-13
8	6,0	4,248E-15	4,194E-15

Tabela 4.1: Erros globais de discretização produzidos pelo Método de Euler na solução numérica de (4.1) para dois passos de integração.

se o Método de Euler

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + hf(t_k, y_k) \end{cases}$$

aplicado ao Problema de Cauchy modelo

$$\begin{cases} \frac{d}{dt}y(t) = \lambda y \\ y(t_0) = y_0 \end{cases},$$

cuja solução exata é dada por $y(t) = y_0 e^{\lambda(t-t_0)}$ (verifique!).

Assim, tem-se que

$$\begin{aligned} y_{k+1} &= y_k + hf(t_k, y_k) \\ &= y_k + \lambda h y_k \\ &= (1 + \lambda h)y_k. \end{aligned} \tag{4.2}$$

Estabelecendo em (4.2) uma dependência da condição inicial y_0 , chega-se a

$$\begin{aligned} y_1 &= (1 + \lambda h) y_0, \\ y_2 &= (1 + \lambda h)^2 y_0, \\ y_3 &= (1 + \lambda h)^3 y_0, \\ &\vdots \\ y_k &= (1 + \lambda h)^k y_0, \end{aligned} \tag{4.3}$$

onde o parâmetro λ pode ser real ou complexo.

Na análise da estabilidade, há duas situações possíveis para $\lambda \in \mathbb{R}$: $\lambda < 0$ ou $\lambda \geq 0$.

- Se $\lambda < 0$, então $\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y_0 e^{\lambda(t-t_0)} = 0$.

A solução numérica y_k terá esse comportamento se

$$|1 + \lambda h| < 1 \Rightarrow -1 < 1 + \lambda h < 1 \Rightarrow -2 < \lambda h < 0.$$

Assim, a solução numérica tem o mesmo comportamento da solução analítica se $\lambda h \in (-2, 0)$. Neste caso, existe uma restrição para os valores de h de modo que haja estabilidade.

- Se $\lambda \geq 0$,

$$\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y_0 e^{\lambda(t-t_0)} = \begin{cases} y_0, & \text{se } \lambda = 0 \\ \infty, & \text{se } \lambda > 0 \end{cases}.$$

Neste caso, a solução numérica

$$y_k = (1 + \lambda h)^k y_0$$

tem o mesmo comportamento da solução analítica independentemente dos valores do passo de integração h .

Portanto, o Método de Euler é um método **condicionalmente estável** cuja região de estabilidade absoluta é dada por $\lambda h \in (-2, 0)$.

Observação

Para $\lambda \in \mathbb{C} \Rightarrow \lambda h \in \mathbb{C}$, tem-se que:

$$\begin{aligned} z &= \lambda h = a + bi, \quad a, b \in \mathbb{R} \\ |1 + \lambda h| < 1 &\Rightarrow |z + 1| < 1 \\ &\Rightarrow |a + bi + 1| < 1 \\ &\Rightarrow |(a + 1) + bi| < 1 \\ &\Rightarrow \sqrt{(a + 1)^2 + b^2} < 1 \\ &\Rightarrow (a + 1)^2 + b^2 < 1 \\ &\Rightarrow [a - (-1)]^2 + b^2 < 1. \end{aligned} \quad (4.4)$$

Em (4.4), tem-se o conjunto dos pontos interiores a um disco centrado no ponto $(-1, 0)$ e de raio 1. O intervalo de estabilidade é definido por $Re(\lambda h) \in (-2, 0)$, enquanto que $Im(\lambda h)$ é responsável apenas por um comportamento oscilatório da solução (Veja Schwarz [17]).

Definição 4.1 (Estabilidade absoluta para os métodos de passo único). *Seja um método de passo único que, aplicado ao Problema de Cauchy modelo*

$$\begin{cases} \frac{d}{dt}y(t) = \lambda y(t) \\ y(t_0) = y_0 \end{cases},$$

conduz a

$$y_{k+1} = \psi(\lambda h)y_k.$$

O conjunto

$$\Omega = \{\mu \in \mathbb{C}; |\psi(\mu)| < 1\}$$

é denominado região de estabilidade absoluta (h fixo) e $\psi(\lambda h)$ é o fator de amplificação. A intersecção da região Ω com a reta real determina o intervalo de estabilidade absoluta do método de passo único.

4.1.1 Exercícios

Exercício 4.1. Explique a estabilidade (ou instabilidade) do Método de Euler aplicado à solução do p.v.i. (4.1) com passos de integração $h = 0,5$ e $h = 0,125$.

Exercício 4.2. Determine o intervalo de estabilidade absoluta para o Método de Euler Modificado.

Exercício 4.3. Determine o intervalo de estabilidade absoluta para o Método de Euler Aprimorado.

Exercício 4.4. A Tabela (4.2) traz o fator de amplificação e o intervalo de estabilidade absoluta para os Métodos de Runge-Kutta de ordem R com R estágios [10].

R	Fator de amplificação	Intervalo de estabilidade absoluta
1	$1 + \lambda h$	$(-2, 0)$
2	$1 + \lambda h + \frac{(\lambda h)^2}{2}$	$(-2, 0)$
3	$1 + \lambda h + \frac{(\lambda h)^2}{2} + \frac{(\lambda h)^3}{6}$	$(-2,51, 0)$
4	$1 + \lambda h + \frac{(\lambda h)^2}{2} + \frac{(\lambda h)^3}{6} + \frac{(\lambda h)^4}{24}$	$(-2,78, 0)$

Tabela 4.2: Intervalos de estabilidade absoluta para os Métodos de Runge-Kutta de ordem R com R estágios.

Calcule o fator de amplificação para o Método de Runge-Kutta Clássico (RK4).

Exercício 4.5. Comprove o efeito da estabilidade (ou instabilidade) usando o Método de Runge-Kutta de quarta ordem com quatro estágios (3.14) para calcular a solução numérica do problema de valor inicial

$$\begin{cases} \frac{d}{dt}y(t) = -5ty^2(t) + \frac{5}{t} - \frac{1}{t^2}, & 1 \leq t \leq 4 \\ y(1) = 1 \end{cases} \quad (4.5)$$

com $h = 0,2$ e $h = 0,4$.

Sugestão: Mostre que a solução exata do problema de valor inicial (4.5) é $y(t) = \frac{1}{t}$. Com a solução exata e a solução numérica, calcule o erro global de discretização.

4.2 Suplemento teórico

4.2.1 Instabilidade inerente

Seja o problema de valor inicial

$$\begin{cases} \frac{d}{dt}y(t) = y(t) - t, & t \in [0, 5] \\ y(0) = 1 \end{cases} \quad (4.6)$$

Em (4.6), tem-se uma equação diferencial ordinária linear, de primeira ordem, não homogênea. A solução exata de (4.6) é dada por $y(t) = t + 1$.

Fator integrante: $e^{\int -dt} = e^{-t}$.

$$\begin{aligned} e^{-t} \left[\frac{d}{dt}y(t) - y(t) \right] &= -te^{-t} \\ e^{-t} \frac{d}{dt}y(t) - e^{-t}y(t) &= -te^{-t} \\ \frac{d}{dt} [e^{-t}y(t)] &= -te^{-t} \\ \int \frac{d}{dt} [e^{-t}y(t)] dt &= - \int te^{-t} dt \\ e^{-t}y(t) &= - \int te^{-t} dt \end{aligned} \quad (4.7)$$

Integrando por partes o lado direito de (4.7), tem-se que:

$$\begin{aligned} u = t &\Rightarrow du = dt \\ dv = e^{-t} dt &\Rightarrow v = -e^{-t} \\ \int te^{-t} dt &= -te^{-t} + \int e^{-t} dt \\ &= -te^{-t} - e^{-t} \\ &= -(t+1)e^{-t}. \end{aligned} \quad (4.8)$$

Substituindo (4.8) em (4.7), chega-se a:

$$\begin{aligned} e^{-t}y(t) &= (t+1)e^{-t} + C; \\ y(t) &= t+1 + Ce^t. \end{aligned} \quad (4.9)$$

Usando a condição inicial em (4.9), obtém-se

$$y(0) = 1 + C \Rightarrow C = 0.$$

Logo,

$$\boxed{y(t) = t + 1.} \quad (4.10)$$

Perturbando a condição inicial em (4.6) em 1%, isto é, $y(0) = 1 \pm 0,01$, tem-se os problemas de valor inicial

$$\begin{cases} \frac{d}{dt}y(t) = y(t) - t, & t \in [0, 5] \\ y(0) = 0,99 \end{cases} \quad (4.11)$$

e

$$\begin{cases} \frac{d}{dt}y(t) = y(t) - t, & t \in [0, 5] \\ y(0) = 1,01 \end{cases} \quad (4.12)$$

As soluções (refaça o cálculo da constante C em (4.9)) de (4.11) e (4.12) são dadas, respectivamente, por

$$y(t) = -0,01e^t + t + 1 \quad e \quad (4.13)$$

$$y(t) = 0,01e^t + t + 1. \quad (4.14)$$

Considere-se $t = 5$ em (4.10), (4.13) e (4.14):

$$t = 5 \text{ em (4.10)} \Rightarrow y(5) = 5 + 1 = 6;$$

$$t = 5 \text{ em (4.13)} \Rightarrow y(5) = -0,01e^5 + 5 + 1 \approx 4,5;$$

$$t = 5 \text{ em (4.14)} \Rightarrow y(5) = 0,01e^5 + 5 + 1 \approx 7,5.$$

Logo, perturbando-se a condição inicial em (4.6) em 1%, a solução varia cerca de 25%.

No exemplo dado, nenhum método numérico será capaz de produzir um erro inferior a 25% se a condição inicial for perturbada em 1%. Este é um problema de estabilidade (ou instabilidade) intrínseco ao problema de valor inicial e, por este motivo, denominado de *instabilidade inerente*.

4.3 Exercícios resolvidos

1. Determine a região de estabilidade absoluta para o Método do Trapézio.

Solução

p.v.i.:

$$\begin{cases} \frac{d}{dt}y(t) = \lambda y \\ y(t_0) = y_0 \end{cases}$$

$$y(t) = y_0 e^{\lambda(t-t_0)}$$

$$\begin{aligned} y_{k+1} &= y_k + h\Phi(t_k, y_k, y_{k+1}, h) \\ y_{k+1} &= y_k + h \frac{f(t_k, y_k) + f(t_{k+1}, y_{k+1})}{2} \\ y_{k+1} &= y_k + h \frac{\lambda y_k + \lambda y_{k+1}}{2} \\ \left(1 - \frac{\lambda h}{2}\right) y_{k+1} &= \left(1 + \frac{\lambda h}{2}\right) y_k \\ y_{k+1} &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_k \end{aligned} \tag{4.15}$$

Estabelecendo-se em (4.15) uma dependência da condição inicial y_0 :

$$\begin{aligned} y_1 &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_0; \\ y_2 &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_1 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right) \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_0 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^2 y_0; \\ y_3 &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_2 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right) \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^2 y_0 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^3 y_0; \\ y_4 &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right) y_3 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right) \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^3 y_0 = \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^4 y_0; \\ &\vdots \\ y_k &= \left(\frac{2 + \lambda h}{2 - \lambda h}\right)^k y_0, \end{aligned} \tag{4.16}$$

onde o parâmetro λ pode ser real ou complexo.

Na análise da estabilidade, há duas situações possíveis para $\lambda \in \mathbb{R}$: $\lambda < 0$ ou $\lambda \geq 0$.

- Se $\lambda < 0$, então $\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y_0 e^{\lambda(t-t_0)} = 0$.

A solução numérica y_k terá esse comportamento se

$$\left| \frac{2 + \lambda h}{2 - \lambda h} \right| < 1 \Rightarrow -1 < \frac{2 + \lambda h}{2 - \lambda h} < 1 \Rightarrow \begin{cases} 2 + \lambda h < 2 - \lambda h \\ -2 + \lambda h < 2 + \lambda h \end{cases} \Rightarrow \lambda h < 0.$$

Assim, a solução numérica tem o mesmo comportamento da solução analítica se $\lambda h \in (-\infty, 0)$. Neste caso, não há nenhuma restrição para os valores de h de modo que haja estabilidade.

- Se $\lambda \geq 0$,

$$\lim_{t \rightarrow \infty} y(t) = \lim_{t \rightarrow \infty} y_0 e^{\lambda(t-t_0)} = \begin{cases} y_0, & \text{se } \lambda = 0 \\ \infty, & \text{se } \lambda > 0 \end{cases}.$$

Neste caso, a solução numérica

$$y_k = \left(\frac{2 + \lambda h}{2 - \lambda h} \right)^k y_0$$

tem o mesmo comportamento da solução analítica independentemente dos valores do passo de integração h .

Como, independentemente de $\lambda < 0$ ou $\lambda \geq 0$, a solução numérica acompanha o comportamento da solução analítica, o Método do Trapézio é **incondicionalmente estável**, ou seja, para qualquer λ não existe restrição para os valores do passo de integração h .

Observação

Para $\lambda \in \mathbb{C} \Rightarrow \lambda h \in \mathbb{C}$, tem-se que:

$$\begin{aligned}
 z &= \lambda h = a + bi, \quad a, b \in \mathbb{R} \\
 \left| \frac{2 + \lambda h}{2 - \lambda h} \right| < 1 &\Rightarrow \left| \frac{2 + z}{2 - z} \right| < 1 \\
 &\Rightarrow \left| \frac{2 + a + bi}{2 - a - bi} \right| < 1 \\
 &\Rightarrow \left| \frac{(a + 2) + bi}{(-a + 2) - bi} \right| < 1 \\
 &\Rightarrow \frac{|(a + 2) + bi|}{|(-a + 2) - bi|} < 1 \\
 &\Rightarrow \frac{\sqrt{(a + 2)^2 + b^2}}{\sqrt{(-a + 2)^2 + (-b)^2}} < 1 \\
 &\Rightarrow \sqrt{(a + 2)^2 + b^2} < \sqrt{(-a + 2)^2 + b^2} \\
 &\Rightarrow a^2 + 4a + 4 + b^2 < a^2 - 4a + 4 + b^2 \\
 &\Rightarrow 8a < 0 \\
 &\Rightarrow a < 0 \\
 &\Rightarrow \operatorname{Re}(\lambda h) < 0.
 \end{aligned} \tag{4.17}$$

Em (4.17), tem-se o conjunto dos pontos do semiplano à esquerda da origem. O intervalo de estabilidade é definido por $\operatorname{Re}(\lambda h) \in (-\infty, 0)$, enquanto que $\operatorname{Im}(\lambda h)$ é responsável apenas por um comportamento oscilatório da solução (Veja Schwarz [17]).

Capítulo 5

Métodos de passo múltiplo lineares

5.1 Definição

Definição 5.1 (Métodos de passo múltiplo lineares). *Um método de passo múltiplo linear (ou método de passo múltiplo linear de n-passos) tem a forma*

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j} \quad (5.1)$$

ou

$$\alpha_n y_{k+n} + \cdots + \alpha_1 y_{k+1} + \alpha_0 y_k = h [\beta_n f_{k+n} + \cdots + \beta_1 f_{k+1} + \beta_0 f_k],$$

onde α_j e β_j são constantes, sendo $\alpha_n \neq 0$ e α_0 ou $\beta_0 \neq 0$.

A relação (5.1) é uma equação de diferenças lineares, cuja solução é uma sequência $\{y_n\}$. O método numérico definido por (5.1) é implícito quando $\beta_n \neq 0$ e explícito quando $\beta_n = 0$. Sem perda de generalidade, assume-se que $\alpha_n = 1$.

Exemplo 5.1 (Método de Adams-Bashforth). *Na notação empregada no Capítulo 1:*

$$\begin{cases} y_0 = y(t_0), y_p \text{ pré-determinado}, 1 \leq p \leq 3 \\ y_{k+1} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}) \end{cases} \quad 3 \leq k \leq n-1.$$

Na notação (5.1):

$$\begin{cases} y_0 = y(t_0), y_p \text{ pré-determinado}, 1 \leq p \leq 3 \\ y_{k+4} = y_{k+3} + \frac{h}{24}(55f_{k+3} - 59f_{k+2} + 37f_{k+1} - 9f_k) \end{cases} \quad k = 0, 1, \dots \quad (5.2)$$

Comparando (5.2) com (5.1), constata-se que

$$\alpha_0 = 0, \alpha_1 = 0, \alpha_2 = 0, \alpha_3 = -1, \alpha_4 = 1$$

e

$$\beta_0 = \frac{-9}{24}, \beta_1 = \frac{37}{24}, \beta_2 = -\frac{59}{24}, \beta_3 = \frac{55}{24}, \beta_4 = 0.$$

O método (5.2) é explícito ($\beta_4 = 0$) de 4-passos e y_p , para $1 \leq p \leq 3$, é obtido numericamente através de um método de passo único de mesma ordem de consistência que a ordem do Método de Adams-Bashforth.

5.2 Dedução

Pode-se deduzir um método de passo múltiplo linear:

- pela forma diferencial do Problema de Cauchy;
- pela forma integral do Problema de Cauchy e quadratura numérica;
- por interpolação polinomial.

Exemplo 5.2 (Método de Simpson). *Considere a tabela de pontos abaixo.*

t	t_k	t_{k+1}	t_{k+2}
$f(t, y(t)) = f$	f_k	f_{k+1}	f_{k+2}

Tabela 5.1: Pontos interpolados no Método de Simpson.

Como

$$\begin{aligned} \frac{d}{dt}y(t) &= f(t, y(t)) \\ \int_{t_k}^{t_{k+2}} \frac{d}{ds}y(s)ds &= \int_{t_k}^{t_{k+2}} f(s, y(s)) ds \Rightarrow [y(s)]_{t_k}^{t_{k+2}} = \int_{t_k}^{t_{k+2}} f(s, y(s)) ds \\ y(t_{k+2}) - y(t_k) &= \int_{t_k}^{t_{k+2}} f(s, y(s)) ds, \end{aligned}$$

tem-se que

$$y(t_{k+2}) - y(t_k) \approx \frac{h}{3} (f_k + 4f_{k+1} + f_{k+2}), \quad (5.3)$$

expressão obtida integrando-se o polinômio de grau 2 que interpola todos os pontos da Tabela (5.1).

A relação (5.3) inspira o método de 2-passos implícito

$$y_{k+2} - y_k = \frac{h}{3} (f_{k+2} + 4f_{k+1} + f_k), \quad (5.4)$$

conhecido como Método de Simpson.

Comparando (5.4) com (5.1), constata-se que

$$\alpha_0 = -1, \alpha_1 = 0, \alpha_2 = 1$$

e

$$\beta_0 = \frac{1}{3}, \beta_1 = \frac{4}{3}, \beta_2 = \frac{1}{3}.$$

5.3 Erro Local de Discretização

Definição 5.2 (Erro local de discretização). *Dado o método numérico de passo múltiplo linear (5.1), o erro local de discretização é definido por*

$$\alpha = \frac{\sum_{j=0}^n \alpha_j y(t_k + jh)}{h} - \sum_{j=0}^n \beta_j f(t_k + jh, y(t_k + jh)). \quad (5.5)$$

Em (5.5), α é o erro produzido pelo método numérico para avançar a solução um passo de integração h partindo de valores exatos $y(t_k + jh)$, $0 \leq j \leq n-1$.

Sejam

$$d = h\alpha = \sum_{j=0}^n \alpha_j y(t_k + jh) - h \sum_{j=0}^n \beta_j f(t_k + jh, y(t_k + jh)) \quad (5.6)$$

e a solução numérica

$$\sum_{j=0}^n \alpha_j y_{k+j} - h \sum_{j=0}^n \beta_j f(t_{k+j}, y_{k+j}) = 0. \quad (5.7)$$

Efetuada-se a diferença (5.6) – (5.7), obtém-se

$$d = \sum_{j=0}^n \alpha_j [y(t_k + jh) - y_{k+j}] + \\ - h \sum_{j=0}^n \beta_j [f(t_k + jh, y(t_k + jh)) - f(t_{k+j}, y_{k+j})]. \quad (5.8)$$

Assumindo valores exatos $y_{k+j} = y(t_k + jh)$, $0 \leq j \leq n-1$, tem-se para $j = n$ em (5.8) que

$$d = y(t_k + nh) - y_{k+n} - h\beta_n [f(t_k + nh, y(t_k + nh)) - f(t_{k+n}, y_{k+n})]. \quad (5.9)$$

Há duas possibilidades em (5.9):

- método explícito ($\beta_n = 0$)

$$d = y(t_{k+n}) - y_{k+n};$$

- método implícito ($\beta_n \neq 0$)

$$d = y(t_k + nh) - y_{k+n} - h\beta_n \underbrace{[f(t_k + nh, y(t_k + nh)) - f(t_{k+n}, y_{k+n})]}_{\text{Teorema do Valor Médio}} \\ = y(t_k + nh) - y_{k+n} - h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \xi_{k+n}) (y(t_k + nh) - y_{k+n}) \\ = \left(1 - h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \xi_{k+n})\right) (y(t_k + nh) - y_{k+n}).$$

5.3.1 Expansão em Série de Taylor

Expandindo-se

$$h\alpha = \sum_{j=0}^n \alpha_j (y(t_k + jh)) - h \sum_{j=0}^n \beta_j \underbrace{f(t_k + jh, y(t_k + jh))}_{y^{(1)}(t_k + jh)}$$

em Série de Taylor com centro em $t = t_k$, tem-se que

$$h\alpha = \sum_{j=0}^n \alpha_j \left[y(t_k) + jhy^{(1)}(t_k) + \frac{j^2 h^2}{2!} y^{(2)}(t_k) + \frac{j^3 h^3}{3!} y^{(3)}(t_k) + \dots + \frac{j^p h^p}{p!} y^{(p)}(t_k) + \dots \right] + \\ - \sum_{j=0}^n \beta_j \left[hy^{(1)}(t_k) + jh^2 y^{(2)}(t_k) + \frac{j^2 h^3}{2!} y^{(3)}(t_k) + \frac{j^3 h^4}{3!} y^{(4)}(t_k) + \dots + \frac{j^{p-1} h^p}{(p-1)!} y^{(p)}(t_k) + \dots \right]$$

$$h\alpha = C_0 y(t_k) + C_1 h y^{(1)}(t_k) + C_2 h^2 y^{(2)}(t_k) + \dots + C_p h^p y^{(p)}(t_k) + \dots, \quad (5.10)$$

onde os coeficientes C_p são

$$\begin{aligned} C_0 &= \sum_{j=0}^n \alpha_j, \\ C_1 &= \sum_{j=0}^n j\alpha_j - \sum_{j=0}^n \beta_j, \\ C_2 &= \sum_{j=0}^n \frac{j^2\alpha_j}{2!} - \sum_{j=0}^n j\beta_j, \\ C_3 &= \sum_{j=0}^n \frac{j^3\alpha_j}{3!} - \sum_{j=0}^n \frac{j^2\beta_j}{2!}, \\ &\vdots \\ C_p &= \sum_{j=0}^n \frac{j^p\alpha_j}{p!} - \sum_{j=0}^n \frac{j^{p-1}\beta_j}{(p-1)!}. \end{aligned}$$

Para que o método numérico seja "útil", espera-se que

$$\lim_{h \rightarrow 0} \alpha(t, h) = 0.$$

Para tanto, deve-se ter em (5.10)

$$C_0 = \sum_{j=0}^n \alpha_j = 0$$

e

$$C_1 = \sum_{j=0}^n j\alpha_j - \sum_{j=0}^n \beta_j = 0.$$

5.4 Consistência

Definição 5.3 (Consistência). *Um método de passo múltiplo linear é consistente com a equação diferencial ordinária (ou simplesmente consistente) se e somente se*

$$\lim_{h \rightarrow 0} \alpha(t, h) = 0,$$

ou seja, se e somente se $C_0 = 0$ e $C_1 = 0$.

5.4.1 Ordem de consistência

Definição 5.4 (Ordem de consistência). *Um método de passo múltiplo linear tem ordem de consistência p se*

$$\alpha = O(h^p),$$

ou seja, se

$$C_0 = C_1 = \dots = C_p = 0 \text{ e } C_{p+1} \neq 0. \quad (5.11)$$

Substituindo-se (5.11) em (5.10), tem-se que

$$h\alpha = C_{p+1}h^{p+1}y^{(p+1)}(\xi) \text{ ou } h\alpha = C_{p+1}h^{p+1}y^{(p+1)}(t_k) + O(h^{p+2}).$$

Observe-se que:

1. O método é consistente $\iff p \geq 1$ ($C_0 = C_1 = 0$);
2. A definição geral de consistência (que também serve para os métodos de passo único) pode ser escrita como
"O método é consistente com a EDO $\iff \lim_{h \rightarrow 0} \alpha(t, h) = 0$ ";

3. Como

$$h\alpha = \left(1 - h\beta_n \frac{\partial f}{\partial y}(\xi_{k+n})\right) (y(t_k + nh) - y_{k+n}) \text{ e}$$

$$h\alpha = C_{p+1}h^{p+1}y^{(p+1)}(t_k) + O(h^{p+2}),$$

então

$$y(t_k + nh) - y_{k+n} = C'_{p+1}h^{p+1}y^{(p+1)}(t_k) + O(h^{p+2}). \quad (5.12)$$

Em (5.12), o termo

$$C'_{p+1}h^{p+1}y^{(p+1)}(t_k)$$

é denominado *erro local de discretização principal*. Para métodos explícitos tem-se

$$C'_{p+1} = C_{p+1}$$

e para métodos implícitos

$$C'_{p+1} = \frac{C_{p+1}}{\left(1 - h\beta_n \frac{\partial f}{\partial y}\right)}.$$

5.5 Convergência

A noção de limite a ser utilizada é

$$\lim_{h \rightarrow 0} y_k = y(t), \quad kh = t - t_0 \text{ (fixado)}$$

sendo y_k a solução aproximada e $y(t)$ a solução exata.

Definição 5.5 (Convergência). *Um método de passo múltiplo linear de n -passos*

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f(t_{k+j}, y_{k+j}) \quad (5.13)$$

é convergente se, para todo Problema de Cauchy

$$\begin{cases} \frac{d}{dt} y(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases},$$

tem-se que

$$\lim_{h \rightarrow 0} y_k = y(t)$$

$\forall t \in [a = t_0, b = t_f]$ e para toda solução da equação de diferenças (5.13) satisfazendo

$$y_p = \psi_p(h)$$

com $\lim_{h \rightarrow 0} \psi_p(h) = \bar{\psi}$ ($= y_0$ em geral), $p = 0, 1, \dots, n-1$.

5.5.1 Condições necessárias à convergência

Em um método de passo múltiplo linear convergente, espera-se que

$$y_{k+j} \xrightarrow{h \rightarrow 0} y(t), \quad j = 0, 1, \dots, n.$$

\downarrow
 $kh = t - t_0$

Equivalentemente,

$$y(t) = y_{k+j} + \underbrace{\theta}_{\text{erro}}(h), \quad j = 0, 1, \dots, n$$

com $\lim_{h \rightarrow 0} \theta_{j,k}(h) = 0$.

Dessa forma,

$$\begin{aligned} \sum_{j=0}^n \alpha_j y(t) &= \sum_{j=0}^n \alpha_j y_{k+j} + \sum_{j=0}^n \alpha_j \theta_{j,k}(h), \\ y(t) \sum_{j=0}^n \alpha_j &= h \sum_{j=0}^n \beta_j f(t_{k+j}, y_{k+j}) + \sum_{j=0}^n \alpha_j \theta_{j,k}(h). \end{aligned} \quad (5.14)$$

Considerando $h \rightarrow 0$ em (5.14), conclui-se que

$$y(t) \sum_{j=0}^n \alpha_j = 0 \Rightarrow \sum_{j=0}^n \alpha_j = C_0 = 0.$$

Ainda supondo que o método seja convergente, tem-se que

$$\begin{aligned} \frac{y_{k+j} - y_k}{jh} &\xrightarrow{h \rightarrow 0} \frac{d}{dt} y(t), \quad j = 1, 2, \dots, n \\ &\quad \downarrow \\ &\quad kh = t - t_0 \\ \frac{d}{dt} y(t) &= \frac{y_{k+j} - y_k}{jh} + \underbrace{\theta_{j,k}(h)}_{\downarrow 0, h \downarrow 0} \\ y_{k+j} - y_k &= jh \frac{d}{dt} y(t) + jh \theta_{j,k}(h) \\ \sum_{j=0}^n \alpha_j y_{k+j} - \sum_{j=0}^n \alpha_j y_k &= h \sum_{j=0}^n j \alpha_j \frac{d}{dt} y(t) + h \sum_{j=0}^n j \alpha_j \theta_{j,k}(h) \\ h \sum_{j=0}^n \beta_j f_{k+j} - y_k \underbrace{\sum_{j=0}^n \alpha_j}_{C_0=0} &= h \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j + h \sum_{j=0}^n j \alpha_j \theta_{j,k}(h) \\ h \sum_{j=0}^n \beta_j f_{k+j} &= h \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j + h \sum_{j=0}^n j \alpha_j \theta_{j,k}(h) \\ \sum_{j=0}^n \beta_j f_{k+j} &= \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j + \sum_{j=0}^n j \alpha_j \theta_{j,k}(h). \end{aligned} \quad (5.15)$$

Como $f_{k+j} \rightarrow f(t, y(t)) = \frac{d}{dt} y(t)$ e $\theta_{j,k}(h) \rightarrow 0$ quando $h \rightarrow 0$, pode-se concluir

de (5.15) que

$$\begin{aligned} \sum_{j=0}^n \beta_j f(t, y(t)) &= \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j \\ f(t, y(t)) \sum_{j=0}^n \beta_j &= \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j \\ \frac{d}{dt} y(t) \sum_{j=0}^n \beta_j &= \frac{d}{dt} y(t) \sum_{j=0}^n j \alpha_j \\ \sum_{j=0}^n \beta_j &= \sum_{j=0}^n j \alpha_j \\ \sum_{j=0}^n \beta_j - \sum_{j=0}^n j \alpha_j &= 0 \\ \sum_{j=0}^n j \alpha_j - \sum_{j=0}^n \beta_j &= 0 \\ C_1 &= 0. \end{aligned}$$

Mostrou-se assim que, se o método for convergente, então $C_0 = 0$ e $C_1 = 0$, isto é, o método é consistente.

Consistência é uma condição necessária à convergência, ou seja, um método de passo múltiplo linear consistente pode ou não ser convergente. Porém, se o método de passo múltiplo linear não for consistente, também não será convergente.

Portanto, algo mais precisa ser satisfeito para garantir a convergência (condições necessárias e suficientes). Isto conduz à noção de estabilidade.

5.5.2 Exercícios

Exercício 5.1. Usando as expansões de Taylor em torno de t_n , determine o método mais exato implícito de passo 2 e o primeiro termo do erro local de truncamento.

Exercício 5.2. Mostre que o método

$$y_{k+2} - y_{k+1} = \frac{h}{3}(3f_{k+1} - 2f_k)$$

é inconsistente.

Exercício 5.3. Mostre que a ordem do método

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4}[(b+3)f_{k+2} + (3b+1)f_k]$$

1. é 2 se $b \neq -1$;
2. é 3 se $b = -1$.

5.6 Exercícios resolvidos

1. Mostre que a ordem do método

$$y_{k+2} - y_{k+1} = \frac{h}{12}(4f_{k+2} + 8f_{k+1} - f_k) \quad (5.16)$$

é zero. Demonstre que o método é divergente usando a solução teórica do p.v.i.

$$\begin{cases} \dot{y}(t) = 1 \\ y(0) = 0 \end{cases} \quad 0 \leq t \leq 2. \quad (5.17)$$

Solução

Tem-se do método (5.16) os seguintes dados:

$$\alpha_0 = 0, \quad \alpha_1 = -1, \quad \alpha_2 = 1, \quad \beta_0 = -\frac{1}{12}, \quad \beta_1 = \frac{2}{3} \quad \text{e} \quad \beta_2 = \frac{1}{3}.$$

Para que a ordem do método seja zero, basta mostrar que $C_0 = 0$ e $C_1 \neq 0$.

$$\begin{aligned} C_0 &= \sum_{j=0}^2 \alpha_j = \alpha_0 + \alpha_1 + \alpha_2 = 0 - 1 + 1 = 0 \\ C_1 &= \sum_{j=0}^2 j\alpha_j - \sum_{j=0}^2 \beta_j = \alpha_1 + 2\alpha_2 - \beta_0 - \beta_1 - \beta_2 \\ &= -1 + 2 + \frac{1}{12} - \frac{2}{3} - \frac{1}{3} = \frac{1}{12} \neq 0 \end{aligned}$$

Como $C_0 = 0$ e $C_1 \neq 0$, conclui-se que o método tem ordem zero.

Um método de passo múltiplo linear é convergente se, para todo Problema de Cauchy

$$\begin{cases} \dot{y}(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases},$$

tem-se que $\lim_{\substack{h \rightarrow 0 \\ kh=t-t_0}} y_k = y(t)$.

No p.v.i. (5.17), $f(t, y(t)) = 1$ e a solução exata é dada por $y(t) = t$ (verifique!).

Para $t = 1$ (fixo), tem-se empregando o método (5.16):

$$\begin{aligned} y_{k+2} - y_{k+1} &= \frac{h}{12}(4 + 8 - 1) = \frac{11}{12}h, \\ y_{k+2} &= y_{k+1} + \frac{11}{12}h. \end{aligned}$$

Logo:

$$\begin{aligned} y_0 &= 0; \\ y_1 &= h \quad (y_1 = y_0 + h = 0 + h); \\ y_2 &= y_1 + \frac{11}{12}h = h + \frac{11}{12}h = \frac{23}{12}h; \\ y_3 &= y_2 + \frac{11}{12}h = \frac{34}{12}h; \\ y_4 &= y_3 + \frac{11}{12}h = \frac{45}{12}h; \\ &\vdots \end{aligned}$$

Assim, $\lim_{\substack{h \rightarrow 0 \\ kh=1}} y_k = 0$. Porém, $y(1) = 1$.

Portanto, $\lim_{\substack{h \rightarrow 0 \\ kh=1}} y_k \neq y(t)$ para o Problema de Cauchy (5.17). Como o limite deve ser verificado para todo Problema de Cauchy, pode-se afirmar que o método de passo múltiplo linear (5.16) é divergente.

Pode-se mostrar também que o método (5.16) é divergente plotando-se simultaneamente a solução analítica e as aproximações geradas pelo método. Os Gráficos (5.1) e (5.2) comparam as soluções considerando passos $h = 0, 1$, $h = 0, 05$ e $h = 0, 025$.

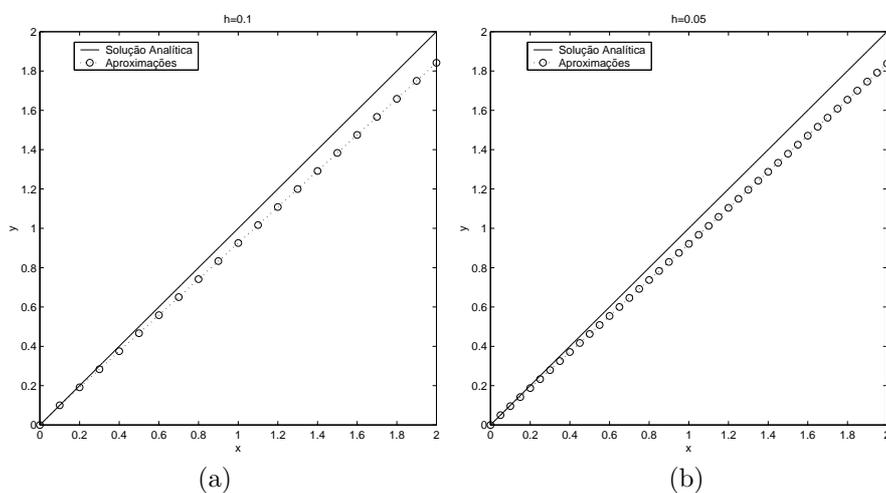


Figura 5.1: Comparação das soluções exata e numérica do Problema de Cauchy (5.17), onde a solução numérica foi obtida através do método (5.16) com (a) $h = 0, 1$ e (b) $h = 0, 05$.

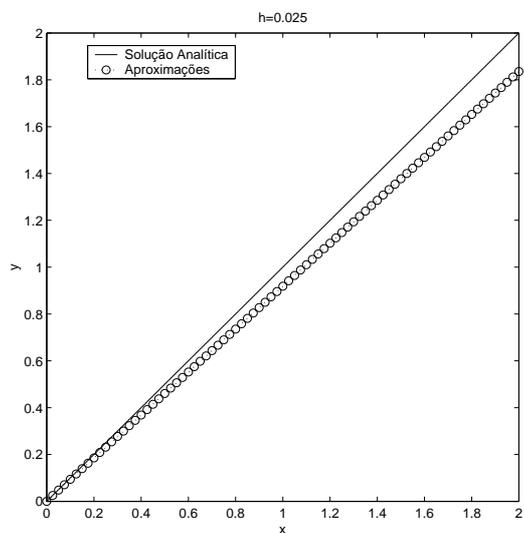


Figura 5.2: Comparação das soluções exata e numérica do Problema de Cauchy (5.17), onde a solução numérica foi obtida através do método (5.16) com $h = 0,025$.

Observa-se nos Gráficos (5.1) e (5.2) o comportamento divergente do método de passo múltiplo linear (5.16). O refinamento do passo de integração não diminui a diferença entre os valores exato e numérico. Pelo fato da ordem do método ser zero, haverá para todo passo h uma diferença quase que constante entre as soluções analítica $y(t_k)$ e numérica y_k em todos os pontos t_k . Fixando-se o ponto $t = 2,0$, o comportamento do erro nas aproximações de $y(2)$ com o passo h sendo reduzido pela metade sucessivamente, por vinte vezes, a partir de $h = 0,5$, pode ser observado no Gráfico (5.3), apresentado em escala logarítmica. Nesse gráfico, nota-se que o erro quase se estabiliza para um passo de integração h menor do que aproximadamente 10^{-2} .

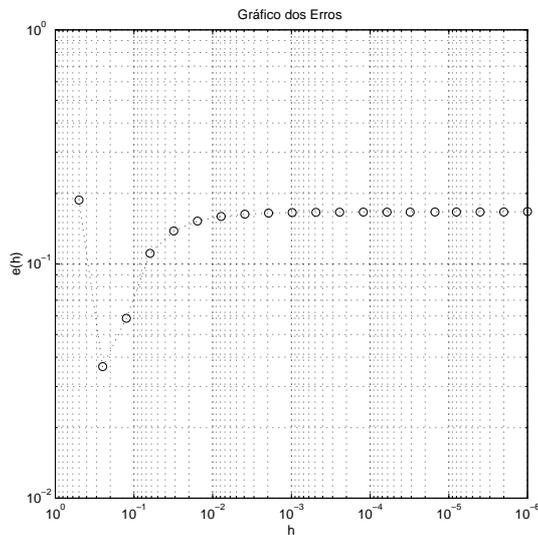


Figura 5.3: Comportamento do erro na solução numérica do Problema de Cauchy (5.17) obtida através do método (5.16) com refinamento do passo de integração h .

2. Aplique o método

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4}[(b+3)f_{k+2} + (3b+1)f_k] \quad (5.18)$$

ao p.v.i.

$$\begin{cases} \dot{y}(t) = y \\ y_0 = y_1 = 0 \end{cases} \quad 0 \leq t \leq 1. \quad (5.19)$$

Comente os resultados obtidos.

Solução

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4}[(b+3)f_{k+2} + (3b+1)f_k]$$

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4}[(b+3)y_{k+2} + (3b+1)y_k]$$

$$y_{k+2} - \frac{h(b+3)}{4}y_{k+2} = -(b-1)y_{k+1} + by_k + \frac{h(3b+1)}{4}y_k$$

$$\left[1 - \frac{h(b+3)}{4}\right]y_{k+2} = -(b-1)y_{k+1} + \left[b + \frac{h(3b+1)}{4}\right]y_k$$

$$y_{k+2} = \frac{-(b-1)y_{k+1} + \left[b + \frac{h(3b+1)}{4} \right] y_k}{1 - \frac{h(b+3)}{4}} \quad (5.20)$$

Observe que no p.v.i. (5.19) tem-se como condições iniciais $y_0 = 0$ e $y_1 = 0$. Aplicando em (5.20) esses valores iniciais, obtém-se $y_2 = 0$. Ao se calcular y_3, y_4, y_5 e assim sucessivamente, o método (5.18) gera uma sequência de zeros. Observe-se também que a solução geral da EDO em (5.19) é dada por $y(t) = Ce^t$. Aplicando a condição inicial $y(0) = 0$, chega-se a $C = 0$ e, conseqüentemente, $y(t) = 0$, a solução do p.v.i.. Como $y(t) = 0$ para todo t e o método numérico (5.18) gera uma sequência de zeros, pode-se concluir que o referido método gera a solução exata do p.v.i. (5.19).

3. Construa um método de 2 passos implícito de ordem máxima contendo um parâmetro livre. Determine a ordem desse método.

Observação: considere o parâmetro livre como sendo $\alpha_0 = a$.

Solução

Método de 2 passos implícito:

$$y_{k+2} + \alpha_1 y_{k+1} + \alpha_0 y_k = h [\beta_2 f_{k+2} + \beta_1 f_{k+1} + \beta_0 f_k], \quad (5.21)$$

com $\beta_2 \neq 0$, $\alpha_2 = 1$ e $\alpha_0 = a$.

Tem-se em (5.21) 4 incógnitas. Logo, são necessárias 4 equações. Para tanto, considere-se o método (5.21) consistente de ordem 3, o que implica $C_0 = C_1 = C_2 = C_3 = 0$ e $C_4 \neq 0$.

$$C_0 = 0 \Rightarrow \sum_{j=0}^2 \alpha_j = 0 \Rightarrow 1 + \alpha_1 + a = 0 \Rightarrow \alpha_1 = -(a+1) \quad (5.22)$$

$$\begin{aligned} C_1 = 0 &\Rightarrow \sum_{j=0}^2 j\alpha_j - \sum_{j=0}^2 \beta_j = 0 \Rightarrow \alpha_1 + 2 - \beta_0 - \beta_1 - \beta_2 = 0 \\ &\Rightarrow 1 - a = \beta_0 + \beta_1 + \beta_2 \end{aligned} \quad (5.23)$$

$$\begin{aligned} C_2 = 0 &\Rightarrow \sum_{j=0}^2 \frac{j^2}{2!} \alpha_j - \sum_{j=0}^2 j\beta_j = 0 \Rightarrow \frac{1}{2} \alpha_1 + 2 - \beta_1 - 2\beta_2 = 0 \\ &\Rightarrow \frac{1}{2} \underbrace{[-(a+1)]}_{(5.22)} + 2 = \beta_1 + 2\beta_2 \\ &\Rightarrow 3 - a = 2\beta_1 + 4\beta_2 \end{aligned} \quad (5.24)$$

$$\begin{aligned}
C_3 = 0 &\Rightarrow \sum_{j=0}^2 \frac{j^3}{3!} \alpha_j - \sum_{j=0}^2 \frac{j^2}{2!} \beta_j = 0 \Rightarrow \frac{1}{6} \alpha_1 + \frac{8}{6} - \frac{1}{2} \beta_1 - 2\beta_2 = 0 \\
&\Rightarrow \frac{1}{6} \underbrace{[-(a+1)]}_{(5.22)} + \frac{8}{6} = \frac{1}{2} \beta_1 + 2\beta_2 \\
&\Rightarrow 7 - a = 3\beta_1 + 12\beta_2
\end{aligned} \tag{5.25}$$

Efetando (5.24) $\times (-3)$ + (5.25), tem-se que

$$2a - 2 = -3\beta_1 \Rightarrow \beta_1 = \frac{2 - 2a}{3}. \tag{5.26}$$

Substituindo (5.26) em (5.24), obtém-se

$$4\beta_2 = 3 - a - 2 \frac{2 - 2a}{3} \Rightarrow \beta_2 = \frac{a + 5}{12}. \tag{5.27}$$

Substituindo (5.26) e (5.27) em (5.23), chega-se a

$$\beta_0 = 1 - a - \frac{2 - 2a}{3} - \frac{a + 5}{12} \Rightarrow \beta_0 = -\frac{5a + 1}{12}.$$

Calculando C_4 , obtém-se

$$\begin{aligned}
C_4 &= \sum_{j=0}^2 \frac{j^4}{4!} \alpha_j - \sum_{j=0}^2 \frac{j^3}{3!} \beta_j = \frac{1}{24} \alpha_1 + \frac{16}{24} - \frac{1}{6} \beta_1 - \frac{8}{6} \beta_2 \\
&= \frac{1}{24} \underbrace{[-(a+1)]}_{(5.22)} + \frac{16}{24} - \frac{1}{6} \underbrace{\frac{2 - 2a}{3}}_{(5.26)} - \frac{8}{6} \underbrace{\frac{a + 5}{12}}_{(5.27)} \\
&= \frac{45 - 3a + 8a - 8 - 8a - 40}{72} \\
&= -\frac{a + 1}{24}
\end{aligned} \tag{5.28}$$

Em (5.28), se $a \neq -1$, então o método (5.21) tem ordem 3 ($C_0 = C_1 = C_2 = C_3 = 0$ e $C_4 \neq 0$); se $a = -1$, então $\alpha_1 = 0$, $\alpha_0 = -1$, $\beta_2 = \frac{1}{3}$,

$\beta_1 = \frac{4}{3}$, $\beta_0 = \frac{1}{3}$ e C_5 é igual a

$$\begin{aligned} C_5 &= \sum_{j=0}^2 \frac{j^5}{5!} \alpha_j - \sum_{j=0}^2 \frac{j^4}{4!} \beta_j = \frac{1}{120} \alpha_1 + \frac{32}{120} - \frac{1}{24} \beta_1 - \frac{16}{24} \beta_2 \\ &= \frac{32}{120} - \frac{1}{24} \frac{4}{3} - \frac{16}{24} \frac{1}{3} \\ &= \frac{4}{15} - \frac{5}{18} = \frac{24 - 25}{90} \\ &= -\frac{1}{90} \neq 0. \end{aligned}$$

Portanto, o método numérico (5.21) tem a forma

$$y_{k+2} - (a+1)y_{k+1} + ay_k = h \left[\frac{a+5}{12} f_{k+2} - \frac{2a-2}{3} f_{k+1} - \frac{5a+1}{12} f_k \right]. \quad (5.29)$$

Em (5.29), se $a \neq -1$, então o método tem ordem 3; se $a = -1$, então o método é o Método de Simpson

$$y_{k+2} = y_k + \frac{h}{3} [f_{k+2} + 4f_{k+1} + f_k],$$

de ordem 4; se $a \neq -5$, então o método é implícito.

Capítulo 6

Estabilidade dos métodos de passo múltiplo

Define-se neste capítulo condições necessárias e suficientes à convergência dos métodos de passo múltiplo lineares baseadas nos conceitos de consistência e estabilidade. Essa análise depende:

- da solução de equações de diferenças lineares (Lima [11, 12]);
- do controle das raízes de um polinômio naturalmente associado aos métodos de passo múltiplo lineares (Stoer [14]);
- de polinômios característicos associados aos métodos de passo múltiplo lineares (Lambert [10]).

6.1 Equações de diferenças lineares

Definição 6.1 (Equação de diferenças linear). *Uma equação de diferenças linear é uma igualdade do tipo*

$$\gamma_n y_{k+n} + \gamma_{n-1} y_{k+n-1} + \gamma_{n-2} y_{k+n-2} + \cdots + \gamma_0 y_k = \phi_k, \quad k = 0, 1, \dots, \quad (6.1)$$

onde $\gamma_j, 0 \leq j \leq n$, são constantes independentes de y_k com $\gamma_0 \neq 0$ e $\gamma_n \neq 0$.

A solução de uma equação de diferenças é uma sequência numérica

$$y_n, y_{n+1}, y_{n+2}, \dots, \text{ isto é, } \{y_j\}_{j \in \mathbb{N}}. \quad (6.2)$$

Em (6.1), se $\phi_k = 0 \forall k \in \mathbb{N}$, então a equação de diferenças é homogênea; caso contrário, é não homogênea.

Seja $\{\widehat{y}_j\}_{j \in \mathbb{N}}$ a solução da equação de diferenças linear homogênea

$$\sum_{j=0}^n \gamma_j y_{k+j} = 0. \quad (6.3)$$

Se $\{\psi_j\}_{j \in \mathbb{N}}$ for uma solução particular da equação de diferenças linear não homogênea

$$\sum_{j=0}^n \gamma_j y_{k+j} = \phi_k \quad k = 0, 1, \dots, \quad (6.4)$$

então a solução de (6.4) é dada por

$$y_j = \widehat{y}_j + \psi_j, \quad j \in \mathbb{N}, \quad (6.5)$$

isto é, a sequência numérica $\{y_j\}_{j \in \mathbb{N}}$.

Assim, a solução (6.5) da equação de diferenças linear não homogênea é dada pela solução da equação homogênea (6.3) acrescida de uma solução particular.

Verifica-se que, se $\{y_j\}_{j \in \mathbb{N}}$ soluciona a equação não homogênea

$$\begin{aligned} \sum_{j=0}^n \gamma_j y_{k+j} &= \sum_{j=0}^n \gamma_j (\widehat{y}_{k+j} + \psi_{k+j}) \\ &= \underbrace{\sum_{j=0}^n \gamma_j \widehat{y}_{k+j}}_{=0} + \sum_{j=0}^n \gamma_j \psi_{k+j} \\ &= \phi_k, \end{aligned}$$

então

$$\sum_{j=0}^n \gamma_j \psi_{k+j} = \phi_k.$$

Um caso particular de interesse da equação não homogênea ocorre quando a sequência $\{\phi_j\}_{j \in \mathbb{N}}$ é constante, isto é,

$$\phi_k = \phi \quad \forall k \in \mathbb{N}.$$

Nestas condições, tem-se como solução particular

$$\psi_k = \frac{\phi}{n} \sum_{j=0}^n \gamma_j, \quad k = 0, 1, \dots,$$

desde que $\sum_{j=0}^n \gamma_j \neq 0$.

Definição 6.2 (Soluções linearmente independentes). *Um conjunto de soluções $\{y_{j,t}\}_{j \in \mathbb{N}, t = 1, 2, \dots, \mathbb{K}}$, é dito linearmente independente se e somente se a combinação linear*

$$a_1 y_{j,1} + a_2 y_{j,2} + a_3 y_{j,3} + \dots + a_{\mathbb{K}} y_{j,\mathbb{K}} = 0, \quad \forall j \in \mathbb{N}$$

implicar em $a_t = 0, t = 1, \dots, \mathbb{K}$.

O conjunto de \mathbb{K} seqüências linearmente independentes, solução da equação de diferenças linear, forma um sistema fundamental e toda solução é dada como

$$\left\{ \sum_{t=1}^{\mathbb{K}} d_t y_{j,t} \right\}_{j \in \mathbb{N}}.$$

Uma das formas de solucionar a equação de diferenças linear homogênea

$$\sum_{j=0}^n \gamma_j y_{k+j} = 0, \quad \gamma_n \neq 0 \text{ e } \gamma_0 \neq 0, \quad (6.6)$$

é analisar soluções da forma

$$y_k = r^k. \quad (6.7)$$

Substituindo (6.7) na equação de diferenças (6.6), obtém-se

$$\sum_{j=0}^n \gamma_j r^{k+j} = 0,$$

$$r^k \sum_{j=0}^n \gamma_j r^j = 0 \Rightarrow \begin{cases} r^k = 0 \Rightarrow r = 0 : \text{solução trivial} \\ \sum_{j=0}^n \gamma_j r^j = 0 \Rightarrow r \text{ é uma raiz do polinômio} \\ p_n(r) = \sum_{j=0}^n \gamma_j r^j \end{cases}.$$

Se todas as raízes do polinômio $p_n(r)$ forem distintas, a solução geral da equação de diferenças linear não homogênea é dada por

$$y_k = \sum_{t=1}^n d_t r_t^k + \psi_k.$$

No caso de raízes múltiplas, a solução deve ser modificada. Por exemplo, se r_1 for uma raiz de multiplicidade 2, tem-se que

$$y_k = d_{1,1}r_1^k + d_{1,2}kr_1^k + \sum_{t=2}^n d_t r_t^k + \psi_k.$$

Em geral, se r_p tem multiplicidade μ_t , a solução geral é escrita na forma

$$\begin{aligned} y_k &= [d_{1,1} + d_{1,2}k + \cdots + d_{1,\mu_1}k(k-1)(k-2)\cdots(k-\mu_1+2)] r_1^k + \\ &+ [d_{2,1} + d_{2,2}k + \cdots + d_{2,\mu_2}k(k-1)(k-2)\cdots(k-\mu_2+2)] r_2^k + \\ &\vdots \\ &+ [d_{p,1} + d_{p,2}k + \cdots + d_{p,\mu_p}k(k-1)(k-2)\cdots(k-\mu_p+2)] r_p^k + \\ &+ \psi_k \end{aligned}$$

onde $\sum_{t=1}^p \mu_t = n$.

Exemplo 6.1. Encontre a solução da equação de diferenças linear

$$y_{k+4} - 4y_{k+3} + 5y_{k+2} - 4y_{k+1} + 4y_k = 4 \quad (6.8)$$

satisfazendo $y_0 = 5$, $y_1 = 0$, $y_2 = -4$ e $y_3 = -12$.

Solução

$$\begin{aligned} p_4(r) &= \sum_{j=0}^4 \gamma_j r^j = \gamma_4 r^4 + \gamma_3 r^3 + \gamma_2 r^2 + \gamma_1 r + \gamma_0 \\ p_4(r) &= r^4 - 4r^3 + 5r^2 - 4r + 4 \\ &= (r^2 + 1)(r - 2)^2 \\ &= (r - i)(r + i)(r - 2)^2 \\ \text{Raízes: } &r_1 = 2 ; r_2 = -i \text{ e } r_3 = i. \end{aligned} \quad (6.9)$$

Solução particular:

$$\psi_k = \frac{\phi}{n} = \frac{4}{1 - 4 + 5 - 4 + 4} = 2, \quad k \in \mathbb{N}.$$

Como $r_1 = 2$ é uma raiz de multiplicidade 2 e a solução particular é $\psi_k = 2$, a solução geral da equação não homogênea é dada por

$$y_k = d_{1,1}r_1^k + d_{1,2}kr_1^k + d_2r_2^k + d_3r_3^k + 2. \quad (6.10)$$

Das condições iniciais

$$\begin{aligned} y_0 &= d_{1,1} + d_2 + d_3 + 2 = 5, \\ y_1 &= d_{1,1}2 + d_{1,2}1 \cdot 2 + d_2(-i) + d_3(i) + 2 = 0, \\ y_2 &= d_{1,1}2^2 + d_{1,2}2 \cdot 2^2 + d_2(-i)^2 + d_3(i)^2 + 2 = -4, \\ y_3 &= d_{1,1}2^3 + d_{1,2}3 \cdot 2^3 + d_2(-i)^3 + d_3(i)^3 + 2 = -12, \end{aligned}$$

obtem-se um sistema linear 4×4 , cujas incógnitas são $d_{1,1}$, $d_{1,2}$, d_2 e d_3 . A solução desse sistema é

$$d_{1,1} = 1, \quad d_{1,2} = -1, \quad d_2 = 1 - i, \quad d_3 = 1 + i. \quad (6.11)$$

Substituindo as raízes (6.9) e os coeficientes (6.11) na equação (6.10), tem-se que

$$\begin{aligned} y_k &= 1 \cdot 2^k + (-1)k2^k + (1 - i)(-i)^k + (1 + i)(i)^k + 2 \\ &= (1 - k)2^k + (1 - i)(-i)^k + (1 + i)(i)^k + 2, \quad k = 0, 1, \dots \end{aligned} \quad (6.12)$$

6.1.1 Exercícios

Exercício 6.1. Calcule os termos y_4 , y_5 e y_6 da solução (6.12) empregando

1. a própria solução (6.12);
2. a equação de diferenças linear não homogênea (6.8).

Exercício 6.2. Como os coeficientes de (6.8) são reais e as condições iniciais também, y_k é necessariamente real. Mostre que $(1 - i)(-i)^k + (1 + i)(i)^k$ é um número real $\forall k \in \mathbb{N}$ ou, equivalentemente, mostre que

$$y_k = 2 \left[\cos\left(\frac{k\pi}{2}\right) - \operatorname{sen}\left(\frac{k\pi}{2}\right) \right] + (1 - k)2^k + 2$$

lembrando que $i = e^{i\frac{\pi}{2}} = \cos\left(\frac{\pi}{2}\right) + i\operatorname{sen}\left(\frac{\pi}{2}\right)$.

Exercício 6.3. A sequência de Fibonacci é uma sequência de números inteiros onde $y_0 = 0$, $y_1 = 1$ e cada um dos demais termos é dado pela soma dos dois termos precedentes.

1. Mencione os dez primeiros termos da sequência de Fibonacci.
2. Escreva a equação de diferenças linear que define a sequência de Fibonacci.
3. Solucione a equação de diferenças linear do item 2 empregando a técnica estudada.

6.2 Divergência

Seja o método de 2-passos explícito

$$y_{k+2} + \alpha_1 y_{k+1} + \alpha_0 y_k = h [\beta_1 f_{k+1} + \beta_0 f_k]$$

de ordem máxima.

Calculando os coeficientes do erro local de discretização, obtêm-se:

$$\begin{aligned} C_0 &= \alpha_0 + \alpha_1 + 1 = 0; \\ C_1 &= 0 \cdot \alpha_0 + 1 \cdot \alpha_1 + 2 \cdot 1 - \beta_0 - \beta_1 = 0; \\ C_2 &= \frac{0^2 \cdot \alpha_0 + 1^2 \cdot \alpha_1 + 2^2 \cdot 1}{2!} - \frac{0 \cdot \beta_0 + 1 \cdot \beta_1}{1!} = 0; \\ C_3 &= \frac{0^3 \cdot \alpha_0 + 1^3 \cdot \alpha_1 + 2^3 \cdot 1}{3!} - \frac{0^2 \cdot \beta_0 + 1^2 \cdot \beta_1}{2!} = 0. \end{aligned} \quad (6.13)$$

A solução do sistema linear (6.13) é $\alpha_0 = -5$, $\alpha_1 = 4$, $\beta_0 = 2$ e $\beta_1 = 4$.

Assim,

$$y_{k+2} + 4y_{k+1} - 5y_k = h [4f_{k+1} + 2f_k]. \quad (6.14)$$

Exemplo 6.2. Solucionar o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = -y(t), & t \in [0, 1] \\ y(0) = 1 \end{cases} \quad (6.15)$$

utilizando o método de terceira ordem (6.14) com $h = 0,01$.

Solução exata do p.v.i.:

$$y(t) = e^{-t}. \quad (6.16)$$

A discretização do Problema de Cauchy (6.15) é dada por

$$\begin{aligned} y_{k+2} &= -4y_{k+1} + 5y_k + h [4f_{k+1} + 2f_k], \\ y_{k+2} &= -4y_{k+1} + 5y_k + h [-4y_{k+1} - 2y_k], \\ y_{k+2} &= -(4 + 4h)y_{k+1} + (5 - 2h)y_k, \\ y_{k+2} + (4 + 4h)y_{k+1} + (-5 + 2h)y_k &= 0. \end{aligned} \quad (6.17)$$

A equação (6.17) é uma equação de diferenças linear homogênea com $\gamma_2 = 1$, $\gamma_1 = 4 + 4h$ e $\gamma_0 = -5 + 2h$, cuja solução é dada por

$$y_k = d_1 r_1^k + d_2 r_2^k, \quad (6.18)$$

onde r_j são as raízes (supondo-as simples) do polinômio

$$\begin{aligned} p_2(r) &= \sum_{j=0}^2 \gamma_j r^j \\ &= \gamma_2 r^2 + \gamma_1 r + \gamma_0 \\ &= r^2 + (4 + 4h)r + (-5 + 2h) \end{aligned}$$

e as constantes d_j são obtidas a partir das condições iniciais

$$y_0 = y(t_0) = 1$$

e

$$y_1 = y(t_0 + h) = y(0 + h) = e^{-h}.$$

Calculando as raízes de $p_2(r)$, obtêm-se

$$r_1 = -2 - 2h + 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2}$$

e

$$r_2 = -2 - 2h - 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2},$$

e calculando d_1 e d_2 chega-se a

$$\begin{cases} 1 = y_0 = d_1 + d_2 \\ e^{-h} = y_1 = d_1 r_1 + d_2 r_2 \end{cases} \Rightarrow d_1 = \frac{r_2 - e^{-h}}{r_2 - r_1} \quad e \quad d_2 = \frac{e^{-h} - r_1}{r_2 - r_1}.$$

Assim,

$$y_k = \frac{r_2 - e^{-h}}{r_2 - r_1} r_1^k + \frac{e^{-h} - r_1}{r_2 - r_1} r_2^k. \quad (6.19)$$

As soluções (6.19) e (6.16) fornecem os resultados presentes na Tabela (6.1). Analisando-a, conclui-se que o método de passo múltiplo linear consistente (6.14) é divergente.

k	$y_k - y(t_k)$
2	$-1,4 \times 10^{-9}$
3	$+5,01 \times 10^{-9}$
4	$-3,00 \times 10^{-8}$
5	$+1,44 \times 10^{-7}$
\vdots	\vdots
98	$-2,57 \times 10^{58}$
99	$+1,29 \times 10^{59}$
100	$-6,52 \times 10^{59}$

Tabela 6.1: Erro global de discretização na solução do problema de valor inicial (6.15) com o método (6.14) e $h = 0,01$.

6.2.1 Exercícios

Exercício 6.4. Mostre que o método de passo múltiplo linear (6.14) é consistente de ordem 3.

Exercício 6.5. Calcule y_2, \dots, y_{100} usando a solução exata (6.19) da equação de diferenças e usando a própria equação de diferenças (6.14).

Os resultados presentes na Tabela (6.1) mostram que consistência e alta ordem não são suficientes à convergência. Usando a solução (6.18) da equação de diferenças, pode-se explicar as causas desse comportamento.

Lembrando que

$$f(h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} h^n,$$

a expansão das raízes r_1 e r_2 , assim como das constantes d_1 e d_2 , em Série de Taylor como função de h em torno de $h = 0$, resulta em:

$$\begin{aligned} \sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2} &= \left(1 + \frac{2}{3}h + \frac{4}{9}h^2\right)^{\frac{1}{2}} = 1 + \frac{1}{3}h + \frac{1}{6}h^2 - \frac{1}{18}h^3 + \frac{1}{216}h^4 + \\ &+ O(h^5); \\ r_1 = -2 - 2h + 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2} &= 1 - h + \frac{1}{2}h^2 - \frac{1}{6}h^3 + \frac{1}{72}h^4 + \\ &+ O(h^5); \\ r_2 = -2 - 2h - 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2} &= -5 - 3h - \frac{1}{2}h^2 + \frac{1}{6}h^3 - \frac{1}{72}h^4 + \\ &+ O(h^5); \end{aligned}$$

$$\begin{aligned}
e^{-h} &= \sum_{n=0}^{\infty} \frac{(-h)^n}{n!} = 1 - h + \frac{1}{2}h^2 - \frac{1}{6}h^3 + \frac{1}{24}h^4 + \\
&\quad + O(h^5); \\
d_1 &= \frac{r_2 - e^{-h}}{r_2 - r_1} = 1 + O(h^2); \\
d_2 &= \frac{e^{-h} - r_1}{r_2 - r_1} = -\frac{1}{216}h^4 + O(h^5).
\end{aligned}$$

Dessa forma,

$$\begin{aligned}
y_k &= d_1 r_1^k + d_2 r_2^k \\
&= \left[1 + O\left(\left(\frac{t}{k}\right)^2\right) \right] \left[1 - \left(\frac{t}{k}\right) + O\left(\left(\frac{t}{k}\right)^2\right) \right]^k + \\
&\quad - \frac{1}{216} \left(\frac{t}{k}\right)^4 \left[1 - O\left(\frac{t}{k}\right) \right] \left[-5 - 3\left(\frac{t}{k}\right) + O\left(\left(\frac{t}{k}\right)^2\right) \right]^k, \\
&= \left[1 + O\left(\left(\frac{t}{k}\right)^2\right) \right] \left[1 - \left(\frac{t}{k}\right) + O\left(\left(\frac{t}{k}\right)^2\right) \right]^k + \\
&\quad - \frac{1}{216} \left(\frac{t}{k}\right)^4 (-5)^k \left[1 - O\left(\frac{t}{k}\right) \right] \left[1 + \frac{3}{5}\left(\frac{t}{k}\right) + O\left(\left(\frac{t}{k}\right)^2\right) \right]^k,
\end{aligned} \tag{6.20}$$

onde y_k representa a solução numérica, ou seja, a solução da equação de diferenças linear, e $kh = t - 0$ está fixado.

Calculando $\lim_{k \rightarrow \infty} y_k$, verifica-se que a primeira parcela do lado direito de (6.20) se comporta como

$$e^{-t},$$

enquanto que a segunda parcela se comporta como

$$-\frac{t^4}{216} \frac{(-5)^k}{k^4} e^{\frac{3t}{5}} \tag{6.21}$$

(Stoer [14])¹.

¹ $\lim_{k \rightarrow \infty} \left(1 + \frac{1}{k}\right)^k = e$, $\lim_{k \rightarrow 0} (1+k)^{\frac{1}{k}} = e$, $\lim_{k \rightarrow \infty} \left(1 + \frac{a}{b} \frac{1}{k}\right)^k = e^{\frac{a}{b}}$

Em (6.21),

$$\lim_{k \rightarrow \infty} \frac{(-5)^k}{k^4} = \pm\infty, \quad (6.22)$$

uma vez que $5^k \gg k^4$ quando $k \rightarrow \infty$.

Portanto, o limite (6.22) explica o comportamento oscilatório e ilimitado da solução y_k quando $k \rightarrow \infty$ ($k \rightarrow \infty \iff h \rightarrow 0$).

De (6.20), conclui-se então que o comportamento da solução da equação de diferenças é controlado pela maior raiz do polinômio

$$p_2(r) = r^2 + (4 + 4h)r + (-5 + 2h),$$

ou seja, $r_1 = -2 - 2h + 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2}$.

Note-se que, quando $h \rightarrow 0$,

$$p_2(r) \rightarrow r^2 + 4r - 5,$$

cujas raízes são $r_1 = 1$ e $r_2 = -5$.

6.3 Polinômios característicos

Definição 6.3. *Dado o método de passo múltiplo linear*

$$y_{k+n} + \sum_{j=0}^{n-1} \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j}, \quad (6.23)$$

o primeiro e o segundo polinômios característicos associados ao método são definidos, respectivamente, por

$$\rho(r) = \sum_{j=0}^n \alpha_j r^j, \quad (6.24)$$

$$\sigma(r) = \sum_{j=0}^n \beta_j r^j, \quad (6.25)$$

sendo $\alpha_n = 1$.

Em relação aos polinômios característicos (6.24) e (6.25), tem-se que o método de passo múltiplo linear (6.23) é consistente se e somente se

$$\rho(1) = 0$$

e

$$\rho'(1) - \sigma(1) = 0.$$

$$\begin{aligned}\rho(r) &= \sum_{j=0}^n \alpha_j r^j \\ &= \alpha_0 r^0 + \alpha_1 r^1 + \alpha_2 r^2 + \alpha_3 r^3 + \cdots + \alpha_n r^n \\ \rho(1) &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n \\ &= \sum_{j=0}^n \alpha_j = C_0\end{aligned}$$

$$\rho(1) = 0 \Rightarrow C_0 = 0$$

$$\begin{aligned}\rho'(r) &= \alpha_1 + 2\alpha_2 r^1 + 3\alpha_3 r^2 + 4\alpha_4 r^3 + \cdots + n\alpha_n r^{n-1} \\ \rho'(1) &= \alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4 + \cdots + n\alpha_n\end{aligned}$$

$$\rho'(1) = \sum_{j=0}^n j\alpha_j$$

$$\begin{aligned}\sigma(r) &= \sum_{j=0}^n \beta_j r^j \\ &= \beta_0 r^0 + \beta_1 r^1 + \beta_2 r^2 + \beta_3 r^3 + \cdots + \beta_n r^n \\ \sigma(1) &= \beta_0 + \beta_1 + \beta_2 + \beta_3 + \cdots + \beta_n\end{aligned}$$

$$= \sum_{j=0}^n \beta_j$$

$$\rho'(1) - \sigma(1) = \sum_{j=0}^n j\alpha_j - \sum_{j=0}^n \beta_j = C_1$$

$$\rho'(1) - \sigma(1) = 0 \Rightarrow C_1 = 0$$

Em um método de passo múltiplo linear, o primeiro polinômio característico $\rho(r)$ tem sempre uma raiz igual a 1. Esta raiz é denominada *raiz principal* e geralmente denotada por r_1 . Assim, a consistência de um método de passo múltiplo linear depende apenas da raiz principal $r_1 = 1$.

6.4 Zero-Estabilidade

A zero-estabilidade de um método de passo múltiplo linear diz respeito ao estudo da estabilidade para $h \rightarrow 0$.

Seja o p.v.i. modelo

$$\begin{cases} \frac{d}{dt}y(t) = 0 & t \in [a, b] \\ y(0) = 0 \end{cases}, \quad (6.26)$$

cuja solução exata é dada por $y(t) = 0$.

Aplicando um método de passo múltiplo linear

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j} \quad (6.27)$$

para discretizar e solucionar o problema de valor inicial (6.26), tem-se que

$$y_{k+n} + \alpha_{n-1}y_{k+n-1} + \cdots + \alpha_0 y_k = 0, \quad (6.28)$$

uma vez que $f(t, y(t)) = 0$. A relação (6.28) é uma equação de diferenças linear homogênea.

Se o método (6.27) é convergente, então $y_k \xrightarrow{h \rightarrow 0} 0 \forall t \in [a, b]$, isto é,

$$\lim_{h \rightarrow 0} y_k = 0, \quad kh = \underbrace{t - 0}_{\text{fixo}}.$$

Para determinar sob quais condições o método é convergente, suponha-se que as raízes do polinômio

$$p_n(r) = \sum_{j=0}^n \alpha_j r^j, \quad (6.29)$$

o qual coincide com o primeiro polinômio característico, sejam reais e distintas. Então,

$$y_k = d_1 r_1^k + d_2 r_2^k + d_3 r_3^k + \cdots + d_n r_n^k.$$

Para a convergência, é necessário que se tenha

$$y_j \xrightarrow{h \rightarrow 0} 0 (= y_0), \quad j = \underbrace{0, 1, \dots, n-1}_{n \text{ condições iniciais}},$$

ou seja, as n condições iniciais $y_0 = y_1 = y_2 = \cdots = y_{n-1} = 0$.

Assim, considere-se y_k da forma

$$y_k = h (d_1 r_1^k + d_2 r_2^k + \cdots + d_n r_n^k).$$

Como

$$\lim_{\substack{h \rightarrow 0 \\ kh=t}} y_k = 0$$

e

$$\lim_{\substack{h \rightarrow 0 \\ kh=t}} h r_j^k = \lim_{k \rightarrow \infty} \frac{t}{k} r_j^k = t \lim_{k \rightarrow \infty} \frac{r_j^k}{k}, \quad (6.30)$$

o método não será convergente se o módulo de qualquer uma das raízes r_j for maior do que 1. De outra maneira, o limite em (6.30) será zero somente se $|r_j| \leq 1$.

Conclusão 1: No caso de raízes reais e distintas, o método (6.27) é convergente se

$$|r_j| \leq 1, \quad j = 1, \dots, n,$$

sendo r_j as raízes do primeiro polinômio característico.

Suponha-se agora que o polinômio (6.29) tem raízes com multiplicidade.

Se r_j tem multiplicidade $m > 1$, então a raiz contribui para a solução com uma parcela da forma

$$[d_{j,1} + d_{j,2}k + \cdots + d_{j,m}k(k-1)\cdots(k-m+2)] r_j^k.$$

Logo,

$$\begin{aligned} y_k &= [d_{1,1} + d_{1,2}k + \cdots + d_{1,m_1}k(k-1)(k-2)\cdots(k-m_1+2)] h r_1^k + \\ &+ [d_{2,1} + d_{2,2}k + \cdots + d_{2,m_2}k(k-1)(k-2)\cdots(k-m_2+2)] h r_2^k + \\ &\vdots \\ &+ [d_{j,1} + d_{j,2}k + \cdots + d_{j,m_j}k(k-1)(k-2)\cdots(k-m_j+2)] h r_j^k, \end{aligned}$$

onde $\sum_{t=1}^j m_t = n$ e

$$\lim_{\substack{h \rightarrow 0 \\ kh=t}} h k^p r_j^k = \lim_{k \rightarrow \infty} \frac{t}{k} k^p r_j^k = t \lim_{k \rightarrow \infty} k^{p-1} r_j^k. \quad (6.31)$$

O método será convergente somente se o limite em (6.31) for zero. Isto ocorre quando $|r_j| < 1$.

Conclusão 2: No caso de raízes com multiplicidade maior que 1, o método (6.27) é convergente se essas raízes tiverem módulo menor que 1.

O exemplo baseado no p.v.i. (6.26) inspira a seguinte definição de estabilidade.

Definição 6.4 (Zero-Estabilidade). *Um método de passo múltiplo linear é zero-estável se nenhuma raiz do primeiro polinômio característico*

$$\rho(r) = \sum_{j=0}^n \alpha_j r^j$$

tiver módulo maior que 1 e toda raiz com módulo 1 for simples (multiplicidade um).

Os resultados a seguir relacionam os conceitos de consistência, zero-estabilidade e convergência.

Teorema 6.1. *As condições necessárias e suficientes para que um método de passo múltiplo linear seja **convergente** são que ele seja **consistente** e **zero-estável**.*

Teorema 6.2. *Nenhum método de passo múltiplo linear de n passos ($n \geq 2$) zero-estável tem ordem maior que $n + 1$ se n for ímpar ou ordem maior que $n + 2$ se n for par.*

6.4.1 Exercícios

Exercício 6.6. *Verifique que todo método de passo único é zero-estável.*

Exercício 6.7. Verifique se o método de passo múltiplo linear

$$y_{k+2} + 4y_{k+1} - 5y_k = h [4f_{k+1} + 2f_k]$$

é zero estável.

Exercício 6.8. Analise a zero-estabilidade do método de passo múltiplo linear

$$y_{k+2} - (1+a)y_{k+1} + ay_k = \frac{h}{2} [(3-a)f_{k+1} - (1+a)f_k]$$

para

1. $a = 0$;
2. $a = -5$.

Exercício 6.9. Mostre que o método de passo múltiplo linear

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4} [(b+3)f_{k+2} + (3b+1)f_k]$$

é zero-instável se $b = -1$.

Exercício 6.10. Empregue o método de passo múltiplo linear

$$y_{k+2} + (b-1)y_{k+1} - by_k = \frac{h}{4} [(b+3)f_{k+2} + (3b+1)f_k],$$

com $b = 0$ e $b = -1$, para solucionar o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = y(t), & t \in [0, 1] \\ y_0 = y_1 = 0 \end{cases}.$$

Comente os resultados obtidos.

É necessário ainda responder algumas perguntas a respeito do emprego de métodos de passo múltiplo lineares.

1. Para $n > 1$, como obter as condições iniciais y_μ , $\mu = 1, 2, \dots, n-1$, necessárias para iniciar um método de passo múltiplo de n -passos?
2. Como escolher o passo de integração h ? (teoria da estabilidade absoluta)
3. Quão precisa é a solução numérica obtida? (delimitação dos erros local e global de discretização)

6.5 Condições iniciais

O método de n -passos

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j}$$

requer as condições y_μ , $\mu = 0, 1, 2, \dots, n-1$, para gerar a sequência y_k, y_{k+1}, \dots . Como $y_0 = y(t_0)$, as demais condições iniciais y_1, y_2, \dots, y_{n-1} são obtidas numericamente com um método de passo único.

É importante salientar que a ordem do método de passo único empregado não pode ser menor do que a ordem do método de passo múltiplo em uso. Caso contrário, o erro introduzido ao se gerar as condições iniciais dominará a solução numérica. Neste caso, o método se comportará como se fosse um método de ordem inferior.

6.6 Erro global de discretização

Seja o método de passo múltiplo linear

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j}. \quad (6.32)$$

Para o método (6.32), o erro local de discretização multiplicado pelo passo de integração h e a solução numérica são dados, respectivamente, por

$$h\alpha = \sum_{j=0}^n \alpha_j y(t_{k+j}) - h \sum_{j=0}^n \beta_j f(t_{k+j}, y(t_{k+j})), \quad (6.33)$$

$$R = \sum_{j=0}^n \alpha_j y_{k+j} - h \sum_{j=0}^n \beta_j f(t_{k+j}, y_{k+j}). \quad (6.34)$$

Efetuada a diferença (6.33) - (6.34), tem-se que

$$\begin{aligned} h\alpha - R &= \sum_{j=0}^n \alpha_j (y(t_{k+j}) - y_{k+j}) + \\ &- h \sum_{j=0}^n \beta_j (f(t_{k+j}, y(t_{k+j})) - f(t_{k+j}, y_{k+j})). \end{aligned} \quad (6.35)$$

Considerando em (6.35) $\phi_k = h\alpha - R$ e empregando o Teorema do Valor Médio,

chega-se a

$$\begin{aligned} \phi_k &= \sum_{j=0}^n \alpha_j (y(t_{k+j}) - y_{k+j}) + \\ &- h \sum_{j=0}^n \beta_j \frac{\partial}{\partial y} f(t_{k+j}, \xi_{k+j}) (y(t_{k+j}) - y_{k+j}). \end{aligned} \quad (6.36)$$

Lembrando que o erro global de discretização no instante t_{k+j} é definido como

$$e_{k+j} = y(t_{k+j}) - y_{k+j},$$

pode-se reescrever (6.36) como

$$\phi_k = \sum_{j=0}^n \alpha_j e_{k+j} - h \sum_{j=0}^n \beta_j \frac{\partial}{\partial y} f(t_{k+j}, \xi_{k+j}) e_{k+j}. \quad (6.37)$$

Aplicando as hipóteses simplificadoras

$$\begin{aligned} \phi_k &= \bar{\phi} = \text{constante}, \\ \frac{\partial}{\partial y} f(t_{k+j}, \xi_{k+j}) &= \lambda = \text{constante}, \end{aligned}$$

à equação (6.37), obtém-se a equação de diferenças linear para o erro global de discretização

$$\sum_{j=0}^n (\alpha_j - h\lambda\beta_j) e_{k+j} = \bar{\phi}, \quad (6.38)$$

cujas soluções são dadas por

$$e_k = \sum_{j=1}^n d_j r_j^k + \frac{\bar{\phi}}{\sum_{j=0}^n (\alpha_j - h\lambda\beta_j)}, \quad (6.39)$$

onde r_j são as raízes (supondo-as simples) do polinômio $p_n(r) = \sum_{j=0}^n (\alpha_j - h\lambda\beta_j) r^j$,

$\sum_{j=1}^n d_j r_j^k$ é a solução da equação de diferenças linear homogênea e $\frac{\bar{\phi}}{\sum_{j=0}^n (\alpha_j - h\lambda\beta_j)}$

é uma solução particular.

Partindo do princípio de que o método de passo múltiplo linear (6.32) é consistente, o que implica que

$$C_0 = \sum_{j=0}^n \alpha_j = 0,$$

e adotando $\bar{h} = h\lambda$, reescreve-se (6.39) como

$$e_k = \sum_{j=1}^n d_j r_j^k + \frac{\bar{\phi}}{-\bar{h} \sum_{j=0}^n \beta_j}, \quad (6.40)$$

solução da equação de diferenças linear para o erro global de discretização

$$\sum_{j=0}^n (\alpha_j - \bar{h}\beta_j) e_{k+j} = \bar{\phi}, \quad (6.41)$$

cujo polinômio característico é

$$\begin{aligned} \pi(r) &= \sum_{j=0}^n (\alpha_j - \bar{h}\beta_j) r^j \\ &= \sum_{j=0}^n \alpha_j r^j - \bar{h} \sum_{j=0}^n \beta_j r^j \\ &= \rho(r) - \bar{h} \sigma(r). \end{aligned} \quad (6.42)$$

Observações

1. Zero-estabilidade: $h \rightarrow 0$.

Observe-se em (6.42) que, se $h \rightarrow 0$, então

$$\pi(r) \rightarrow \rho(r).$$

2. Estabilidade absoluta ou estabilidade fraca: h fixo.

6.7 Estabilidade absoluta

Definição 6.5 (Estabilidade absoluta). *Um método de passo múltiplo linear é absolutamente estável se, para $\bar{h} = h\lambda$ dado (\bar{h} fixo), todas as raízes do polinômio*

$$\pi(r) = \rho(r) - \bar{h} \sigma(r)$$

associado à equação de diferenças linear

$$\sum_{j=0}^n (\alpha_j - \bar{h}\beta_j) e_{k+j} = \bar{\phi}$$

satisfazem

$$|r_j| < 1, \quad j = 1, 2, \dots, n. \quad (6.43)$$

Caso a desigualdade (6.43) não seja observada, o método de passo múltiplo linear é *absolutamente instável*. Além disso, o intervalo (α, β) da reta real é denominado *intervalo de estabilidade absoluta* se o método de passo múltiplo linear for absolutamente estável para todo $\bar{h} \in (\alpha, \beta)$. Se para todo $\bar{h} \in (\alpha, \beta)$ o método de passo múltiplo linear for absolutamente instável, então não há intervalo de estabilidade absoluta.

Sabe-se que $r_1 = 1$ é a raiz principal do polinômio característico $\rho(r)$. Logo,

$$\bar{h} \rightarrow 0 \Rightarrow \pi(r) \rightarrow \rho(r) \Rightarrow r_1 \rightarrow 1.$$

Pode-se mostrar que

$$r_1 = e^{\bar{h}} + O(\bar{h}^{p+1}). \quad (6.44)$$

Considerando $\bar{h} > 0$ em (6.44), conclui-se que

$$\bar{h} > 0 \Rightarrow \lambda > 0 \Rightarrow r_1 > 1,$$

ou seja, o método de passo múltiplo linear é absolutamente instável.

Conclusão: Todo método de passo múltiplo linear zero-estável não possui intervalo de estabilidade absoluta contido na parte positiva do eixo real.

Na prática, pode-se determinar intervalos de estabilidade absoluta da seguinte forma:

1. Calcular as raízes r_j do polinômio $\pi(r) = \rho(r) - \bar{h} \sigma(r)$ para um conjunto de valores de \bar{h} nas vizinhanças da origem;
2. Representar graficamente $|r_j|$ ($n \times \bar{h}$);
3. Observar os intervalos para os quais $|r_j| < 1$.

6.7.1 Exercícios

Exercício 6.11. Reveja as noções de estabilidade absoluta e de zero-estabilidade e suas relações com a convergência.

Exercício 6.12. Veja em Lambert [10], Capítulo 3, as noções de estabilidade relativa.

Exercício 6.13. *Seja o método de passo múltiplo linear*

$$y_{k+2} - (1+a)y_{k+1} + ay_k = \frac{h}{12} [(5+a)f_{k+2} + 8(1-a)f_{k+1} - (1+5a)f_k] \quad (6.45)$$

onde $-1 \leq a < 1$.

1. *Mostre que o intervalo de estabilidade absoluta do método (6.45) é*

$$\left(\frac{6(a+1)}{a-1}, 0 \right).$$

2. *Para $a = -0,9$, use o método (6.45) para solucionar o p.v.i.*

$$\begin{cases} \frac{d}{dt}y(t) = -20y(t), & t \in [0, 1] \\ y(0) = 1 \end{cases} \quad (6.46)$$

com $h = 0,01$, $h = 0,02$ e $h = 0,04$.

6.8 Limitante para o erro local de discretização

O produto do passo de integração h pelo erro local de discretização de um método de passo múltiplo linear de n -passos de ordem p se escreve como

$$\begin{aligned} h\alpha &= \sum_{j=0}^n \alpha_j y(t_k + jh) - h \sum_{j=0}^n \beta_j f(t_k + jh, y(t_k + jh)) \\ &= \sum_{j=0}^n \alpha_j y(t_k + jh) - h \sum_{j=0}^n \beta_j y^{(1)}(t_k + jh) \end{aligned} \quad (6.47)$$

ou

$$h\alpha = C_{p+1} h^{p+1} y^{(p+1)}(\xi). \quad (6.48)$$

Na Série de Taylor

$$y(t+h) = y(t) + hy^{(1)}(t) + \frac{h^2}{2!}y^{(2)}(t) + \dots + \frac{h^p}{p!}y^{(p)}(t) + R_{p+1},$$

o resto R_{p+1} pode assumir as seguintes formas:

- diferencial

$$\frac{h^{p+1}}{(p+1)!}y^{(p+1)}(\xi) \text{ com } \xi \in (a, b); \quad (6.49)$$

- integral

$$\frac{1}{p!} \int_0^h (h - \xi)^p y^{(p+1)}(t_k + \xi) d\xi. \quad (6.50)$$

Efetuada a substituição

$$\xi = sh \quad (\xi \rightarrow 0 \Rightarrow s \rightarrow 0, \quad \xi \rightarrow jh \Rightarrow s \rightarrow j) \Rightarrow \frac{d\xi}{ds} = h \Rightarrow d\xi = hds$$

no resto na forma integral (6.50), tem-se que

$$\frac{1}{p!} \int_0^{jh} (jh - \xi)^p y^{(p+1)}(t_k + \xi) d\xi = \frac{h^{p+1}}{p!} \int_0^j (j - s)^p y^{(p+1)}(t_k + sh) ds.$$

Substituindo as expansões em Série de Taylor

$$\begin{aligned} y(t_k + jh) &= y(t_k) + jhy^{(1)}(t_k) + \frac{(jh)^2}{2!} y^{(2)}(t_k) + \dots + \frac{(jh)^p}{p!} y^{(p)}(t_k) + \\ &+ \frac{h^{p+1}}{p!} \int_0^j (j - s)^p y^{(p+1)}(t_k + sh) ds, \end{aligned} \quad (6.51)$$

$$\begin{aligned} y^{(1)}(t_k + jh) &= y^{(1)}(t_k) + jhy^{(2)}(t_k) + \frac{(jh)^2}{2!} y^{(3)}(t_k) + \dots + \frac{(jh)^{p-1}}{(p-1)!} y^{(p)}(t_k) + \\ &+ \frac{h^p}{(p-1)!} \int_0^j (j - s)^{p-1} y^{(p+1)}(t_k + sh) ds, \end{aligned} \quad (6.52)$$

em (6.47), obtém-se

$$\begin{aligned} h\alpha &= \frac{h^{p+1}}{p!} \sum_{j=0}^n \left(\alpha_j \int_0^j (j - s)^p y^{(p+1)}(t_k + sh) ds \right) + \\ &- \frac{h^{p+1}}{(p-1)!} \sum_{j=0}^n \left(\beta_j \int_0^j (j - s)^{p-1} y^{(p+1)}(t_k + sh) ds \right), \end{aligned} \quad (6.53)$$

uma vez que o método tem ordem p .

Como $j - s \geq 0$, uma vez que $0 \leq s \leq j$, a expressão (6.53) pode ser reescrita de uma forma mais compacta usando-se a função de corte

$$z_+ = \begin{cases} z, & \text{se } z \geq 0 \\ 0, & \text{se } z < 0 \end{cases}. \quad (6.54)$$

Assim,

$$\begin{aligned}
h\alpha &= \frac{h^{p+1}}{p!} \sum_{j=0}^n \left(\alpha_j \int_0^n (j-s)_+^p y^{(p+1)}(t_k + sh) ds \right) + \\
&\quad - \frac{h^{p+1}}{(p-1)!} \sum_{j=0}^n \left(\beta_j \int_0^n (j-s)_+^{p-1} y^{(p+1)}(t_k + sh) ds \right) \\
&= \frac{h^{p+1}}{p!} \int_0^n \left\{ \sum_{j=0}^n \left[\alpha_j (j-s)_+^p - p\beta_j (j-s)_+^{p-1} \right] y^{(p+1)}(t_k + sh) \right\} ds \\
&= \frac{h^{p+1}}{p!} \int_0^n G(s) y^{(p+1)}(t_k + sh) ds, \tag{6.55}
\end{aligned}$$

sendo

$$G(s) = \sum_{j=0}^n \left[\alpha_j (j-s)_+^p - p\beta_j (j-s)_+^{p-1} \right] \tag{6.56}$$

a função influência do método de passo múltiplo linear.

Pode-se então estabelecer a delimitação

$$\begin{aligned}
|h\alpha| &= \left| \frac{h^{p+1}}{p!} \int_0^n G(s) y^{(p+1)}(t_k + sh) ds \right| \\
&\leq \frac{h^{p+1}}{p!} \int_0^n |G(s)| \left| y^{(p+1)}(t_k + sh) \right| ds. \tag{6.57}
\end{aligned}$$

Empregando o Teorema do Valor Médio para Integrais, pode-se reescrever a desigualdade (6.57) como

$$\begin{aligned}
|h\alpha| &\leq \frac{h^{p+1}}{p!} \left| y^{(p+1)}(\xi) \right| \int_0^n |G(s)| ds \\
&\leq \frac{h^{p+1}}{p!} \mathbb{G} \mathbb{Y}, \tag{6.58}
\end{aligned}$$

onde

$$\mathbb{G} = \int_0^n |G(s)| ds$$

e

$$\mathbb{Y} = \max_{\xi \in [a, b]} \left| y^{(p+1)}(\xi) \right|.$$

Aplicando o Teorema do Valor Médio para Integrais à igualdade (6.55) (supondo que $y^{(p+1)}(t_k + sh)$ mantenha o sinal para $s \in [0, n]$), constata-se que

$$\begin{aligned} h\alpha &= \frac{h^{p+1}}{p!} \int_0^n G(s)y^{(p+1)}(t_k + sh)ds \\ &= \frac{h^{p+1}}{p!} y^{(p+1)}(\xi) \int_0^n G(s)ds. \end{aligned} \quad (6.59)$$

Como

$$h\alpha = C_{p+1}h^{p+1}y^{(p+1)}(\xi), \quad (6.60)$$

ao se comparar (6.59) e (6.60) conclui-se que

$$C_{p+1} = \frac{1}{p!} \int_0^n G(s)ds. \quad (6.61)$$

6.8.1 Exercícios

Exercício 6.14. *Seja o método de passo múltiplo linear de 2-passos consistente zero-estável definido por*

$$\alpha_2 = 1, \alpha_1 = -1, \alpha_0 = 0, \beta_2 = 0, \beta_1 = \frac{3}{2} \text{ e } \beta_0 = -\frac{1}{2}. \quad (6.62)$$

1. *Construa o gráfico da função influência $G(s)$ do método de passo múltiplo linear (6.62).*
2. *Calcule $\int_0^n G(s)ds$.*

Observação: Use aplicativos (Maple, Matlab ou outros) ou faça um programa em linguagem C para solucionar o exercício.

6.9 Limitante para o erro global de discretização

Um método de passo múltiplo linear

$$\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j} \quad (6.63)$$

é de classe A se

$$\begin{aligned} \alpha_n &= 1, \\ \alpha_j &\leq 0, \quad j = 0, 1, 2, \dots, n-1, \\ \sum_{j=0}^n \alpha_j &= 0. \end{aligned}$$

Para um método de passo múltiplo linear de classe A explícito, pode-se mostrar (Lambert [10]) que o erro global de discretização é limitado por

$$|e_k| \leq [\delta + (t_k - t_0)(h^p \mathbb{G}\mathbb{Y} + h^q K_1)] e^{LB(t_k - t_0)}, \quad (6.64)$$

onde

$$\begin{aligned} \delta &= \max_{\mu=0,1,\dots,n-1} |e_\mu|, \\ t_k &: \text{ instante de estudo,} \\ t_0 &: \text{ instante inicial,} \\ h &: \text{ passo de integração,} \\ p &: \text{ ordem do método,} \\ \mathbb{G} \text{ e } \mathbb{Y} &: \text{ constantes de delimitação do erro local de discretização,} \\ q \text{ e } K_1 &: \text{ constantes do erro de truncamento da solução numérica,} \\ L &: \text{ constante de Lipschitz,} \\ B &= \sum_{j=0}^{n-1} |\beta_j|. \end{aligned}$$

Para um método de passo múltiplo linear qualquer, se

$$|h\beta_n\alpha_n^{-1}|L < 1, \quad (6.65)$$

então o erro global de discretização é limitado por

$$|e_k| \leq \Gamma^* [An\delta + (t_k - t_0)(h^p \mathbb{G}\mathbb{Y} + h^q K_1)] e^{\Gamma^* LB(t_k - t_0)}, \quad (6.66)$$

onde

$$\begin{aligned} \Gamma^* &= \frac{\Gamma}{1 - h|\beta_n\alpha_n^{-1}|L}, \\ \Gamma &= \sup_{\ell=0,1,\dots} |\gamma_\ell|, \\ A &= \sum_{j=0}^n |\alpha_j|, \\ n &: \text{ número de passos do método,} \\ \delta &= \max_{\mu=0,1,\dots,n-1} |e_\mu|, \\ t_k &: \text{ instante de estudo,} \\ t_0 &: \text{ instante inicial,} \\ h &: \text{ passo de integração,} \end{aligned}$$

- p : ordem do método,
 \mathbb{G} e \mathbb{Y} : constantes de delimitação do erro local de discretização,
 q e K_1 : constantes do erro de truncamento da solução numérica,
 L : constante de Lipschitz,
 $B = \sum_{j=0}^n |\beta_j|$,

com

$$\frac{1}{\alpha_n + \alpha_{n-1}\xi + \dots + \alpha_0\xi^n} = \gamma_0 + \gamma_1\xi + \gamma_2\xi^2 + \dots.$$

6.9.1 Exercícios

Exercício 6.15. *Dê exemplos de métodos de passo múltiplo lineares de classe A explícitos ($\beta_n = 0$).*

Exercício 6.16. *Deduza a limitação para o erro global de discretização*

$$|e_k| \leq [\delta + (t_k - t_0)(h^p\mathbb{G}\mathbb{Y} + h^qK_1)] e^{LB(t_k - t_0)}$$

em um método de passo múltiplo linear de classe A explícito (Veja Lambert [10], página 55).

6.10 Suplemento teórico

6.10.1 Teorema do Valor Médio para Integrais

Teorema 6.3. *Se $\phi(t)$ é uma função contínua e $g(t)$ é uma função integrável de sinal constante no intervalo $[a, b]$, então existe $\xi \in (a, b)$ tal que*

$$\int_a^b \phi(t)g(t)dt = \phi(\xi) \int_a^b g(t)dt.$$

6.11 Exercícios resolvidos

1. Use o método de passo múltiplo linear

$$y_{k+2} + 4y_{k+1} + y_k = \frac{h}{2} [8f_{k+1} + 4f_k] \quad (6.67)$$

para calcular a solução numérica do p.v.i.

$$\begin{cases} \dot{y} = 4t\sqrt{y} \\ y(0) = 1 \end{cases} \quad 0 \leq t \leq 2 \quad (6.68)$$

com $h = 0,1$, $h = 0,05$ e $h = 0,025$. Comente os resultados obtidos.

Solução

Solução analítica do p.v.i.:

$$\begin{aligned} \frac{dy}{dt} &= 4t\sqrt{y} \\ \frac{1}{\sqrt{y}} \frac{dy}{dt} &= 4t \\ \int \frac{1}{\sqrt{y}} \frac{dy}{dt} dt &= \int 4t dt \\ \frac{y^{\frac{1}{2}}}{\frac{1}{2}} &= \frac{4t^2}{2} + C \\ y^{\frac{1}{2}} &= t^2 + \frac{C}{2} \\ y &= \left(t^2 + \frac{C}{2}\right)^2 \\ y(0) = 1 &\Rightarrow C = 2 \\ y(t) &= (t^2 + 1)^2. \end{aligned} \quad (6.69)$$

O método (6.67) foi implementado utilizando a solução analítica (6.69) para obter y_1 , ou seja,

$$y_1 = y(t_0 + h) = y(h) = (h^2 + 1)^2.$$

Este dado é necessário para iniciar o cálculo das aproximações em um método de 2-passos.

O método (6.67) não é consistente, nem zero-estável (verifique!). Logo, divergente. A redução no passo de integração h aumenta o erro global de discretização. O método (6.67) gera $y_{k's}$ negativos a partir de certos valores de t_k , para os quais $f(t_k, y_k)$ tem imagens complexas.

Considerando um ponto fixo $t = 0,3$, pode-se perceber que a redução em h provoca aumento no erro. As Tabelas (6.2), (6.3) e (6.4) apresentam os valores gerados até o ponto em que y_k se torna negativo.

h = 0,1			
t_k	$y(t_k)$	y_k	$ y(t_k) - y_k $
0,000	1,000000	1,000000	0,000000
0,100	1,020100	1,020100	0,000000
0,200	1,081600	1,081200	0,000400
0,300	1,188100	1,189238	0,001138
0,400	1,345600	1,338866	0,006734
0,500	1,562500	1,592994	0,030494
0,600	1,849600	1,702337	0,147263
0,700	2,220100	2,913023	0,692923
0,800	2,689600	-0,602567	3,292167

Tabela 6.2: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,1$.

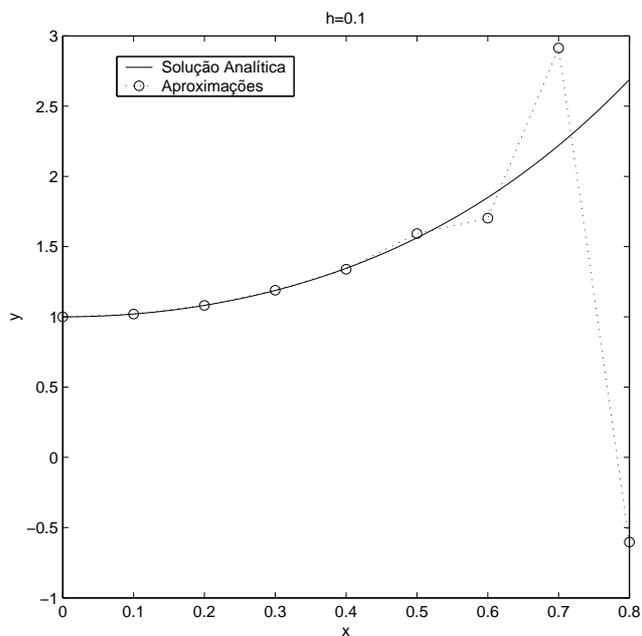


Figura 6.1: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,1$.

h = 0,05			
t_k	$y(t_k)$	y_k	$ y(t_k) - y_k $
0,000	1,000000	1,000000	0,000000
0,050	1,005006	1,005006	0,000000
0,100	1,020100	1,020075	0,000025
0,150	1,045506	1,045580	0,000074
0,200	1,081600	1,081158	0,000442
0,250	1,128906	1,130988	0,002082
0,300	1,188100	1,177715	0,010385
0,350	1,260006	1,310883	0,050877
0,400	1,345600	1,095852	0,249748
0,450	1,446006	2,666284	1,220278
0,500	1,562500	-4,430548	5,993048

Tabela 6.3: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,05$.

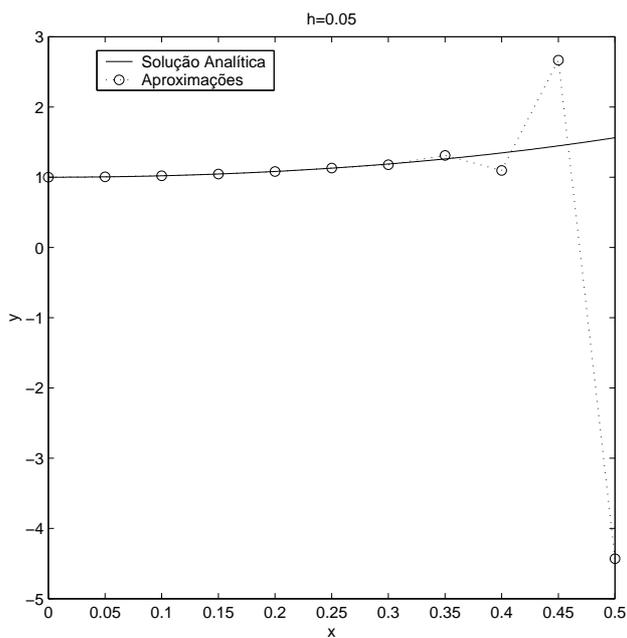


Figura 6.2: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,05$.

h = 0,025			
t_k	$y(t_k)$	y_k	$ y(t_k) - y_k $
0,000	1,000000	1,000000	0,000000
0,025	1,001250	1,001250	0,000000
0,050	1,005006	1,005005	0,000002
0,075	1,011282	1,011286	0,000005
0,100	1,020100	1,020072	0,000028
0,125	1,031494	1,031627	0,000133
0,150	1,045506	1,044835	0,000672
0,175	1,062188	1,065521	0,003334
0,200	1,081600	1,065009	0,016591
0,225	1,103813	1,186258	0,082445
0,250	1,128906	0,719318	0,409588
0,275	1,156969	3,187841	2,030872
0,300	1,188100	-8,915969	10,104069

Tabela 6.4: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,025$.

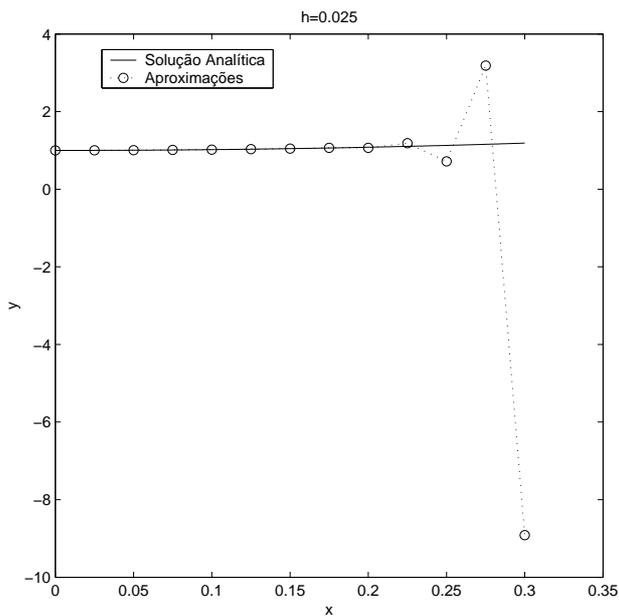


Figura 6.3: Comparação das soluções exata e numérica do problema de valor inicial (6.68), onde a solução numérica foi obtida através do método (6.67) com $h = 0,025$.

2. Constate a instabilidade do método de passo múltiplo linear

$$y_{k+2} - y_{k+1} = \frac{h}{3} [3f_{k+1} - 2f_k] \quad (6.70)$$

empregando-o para solucionar numericamente o p.v.i.

$$\begin{cases} \dot{y} = 4t\sqrt{y} \\ y(0) = 1 \end{cases} \quad 0 \leq t \leq 2, \quad (6.71)$$

cuja solução exata é $y(t) = (t^2 + 1)^2$.

Solução

O método (6.70) é zero-estável, porém inconsistente (verifique!). Logo, divergente. As Figuras (6.4) e (6.5) foram geradas com passos de integração $h = 0,1$, $h = 0,05$ e $h = 0,025$. Note-se que a distância entre as aproximações e a solução analítica aumenta à medida que o passo de integração é reduzido, fato que caracteriza a instabilidade do método numérico (6.70).

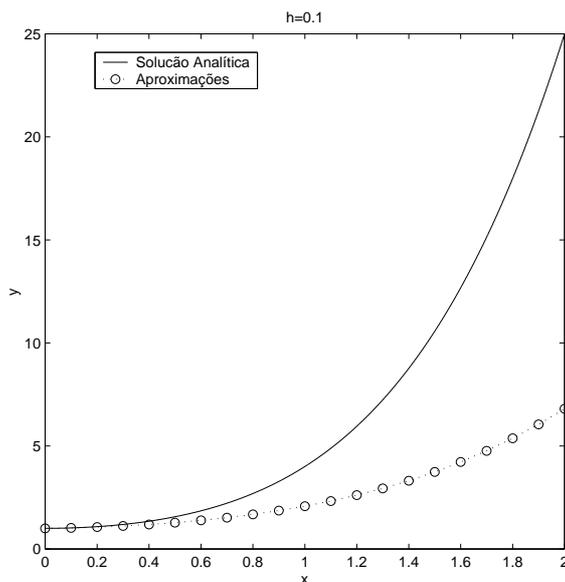


Figura 6.4: Comparação das soluções exata e numérica do problema de valor inicial (6.71), onde a solução numérica foi obtida através do método (6.70) com $h = 0,1$.

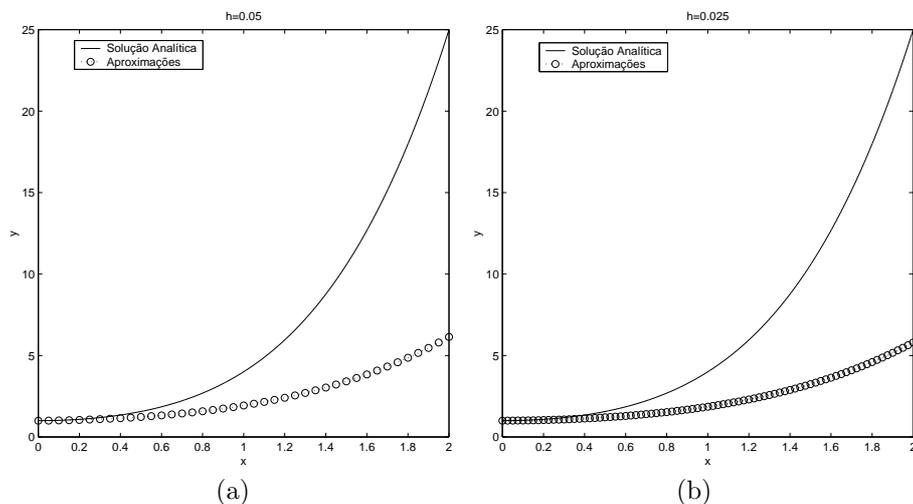


Figura 6.5: Comparação das soluções exata e numérica do problema de valor inicial (6.71), onde a solução numérica foi obtida através do método (6.70) com (a) $h = 0,05$ e (b) $h = 0,025$.

3. Ilustre o efeito da instabilidade usando o Método de Runge-Kutta Clássico para calcular a solução numérica do p.v.i.

$$\begin{cases} \dot{y} = -5ty^2 + \frac{5}{t} - \frac{1}{t^2}, & 1 \leq t \leq 4 \\ y(1) = 1 \end{cases} \quad (6.72)$$

com:

(a) $h = 0,2$;

(b) $h = 0,4$.

Solução

Solução exata do p.v.i.: $y = \frac{1}{t}$.

Observação: Busque informações em Boyce & Diprima [15] a respeito da Equação de Riccati.

A Tabela (6.5) apresenta o erro global de discretização cometido ao se empregar o Método de Runge-Kutta Clássico com $h = 0,2$ e $h = 0,4$ para solucionar o p.v.i. (6.74).

t_k	$y(t_k)$	h=0,2	h=0,4
		$ y(t_k) - y_k $	$ y(t_k) - y_k $
1,0	1,00000	0,00E+00	0,00E+00
1,4	0,71429	4,94E-03	9,99E-02
1,8	0,55556	2,64E-03	8,88E-01
2,2	0,45455	1,37E-03	1,64E-01
2,6	0,38462	7,81E-04	7,78E-01
3,0	0,33333	4,84E-04	4,39E+00
3,4	0,29412	3,21E-04	3,05E+17
3,8	0,26316	2,23E-04	1,12E+288
4,2	0,23810	1,62E-04	overflow

Tabela 6.5: Comparação das soluções exata e numérica do problema de valor inicial (6.74), onde a solução numérica foi obtida através do Método de Runge-Kutta Clássico com $h = 0,2$ e $h = 0,4$.

Estabilidade absoluta:

$$\begin{aligned}
 \lambda &= \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} \left[-5ty^2 + \frac{5}{t} - \frac{1}{t^2} \right] \\
 &= -10yt, \\
 \lambda &= 10 \left(\frac{1}{t} \right) t \\
 &= -10.
 \end{aligned} \tag{6.73}$$

O intervalo de estabilidade absoluta para o Método de Runge-Kutta Clássico é $(-2, 78, 0)$.

- Para $h = 0,2$, $\lambda h = f_y h = -10 \cdot 0,2 = -2,0 \in (-2, 78, 0)$. O método é estável para $h = 0,2$.
- Para $h = 0,4$, $\lambda h = f_y h = -10 \cdot 0,4 = -4,0 \notin (-2, 78, 0)$. O método é instável para $h = 0,4$.

A Figura (6.6) ilustra a estabilidade do método para o passo de integração $h = 0,2$ e a instabilidade do método para $h = 0,4$.

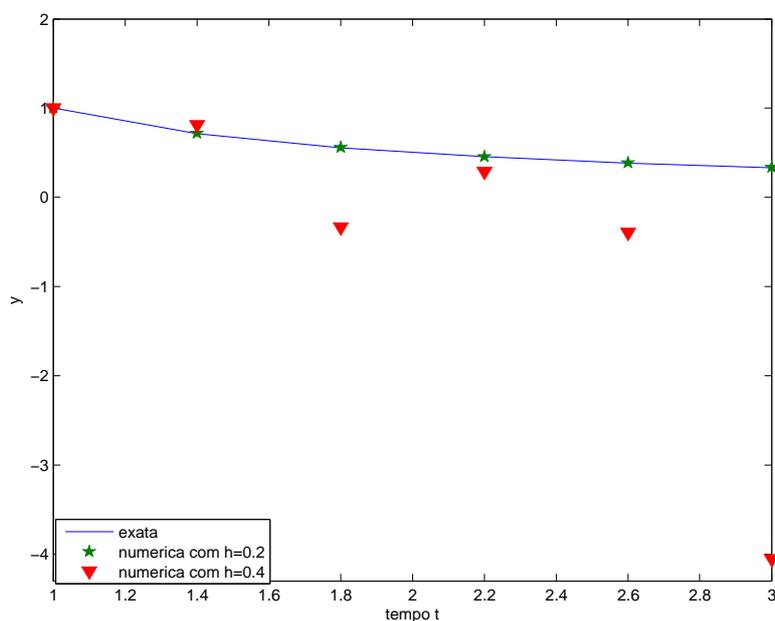


Figura 6.6: Comparação das soluções exata e numérica do problema de valor inicial (6.74), onde a solução numérica foi obtida através do Método de Runge-Kutta Clássico com $h = 0,2$ e $h = 0,4$.

4. O Método de Quade (“Quade’s method”) é um método de passo múltiplo linear definido por

$$y_{k+4} - \frac{8}{19}(y_{k+3} - y_{k+1}) - y_k = \frac{6h}{19}(f_{k+4} + 4f_{k+3} + 4f_{k+1} + f_k). \quad (6.74)$$

- (a) Determine a ordem de consistência e a constante do erro local de truncamento do método (6.74).

Solução

$$\alpha_4 = 1, \quad \alpha_3 = -\frac{8}{19}, \quad \alpha_2 = 0, \quad \alpha_1 = \frac{8}{19}, \quad \alpha_0 = -1$$

$$\beta_4 = \frac{6}{19}, \quad \beta_3 = \frac{24}{19}, \quad \beta_2 = 0, \quad \beta_1 = \frac{24}{19}, \quad \beta_0 = \frac{6}{19}$$

$$C_0 = \sum_{j=0}^4 \alpha_j = -1 + \frac{8}{19} - \frac{8}{19} + 1 = 0$$

$$\begin{aligned}
C_1 &= \sum_{j=0}^4 j\alpha_j - \sum_{j=0}^4 \beta_j \\
&= \left(\frac{8}{19} - \frac{24}{19} + 4 \right) - \left(\frac{12}{19} + \frac{48}{19} \right) \\
&= \frac{60}{19} - \frac{60}{19} = 0 \\
C_2 &= \sum_{j=0}^4 \frac{j^2}{2!} \alpha_j - \sum_{j=0}^4 j\beta_j \\
&= \frac{1}{2} \left(\frac{8}{19} - \frac{72}{19} + 16 \right) - \left(\frac{24}{19} + \frac{72}{19} + \frac{24}{19} \right) \\
&= \frac{120}{19} - \frac{120}{19} = 0 \\
C_3 &= \sum_{j=0}^4 \frac{j^3}{3!} \alpha_j - \sum_{j=0}^4 \frac{j^2}{2!} \beta_j \\
&= \frac{1}{6} \left(\frac{8}{19} - \frac{216}{19} + 64 \right) - \frac{1}{2} \left(\frac{24}{19} + \frac{216}{19} + \frac{96}{19} \right) \\
&= \frac{168}{19} - \frac{168}{19} = 0 \\
C_4 &= \sum_{j=0}^4 \frac{j^4}{4!} \alpha_j - \sum_{j=0}^4 \frac{j^3}{3!} \beta_j \\
&= \frac{1}{24} \left(\frac{8}{19} - \frac{648}{19} + 256 \right) - \frac{1}{6} \left(\frac{24}{19} + \frac{648}{19} + \frac{384}{19} \right) \\
&= \frac{176}{19} - \frac{176}{19} = 0 \\
C_5 &= \sum_{j=0}^4 \frac{j^5}{5!} \alpha_j - \sum_{j=0}^4 \frac{j^4}{4!} \beta_j = \\
&= \frac{1}{120} \left(\frac{8}{19} - \frac{1944}{19} + 1024 \right) - \frac{1}{24} \left(\frac{24}{19} + \frac{1944}{19} + \frac{1536}{19} \right) \\
&= \frac{146}{19} - \frac{146}{19} = 0 \\
C_6 &= \sum_{j=0}^4 \frac{j^6}{6!} \alpha_j - \sum_{j=0}^4 \frac{j^5}{5!} \beta_j = \\
&= \frac{1}{720} \left(\frac{8}{19} - \frac{5832}{19} + 4096 \right) - \frac{1}{120} \left(\frac{24}{19} + \frac{5832}{19} + \frac{6144}{19} \right) \\
&= \frac{100}{19} - \frac{100}{19} = 0
\end{aligned}$$

$$\begin{aligned}
C_7 &= \sum_{j=0}^4 \frac{j^7}{7!} \alpha_j - \sum_{j=0}^4 \frac{j^6}{6!} \beta_j = \\
&= \frac{1}{5040} \left(\frac{8}{19} - \frac{17496}{19} + 16384 \right) + \\
&\quad - \frac{1}{720} \left(\frac{24}{19} + \frac{17496}{19} + \frac{24576}{19} \right) \\
&= \frac{6121}{1995} - \frac{877}{285} = -\frac{6}{665}
\end{aligned}$$

O Método de Quade (6.74) é consistente ($C_0 = C_1 = 0$) de ordem 6. O erro local de truncamento do método é

$$\alpha = C_7 h^6 y^{(7)}(\xi)$$

com constante $C_7 = -\frac{6}{665}$.

(b) Verifique que o método (6.74) é convergente.

Solução

Primeiro polinômio característico:

$$\begin{aligned}
r^4 - \frac{8}{19}r^3 + \frac{8}{19}r - 1 &= r^4 - 1 - \frac{8}{19}r(r^2 - 1) \\
&= (r^2 - 1)(r^2 + 1) - \frac{8}{19}r(r^2 - 1) \\
&= (r^2 - 1) \left(r^2 - \frac{8}{19}r + 1 \right) \\
&= (r + 1)(r - 1) \left(r^2 - \frac{8}{19}r + 1 \right).
\end{aligned} \tag{6.75}$$

Raízes do polinômio (6.75):

$$\begin{aligned}
r + 1 = 0 &\Rightarrow r_1 = -1; \\
r - 1 = 0 &\Rightarrow r_2 = 1; \\
r^2 - \frac{8}{19}r + 1 = 0 &\Rightarrow r_3 = \frac{4}{19} + \frac{i}{19}\sqrt{345}; \\
&\quad r_4 = \frac{4}{19} - \frac{i}{19}\sqrt{345}.
\end{aligned}$$

Módulo das raízes r_3 e r_4 :

$$|r_3| = |r_4| = \left[\left(\frac{4}{19} \right)^2 + \left(\frac{\sqrt{345}}{19} \right)^2 \right]^{\frac{1}{2}} = \left(\frac{16}{361} + \frac{345}{361} \right)^{\frac{1}{2}} = 1.$$

Como as quatro raízes do polinômio (6.75) são distintas e têm módulo 1, o método é zero-estável.

O Método de Quade (6.74) é convergente, uma vez que é consistente e zero-estável.

5. Seja o método de passo múltiplo linear

$$y_{k+3} + \alpha(y_{k+2} - y_{k+1}) - y_k = \frac{h}{2}(3 + \alpha)(f_{k+2} + f_{k+1}). \quad (6.76)$$

(a) Determine para quais valores de α o método (6.76) é zero-estável.

Solução

$$\alpha_3 = 1, \alpha_2 = \alpha, \alpha_1 = -\alpha, \alpha_0 = -1$$

Primeiro polinômio característico:

$$\begin{aligned} r^3 + \alpha r^2 - \alpha r - 1 &= (r^3 - 1) + \alpha(r^2 - r) \\ &= (r - 1)(r^2 + r + 1) + \alpha r(r - 1) \\ &= (r - 1)[r^2 + (1 + \alpha)r + 1]. \end{aligned} \quad (6.77)$$

Uma das raízes do polinômio (6.77) é $r_1 = 1$. As outras duas raízes são dadas por

$$r_{2,3} = \frac{-(1 + \alpha) \pm \sqrt{(1 + \alpha)^2 - 4}}{2}. \quad (6.78)$$

i. Considerando $\alpha = 1$ ou $\alpha = -3$ em (6.78):

$$\begin{aligned} \alpha = 1 &\Rightarrow (1 + \alpha)^2 - 4 = 0 \Rightarrow r_2 = r_3 = -1; \\ \alpha = -3 &\Rightarrow (1 + \alpha)^2 - 4 = 0 \Rightarrow r_2 = r_3 = 1. \end{aligned}$$

Como em ambos os casos o polinômio tem raiz de multiplicidade 2 e módulo 1, o método não é zero-estável.

ii. Considerando $\alpha > 1$ ou $\alpha < -3$ em (6.78):

$$\begin{aligned} \alpha > 1 &\Rightarrow 1 + \alpha > 2 \Rightarrow -(1 + \alpha) < -2 \\ &\Rightarrow -\frac{1 + \alpha}{2} < -1 \\ &\Rightarrow -\frac{1 + \alpha}{2} - \frac{\sqrt{(1 + \alpha)^2 - 4}}{2} < -1 \\ &\Rightarrow r_3 \text{ é menor que } -1; \end{aligned}$$

$$\begin{aligned}
\alpha < -3 &\Rightarrow 1 + \alpha < -2 \Rightarrow -(1 + \alpha) > 2 \\
&\Rightarrow -\frac{1 + \alpha}{2} > 1 \\
&\Rightarrow -\frac{1 + \alpha}{2} + \frac{\sqrt{(1 + \alpha)^2 - 4}}{2} > 1 \\
&\Rightarrow r_2 \text{ é maior que } 1.
\end{aligned}$$

Como em ambos os casos o polinômio tem raiz de módulo maior que 1, o método não é zero-estável.

iii. Considerando $-3 < \alpha < 1$ em (6.78):

$$\begin{aligned}
-3 < \alpha < 1 &\Rightarrow -2 < 1 + \alpha < 2 \Rightarrow (1 + \alpha)^2 < 4 \\
&\Rightarrow r_{2,3} = -\frac{1 + \alpha}{2} \pm \frac{i}{2} \sqrt{4 - (1 + \alpha)^2}; \\
|r_2| = |r_3| &= \left[\frac{(1 + \alpha)^2}{4} + \frac{4 - (1 + \alpha)^2}{4} \right]^{\frac{1}{2}} = 1.
\end{aligned}$$

Para $\alpha \in (-3, 1)$ o método (6.76) é zero-estável, uma vez que as três raízes do primeiro polinômio característico (6.77) são distintas de módulo 1.

(b) Mostre que existe um valor de α para o qual o método (6.76) tem ordem 4, mas para que ele seja zero-estável sua ordem não pode exceder 2.

Solução

$$\begin{aligned}
\alpha_3 &= 1, \alpha_2 = \alpha, \alpha_1 = -\alpha, \alpha_0 = -1 \\
\beta_3 &= 0, \beta_2 = \frac{3 + \alpha}{2}, \beta_1 = \frac{3 + \alpha}{2}, \beta_0 = 0
\end{aligned}$$

$$C_0 = \sum_{j=0}^3 \alpha_j = -1 - \alpha + \alpha + 1 = 0$$

$$\begin{aligned}
C_1 &= \sum_{j=0}^3 j\alpha_j - \sum_{j=0}^3 \beta_j \\
&= -\alpha + 2\alpha + 3 - \left(\frac{3+\alpha}{2} + \frac{3+\alpha}{2} \right) \\
&= \alpha + 3 - 3 - \alpha = 0 \\
C_2 &= \sum_{j=0}^3 \frac{j^2}{2!} \alpha_j - \sum_{j=0}^3 j\beta_j \\
&= \frac{1}{2}(-\alpha + 4\alpha + 9) - \left(\frac{3+\alpha}{2} + 3 + \alpha \right) \\
&= \frac{1}{2}(9 + 3\alpha) - \frac{1}{2}(9 + 3\alpha) = 0 \\
C_3 &= \sum_{j=0}^3 \frac{j^3}{3!} \alpha_j - \sum_{j=0}^3 \frac{j^2}{2!} \beta_j \\
&= \frac{1}{6}(-\alpha + 8\alpha + 27) - \frac{1}{2} \left(\frac{3+\alpha}{2} + 4\frac{3+\alpha}{2} \right) \\
&= \frac{1}{6}(27 + 7\alpha) - \frac{1}{4}(15 + 5\alpha) \\
&= \frac{18 - 2\alpha}{24} = \frac{9 - \alpha}{12}
\end{aligned}$$

$$C_3 \neq 0 \Rightarrow \alpha \neq 9$$

Para que o método (6.76) seja zero-estável é necessário que a ordem do mesmo seja 2, uma vez que $\alpha = 9 \notin (-3, 1)$.

$$\begin{aligned}
C_4 &= \sum_{j=0}^3 \frac{j^4}{4!} \alpha_j - \sum_{j=0}^3 \frac{j^3}{3!} \beta_j \\
&= \frac{1}{24}(-\alpha + 16\alpha + 81) - \frac{1}{6} \left(\frac{3+\alpha}{2} + 8\frac{3+\alpha}{2} \right) \\
&= \frac{1}{24}(81 + 15\alpha) - \frac{1}{12}(27 + 9\alpha) \\
&= \frac{27 - 3\alpha}{24} = \frac{9 - \alpha}{8}
\end{aligned}$$

$$\alpha = 9 \Rightarrow C_4 = 0$$

$$\begin{aligned}C_5 &= \sum_{j=0}^3 \frac{j^5}{5!} \alpha_j - \sum_{j=0}^3 \frac{j^4}{4!} \beta_j \\&= \frac{1}{120} (-\alpha + 32\alpha + 243) - \frac{1}{24} \left(\frac{3+\alpha}{2} + 16 \frac{3+\alpha}{2} \right) \\&= \frac{1}{120} (243 + 31\alpha) - \frac{1}{48} (51 + 17\alpha) \\&= \frac{231 - 23\alpha}{240} \\ \alpha = 9 &\Rightarrow C_5 = \frac{1}{10}\end{aligned}$$

Para $\alpha = 9$, o método (6.76) tem ordem 4, porém não é zero-estável.

Capítulo 7

Métodos Preditores-Corretores

Para obter y_{k+n} em um método de passo múltiplo linear implícito, pode-se comparar métodos de passo múltiplo lineares explícitos e implícitos.

Resultado: Dado um método de passo múltiplo linear de n -passos ($n \geq 2$) zero-estável, a maior ordem que pode ser obtida é $n + 1$ se n for ímpar e $n + 2$ se n for par.

Método de passo múltiplo linear:

$$\begin{aligned}\sum_{j=0}^n \alpha_j y_{k+j} &= h \sum_{j=0}^n \beta_j f_{k+j}; \\ y_{k+n} + \sum_{j=0}^{n-1} \alpha_j y_{k+j} &= h \sum_{j=0}^n \beta_j f_{k+j}; \\ y_{k+n} &= -\sum_{j=0}^{n-1} \alpha_j y_{k+j} + h \sum_{j=0}^n \beta_j f_{k+j}.\end{aligned}\quad (7.1)$$

- Métodos explícitos:

$2n$ coeficientes ($\beta_n = 0$; impor $C_j = 0, 0 \leq j \leq 2n - 1$).

- Métodos implícitos:

$2n + 1$ coeficientes ($\beta_n \neq 0$; impor $C_j = 0, 0 \leq j \leq 2n$).

Como para métodos implícitos tem-se sempre um coeficiente β_j a mais, o método implícito terá ordem 1 a mais do que a ordem do método explícito com o

mesmo número de passos.

Parece não haver, baseando-se na ordem do método fixado o número de passos, uma real vantagem na utilização de um método implícito de n passos, pois bastaria considerar um método explícito de $n + 1$ passos para se obter um método de mesma ordem sem a necessidade de solucionar uma equação para y_{k+n} . Porém, há.

A real vantagem está no fato de que a amplitude do intervalo de estabilidade absoluta dos métodos implícitos é maior que a dos métodos explícitos de mesma ordem. As Tabelas (7.1) e (7.2) mostram o intervalo de estabilidade absoluta para os métodos explícitos de Adams-Bashforth e para os métodos implícitos de Adams-Moulton, respectivamente, para os quais o primeiro polinômio característico¹ é da forma

$$\rho(r) = r^j - r^{j-1}.$$

Nº de passos n	1	2	3	4
Ordem p	1	2	3	4
Coefficiente C_{p+1}	1/2	5/12	3/8	251/720
Estabilidade absoluta	$(-2, 0)$	$(-1, 0)$	$(-6/11, 0)$	$(-3/10, 0)$

Tabela 7.1: Características dos Métodos de Adams-Bashforth.

Nº de passos n	1	2	3	4
Ordem p	2	3	4	5
Coefficiente C_{p+1}	$-1/12$	$-1/24$	$-19/720$	$-3/160$
Estabilidade absoluta	$(-\infty, 0)$	$(-6, 0)$	$(-3, 0)$	$(-90/49, 0)$

Tabela 7.2: Características dos Métodos de Adams-Moulton.

Observando as Tabelas (7.1) e (7.2), constata-se que as vantagens dos métodos implícitos são muitas, entre elas menores valores absolutos para os coeficientes do erro local de discretização C_{p+1} e intervalos de estabilidade absoluta de maior amplitude. Vantagens como essas compensam o fato de que é necessário muitas vezes determinar y_{k+n} através de um método numérico para calcular as raízes de uma equação (dicotomia, ponto fixo e outros).

¹Métodos de passo múltiplo lineares para os quais o primeiro polinômio característico é da forma

$$\rho(r) = r^j - r^{j-2}$$

são chamados Métodos de Nyström se explícitos, e Métodos de Milne-Simpson se implícitos.

7.1 Definição

Considere o método de passo múltiplo linear implícito

$$y_{k+n} + \sum_{j=0}^{n-1} \alpha_j y_{k+j} = h\beta_n f(t_{k+n}, y_{k+n}) + h \sum_{j=0}^{n-1} \beta_j f_{k+j}. \quad (7.2)$$

Ao calcular y_{k+n} usando o Método do Ponto Fixo, tem-se que

$$y_{k+n}^{[s+1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j} = h\beta_n f(t_{k+n}, y_{k+n}^{[s]}) + h \sum_{j=0}^{n-1} \beta_j f_{k+j} \quad s = 0, 1, 2, \dots, \quad (7.3)$$

onde $y_{k+n}^{[0]}$ é uma aproximação inicial ("chute inicial"). As iterações em (7.3) convergem para y_{k+n} se

$$h < \frac{1}{L|\beta_n|}, \quad (7.4)$$

sendo L a constante de Lipschitz.

7.1.1 Exercícios

Exercício 7.1. Mostre a validade da restrição (7.4): $h < \frac{1}{L|\beta_n|}$.

Há duas opções para solucionar as iterações advindas da aplicação do Método do Ponto Fixo ao esquema implícito (7.2):

1^a iterar até que $|y_{k+n}^{[s+1]} - y_{k+n}^{[s]}| < \epsilon$, onde ϵ é uma precisão pré-fixada;

2^a iterar m vezes, com m fixo e $s = 0, 1, 2, \dots, m - 1$.

Para diminuir o número total de iterações na primeira opção, é necessário utilizar boas aproximações iniciais $y_{k+n}^{[0]}$. Observe que, dependendo de $y_{k+n}^{[0]}$ e ϵ , muitas avaliações (cálculos) da função f podem ser necessárias.

Uma forma conveniente de se obter $y_{k+n}^{[0]}$ é usar um método de passo múltiplo linear explícito de ordem p^* , denominado *preditor*, e depois calcular $y_{k+n}^{[1]}$ através de um método de passo múltiplo linear implícito de ordem p , denominado *corretor*. Pode-se denotar a estratégia da seguinte maneira:

P : uma aplicação do método explícito preditor, o qual gera uma boa aproximação para $y_{k+n}^{[0]}$;

E : um cálculo de f (evaluation);

C : uma aplicação do método implícito corretor, o qual gera a solução numérica para y_{k+n} através do Método do Ponto Fixo.

Em geral, a forma de aplicação do Método Predictor-Corretor é descrita como

$$P(EC)^m \quad (7.5)$$

ou

$$P(EC)^m E, \quad (7.6)$$

sendo

$$\begin{aligned} P &= y_{k+n}^{[0]}, \\ E &= f(t_{k+n}, y_{k+n}^{[0]}), \\ C &= y_{k+n}^{[1]}, \\ E &= f(t_{k+n}, y_{k+n}^{[1]}), \end{aligned}$$

e m o número fixo de vezes que $y_{k+n}^{[m]}$ será calculado. O modo (7.6) difere do modo (7.5) por uma avaliação a mais da função f . Obviamente, optar pelo modo (7.5) ou (7.6) afeta o próximo passo de integração.

O corretor, para m geral, é dado por

$$y_{k+n}^{[s+1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h\beta_n f(t_{k+n}, y_{k+n}^{[s]}) + h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y_{k+j}^{[m-t]}) \quad t = 0, 1. \quad (7.7)$$

Para $s = 0$ em (7.7), tem-se que

$$y_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h\beta_n f(t_{k+n}, y_{k+n}^{[0]}) + h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y_{k+j}^{[m-t]}),$$

onde $t = 1 \Rightarrow P(EC)^m$ e $t = 0 \Rightarrow P(EC)^m E$.

Para $s = 1$ em (7.7), tem-se que

$$y_{k+n}^{[2]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h\beta_n f(t_{k+n}, y_{k+n}^{[1]}) + h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y_{k+j}^{[m-t]}),$$

onde $t = 1 \Rightarrow P(EC)^m$ e $t = 0 \Rightarrow P(EC)^m E$.

Definição 7.1 (Método Preditor-Corretor). *Sejam os métodos de passo múltiplo lineares usados como preditor e corretor, respectivamente, definidos pelos polinômios característicos*

$$\rho^*(r) = \sum_{j=0}^n \alpha_j^* r^j, \quad \alpha_n^* = 1, \quad \sigma^*(r) = \sum_{j=0}^n \beta_j^* r^j$$

e

$$\rho(r) = \sum_{j=0}^n \alpha_j r^j, \quad \alpha_n = 1, \quad \sigma(r) = \sum_{j=0}^n \beta_j r^j.$$

Os modos $P(EC)^m E$ e $P(EC)^m$ são definidos como:

$P(EC)^m E$

$$y_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* y_{k+j}^{[m]} = h \sum_{j=0}^{n-1} \beta_j^* f_{k+j}^{[m]} \quad \text{prediz}$$

Para $s = 0, 1, \dots, m-1$

$$f_{k+n}^{[s]} = f(t_{k+n}, y_{k+n}^{[s]}) \quad \text{avalia}$$

$$y_{k+n}^{[s+1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h \beta_n f_{k+n}^{[s]} + h \sum_{j=0}^{n-1} \beta_j f_{k+j}^{[m]} \quad \text{corrige}$$

$$f_{k+n}^{[m]} = f(t_{k+n}, y_{k+n}^{[m]}) \quad \text{avalia};$$

$P(EC)^m$

$$y_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* y_{k+j}^{[m]} = h \sum_{j=0}^{n-1} \beta_j^* f_{k+j}^{[m-1]} \quad \text{prediz}$$

Para $s = 0, 1, \dots, m-1$

$$f_{k+n}^{[s]} = f(t_{k+n}, y_{k+n}^{[s]}) \quad \text{avalia}$$

$$y_{k+n}^{[s+1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h \beta_n f_{k+n}^{[s]} + h \sum_{j=0}^{n-1} \beta_j f_{k+j}^{[m-1]} \quad \text{corrige.}$$

Observações

1. Para $m \rightarrow \infty$, as características do Método Preditor-Corretor são aquelas do corretor (ordem de consistência, intervalo de estabilidade absoluta, ...);
2. Como considerar $m \rightarrow \infty$ é computacionalmente caro, em geral se considera $m = 2$.

7.2 Erro local de discretização

Sejam P o método preditor e C o método corretor, definidos por

$$P: y_{k+n}^{[s]} + \sum_{j=0}^{n-1} \alpha_j^* y_{k+j}^{[m]} = h \sum_{j=0}^{n-1} \beta_j^* f_{k+j}^{[m-t]}, \quad (7.8)$$

$$C: y_{k+n}^{[s+1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[m]} = h \beta_n f(t_{k+n}, y_{k+n}^{[s]}) + h \sum_{j=0}^{n-1} \beta_j f_{k+j}^{[m-t]}, \quad (7.9)$$

onde $s = 0, 1, \dots, m-1$, $t = 0 \Rightarrow P(EC)^m E$ e $t = 1 \Rightarrow P(EC)^m$.

O erro local de discretização principal dos métodos preditor e corretor é dado, respectivamente, por

$$P: y(t_{k+n}) - y_{k+n}^{[s]} = C_{p^*+1}^* h^{p^*+1} y^{(p^*+1)}(t_k) + O(h^{p^*+2}) = h\alpha^*, \quad (7.10)$$

$$C: y(t_{k+n}) - y_{k+n}^{[s+1]} = C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}) = h\alpha. \quad (7.11)$$

O produto do erro local de discretização pelo passo de integração h (suficientemente pequeno) para um método de passo múltiplo linear é dado por

$$\begin{aligned} h\alpha &= y(t_{k+n}) + \sum_{j=0}^{n-1} \alpha_j y(t_{k+j}) - h\beta_n f(t_{k+n}, y(t_{k+n})) + \\ &- h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y(t_{k+j})). \end{aligned} \quad (7.12)$$

Como até $j = n$ não se cometeu qualquer erro, a diferença (7.12) - (7.9) equivale a

$$h\alpha = y(t_{k+n}) - y_{k+n}^{[s+1]} - h\beta_n \left[f(t_{k+n}, y(t_{k+n})) - f(t_{k+n}, y_{k+n}^{[s]}) \right]. \quad (7.13)$$

Aplicando o Teorema do Valor Médio a (7.13), chega-se a

$$h\alpha = y(t_{k+n}) - y_{k+n}^{[s+1]} - h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(y(t_{k+n}) - y_{k+n}^{[s]} \right),$$

$$y(t_{k+n}) - y_{k+n}^{[s+1]} = h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(y(t_{k+n}) - y_{k+n}^{[s]} \right) + h\alpha, \quad (7.14)$$

onde $\eta_{k+n,s}$ pertence ao intervalo de extremos $y_{k+n}^{[s]}$ e $y(t_{k+n})$.

(a) Caso $p^* \geq p$ Substituindo (7.10) e (7.11) em (7.14) e considerando $s = 0$, tem-se que

$$\begin{aligned} y(t_{k+n}) - y_{k+n}^{[1]} &= h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(C_{p^*+1}^* h^{p^*+1} y^{(p^*+1)}(t_k) + O(h^{p^*+2}) \right) + \\ &+ C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}), \\ y(t_{k+n}) - y_{k+n}^{[1]} &= C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}). \end{aligned}$$

Pode-se agora substituir a expressão anterior e (7.11) em (7.14) com $s = 1$. Repetindo a substituição para $s = 2, 3, \dots, m-1$, conclui-se que

$$y(t_{k+n}) - y_{k+n}^{[m]} = C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}).$$

Logo, o erro local de discretização principal do Método Predictor-Corretor é o mesmo do corretor quando $p^* \geq p \quad \forall m \geq 1$.

(b) Caso $p^* = p - 1$ Substituindo (7.10) e (7.11) em (7.14) e considerando $s = 0$, tem-se que

$$\begin{aligned} y(t_{k+n}) - y_{k+n}^{[1]} &= h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(C_p^* h^p y^{(p)}(t_k) + O(h^{p+1}) \right) + \\ &+ C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}), \\ y(t_{k+n}) - y_{k+n}^{[1]} &= \beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(C_p^* h^{p+1} y^{(p)}(t_k) + O(h^{p+2}) \right) + \\ &+ C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}), \\ y(t_{k+n}) - y_{k+n}^{[1]} &= \left[\beta_n \frac{\partial f}{\partial y} C_p^* y^{(p)}(t_k) + C_{p+1} y^{(p+1)}(t_k) \right] h^{p+1} + O(h^{p+2}). \end{aligned}$$

Para $m = 1$, o erro local de discretização principal do Método Predictor-Corretor é da mesma ordem do corretor, porém não idêntico. Entretanto, com sucessivas substituições em (7.14), tem-se para $m \geq 2$

$$y(t_{k+n}) - y_{k+n}^{[m]} = C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}),$$

ou seja, o erro local de discretização principal do Método Predictor-Corretor se torna igual ao do corretor.

(c) Caso $p^* = p - 2$ Substituindo (7.10) e (7.11) em (7.14) e considerando $s = 0$, tem-se que

$$\begin{aligned}
y(t_{k+n}) - y_{k+n}^{[1]} &= h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(C_{p-1}^* h^{p-1} y^{(p-1)}(t_k) + O(h^p) \right) + \\
&\quad + C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}), \\
y(t_{k+n}) - y_{k+n}^{[1]} &= \beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n,s}) \left(C_{p-1}^* h^p y^{(p-1)}(t_k) + O(h^{p+1}) \right) + \\
&\quad + C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}), \\
y(t_{k+n}) - y_{k+n}^{[1]} &= \left[\beta_n \frac{\partial f}{\partial y} C_{p-1}^* y^{(p-1)}(t_k) + C_{p+1} h y^{(p+1)}(t_k) \right] h^p + O(h^{p+1}).
\end{aligned}$$

Para $m = 1$, o erro local de discretização principal do Método Predictor-Corretor tem ordem um a menos do que a ordem do corretor. Substituindo a expressão anterior e (7.11) em (7.14) com $s = 1$, verifica-se que

$$y(t_{k+n}) - y_{k+n}^{[2]} = \left[\left(\beta_n \frac{\partial f}{\partial y} \right)^2 C_{p-1}^* y^{(p-1)}(t_k) + \left(1 + h^2 \beta_n \frac{\partial f}{\partial y} \right) C_{p+1} y^{(p+1)}(t_k) \right] h^{p+1} + O(h^{p+2}).$$

Para $m = 2$, o erro local de discretização principal do Método Predictor-Corretor é da mesma ordem do corretor, porém não idêntico. Entretanto, com sucessivas substituições em (7.14), tem-se para $m \geq 3$

$$y(t_{k+n}) - y_{k+n}^{[m]} = C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}),$$

ou seja, o erro local de discretização principal do Método Predictor-Corretor se torna igual ao do corretor.

(d) Caso geral: $p^* = p - q$, $0 < q \leq p$.

Teorema 7.1 (Erro local de discretização principal). *Seja o Método Predictor-Corretor, onde o predictor tem ordem p^* e o corretor tem ordem p , aplicado no modo $P(EC)^m$ ou $P(EC)^m E$, com $p^*, p, m \in \mathbb{N}$, $p^* \geq 0$, $p \geq 1$ e $m \geq 1$.*

- Se $p^* \geq p$, então o erro local de discretização principal do Método Predictor-Corretor é igual ao do corretor.
- Se $p^* = p - q$, $0 < q \leq p$, então o erro local de discretização principal do Método Predictor-Corretor

(a) é igual ao do corretor quando $m \geq q + 1$;

(b) é da mesma ordem que a do corretor, porém não idêntico, quando $m = q$;

(c) é da forma $K h^{p-q+m+1} + O(h^{p-q+m+2})$ quando $m \leq q - 1$.

Observações

1^a: O erro local de discretização principal do Método Preditor-Corretor é determinado pela ordem do corretor.

2^a: Adotar $p^* > p$ requer geralmente um passo de integração desnecessariamente grande e a precisão (ordem elevada) do preditor não se reflete no erro local de discretização do Método Preditor-Corretor. Assim, a melhor escolha é dada por $p^* = p - m \geq 0$. Entretanto, $p^* = p$ é adequado para estimar o erro local de discretização do Método Preditor-Corretor.

7.3 Dispositivo de Milne

Esse dispositivo tem por finalidade fornecer uma estimativa para o erro local de discretização do Método Preditor-Corretor.

Sejam

$$P : C_{p^*+1}^* h^{p^*+1} y^{(p^*+1)}(t_k) + O(h^{p^*+2}) = y(t_{k+n}) - y_{k+n}^{[0]}, \quad (7.15)$$

$$C : C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2}) = y(t_{k+n}) - y_{k+n}^{[m]}. \quad (7.16)$$

Suponhamos $p^* = p$. Efetuando a diferença (7.16) - (7.15), tem-se, em primeira aproximação, que

$$(C_{p+1}^* - C_{p+1}) h^{p+1} y^{(p+1)}(t_k) = y_{k+n}^{[m]} - y_{k+n}^{[0]}. \quad (7.17)$$

Multiplicando (7.17) por C_{p+1} e por C_{p+1}^* , obtém-se

$$C_{p+1} h^{p+1} y^{(p+1)}(t_k) = \frac{C_{p+1}}{C_{p+1}^* - C_{p+1}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right), \quad (7.18)$$

$$C_{p+1}^* h^{p+1} y^{(p+1)}(t_k) = \frac{C_{p+1}^*}{C_{p+1}^* - C_{p+1}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right), \quad (7.19)$$

uma estimativa para o erro local de discretização principal do Método Preditor-Corretor e do preditor, respectivamente.

Pode-se melhorar $y_{k+n}^{[m]}$ com base na estimativa (7.18).

$$y(t_{k+n}) - y_{k+n}^{[m]} = C_{p+1} h^{p+1} y^{(p+1)}(t_k) + O(h^{p+2})$$

$$y(t_{k+n}) - \underbrace{\left(y_{k+n}^{[m]} + C_{p+1} h^{p+1} y^{(p+1)}(t_k) \right)}_{\hat{y}_{k+n}^{[m]}} = O(h^{p+2})$$

$$\hat{y}_{k+n}^{[m]} = y_{k+n}^{[m]} + C_{p+1} h^{p+1} y^{(p+1)}(t_k) \quad (7.20)$$

Substituindo (7.18) em (7.20), obtém-se

$$\hat{y}_{k+n}^{[m]} = y_{k+n}^{[m]} + \frac{C_{p+1}}{C_{p+1}^* - C_{p+1}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right). \quad (7.21)$$

Em (7.21), $\hat{y}_{k+n}^{[m]}$ é o *Modificador M* do corretor.

O dispositivo de Milne também pode ser empregado para aprimorar a solução do preditor. A expressão (7.19) não pode ser empregada pois, no momento de aplicar o preditor, $y_{k+n}^{[m]}$ não é conhecido. Nesse caso,

$$C_{p+1}^* h^{p+1} y^{(p+1)}(t_k) = \underbrace{C_{p+1}^* h^{p+1} y^{(p+1)}(t_{k-1})}_{\hat{y}_{k+n}^{[0]} - y_{k+n}^{[0]}} + O(h^{p+2}),$$

$$\hat{y}_{k+n}^{[0]} - y_{k+n}^{[0]} = C_{p+1}^* h^{p+1} y^{(p+1)}(t_{k-1}). \quad (7.22)$$

Substituindo (7.19) em (7.22), tem-se que

$$\hat{y}_{k+n}^{[0]} = y_{k+n}^{[0]} + \frac{C_{p+1}^*}{C_{p+1}^* - C_{p+1}} \left(y_{k+n-1}^{[m]} - y_{k+n-1}^{[0]} \right). \quad (7.23)$$

Em (7.23), $\hat{y}_{k+n}^{[0]}$ é o *Modificador M* do preditor.

Os modos $P(EC)^m E$ e $P(EC)^m$ podem ser agora reescritos com os modificadores do preditor e do corretor como

$$PM(ECM)^m E$$

e

$$PM(ECM)^m.$$

Observação: Para $p^* = p$, o dispositivo de Milne pode ser utilizado ou para aprimorar a solução $y_{k+n}^{[m]}$ ou para o controle automático do passo de integração, não para ambos. Caso se queira valores aproximados mais precisos, é preferível aumentar a ordem do método e usar o dispositivo de Milne para controlar o passo de integração.

7.4 Estabilidade absoluta

Sabe-se que o polinômio cujas raízes controlam as propriedades de estabilidade absoluta de um método de passo múltiplo linear é o polinômio

$$\pi(r) = \rho(r) - \bar{h}\sigma(r),$$

onde

$$\rho(r) = \sum_{j=0}^n \alpha_j r^j$$

é o primeiro polinômio característico,

$$\sigma(r) = \sum_{j=0}^n \beta_j r^j$$

é o segundo polinômio característico e

$$\bar{h} = h\lambda = h \frac{\partial f}{\partial y}$$

é uma constante.

Para determinar o polinômio característico $\pi(r)$ de um Método Preditor-Corretor aplicado ao modo PECE ($m = 1$), considere-se

$$P: y_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* y_{k+j}^{[1]} = h \sum_{j=0}^{n-1} \beta_j^* f(t_{k+j}, y_{k+j}^{[1]}), \quad (7.24)$$

$$C: y_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j y_{k+j}^{[1]} = h\beta_n f(t_{k+n}, y_{k+n}^{[0]}) + h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y_{k+j}^{[1]}), \quad (7.25)$$

$$P: y(t_{k+n}) + \sum_{j=0}^{n-1} \alpha_j^* y(t_{k+j}) = h \sum_{j=0}^{n-1} \beta_j^* f(t_{k+j}, y(t_{k+j})) + \alpha^* h, \quad (7.26)$$

$$C: y(t_{k+n}) + \sum_{j=0}^{n-1} \alpha_j y(t_{k+j}) = h\beta_n f(t_{k+n}, y(t_{k+n})) + h \sum_{j=0}^{n-1} \beta_j f(t_{k+j}, y(t_{k+j})) + \alpha h. \quad (7.27)$$

Efetuando-se a diferença (7.26) - (7.24), obtém-se

$$\begin{aligned} y(t_{k+n}) - y_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* [y(t_{k+j}) - y_{k+j}^{[1]}] &= h \sum_{j=0}^{n-1} \beta_j^* [f(t_{k+j}, y(t_{k+j})) - f(t_{k+j}, y_{k+j}^{[1]})] + \\ &+ \alpha^* h, \\ e_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* e_{k+j}^{[1]} &= h \sum_{j=0}^{n-1} \beta_j^* [f(t_{k+j}, y(t_{k+j})) - f(t_{k+j}, y_{k+j}^{[1]})] + \\ &+ \alpha^* h. \end{aligned} \quad (7.28)$$

Efetuada-se a diferença (7.27) - (7.25), chega-se a

$$\begin{aligned}
y(t_{k+n}) - y_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j [y(t_{k+j}) - y_{k+j}^{[1]}] &= h\beta_n [f(t_{k+n}, y(t_{k+n})) - f(t_{k+n}, y_{k+n}^{[0]})] + \\
&+ \sum_{j=0}^{n-1} \beta_j [f(t_{k+j}, y(t_{k+j})) - f(t_{k+j}, y_{k+j}^{[1]})] + \alpha h, \\
e_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j e_{k+j}^{[1]} &= h\beta_n [f(t_{k+n}, y(t_{k+n})) - f(t_{k+n}, y_{k+n}^{[0]})] + \\
&+ h \sum_{j=0}^{n-1} \beta_j [f(t_{k+j}, y(t_{k+j})) - f(t_{k+j}, y_{k+j}^{[1]})] + \alpha h.
\end{aligned} \tag{7.29}$$

Utilizando o Teorema do Valor Médio nas equações de diferenças lineares (7.28) e (7.29), obtêm-se

$$e_{k+n}^{[0]} + \sum_{j=0}^{n-1} \alpha_j^* e_{k+j}^{[1]} = h \sum_{j=0}^{n-1} \beta_j^* \frac{\partial f}{\partial y}(t_{k+j}, \eta_{k+j}) e_{k+j}^{[1]} + \alpha^* h \tag{7.30}$$

e

$$\begin{aligned}
e_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j e_{k+j}^{[1]} &= h\beta_n \frac{\partial f}{\partial y}(t_{k+n}, \eta_{k+n}) e_{k+n}^{[0]} + \\
&+ h \sum_{j=0}^{n-1} \beta_j \frac{\partial f}{\partial y}(t_{k+j}, \eta_{k+j}) e_{k+j}^{[1]} + \alpha h.
\end{aligned} \tag{7.31}$$

Sob as condições de que $\frac{\partial f}{\partial y} = \lambda$ e $h\lambda = \bar{h}$, substituindo (7.30) em (7.31) chega-se a

$$\begin{aligned}
e_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j e_{k+j}^{[1]} &= \bar{h}\beta_n \left[- \sum_{j=0}^{n-1} \alpha_j^* e_{k+j}^{[1]} + \bar{h} \sum_{j=0}^{n-1} \beta_j^* e_{k+j}^{[1]} + \alpha^* h \right] + \\
&+ \bar{h} \sum_{j=0}^{n-1} \beta_j e_{k+j}^{[1]} + \alpha h, \\
e_{k+n}^{[1]} + \sum_{j=0}^{n-1} \alpha_j e_{k+j}^{[1]} - \bar{h} \sum_{j=0}^{n-1} \beta_j e_{k+j}^{[1]} &= \bar{h}\beta_n \left[- \sum_{j=0}^{n-1} \alpha_j^* e_{k+j}^{[1]} + \bar{h} \sum_{j=0}^{n-1} \beta_j^* e_{k+j}^{[1]} \right] + \\
&+ \bar{h}\beta_n \alpha^* h + \alpha h, \\
\sum_{j=0}^n \alpha_j e_{k+j}^{[1]} - \bar{h} \sum_{j=0}^{n-1} \beta_j e_{k+j}^{[1]} &= -\bar{h}\beta_n \sum_{j=0}^{n-1} (\alpha_j^* - \bar{h}\beta_j^*) e_{k+j}^{[1]} + \phi,
\end{aligned} \tag{7.32}$$

onde $\bar{h} = h\lambda = h\frac{\partial f}{\partial y}$ e $\phi = \bar{h}\beta_n\alpha^*h + \alpha h$.

Adicionado $-\bar{h}\beta_n e_{k+n}^{[1]}$ a ambos os lados de (7.32) e considerando $\alpha_n^* = 1$ e $\beta_n^* = 0$, obtém-se

$$\begin{aligned} \sum_{j=0}^n (\alpha_j - \bar{h}\beta_j) e_{k+j}^{[1]} &= -\bar{h}\beta_n \sum_{j=0}^n (\alpha_j^* - \bar{h}\beta_j^*) e_{k+j}^{[1]} + \phi, \\ \sum_{j=0}^n (\alpha_j - \bar{h}\beta_j) e_{k+j}^{[1]} + \bar{h}\beta_n \sum_{j=0}^n (\alpha_j^* - \bar{h}\beta_j^*) e_{k+j}^{[1]} &= \phi, \\ \sum_{j=0}^n [(\alpha_j - \bar{h}\beta_j) + \bar{h}\beta_n (\alpha_j^* - \bar{h}\beta_j^*)] e_{k+j}^{[1]} &= \phi. \end{aligned} \quad (7.33)$$

Em (7.33), tem-se uma equação de diferenças linear cuja solução é dada por

$$e_k^{[1]} = \sum_{j=1}^n d_j r_j^k + \frac{\phi}{\sum_{j=0}^n [(\alpha_j - \bar{h}\beta_j) + \bar{h}\beta_n (\alpha_j^* - \bar{h}\beta_j^*)]},$$

onde r_j são as raízes do polinômio

$$\pi(r) = \sum_{j=0}^n [(\alpha_j - \bar{h}\beta_j) + \bar{h}\beta_n (\alpha_j^* - \bar{h}\beta_j^*)] r^j. \quad (7.34)$$

Pode-se reescrever o polinômio (7.34) em função dos polinômios característicos do preditor e do corretor. Assim,

$$\pi(r) = \rho(r) - \bar{h}\sigma(r) + \bar{h}\beta_n [\rho^*(r) - \bar{h}\sigma^*(r)], \quad (7.35)$$

onde $\rho^*(r) = \sum_{j=0}^n \alpha_j^* r^j$, $\rho(r) = \sum_{j=0}^n \alpha_j r^j$, $\sigma^*(r) = \sum_{j=0}^n \beta_j^* r^j$ e $\sigma(r) = \sum_{j=0}^n \beta_j r^j$.

Caso as raízes r_j do polinômio (7.35) satisfaçam a condição

$$|r_j| < 1, \quad j = 1, 2, \dots, n,$$

o Método Preditor-Corretor será absolutamente estável. O intervalo de estabilidade absoluta do Método Preditor-Corretor será (α, β) $\alpha, \beta \in \mathbb{R}$ tal que $\bar{h} = h\lambda \in (\alpha, \beta)$.

7.4.1 Exercícios

Exercício 7.2. Veja em Lambert [10], Capítulo 3, página 97, qual é o polinômio característico para um Método Preditor-Corretor no modo $P(EC)^m E$ e no modo $P(EC)^m$.

7.5 Controle do passo de integração

Para controlar automaticamente o passo de integração em um Método Preditor-Corretor tem-se que usar simultaneamente três critérios de seleção:

- 1 Dispositivo de Milne ($p^* = p$) para estimar o erro local de discretização principal;
- 2 $\bar{h} = h\lambda = h \frac{\partial f}{\partial y}$ pertencente ao intervalo de estabilidade absoluta;
- 3 Condição de convergência do Método do Ponto Fixo

$$h < \frac{1}{L|\beta_n|}.$$

No segundo critério, faz-se necessário estimar $\frac{\partial f}{\partial y}$. Uma das estimativas possíveis (Lambert [10]) é dada por

$$\bar{h} = h\lambda = h \frac{\partial f}{\partial y} \approx h \frac{f(t_{k+n}, y_{k+n}^{[1]}) - f(t_{k+n}, y_{k+n}^{[0]})}{y_{k+n}^{[1]} - y_{k+n}^{[0]}}$$

quando $m = 1$ e

$$\bar{h} = h\lambda = h \frac{\partial f}{\partial y} \approx \frac{y_{k+n}^{[m]} - y_{k+n}^{[m-1]}}{\beta_n (y_{k+n}^{[m-1]} - y_{k+n}^{[m-2]})}$$

quando $m \geq 2$.

7.5.1 Exercícios

Exercício 7.3. Deduza as fórmulas que estimam a parte principal do erro local de truncamento dos métodos que usam como preditor Adams-Bashforth e como corretor Adams-Moulton, ambos de ordem 4.

Exercício 7.4. Defina formalmente os algoritmos de 4 passos que usam o par preditor-corretor de Adams-Bashforth-Moulton de quarta ordem nos modos

1. PEC;
2. PECE;

3. *PMEC*;

4. *PMECE*.

Exercício 7.5. Aplique o Método Preditor-Corretor definido por

$$y_{n+4} - y_{n+3} = \frac{h}{24}(55f_{n+3} - 59f_{n+2} + 37f_{n+1} - 9f_n)$$

$$y_{n+4} - y_{n+3} = \frac{h}{24}(9f_{n+4} + 19f_{n+3} - 5f_{n+2} + f_{n+1})$$

no modo *PECE*, ao p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = -20y(t) & 0 \leq t \leq 1 \\ y(0) = 1 \end{cases}$$

Ilustre o efeito da estabilidade ou instabilidade usando

1. $h = 0,1$;

2. $h = 0,01$.

Exercício 7.6. Seja o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = \frac{2t}{y^2} & 0 \leq t \leq 3 \\ y(0) = 1 \end{cases} \quad (7.36)$$

1. Solucione o p.v.i. (7.36) usando $h = 0,1$ e o Método Preditor-Corretor

$$y_{n+1} - y_n = \frac{h}{12}(23f_n - 16f_{n-1} + 5f_{n-2})$$

$$y_{n+1} - y_n = \frac{h}{24}(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$$

2. Solucione o p.v.i. (7.36) usando $h = 0,1$ e o Método Preditor-Corretor

$$y_{n+1} - y_n = \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$$

$$y_{n+1} - y_n = \frac{h}{24}(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$$

3. Compare os dois métodos usando y_1 , y_2 e y_3 obtidos por RK44.

Exercício 7.7. Considere o preditor P e o corretor C definidos pelos polinômios característicos presentes na Tabela (7.3).

P	$\rho^*(\zeta) = \zeta^2 - 1$	$\sigma^*(\zeta) = (3/2)(\zeta - 1)$
C	$\rho(\zeta) = \zeta^2 - 1$	$\sigma(\zeta) = (1/2)(\zeta^2 + \zeta)$

Tabela 7.3: Polinômios característicos que definem o Método Predictor-Corretor.

1. *Determine a ordem do método.*
2. *Calcule o coeficiente da parte principal do erro local de discretização.*
3. *Mostre que o coeficiente determinado no item anterior é igual ao do corretor.*

7.6 Suplemento teórico

7.6.1 Teorema do Ponto Fixo

Definição 7.2 (Ponto fixo). *Um número p é um ponto fixo de uma função $g(t)$ se*

$$g(p) = p,$$

ou equivalentemente, se

$$g(p) - p = 0.$$

Teorema 7.2. *Seja $g(t) \in C^{(1)}[a, b]$ tal que $g(t) \in [a, b] \forall t \in [a, b]$. Suponha ainda que $g'(t)$ exista em (a, b) com*

$$|g'(t)| \leq k \forall t \in (a, b),$$

onde $k < 1$ é uma constante positiva. Então, para um número qualquer $p_0 \in [a, b]$, a sequência definida por

$$p_{n+1} = g(p_n), \quad n \geq 0$$

converge para o único ponto fixo $p \in [a, b]$.

Observação: O Teorema do Ponto Fixo permite determinar a raiz p de $f(t) = 0$, ou seja, $f(p) = 0$. Para empregá-lo, é necessário reescrever $f(t) = 0$ como $t = g(t)$.

7.7 Exercícios resolvidos

1. Considere o preditor P e dois corretores $C^{(1)}$ e $C^{(2)}$, definidos pelos polinômios característicos presentes na Tabela (7.4).

P	$\rho^*(\zeta) = \zeta^4 - 1$	$\sigma^*(\zeta) = (4/3)(2\zeta^3 - \zeta^2 + 2\zeta)$
$C^{(1)}$	$\rho_1(\zeta) = \zeta^2 - 1$	$\sigma_1(\zeta) = (1/3)(\zeta^2 + 4\zeta + 1)$
$C^{(2)}$	$\rho_2(\zeta) = \zeta^3 - (9/8)\zeta^2 + (1/8)$	$\sigma_2(\zeta) = (3/8)(\zeta^3 + 2\zeta^2 - \zeta)$

Tabela 7.4: Polinômios característicos que definem o Método Preditor-Corretor.

- (a) Deduza uma estimativa do erro local de discretização
- i. para um Método Preditor-Corretor que use P e $C^{(1)}$;
 - ii. para um Método Preditor-Corretor que use P e $C^{(2)}$.
- (b) Escreva o algoritmo que utiliza P e $C^{(1)}$ no modo *PECE*.
- (c) Escreva o algoritmo que utiliza P e $C^{(2)}$ no modo *PMECME*.

Solução

Como $\rho^*(r) = \sum_{j=0}^n \alpha_j^* r^j$, $\rho(r) = \sum_{j=0}^n \alpha_j r^j$, $\sigma^*(r) = \sum_{j=0}^n \beta_j^* r^j$, $\sigma(r) = \sum_{j=0}^n \beta_j r^j$ e $\sum_{j=0}^n \alpha_j y_{k+j} = h \sum_{j=0}^n \beta_j f_{k+j}$, os polinômios característicos da

Tabela (7.4) definem os seguintes métodos de passo múltiplo lineares:

$$P: \quad y_{k+4} - y_k = \frac{4}{3}h [2f_{k+3} - f_{k+2} + 2f_{k+1}]; \quad (7.37)$$

$$C^{(1)}: \quad y_{k+2} - y_k = \frac{1}{3}h [f_{k+2} + 4f_{k+1} + f_k]; \quad (7.38)$$

$$C^{(2)}: \quad y_{k+3} - \frac{9}{8}y_{k+2} + \frac{1}{8}y_k = \frac{3}{8}h [f_{k+3} + 2f_{k+2} - f_{k+1}]. \quad (7.39)$$

Pode-se mostrar que os métodos P , $C^{(1)}$ e $C^{(2)}$ têm ordem de consistência 4 (verifique que $C_0 = C_1 = C_2 = C_3 = C_4 = 0$ e $C_5 \neq 0$), o que possibilita o uso do dispositivo de Milne. Ao calcular o coeficiente C_{p+1} do erro, obtêm-se

$$C_5^* = \frac{14}{45}, \quad (7.40)$$

$$C_5^{(1)} = -\frac{1}{90}, \quad (7.41)$$

$$C_5^{(2)} = -\frac{1}{40}. \quad (7.42)$$

Substituindo (7.40), (7.41) e (7.42) em

$$C_{p+1}h^{p+1}y^{(p+1)}(t_k) = \frac{C_{p+1}}{C_{p+1}^* - C_{p+1}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right),$$

conclui-se que:

1(a)(i)

$$\begin{aligned} C_5^{(1)}h^5y^{(5)}(t_k) &= \frac{-\frac{1}{90}}{\frac{14}{45} + \frac{1}{90}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right) \\ &= -\frac{1}{29} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right); \end{aligned}$$

1(a)(ii)

$$\begin{aligned} C_5^{(2)}h^5y^{(5)}(t_k) &= \frac{-\frac{1}{40}}{\frac{14}{45} + \frac{1}{40}} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right) \\ &= -\frac{9}{121} \left(y_{k+n}^{[m]} - y_{k+n}^{[0]} \right). \end{aligned}$$

Como o preditor é um método de 4 passos, os corretores $C^{(1)}$ e $C^{(2)}$ podem ser reescritos na forma

$$C^{(1)}: y_{k+4} - y_{k+2} = \frac{1}{3}h [f_{k+4} + 4f_{k+3} + f_{k+2}], \quad (7.43)$$

$$C^{(2)}: y_{k+4} - \frac{9}{8}y_{k+3} + \frac{1}{8}y_{k+1} = \frac{3}{8}h [f_{k+4} + 2f_{k+3} - f_{k+2}]. \quad (7.44)$$

Assim:

1(b)

$$P: y_{k+4}^{[0]} - y_k^{[1]} = \frac{4}{3}h [2f_{k+3}^{[1]} - f_{k+2}^{[1]} + 2f_{k+1}^{[1]}]$$

$$E: f_{k+4}^{[0]} = f(t_{k+4}, y_{k+4}^{[0]})$$

$$C: y_{k+4}^{[1]} - y_{k+2}^{[1]} = \frac{1}{3}h [f_{k+4}^{[0]} + 4f_{k+3}^{[1]} + f_{k+2}^{[1]}]$$

$$E: f_{k+4}^{[1]} = f(t_{k+4}, y_{k+4}^{[1]});$$

1(c)

$$P : \quad y_{k+4}^{[0]} - \hat{y}_k^{[1]} = \frac{4}{3}h \left[2\hat{f}_{k+3}^{[1]} - \hat{f}_{k+2}^{[1]} + 2\hat{f}_{k+1}^{[1]} \right]$$

$$M : \quad \hat{y}_{k+4}^{[0]} = y_{k+4}^{[0]} + \frac{112}{121} \left(y_{k+3}^{[1]} - y_{k+3}^{[0]} \right)$$

$$E : \quad \hat{f}_{k+4}^{[0]} = f \left(t_{k+4}, \hat{y}_{k+4}^{[0]} \right)$$

$$C : \quad y_{k+4}^{[1]} - \frac{9}{8}\hat{y}_{k+3}^{[1]} + \frac{1}{8}\hat{y}_{k+1}^{[1]} = \frac{3}{8}h \left[\hat{f}_{k+4}^{[0]} + 2\hat{f}_{k+3}^{[1]} - \hat{f}_{k+2}^{[1]} \right]$$

$$M : \quad \hat{y}_{k+4}^{[1]} = y_{k+4}^{[1]} - \frac{9}{121} \left(y_{k+4}^{[1]} - y_{k+4}^{[0]} \right)$$

$$E : \quad \hat{f}_{k+4}^{[1]} = f \left(t_{k+4}, \hat{y}_{k+4}^{[1]} \right).$$

Observação

$$\hat{y}_{k+n}^{[0]} = y_{k+n}^{[0]} + \frac{C_{p+1}^*}{C_{p+1}^* - C_{p+1}} \left(y_{k+n-1}^{[m]} - y_{k+n-1}^{[0]} \right)$$

$$\hat{y}_{k+4}^{[0]} = y_{k+4}^{[0]} + \frac{\frac{14}{45}}{\frac{14}{45} + \frac{1}{40}} \left(y_{k+3}^{[1]} - y_{k+3}^{[0]} \right)$$

$$\hat{y}_{k+4}^{[0]} = y_{k+4}^{[0]} + \frac{112}{121} \left(y_{k+3}^{[1]} - y_{k+3}^{[0]} \right)$$

Capítulo 8

Equações diferenciais rígidas

8.1 Definição

Uma equação diferencial é dita *rígida* quando a sua solução exata tem um termo da forma

$$e^{-ct}, \quad (8.1)$$

onde c é uma constante positiva elevada ($c \gg 0$).¹ O termo (8.1) é a *parte transiente* da solução, o qual decai rapidamente para zero quando t aumenta. A rigidez é então caracterizada pela existência de duas ou mais escalas diferentes para a variável independente t nas quais as variáveis dependentes variam.

A n -ésima derivada de (8.1) é dada pelo termo

$$\pm c^n e^{-ct}, \quad (8.2)$$

o qual não decai rapidamente para zero como (8.1). Ainda, como a derivada está presente no termo do erro e é calculada em um número entre zero e t , o termo (8.2) pode crescer rapidamente à medida que t também cresce. Dessa forma, se a magnitude da derivada (8.2) aumenta porém a solução não, o erro pode se tornar muito grande e dominar a solução.

Problemas de valor inicial com equações rígidas são relativamente comuns, particularmente no estudo de vibrações, reações químicas e circuitos elétricos. O nome sistema rígido deriva do movimento do sistema massa-mola com constante de elasticidade elevada.

¹Ou (8.1) tem a forma $e^{\lambda t}$, sendo $\lambda = a + bi$ uma constante complexa com parte real negativa ($a \ll 0$).

Exemplo 8.1. *O p.v.i.*

$$\begin{cases} \frac{d}{dt}y(t) = -30y(t), & t \in \left[0, \frac{3}{2}\right] \\ y(0) = \frac{1}{3} \end{cases} \quad (8.3)$$

tem solução exata dada por $y(t) = \frac{1}{3}e^{-30t}$ (verifique!). A Tabela (8.1) mostra, para $t = \frac{3}{2}$, o erro global de discretização cometido na solução do p.v.i. (8.3) com $h = 0,1$ e os Métodos de Euler, Runge-Kutta Clássico e o Preditor-Corretor Adams-Bashforth-Moulton de quarta ordem.

Solução exata	9,54173E-21
Euler	-1,09225E+04
Runge-Kutta Clássico	3,95730E+01
Preditor-Corretor	8,03840E+05

Tabela 8.1: Erro global de discretização do problema de valor inicial (8.3) em $t = \frac{3}{2}$, com $h = 0,1$ e os Métodos de Euler, Runge-Kutta Clássico e Adams-Bashforth-Moulton de quarta ordem - Burden [7].

Exemplo 8.2. *O sistema de equações diferenciais*

$$\begin{cases} \frac{d}{dt}y_1(t) = 9y_1(t) + 24y_2(t) + 5\cos(t) - \frac{1}{3}\sin(t) \\ \frac{d}{dt}y_2(t) = -24y_1(t) - 51y_2(t) - 9\cos(t) + \frac{1}{3}\sin(t) \end{cases} \quad (8.4)$$

com condições iniciais

$$\begin{cases} y_1(0) = \frac{4}{3} \\ y_2(0) = \frac{2}{3} \end{cases} \quad (8.5)$$

tem solução exata dada por

$$\begin{aligned} y_1(t) &= 2e^{-3t} - e^{-39t} + \frac{1}{3}\cos(t) \\ y_2(t) &= -e^{-3t} + 2e^{-39t} - \frac{1}{3}\cos(t) \end{aligned} \quad (8.6)$$

O termo transiente e^{-39t} na solução (8.6) torna o sistema (8.4)-(8.5) rígido. As Tabelas (8.2) e (8.3) ilustram os resultados obtidos na solução do sistema (8.4)-(8.5) com o Método de Runge-Kutta Clássico e os passos de integração $h = 0,05$ e $h = 0,1$.

k	t_k	$y_1(t_k)$	$y_{1,k}$	$y_2(t_k)$	$y_{2,k}$
1	0,1	1,7931E+00	1,7122E+00	-1,0320E+00	-8,7032E-01
2	0,2	1,4239E+00	1,4141E+00	-8,7468E-01	-8,5501E-01
3	0,3	1,1316E+00	1,1305E+00	-7,2500E-01	-7,2289E-01
4	0,4	9,0941E-01	9,0928E-01	-6,0821E-01	-6,0795E-01
5	0,5	7,3879E-01	7,3875E-01	-5,1566E-01	-5,1558E-01
6	0,6	6,0571E-01	6,0568E-01	-4,4041E-01	-4,4036E-01
7	0,7	4,9986E-01	4,9984E-01	-3,7740E-01	-3,7735E-01
8	0,8	4,1367E-01	4,1365E-01	-3,2295E-01	-3,2291E-01
9	0,9	3,4161E-01	3,4159E-01	-2,7441E-01	-2,7437E-01
10	1,0	2,7967E-01	2,7966E-01	-2,2989E-01	-2,2985E-01

Tabela 8.2: Solução numérica do sistema rígido (8.4)-(8.5) pelo Método de Runge-Kutta Clássico com $h = 0,05$ - Burden [7].

k	t_k	$y_1(t_k)$	$y_{1,k}$	$y_2(t_k)$	$y_{2,k}$
1	0,1	1,7931E+00	-2,6452E+00	-1,0320E+00	7,8445E+00
2	0,2	1,4239E+00	-1,8452E+01	-8,7468E-01	3,8876E+01
3	0,3	1,1316E+00	-8,7472E+01	-7,2500E-01	1,7648E+02
4	0,4	9,0941E-01	-9,3407E+02	-6,0821E-01	7,8935E+02
5	0,5	7,3879E-01	-1,7600E+03	-5,1566E-01	3,5210E+03
6	0,6	6,0571E-01	-7,8486E+03	-4,4041E-01	1,5698E+04
7	0,7	4,9986E-01	-3,4990E+04	-3,7740E-01	6,9980E+04
8	0,8	4,1367E-01	-1,5598E+05	-3,2295E-01	3,1196E+05
9	0,9	3,4161E-01	-6,9533E+05	-2,7441E-01	1,3907E+06
10	1,0	2,7967E-01	-3,0997E+06	-2,2989E-01	6,1994E+06

Tabela 8.3: Solução numérica do sistema rígido (8.4)-(8.5) pelo Método de Runge-Kutta Clássico com $h = 0,1$ - Burden [7].

8.2 Métodos numéricos para equações rígidas

Há três importantes classes de métodos de alta ordem para equações diferenciais rígidas [19, 20]:

- Generalizações dos Métodos de Runge-Kutta: dentre estes, os mais usados são os Métodos de Rosenbrock (Numerical Recipes [19, 20]);
- Generalizações do Método de Bulirsch-Stoer (Stoer [14]);
- Métodos Preditores-Corretores derivados do Método de Gear (Lambert [10]).

8.2.1 Exercícios

Exercício 8.1. Para o p.v.i. (8.3), verifique que λh , onde $\lambda = -30$ e $h = 0,1$, não pertence ao intervalo de estabilidade absoluta dos Métodos de Euler, Runge-Kutta Clássico e o Preditor-Corretor Adams-Bashforth-Moulton de quarta ordem.

Exercício 8.2. Solucione os problemas de valor inicial abaixo usando $h = 0,1$ e os Métodos de Euler, Runge-Kutta Clássico e Trapézio, este com $\epsilon = 10^{-5}$. Compare os resultados numéricos com a solução exata.

(a)

$$\begin{cases} \frac{d}{dt}y(t) = -20(y(t) - t^2) + 2t, & t \in [0, 1] \\ y(0) = \frac{1}{3} \end{cases} \quad (8.7)$$

Solução exata: $y(t) = t^2 + \frac{1}{3}e^{-20t}$

(b)

$$\begin{cases} \frac{d}{dt}y(t) = -20y(t) + 20\text{sen}(t) + \cos(t), & t \in [0, 2] \\ y(0) = 1 \end{cases} \quad (8.8)$$

Solução exata: $y(t) = \text{sen}(t) + e^{-20t}$

(c)

$$\begin{cases} \frac{d}{dt}y(t) = \frac{50}{y(t)} - 50y(t), & t \in [0, 1] \\ y(0) = \sqrt{2} \end{cases} \quad (8.9)$$

Solução exata: $y(t) = (1 + e^{-100t})^{\frac{1}{2}}$

8.3 Exercícios resolvidos

1. O p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = 5e^{5t}(y(t) - t)^2 + 1, & t \in [0, 1] \\ y(0) = -1 \end{cases} \quad (8.10)$$

tem solução exata dada por $y(t) = t - e^{-5t}$.

Verifique os efeitos da rigidez solucionando o p.v.i. (8.10) com $h = 0,2$ e $h = 0,25$, empregando os Métodos de Runge-Kutta Clássico e do Trapézio, este último com $\epsilon = 10^{-6}$.

Solução

As Tabelas (8.4) e (8.5) mostram o erro global de discretização cometido em $t = 1$ ao se empregar os Métodos de Runge-Kutta Clássico e do Trapézio.

		Runge-Kutta		Trapézio	
k	t_k	y_k	$ y(t_k) - y_k $	y_k	$ y(t_k) - y_k $
0	0,0	-1,0000E+00	0,0000E+00	-1,0000E+00	0,0000E+00
1	0,2	-1,4885E-01	1,9027E-02	-1,4150E-01	2,6383E-02
2	0,4	2,6849E-01	3,8237E-03	2,7486E-01	1,0197E-02
3	0,6	5,5199E-01	1,7798E-03	5,5398E-01	3,7700E-03
4	0,8	7,8229E-01	6,0131E-04	7,8307E-01	1,3876E-03
5	1,0	9,9349E-01	2,2845E-04	9,9377E-01	5,1050E-04

Tabela 8.4: Erro global de discretização cometido em $t = 1$ na solução do problema de valor inicial (8.10), onde a solução numérica foi obtida através dos Métodos de Runge-Kutta Clássico e do Trapézio com $h = 0,2$ - Burden [7].

		Runge-Kutta		Trapézio	
k	t_k	y_k	$ y(t_k) - y_k $	y_k	$ y(t_k) - y_k $
0	0,0	-1,0000E+00	0,0000E+00	-1,0000E+00	0,0000E+00
1	0,25	4,0143E-01	4,3794E-01	5,4557E-03	4,1961E-02
2	0,5	3,4375E+00	3,0196E+00	4,2676E-01	8,8422E-03
3	0,75	1,4464E+23	1,4464E+23	7,2915E-01	2,6706E-03
4	1,0	overflow		9,9402E-01	7,5790E-04

Tabela 8.5: Erro global de discretização cometido em $t = 1$ na solução do problema de valor inicial (8.10), onde a solução numérica foi obtida através dos Métodos de Runge-Kutta Clássico e do Trapézio com $h = 0,25$ - Burden [7].

Apêndice A

Exercícios complementares

1. Seja o método de passo único descrito abaixo.

Método de Runge-Kutta de 2ª ordem com dois estágios

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h) \end{cases}, \quad (\text{A.1})$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$ e

$$\Phi(t_k, y_k, h) = \frac{1}{4}(\kappa_1 + 3\kappa_2),$$

sendo

$$\begin{cases} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f\left(t_k + \frac{2}{3}h, y_k + \frac{2}{3}h\kappa_1\right) \end{cases}.$$

1.1 O método de passo simples (A.1) é consistente e convergente? Justifique.

1.2 Comprove a ordem de consistência do método de passo simples (A.1) verificando que ele concorda com o Método da Série de Taylor (A.2) até o termo de segunda ordem.

Método da Série de Taylor

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h\Phi(t_k, y_k, h) \end{cases}, \quad (\text{A.2})$$

onde

$$\Phi(t, y, h) = f(t, y) + \frac{h}{2!} Df(t, y) + \frac{h^2}{3!} D^2 f(t, y) + \cdots + \frac{h^{q-1}}{q!} D^{q-1} f(t, y),$$

$$\text{com } D = \frac{\partial}{\partial t} + f \frac{\partial}{\partial y}.$$

- 1.3 Determine a região de estabilidade absoluta do método de passo simples (A.1). Faça a análise para $\lambda \in \mathbb{R}$.
- 1.4 Use o método de passo simples (A.1) para solucionar o Problema de Cauchy

$$\begin{cases} \frac{d}{dt} y(t) = te^{3t} - 2y(t), & t \in [0, 1] \\ y(0) = 0 \end{cases}, \quad (\text{A.3})$$

com $h = 0,5$. Calcule o erro global de discretização no instante $t = 1$ e comente o resultado obtido.

- 1.5 Ao solucionar o Problema de Cauchy numericamente, é mais vantajoso computacionalmente usar o método de passo simples (A.1) ou o Método de Euler Aprimorado (A.4)? Justifique.
-

Método de Euler Aprimorado

$$\begin{cases} y_0 = y(t_0) \\ y_{k+1} = y_k + h \Phi(t_k, y_k, h) \end{cases} \quad (\text{A.4})$$

com $t_{k+1} = t_k + h$, $0 \leq k \leq n-1$, $h = \frac{b-a}{n}$, $\Phi(t_k, y_k, h) = \frac{1}{2}(\kappa_1 + \kappa_2)$ e

$$\begin{cases} \kappa_1 = f(t_k, y_k) \\ \kappa_2 = f(t_k + h, y_k + h \kappa_1) \end{cases}.$$

2 Solucione a equação de diferenças linear

$$y_{k+2} + 4y_k = 15$$

com condições iniciais $y_0 = 1$ e $y_1 = 2$.

3 Calcule φ , δ e γ para que o método de passo múltiplo linear

$$y_{k+4} - y_k + \varphi(y_{k+3} - y_{k+1}) = h[\delta(f_{k+3} - f_{k+1}) + \gamma f_{k+2}] \quad (\text{A.5})$$

seja consistente de ordem 3.

4 O método de passo múltiplo linear

$$y_{k+3} = \frac{1}{8}(9y_{k+2} - y_k) + \frac{3h}{8}(f_{k+3} + 2f_{k+2} - f_{k+1}) \quad (\text{A.6})$$

é convergente? Justifique.

5 Considere os Métodos Preditor P e Corretor C definidos por

$$\begin{aligned} y_{k+2} + 4y_{k+1} - 5y_k &= h(4f_{k+1} + 2f_k), \\ y_{k+2} - y_k &= \frac{h}{3}(f_{k+2} + 4f_{k+1} + f_k). \end{aligned}$$

5.1 Determine o erro local de discretização principal do Método Preditor-Corretor.

5.2 Escreva o polinômio

$$\pi(r) = \rho(r) - \bar{h}\sigma(r) + \bar{h}\beta_n [\rho^*(r) - \bar{h}\sigma^*(r)] \quad (\text{A.7})$$

associado à equação de diferenças linear para o erro global de discretização do Método Preditor-Corretor no modo *PECE*. Qual é a aplicabilidade do polinômio (A.7)?

5.3 Escreva o algoritmo que utiliza o Método Preditor-Corretor no modo *PECE*.

Apêndice B

Exercícios computacionais

1 Considere o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = -20y(t), & t \in [0, 1] \\ y(0) = 1 \end{cases} . \quad (\text{B.1})$$

1.1 Solucione numericamente o p.v.i. (B.1) usando os Métodos de Euler e de Runge-Kutta de 4^a ordem com 4 estágios e os seguintes passos de integração:

- A $h = 0,05$;
- B $h = 0,1$;
- C $h = 0,2$.

Implemente os métodos numéricos em linguagem C ou Fortran.

1.2 Analise e justifique os resultados obtidos, comparando os dois métodos numéricos implementados em relação a:

- A convergência;
- B estabilidade;
- C consistência.

2 Seja o Método de Runge-Kutta-Fehlberg, definido por

$$\alpha_k^{RKF} \approx \frac{1}{360}\kappa_1 - \frac{128}{4275}\kappa_3 - \frac{2197}{75240}\kappa_4 + \frac{1}{50}\kappa_5 + \frac{2}{55}\kappa_6.$$

2.1 Valide o Método de Runge-Kutta-Fehlberg, em linguagem C ou Fortran, empregando o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = \lambda y(t), & t \in [a, b] \\ y(a) = y(t_0) = y_0 \end{cases} , \quad (\text{B.2})$$

onde λ é uma constante.

2.2 Considere o p.v.i.

$$\begin{cases} \frac{d}{dt}y(t) = -2x(t) - y(t), & t \in [0, 10] \\ y(0) = -1 \end{cases} \quad (\text{B.3})$$

- A** Solucione o p.v.i. (B.3) utilizando o Método de Runge-Kutta-Fehlberg, em linguagem C ou Fortran, com $h = 0,2$ e
- (a) $\epsilon = 10^{-6}$;
- (b) $\epsilon = 10^{-3}$.
- B** Compare o erro local de discretização e o passo de integração. Organize os dados em uma tabela.
- C** Calcule o erro global de discretização no instante $t = 10$ e comente os resultados obtidos.

3 Modelagem da propagação de doenças contagiosas

3.1 Problema

Na teoria de propagação de doenças contagiosas, uma equação diferencial ordinária não linear de primeira ordem homogênea pode ser usada para prever o número de indivíduos infectados em um tempo qualquer (em dias), desde que simplificações adequadas sejam adotadas. Em particular, considera-se que todos os indivíduos de uma população fixa podem ser contaminados e que, uma vez infectados, permanecem nessa condição.

Sejam $x(t)$ o número de indivíduos suscetíveis à infecção e $y(t)$ o número de indivíduos infectados. É razoável supor que a taxa de variação temporal do número de infectados seja proporcional ao produto do número de indivíduos suscetíveis pelo número de indivíduos infectados. Assim, tem-se que

$$\frac{d}{dt}y(t) = k x(t)y(t), \quad (\text{B.4})$$

onde k é uma constante e

$$x(t) + y(t) = m, \quad (\text{B.5})$$

sendo m o tamanho da população. Como $x(t) = m - y(t)$, a equação (B.4) pode ser reescrita como

$$\frac{d}{dt}y(t) = k[m - y(t)]y(t). \quad (\text{B.6})$$

A equação (B.6) é chamada *Equação de Bernoulli*. Esta equação pode, pela substituição

$$u(t) = \frac{1}{y(t)}, \quad (\text{B.7})$$

ser reescrita como a equação diferencial ordinária linear de primeira ordem não homogênea (verifique!)

$$\frac{d}{dt}u(t) = k[1 - m u(t)]. \quad (\text{B.8})$$

3.2 Questões

- A** Calcule a solução exata (família de soluções) da equação (B.6) e determine $\lim_{t \rightarrow \infty} y(t)$. Esse limite é aceitável? Justifique.
- B** Implemente computacionalmente, em linguagem C ou Fortran, os seguintes métodos de aproximação da solução do Problema de Cauchy

$$\begin{cases} \dot{y}(t) = \frac{d}{dt}y(t) = f(t, y(t)), & t \in [a, b] \\ y(t_0) = y(a) = y_0 \end{cases} :$$

- Euler;
- Euler Aprimorado;
- Trapézio;
- Runge-Kutta-Fehlberg;
- Adams-Bashforth de 4 passos;
- Adams-Moulton de 4 passos.

C Considere na equação (B.6):

$$m = 10^5;$$

$$k = 2 \times 10^{-6};$$

$$y(0) = 10^3;$$

$$t \in [0, 30].$$

- (a) Use os métodos implementados no item B para aproximar a solução da equação (B.6) no instante de tempo $t = 30$ dias. Otimize o passo temporal em cada método e calcule o erro global de discretização.
- (b) Compare os métodos empregados. Use tabelas e gráficos (plote o gráfico da solução exata e da solução numérica empregando aplicativos como o winplot, o octave, o maple, o matlab, o mathematica, etc.).

Apêndice C

Algoritmos

1. Algoritmo para o Método de Euler

```
Entre com o intervalo  $[a, b]$   
Entre com o número de divisões  $n$   
Entre com o valor inicial  $y_0$  em  $t = a$   
 $h \leftarrow (b - a)/n$   
 $y \leftarrow y_0$   
Para  $i = 0$  a  $n - 1$  faça  
     $t \leftarrow a + ih$   
     $y \leftarrow y + hf(t, y)$   
Fim para
```

algoritmo C.0.1: Algoritmo para o Método de Euler.

2. Algoritmo para o Método de Euler Aprimorado

```
 $y \leftarrow y_0;$   
 $t \leftarrow t_0;$   
 $h \leftarrow \frac{t_{final} - t_0}{n};$   
Para  $k = 1$  até  $n$  faça  
     $\kappa_1 \leftarrow f(t, y);$   
     $\kappa_2 \leftarrow f(t + h, y + h\kappa_1);$   
     $y \leftarrow y + \frac{h}{2}(\kappa_1 + \kappa_2);$   
    Imprima  $t, y;$   
     $t \leftarrow t_0 + kh;$ 
```

algoritmo C.0.2: Algoritmo para o Método de Euler Aprimorado.

Apêndice D

O Teorema de Picard

D.1 Teoremas auxiliares

Teorema D.1. *Seja $X \subset \mathbb{R}$. Se uma sequência de funções $f_n : X \rightarrow \mathbb{R}$ converge uniformemente para $f : X \rightarrow \mathbb{R}$ e cada f_n é contínua no ponto $a \in X$, então f é contínua no ponto a .*

Teorema D.2 (Teste de Weierstrass). *Dada a sequência de funções $f_n : X \rightarrow \mathbb{R}$, seja $\sum a_n$ uma série convergente de números reais $a_n \geq 0$ tais que $|f_n(t)| \leq a_n$ para todo $n \in \mathbb{N}$ e todo $t \in X$. Nestas condições, as séries $\sum |f_n|$ e $\sum f_n$ são uniformemente convergentes.*

D.2 O Teorema de Picard

Sejam $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}$ e

$$I_a = [t_0 - a, t_0 + a]$$
$$B_b = [y_0 - b, y_0 + b]$$

onde $a, b \in \mathbb{R}_+^*$, de maneira que $\Omega = I_a \times B_b \subset \mathbb{R} \times \mathbb{R}$ é um conjunto compacto contendo (t_0, y_0) . Seja $f : \Omega \subset \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ não identicamente nula e contínua em Ω e, portanto, limitada em Ω . Suponha-se ainda que f é lipschitziana com respeito à segunda variável em Ω .

Nesses termos, denota-se

$$M = \sup\{|f(t, y)|; (t, y) \in \Omega\} > 0$$

e L a constante de Lipschitz de f em Ω , de forma que para todo $t \in I_a$

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \quad \forall y_1, y_2 \in B_b.$$

Definam-se por fim

$$\alpha = \min \left\{ a, \frac{b}{M} \right\} > 0$$

e

$$I_\alpha = [t_0 - \alpha, t_0 + \alpha] \subset I_a.$$

Teorema D.3 (Picard). *Existe uma única função $\phi : I_\alpha \rightarrow B_b$ que é solução do Problema de Cauchy*

$$y' = f(t, y), \quad y(t_0) = y_0. \quad (\text{D.1})$$

Demonstração. Seja $\psi : I_\alpha \rightarrow B_b$ uma função contínua qualquer (elas existem, já que a função constante igual a y_0 o é). Define-se então

$$\psi^*(t) = y_0 + \int_{t_0}^t f(s, \psi(s)) ds \quad \forall t \in I_\alpha.$$

Deve-se mostrar que ψ^* é uma função contínua definida de I_α em B_b .

De fato, ψ^* está bem definida em I_α já que ψ é contínua em I_α , o que implica que $f(t, \psi(t))$ está bem definida e é contínua em I_α e, portanto, integrável. Além disso, segue da definição de ψ^* e do Teorema Fundamental do Cálculo que ψ^* é derivável em I_α e, portanto, contínua.

Resta mostrar que $\psi^*(t) \in B_b$ qualquer que seja $t \in I_\alpha$. De fato, dado $s \in I_\alpha$ tem-se que $\psi(s) \in B_b$, donde $(s, \psi(s)) \in \Omega$ e então $|f(s, \psi(s))| \leq M$. Assim, dado $t \in I_\alpha$,

$$\begin{aligned} |\psi^*(t) - y_0| &= \left| y_0 + \int_{t_0}^t f(s, \psi(s)) ds - y_0 \right| = \left| \int_{t_0}^t f(s, \psi(s)) ds \right| \\ &\leq \left| \int_{t_0}^t |f(s, \psi(s))| ds \right| \\ &\leq \left| \int_{t_0}^t M ds \right| = M|t - t_0| \\ &\leq M\alpha = b, \end{aligned}$$

o que mostra que $\psi^*(t) \in B_b$.

Define-se agora a *sequência de aproximações sucessivas*

$$\begin{aligned} \phi_0(t) &= y_0, \quad \forall t \in I_\alpha \\ \phi_{n+1}(t) &= y_0 + \int_{t_0}^t f(s, \phi_n(s)) ds, \quad \forall t \in I_\alpha, n \geq 1. \end{aligned}$$

Observe-se que isto define uma seqüência $(\phi_n)_{n \in \mathbb{N}}$ de funções contínuas definidas de I_α em B_b , em vista da discussão acima e levando-se em consideração que ϕ_0 é uma função contínua definida de I_α em B_b .

Mostra-se a seguir que a seqüência de funções $(\phi_n)_{n \in \mathbb{N}}$ é uniformemente convergente. Primeiramente, observe que para todo $n \in \mathbb{N}$ pode-se escrever

$$\begin{aligned}\phi_n(t) &= \phi_n(t) - \phi_{n-1}(t) + \phi_{n-1}(t) - \phi_{n-2}(t) + \cdots + \phi_1(t) - \phi_0(t) + \phi_0(t) \\ &= y_0 + \sum_{k=0}^n (\phi_{k+1}(t) - \phi_k(t)),\end{aligned}$$

de maneira que se pode mostrar que a série de funções $\sum_k (\phi_{k+1} - \phi_k)$ é uniformemente convergente.

Observe-se agora que

$$|\phi_{k+1}(t) - \phi_k(t)| \leq b \frac{L^k |t - t_0|^k}{k!} \quad (\text{D.2})$$

para todo $k \in \mathbb{N}$, o que é possível mostrar por indução sobre k . Para $k = 0$, tem-se que

$$\begin{aligned}|\phi_1(t) - \phi_0(t)| &= \left| \int_{t_0}^t f(s, y_0) ds \right| \\ &\leq \left| \int_{t_0}^t |f(s, y_0)| ds \right| \\ &\leq M\alpha = b.\end{aligned}$$

Supondo por indução que (D.2) é verificada para algum $m \geq 0$, tem-se que

$$\begin{aligned}|\phi_{m+1}(t) - \phi_m(t)| &= \left| \int_{t_0}^t [f(s, \phi_m(s)) - f(s, \phi_{m-1}(s))] ds \right| \\ &\leq \left| \int_{t_0}^t |f(s, \phi_m(s)) - f(s, \phi_{m-1}(s))| ds \right| \\ &\leq \left| \int_{t_0}^t L |\phi_m(s) - \phi_{m-1}(s)| ds \right| = L \left| \int_{t_0}^t |\phi_m(s) - \phi_{m-1}(s)| ds \right| \\ &\leq L \left| \int_{t_0}^t b \frac{L^m |s - t_0|^m}{m!} ds \right| \\ &\leq b \frac{L^{m+1}}{m!} \left| \int_{t_0}^t |s - t_0|^m ds \right| = b \frac{L^{m+1} |t - t_0|^{m+1}}{(m+1)!},\end{aligned}$$

isto é, (D.2) também é verificada para $m + 1$. Isso conclui a prova por indução de que (D.2) se verifica para todo $k \in \mathbb{N}$. Mais que isso,

$$|\phi_{k+1}(t) - \phi_k(t)| \leq b \frac{L^k \alpha^k}{k!} \quad \forall k \in \mathbb{N}, t \in I_\alpha.$$

Observe-se agora que $b \frac{L^k \alpha^k}{k!} \geq 0$ para todo $k \in \mathbb{N}$ e que este é o termo geral da série numérica que converge para $be^{L\alpha}$. Pelo Teste de Weierstrass (Teorema D.2), tem-se que a série $\sum_k (\phi_{k+1} - \phi_k)$ converge uniformemente em I_α . Logo, a sequência $(\phi_n)_{n \in \mathbb{N}}$ também converge uniformemente em I_α .

Denote-se

$$\phi = \lim_{n \rightarrow \infty} \phi_n,$$

de maneira que a convergência uniforme implica pelo Teorema D.1 que ϕ é uma função contínua. Fazendo-se $n \rightarrow \infty$ nos dois lados da equação

$$\phi_{n+1}(t) = y_0 + \int_{t_0}^t f(s, \phi_n(s)) ds,$$

tem-se que

$$\phi(t) = y_0 + \int_{t_0}^t f(s, \phi(s)) ds,$$

ou seja, ϕ é uma solução de (D.1).

Para demonstrar a unicidade da solução de (D.1), suponha-se que $\chi : I_\alpha \rightarrow B_b$ seja uma dessas soluções. Ou seja,

$$\chi(t) = y_0 + \int_{t_0}^t f(s, \chi(s)) ds.$$

Então, de maneira semelhante ao que foi feito anteriormente, pode-se mostrar por indução que

$$|\chi(t) - \phi_n(t)| \leq b \frac{L^n \alpha^n}{n!} \quad \forall t \in I_\alpha$$

para todo $n \in \mathbb{N}$. Como $b \frac{L^n \alpha^n}{n!} \rightarrow 0$ quando $n \rightarrow \infty$, então $\phi_n \rightarrow \chi$, onde a convergência é evidentemente uniforme. Segue da unicidade do limite de $(\phi_n)_{n \in \mathbb{N}}$ que $\phi = \chi$. \square

Referências Bibliográficas

- [1] D. G. Figueiredo A. F. Neves. *Equações diferenciais aplicadas*. IMPA, 1997.
- [2] C. I. Doering A. O. Lopes. *Equações diferenciais ordinárias*. IMPA, 2010.
- [3] Carl B. Boyer. *História da Matemática*. Edgard Blücher, 1987.
- [4] M. Braun. *Equações diferenciais e suas aplicações*. Campus, 1979.
- [5] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2003.
- [6] B. W. Kernighan D. M. Ritchie. *C - A linguagem de programação*. Campus, 1986.
- [7] R. L. Burden J. D. Faires. *Numerical analysis*. Brooks/Cole Publishing Company, 1997.
- [8] A. M. Roma J. S. Bevilacqua. *Resolução numérica de equações diferenciais ordinárias com condições iniciais*. SBMAC, 1997.
- [9] W. Kaplan. *Cálculo avançado. Volumes 1 e 2*. Edgard Blücher, 2006.
- [10] J. D. Lambert. *Computational methods in ordinary differential equations*. John Wiley & Sons, 1973.
- [11] E. L. Lima. *Álgebra linear*. IMPA, 2008.
- [12] E. L. Lima. *Álgebra linear - Exercícios e soluções*. IMPA, 2009.
- [13] C. F. Gerald P. O. Wheatley. *Applied numerical analysis*. Addison-Wesley Publishing Company, 1994.
- [14] J. Stoer R. Bulirsch. *Introduction to numerical analysis*. Springer, 1992.
- [15] W. E. Boyce R. C. Diprima. *Equações diferenciais elementares e problemas de valores de contorno*. LTC, 2010.
- [16] H. Schildt. *C completo e total*. Makron Books, 1997.
- [17] H. R. Schwarz. *Numerical Analysis - A comprehensive introduction*. John Wiley & Sons, 1989.

- [18] J. Sotomayor. *Lições de equações diferenciais ordinárias*. IMPA, 1979.
- [19] W. T. Vetterling B. P. Flannery W. H. Press, S. A. Teukolsky. *Numerical recipes in C - The art of scientific computing*. Cambridge University Press, 1992.
- [20] W. T. Vetterling B. P. Flannery W. H. Press, S. A. Teukolsky. *Numerical recipes in Fortran - The art of scientific computing*. Cambridge University Press, 1992.

Índice Remissivo

- Algoritmo para o Método de Euler Aprimorado, 178
- Cauchy, 21
- Condição de Lipschitz, 6, 7, 16
- Conjunto convexo, 15
- Diabetes, 1
- Discretização
 - do Problema de Cauchy, 8
 - método explícito, 10, 12, 23, 29, 85
 - método implícito, 12, 29, 85
 - métodos de p passos, 12
 - métodos de passo único, 10, 24, 29, 75
 - métodos de passo múltiplo, 12
 - métodos de passo múltiplo lineares, 85
 - passo de integração, 8
- Equação de diferenças linear, 85
 - homogênea, 104
 - não homogênea, 104
- Equações diferenciais rígidas, 163
 - métodos numéricos, 165
- Euler, 44
- Kutta, 60
- Métodos
 - condicionalmente estáveis, 77
 - da Série de Taylor, 51
 - de Adams-Bashforth, 12, 14
 - de Adams-Moulton, 12
 - de Euler, 8, 23, 24, 29, 32, 75, 76
 - de Euler Aprimorado, 10, 14, 31, 33, 55
 - de Euler Implícito, 10, 33
 - de Euler Modificado, 53, 54
 - de Runge-Kutta, 52
 - clássico, 55
 - de ordens superiores, 55
 - de quarta ordem com cinco estágios, 56
 - de quarta ordem com quatro estágios, 55
 - de quinta ordem com seis estágios, 56
 - de R estágios, 54
 - de terceira ordem com 3 estágios, 55
 - Fehlberg, 57
 - de Simpson, 12, 13, 25, 87
 - do Ponto Médio, 53
 - do Trapézio, 10, 14, 25, 29, 33, 46
 - incondicionalmente estáveis, 82
- Métodos de passo único, 10, 24, 29, 75
 - consistência, 31
 - ordem, 32
 - convergência, 34, 37
 - estimativa da ordem, 39
 - depuração, 39
 - erro global de discretização, 33
 - delimitação, 36
 - estimativa, 38
 - expansão assintótica, 37
 - erro local de discretização, 30
 - estabilidade absoluta, 75
 - instabilidade inerente, 80
- Métodos de passo múltiplo lineares, 12, 85
 - consistência, 89
 - ordem, 90

- convergência, 91
- dedução, 86
- delimitante para o erro local de discretização
 - função influência, 124
- divergência, 108
- erro global de discretização, 118
 - polinômio característico, 120
- erro local de discretização, 87
 - principal, 90
- estabilidade, 103
 - estabilidade absoluta, 120
 - estabilidade fraca, 120
 - zero-estabilidade, 114
- limitante para o erro global de discretização, 125
- limitante para o erro local de discretização, 122
- polinômios característicos, 113
 - raiz principal, 113
- Métodos Preditores-Corretores, 143
 - controle do passo de integração, 156
 - corretor, 145
 - dispositivo de Milne, 151
 - erro local de discretização, 148
 - estabilidade absoluta, 152
 - métodos de Adams-Bashforth, 144
 - métodos de Adams-Moulton, 144
 - métodos de Milne-Simpson, 144
 - métodos de Nyström, 144
 - modificador M, 152
 - polinômio característico, 153
 - preditor, 145
- Polinômio interpolador de Lagrange, 13, 26
- Problema de Cauchy, 3, 14, 29, 76
 - forma integral, 7
- Problema de valor inicial, 16
- Quadratura numérica, 12, 13, 25
- Regra da cadeia, 59
- Runge, 59
- Série de Taylor, 8, 12, 32
- de uma função de duas variáveis, 19
- de uma função de uma variável, 17
- Solução manufaturada, 39
- Teorema
 - de existência e unicidade de solução do Problema de Cauchy, 6
 - de Picard, 17, 179
 - do Confronto, 36
 - do Ponto Fixo, 158
 - do Valor Médio, 7, 18, 88, 118, 148, 154
 - para Integrais, 127
 - Fundamental do Cálculo, 7
- Unicidade de solução
 - condições de contorno, 4
 - condições iniciais, 4