Exemplo de normas induzidas : → não dar a função antes, causa confusão.

1) $|||A|||_1 \triangleq \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$ (máxima soma de colunas)

esta norma é induzida pela norma vetorial ℓ_1:

Seja $A = [\,a_1 \;\; a_2 \; \cdots \; a_n\,] \in \mathbb{C}^{n\times n}$ e $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Então

$$\|Ax\|_1 = \|x_1 a_1 + \cdots + x_n a_n\|_1 \leq \sum_{i=1}^{n} \|x_i a_i\|_1 = \sum_{i=1}^{n} |x_i| \, \|a_i\|_1$$

$$\leq \left[\sum_{i=1}^{n} |x_i|\right]\left[\max_{1\le k\le n} \|a_k\|_1\right] \left(\cancel{= \sum_{i=1}^{n} |x_i| \, |||A|||_1 = \|x\|_1 \cdot |||A|||_1} \right),$$

e portanto, $\max_{\|x\|_1=1} \|Ax\|_1 \leq \cancel{|||A|||_1}\ \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$. Por outro lado, se

escolhermos $x = e_k$ onde k é tal que $\|a_k\|_1 = \max_{1\le i\le n} \|a_i\|_1$,

então $\|Ax\|_1 = \|Ae_k\|_1 = \|a_k\|_1 = |||A|||_1$. Com isto, provamos

que $|||A|||_1 \triangleq \max_{\|x\|_1=1} \|Ax\|_1 = \cancel{|||A|||_1}\ \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| = \max_{1\le j\le n} \|a_j\|_1$.

2) $|||A|||_\infty \triangleq \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ji}|$. Esta norma é induzida pela norma vetorial ℓ_∞, o que pode ser provado de forma semelhante ao resultado anterior.

3) $|||A|||_2$ ou norma espectral:

$$|||A|||_2 \triangleq \max \{ \sqrt{\lambda} : \lambda \text{ é um autovalor de } A^H A \}.$$

(note que se $A^H A x = \lambda x$ e $x \neq 0$, então $x^H A^H A x = \lambda \|x\|^2 = \underbrace{\|Ax\|^2}_{\in \mathbb{R}} \geq 0$, e portanto $\lambda \in \mathbb{R}$, $\lambda \geq 0$)

Vamos provar nas próximas aulas que a norma espectral é induzida pela norma vetorial Euclideana:

$$|||A|||_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 .$$

4) Construindo novas normas: Teorema: Se $|||\cdot|||$ é uma norma matricial em $\mathbb{C}^{n\times n}$ e $S \in \mathbb{C}^{n\times n}$ é não-singular, então $|||A|||_S = |||S^{-1}AS|||$ é uma norma matricial.

Prova: Exercício.

Como vimos, existe uma grande variedade de normas matriciais. Como escolher qual usar em cada situação? Dentre as normas que podem ser usadas com matrizes, aquelas que satisfazem o axioma submultiplicativo (e que chamamos de "normas matriciais") satisfazem $|||I||| \geq 1$, o que é uma espécie de normalização (note que se $|||\cdot|||$ é uma norma matricial, a função $|||\cdot|||_\alpha = \alpha\, |||\cdot|||$ não satisfaz necessariamente ao axioma submultiplicativo). Existiria então, dentre todas as normas submultiplicativas, uma norma mínima, tal que $|||A||| \leq |||A|||_\alpha$ para todo $A \in \mathbb{C}^{n\times n}$ e todas as outras normas matriciais $|||\cdot|||_\alpha$? (nota: o índice α representa aqui apenas "uma outra norma", trata-se somente de uma forma de diferenciar os símbolos $|||\cdot|||$ e $|||\cdot|||_\alpha$ para normas diferentes).

Veremos agora que existe entre as normas matriciais uma classe de normas ótimas, num certo sentido. Dentro desta classe, não há normas melhores ou piores consistentemente. Para este estudo, vamos precisar da seguinte definição:

Definição: Denomina-se raio espectral de uma matriz $A \in \mathbb{C}^{n\times n}$ a quantidade

$$\rho(A) \triangleq \max_{1 \leq i \leq A} |\lambda_i(A)|.$$

Teorema: Se $|||\cdot|||$ é uma norma matricial, então

$$|||A||| \geq \rho(A) \quad \text{para todo } A \in \mathbb{C}^{n\times n}.$$

Prova: Seja λ o autovalor de A tal que $|\lambda| = \rho(A)$, e $\underset{\sim}{x}$ um autovetor correspondente a λ. Defina

$$X = [\underset{\sim}{x} \;\; \underset{\sim}{x} \;\; \cdots \;\; \underset{\sim}{x}] \in \mathbb{C}^{n\times n}.$$

Então,

$$|\lambda| \cdot |||X||| = |||\lambda X||| = |||AX||| \leq |||A||| \cdot |||X|||.$$

Como $\underset{\sim}{x} \neq 0$, $|||X||| \neq 0$ e o resultado segue. ∎

Note que a propriedade submultiplicativa é essencial para a demonstração do teorema!

Na verdade, o raio espectral não é apenas um limite inferior para todas as normas possíveis de uma matriz; $\rho(A)$ é o <u>maior</u> desses limites inferiores. Ou seja,

$$\rho(A) = \inf_{\substack{|||\cdot|||_\alpha \text{ é uma} \\ \text{norma matricial}}} |||A|||_\alpha \,, \quad \text{para todo } A \in \mathbb{C}^{n\times n}.$$

<u>Teorema</u>: Sejam $A \in \mathbb{C}^{n\times n}$ e $\varepsilon > 0$ dados. Então existe uma norma matricial $|||\cdot|||$ tal que

$$\rho(A) \leq |||A||| \leq \rho(A) + \varepsilon.$$

<u>Prova</u>: Seja S uma transformação de similaridade que coloca A na forma de Jordan:

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & a(1) & & \\ & \lambda_2 & a(2) & \\ & & \ddots & a(n-1) \\ 0 & & & \lambda_n \end{bmatrix},$$

onde $a(i) = 0$ ou 1, dependendo da estrutura dos blocos de Jordan que compõe $S^{-1}AS$.

Seja $D = \begin{bmatrix} \varepsilon & & 0 \\ & \varepsilon^2 & \\ & \ddots & \\ 0 & & \varepsilon^n \end{bmatrix}$, e defina $|||B||| = |||D^{-1}S^{-1}BSD|||_1$.

Então $|||A||| = \left|\left|\left| \begin{bmatrix} \lambda_1 & \varepsilon a(1) & & \\ & \lambda_2 & \varepsilon a(2) & \\ & & \ddots & \varepsilon a(n-1) \\ & & & \lambda_n \end{bmatrix} \right|\right|\right|_1 \leq \rho(A) + \varepsilon.$ ∎

Note no entanto que $\rho(A)$ não é uma norma matricial (nem se relaxarmos a definição de norma matricial eliminando o axioma submultiplicativo), pois $\rho(A)$ não obedece à desigualdade triangular, e $\rho(A)$ pode ser igual a zero para $A \neq 0$:

$$\underbrace{\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}}_{\rho(A)=|a|} + \underbrace{\begin{bmatrix} -a & 1 \\ 0 & -a \end{bmatrix}}_{\rho(B)=|a|} = \underbrace{\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}}_{\rho(A+B)=0.}$$

Como veremos em breve ao estudar convergência de séries de matrizes, é conveniente, em muitas situações, usar uma norma tal que $|||A|||$ seja o menor possível. Vamos mostrar agora que não existe uma tal norma uniformemente melhor que as outras, mas que as <u>normas induzidas</u> se aproximam deste ideal o melhor possível.

Teorema: Seja $|||\cdot|||$ uma norma matricial em $\mathbb{C}^{n\times n}$, e seja $|||\cdot|||_\alpha$ uma dada norma induzida, também em $\mathbb{C}^{n\times n}$. Então

a) Existe uma norma induzida $N(\cdot)$ em $\mathbb{C}^{n\times n}$ tal que $N(A) \leq |||A|||$ para todo $A \in \mathbb{C}^{n\times n}$, e

b) $|||A||| \leq |||A|||_\alpha$ para todo $A \in \mathbb{C}^{n\times n}$ se e somente se $|||A||| = |||A|||_\alpha$ para todo $A \in \mathbb{C}^{n\times n}$.

Prova: Defina a norma vetorial $\|\cdot\|$ em $\mathbb{C}^n$ por
$\|x\| = |||X|||$, onde $X = [x \; x, \dots \; x] \in \mathbb{C}^{n\times n}$,
e defina $N(\cdot)$ como a norma matricial induzida por $\|\cdot\|$. Então, para qualquer $A \in \mathbb{C}^{n\times n}$, temos

$$N(A) = \max_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x\neq 0} \frac{|||[Ax \; Ax \cdots Ax]|||}{|||[x \; x \cdots x]|||} = \max_{x\neq 0} \frac{|||AX|||}{|||X|||} \leq$$

$$\leq |||A||| \max_{x\neq 0} \frac{|||X|||}{|||X|||} = |||A|||.$$

Para provar (b), note que $N(A) \leq |||A||| \leq |||A|||_\alpha$ para todo $A \in \mathbb{C}^{n\times n}$, e use o teorema ~~e o corolário~~ abaixo ∎

Teorema: Sejam $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$ duas normas vetoriais em $\mathbb{C}^n$, e sejam $|||\cdot|||_\alpha$ e $|||\cdot|||_\beta$ as respectivas normas induzidas em $\mathbb{C}^{n\times n}$. Defina

$$R_{\alpha\beta} = \max_{x\neq 0} \frac{\|x\|_\alpha}{\|x\|_\beta}, \qquad R_{\beta\alpha} = \max_{x\neq 0} \frac{\|x\|_\beta}{\|x\|_\alpha}.$$

Então, $$\max_{A\neq 0} \frac{|||A|||_\alpha}{|||A|||_\beta} = R_{\alpha\beta} R_{\beta\alpha} = \max_{A\neq 0} \frac{|||A|||_\beta}{|||A|||_\alpha}.$$

Prova: Sejam $A \in \mathbb{C}^{n\times n}$ e $x \in \mathbb{C}^n$ dados, e assuma $x \neq 0$ e $Ax \neq 0$. Então

$$\frac{\|Ax\|_\alpha}{\|x\|_\alpha} = \frac{\|Ax\|_\alpha}{\|Ax\|_\beta} \cdot \frac{\|Ax\|_\beta}{\|x\|_\beta} \cdot \frac{\|x\|_\beta}{\|x\|_\alpha} \leq R_{\alpha\beta} \cdot \frac{\|Ax\|_\beta}{\|x\|_\beta} \cdot R_{\beta\alpha} \qquad (2)$$

(note que a desigualdade final também vale se $Ax = 0$).

Então
$$|||A|||_\alpha = \max_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\alpha} \leq \max_{x\neq 0} \frac{\|Ax\|_\beta}{\|x\|_\beta} \cdot R_{\alpha\beta} \cdot R_{\beta\alpha} = R_{\alpha\beta} R_{\beta\alpha} |||A|||_\beta.$$

Ou seja: mostramos que $\frac{|||A|||_\alpha}{|||A|||_\beta} \leq R_{\alpha\beta} \cdot R_{\beta\alpha}$. Falta agora mostrar que existe A_0 para o qual a desigualdade se torna igualdade.

Os dois máximos em (2) são atingidos para certos vetores em $\mathbb{C}^n$:

Sejam $\underset{\sim}{y}, \underset{\sim}{z}$ tais que $\|\underset{\sim}{y}\|_2 = \|\underset{\sim}{z}\|_2 = 1$, $\|\underset{\sim}{y}\|_\alpha = R_{\alpha\beta}\|\underset{\sim}{y}\|_\beta$, e $\|\underset{\sim}{z}\|_\beta = R_{\beta\alpha}\|\underset{\sim}{z}\|_\alpha$.

Como vimos ao estudar normas duais, existe um vetor $\underset{\sim}{z_0} \in \mathbb{C}^n$ tal que

(a) $|\underset{\sim}{z_0}^H \underset{\sim}{x}| \leq \|\underset{\sim}{x}\|_\beta$ para todo $\underset{\sim}{x} \in \mathbb{C}^n$

(b) $\underset{\sim}{z_0}^H \underset{\sim}{z} = \|\underset{\sim}{z}\|_\beta$.

Defina a matriz $A_0 = \underset{\sim}{y}\,\underset{\sim}{z_0}^H$. Usando (b), temos

$$\frac{\|A_0\underset{\sim}{z}\|_\alpha}{\|\underset{\sim}{z}\|_\alpha} = \frac{\|\underset{\sim}{y}\,\underset{\sim}{z_0}^H\underset{\sim}{z}\|_\alpha}{\|\underset{\sim}{z}\|_\alpha} = \frac{\|\underset{\sim}{y}\|_\alpha\,|\underset{\sim}{z_0}^H\underset{\sim}{z}|}{\|\underset{\sim}{z}\|_\alpha} = \|\underset{\sim}{y}\|_\alpha \cdot \frac{\|\underset{\sim}{z}\|_\beta}{\|\underset{\sim}{z}\|_\alpha},$$

e portanto, $$|||A_0|||_\alpha \geq \frac{\|\underset{\sim}{y}\|_\alpha}{\|\underset{\sim}{y}\|_\beta}\frac{\|\underset{\sim}{z}\|_\beta}{\|\underset{\sim}{z}\|_\alpha}\cdot\|\underset{\sim}{y}\|_\beta = R_{\alpha\beta}R_{\beta\alpha}\|\underset{\sim}{y}\|_\beta.$$

Por outro lado,

$$\frac{\|A_0\underset{\sim}{x}\|_\beta}{\|\underset{\sim}{x}\|_\beta} = \frac{\|\underset{\sim}{y}\,\underset{\sim}{z_0}^H\underset{\sim}{x}\|_\beta}{\|\underset{\sim}{x}\|_\beta} = \frac{\|\underset{\sim}{y}\|_\beta\,|\underset{\sim}{z_0}^H\underset{\sim}{x}|}{\|\underset{\sim}{x}\|_\beta} \leq \frac{\|\underset{\sim}{y}\|_\beta\|\underset{\sim}{x}\|_\beta}{\|\underset{\sim}{x}\|_\beta} = \|\underset{\sim}{y}\|_\beta,$$

e portanto, $|||A_0|||_\beta \leq \|\underset{\sim}{y}\|_\beta$.

Combinando as duas desigualdades, vem

$$\frac{|||A_0|||_\alpha}{|||A_0|||_\beta} \geq R_{\alpha\beta}\cdot R_{\beta\alpha}.$$

Pelo resultado anterior, tínhamos a desigualdade no sentido inverso para todo $A \in \mathbb{C}^{n\times n}$. A conclusão é que

$$\frac{|||A_0|||_\alpha}{|||A_0|||_\beta} = R_{\alpha\beta}R_{\beta\alpha}.$$ ∎

<u>Corolário</u>: Sejam $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$ duas normas vetoriais em $\mathbb{C}^n$; com as respectivas normas matriciais induzidas $|||\cdot|||_\alpha$ e $|||\cdot|||_\beta$. Então $|||A|||_\alpha \leq |||A|||_\beta$ para todo $A \in \mathbb{C}^{n\times n}$ se e somente se $|||A|||_\alpha = |||A|||_\beta$ para todo $A \in \mathbb{C}^{n\times n}$.

<u>Prova</u>: Observe que

$$R_{\alpha\beta} = \max_{\underset{\sim}{x}\neq 0}\frac{\|\underset{\sim}{x}\|_\alpha}{\|\underset{\sim}{x}\|_\beta} = \left[\min_{\underset{\sim}{x}\neq 0}\frac{\|\underset{\sim}{x}\|_\beta}{\|\underset{\sim}{x}\|_\alpha}\right]^{-1} \geq \left[\max_{\underset{\sim}{x}\neq 0}\frac{\|\underset{\sim}{x}\|_\beta}{\|\underset{\sim}{x}\|_\alpha}\right]^{-1} = \frac{1}{R_{\beta\alpha}}$$

$$\Rightarrow \boxed{R_{\alpha\beta}R_{\beta\alpha} \geq 1}$$

Horn, Matrix Analysis, cap. 1 e 3.

<u>Definições</u>: Seja $A \in \mathbb{C}^{n \times n}$. <u>Um autovalor</u> de A é um escalar λ tal que existe $\underset{\sim}{x} \in \mathbb{C}^n$, <u>$\underset{\sim}{x} \neq 0$</u> satisfazendo

$$A\underset{\sim}{x} = \lambda \underset{\sim}{x}.$$

$\underset{\sim}{x}$ é denominado <u>autovetor</u> de A, relativo ao autovalor λ.

O conjunto dos autovalores de uma matriz é o seu <u>espectro</u>:

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda \text{ é um autovalor de } A\}.$$

O <u>raio espectral</u> de A é o máximo módulo dos autovalores de A:

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$$

Se $(\lambda, \underset{\sim}{x})$ é um par autovalor / autovetor de A, então

$$(\lambda I - A)\underset{\sim}{x} = \lambda \underset{\sim}{x} - A\underset{\sim}{x} = \underset{\sim}{0}$$

$\Rightarrow$ λ é autovalor de A se e somente se $\lambda I - A$ é singular, ou seja, se $\det(\lambda I - A) = 0$.

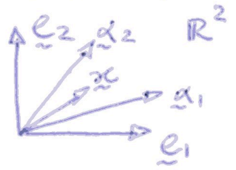
Define-se <u>polinômio característico</u> de A como o polinômio (de grau n)

$$p_A(s) = \det(sI - A).$$

Claramente, $\lambda \in \sigma(A)$ se e somente se é uma raiz de $p_A(\lambda)$. Como $p_A(s)$ tem grau n, existem exatamente <u>n</u> autovalores de A (possivelmente com algumas repetições).

<u>SIMILARIDADE E MUDANÇA DE BASE</u>

Recordação: mudança de base



$\mathcal{B}_1 = \{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ $\quad [\underset{\sim}{x}]_{\mathcal{B}_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\underset{\sim}{x} = x_1 \underset{\sim}{e}_1 + x_2 \underset{\sim}{e}_2$

outra base: $\mathcal{B}_2 = \{\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2\}$

Como escrever $\underset{\sim}{x}$ em função de $\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2$? $\underset{\sim}{x} = x_1^\alpha \underset{\sim}{\alpha}_1 + x_2^\alpha \underset{\sim}{\alpha}_2$

Matriz de mudança de base

$$\begin{aligned} \underset{\sim}{\alpha}_1 &= \alpha_{11}\underset{\sim}{e}_1 + \alpha_{21}\underset{\sim}{e}_2 \\ \underset{\sim}{\alpha}_2 &= \alpha_{12}\underset{\sim}{e}_1 + \alpha_{22}\underset{\sim}{e}_2 \end{aligned} \Rightarrow \begin{aligned} \underset{\sim}{x} = {} & x_1^\alpha(\alpha_{11}\underset{\sim}{e}_1 + \alpha_{21}\underset{\sim}{e}_2) + \\ & + x_2^\alpha(\alpha_{12}\underset{\sim}{e}_1 + \alpha_{22}\underset{\sim}{e}_2) \end{aligned}$$

Por outro lado, se existir S invertível tal que $S^{-1} A S = \Lambda$, então $AS = S\Lambda \Rightarrow$ as colunas de S são os autovetores de A, e formam um conjunto l.i. porque S é invertível. ∎

Falta saber quando os autovetores de uma matriz formam uma base para o espaço todo. Um caso que é relativamente fácil de verificar é dado no seguinte teorema.

<u>Teorema</u>: Se $\lambda_1, \ldots, \lambda_k$ forem autovalores de $A \in \mathbb{C}^{n \times n}$, dois a dois <u>distintos</u>, então os autovetores correspondentes $\{x^{(1)}, \ldots, x^{(k)}\}$ formam um conjunto l.i..

<u>Prova</u>: Suponha que $\{x^{(1)}, \ldots, x^{(k)}\}$ ~~sejam~~ l.d.. Então existem $\alpha_1 \ldots \alpha_k$ tais que $\sum_{i=1}^{k} \alpha_i x^{(i)} = 0$.

Em geral, haverá mais de uma seqüência $\{\alpha_1 \ldots \alpha_k\}$ como acima, e haverá também uma com o menor número de elementos $\{\beta_i\}_{i=1}^{r}$ (todos não-nulo) (que vamos assumir, sem perda de generalidade, correspondentes a $x^{(1)} \ldots x^{(r)}$): $\sum_{i=1}^{r} \beta_i x^{(i)} = 0$, $r \leq k$, $r \geq 1$.

Então, vale também $0 = A . \sum_{i=1}^{r} \beta_i x^{(i)} = \sum_{i=1}^{r} \beta_i \lambda_i x^{(i)}$,

e portanto, $0 = \sum_{i=1}^{r} \beta_i \lambda_i x^{(i)} - \lambda_r \sum_{i=1}^{r} \beta_i x^{(i)} =$

$$= \sum_{i=1}^{r} \beta_i (\lambda_i - \lambda_r) x^{(i)} = \sum_{i=1}^{r-1} \beta_i (\lambda_i - \lambda_r) x^{(i)}$$

Portanto, $\{\beta_i (\lambda_i - \lambda_r)\}_{i=1}^{r-1}$ é outra seqüência com menos elementos $\beta_i (\lambda_i - \lambda_r) \neq 0$, que $\{\beta_i\}_{i=1}^{r}$, e que também satisfaz $\sum_{i=1}^{r-1} (\lambda_i - \lambda_r) \beta_i x^{(i)} = 0$, contradizendo nossa hipótese.

Concluímos que os autovalores $\{x^{(1)} \ldots x^{(k)}\}$ devem ser linearmente independentes. ∎

$$\therefore \quad \underset{\sim}{x} = (x_1^\alpha \alpha_{11} + x_2^\alpha \alpha_{12})\, \underset{\sim}{e}_1 + (x_1^\alpha \alpha_{21} + x_2^\alpha \alpha_{22})\, \underset{\sim}{e}_2$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{22} & \alpha_{22} \end{bmatrix}}_{S} \begin{bmatrix} x_1^\alpha \\ x_2^\alpha \end{bmatrix} \rightarrow [x]_{B_2} = S^{-1}[x]_{B_1}$$

S = matriz de mudança de base.

Transformação linear $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$

$$\underset{\sim}{y} = T\underset{\sim}{x} \Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = [y]_{B_1} = T. \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = T. [\underset{\sim}{x}]_{B_1}$$

Como reescrever a transformação acima com $\underset{\sim}{y}$ e $\underset{\sim}{x}$ na base B_2?

$$S^{-1}[\underset{\sim}{y}]_{B_1} = S^{-1}T\underbrace{S.\,S^{-1}}_{=I}\,[\underset{\sim}{x}]_{B_1} \Rightarrow [y]_{B_2} = S^{-1}TS.\,[x]_{B_2}$$

A transformação $S^{-1}TS$ é chamada <u>transformação de similaridade</u>; diz-se que T e $S^{-1}TS$ são matrizes similares.

Uma transformação de similaridade não altera os autovalores de uma matriz:

$$\det(sI - A) = \det(sI - A).\det(S^{-1}S) = \det(S^{-1}(sI-A)S) =$$
$$= \det(sI - S^{-1}AS) \blacksquare$$

<u>Definição</u>: A matriz A é dita diagonalizável se existir S inversível tal que

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_n \end{bmatrix} = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

<u>Teorema</u>: Seja $A \in \mathbb{C}^{n\times n}$. A é diagonalizável se e somente se existir um conjunto de n autovetores l.i..

Prova: Se $Ax^{(i)} = \lambda_i$, e $\{x^{(1)}, \dots, x^{(n)}\}$ for l.i., a matriz $S = [x^{(1)} \dots x^{(n)}]$ é não singular (pois tem colunas l.i.), e

$$AS = [Ax^{(1)}\ Ax^{(2)} \dots Ax^{(n)}] = [\lambda_1 x^{(1)}\ \lambda_2 x^{(2)} \dots \lambda_n x^{(n)}] =$$
$$= [x^{(1)}\ x^{(2)} \dots x^{(n)}] \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_n \end{bmatrix} = S.\Lambda$$

$$\Rightarrow S^{-1}AS = \Lambda = \mathrm{diag}(\lambda_i)$$

Segue destes resultados o seguinte teorema:

<u>Teorema</u>: Se $A \in \mathbb{C}^{n \times n}$ tiver n autovalores distintos, então A é diagonalizável.

Multiplicidade de autovalores

Alguns autovalores de uma matriz podem ser repetidos, se o polinômio característico tiver raízes múltiplas.

$$p(s) = (s-\lambda_1)^{k_1}(s-\lambda_2)^{k_2}\cdots(s-\lambda_r)^{k_r}$$

$$k_1 + k_2 + \cdots + k_r = n$$

k_i é a <u>multiplicidade algébrica</u> do autovalor λ_i

Define-se também a <u>multiplicidade geométrica</u> de λ_i como o número de autovetores independentes relacionados a λ_i.

A multiplicidade geométrica é sempre menor ou igual à algébrica. Por exemplo,

$$A = \begin{bmatrix} 0{,}5 & 1 \\ 0 & 0{,}5 \end{bmatrix} \Rightarrow p(s) = (s-0{,}5)^2;$$ mult. algébrica de $\lambda = 0{,}5$ é 2.

Os autovetores relativos a $\lambda = 0{,}5$ devem satisfazer

$$(0{,}5I - A)\underset{\sim}{x} = \underset{\sim}{0}$$

$$\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underset{\sim}{0} \Rightarrow \begin{array}{l} x_2 = 0 \\ x_1 = \text{qualquer} \end{array}$$

$\therefore \underset{\sim}{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, não há outro autovetor independente.

Por outro lado, $A_2 = \begin{bmatrix} 0{,}5 & 0 \\ 0 & 0{,}5 \end{bmatrix}$ tem 2 autovetores l.i.: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Formas Canônicas (Forma de Jordan)

Definição: Um bloco de Jordan $J_k(\lambda)$ é uma matriz triangular superior de dimensão k

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

Uma matriz de Jordan $J \in \mathbb{C}^{n\times n}$ é uma soma direta de blocos de Jordan:

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_k}(\lambda_k) \end{bmatrix}$$

(alguns λ_i podem ser repetidos)

Teorema: Toda matriz $A \in \mathbb{C}^{n\times n}$ é similar a uma matriz de Jordan. Esta última é única no sentido que os blocos de Jordan que a compõe são únicos, apesar da ordem não ser pré-fixada.

Prova: Horn, Matrix Analysis, pp. 122-128. ∎

Este resultado é extremamente importante. Apesar de ser quase impossível se calcular na prática a forma de Jordan de uma matriz (trata-se de um problema mal condicionado, como veremos a seguir), a forma de Jordan é muito útil para análises teóricas. Muitas vezes, para se determinar se uma certa hipótese (idéia) vale para todas as matrizes $A \in \mathbb{C}^{n\times n}$, é mais fácil pensar se a idéia funciona para qualquer possível matriz na forma de Jordan.

A forma de Jordan não é uma "função" contínua dos elementos de uma matriz (daí a dificuldade de calcular numericamente a forma de Jordan). Por exemplo, considere

$$A_\varepsilon = \begin{bmatrix} \varepsilon & 0 \\ 1 & 0 \end{bmatrix}.$$ Para $\varepsilon \neq 0$, então

$$A_\varepsilon = \underbrace{\begin{bmatrix} 0 & \varepsilon \\ 1 & 1 \end{bmatrix}}_{= S_\varepsilon} \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & \varepsilon \end{bmatrix}}_{= J_\varepsilon} \underbrace{\begin{bmatrix} -1/\varepsilon & 1 \\ 1/\varepsilon & 0 \end{bmatrix}}_{= S_\varepsilon^{-1}}$$

J_ε tem dois blocos: $J_1(0)$ e $J_1(\varepsilon)$ (A_ε é diagonalizável).

No entanto, para $\varepsilon = 0$ $A_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, que tem forma de Jordan $J_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (use $S_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$). J_0 possui apenas um bloco de Jordan, $J_2(0)$.

<u>Fato</u>: Os autovalores de uma matriz $A \in \mathbb{C}^{n \times n}$ são funções <u>contínuas</u> dos elementos de A. (A prova requer análise complexa – não é tão difícil ver que os <u>coeficientes</u> do polinômio característico são funções contínuas dos elementos da matriz).

$\rightarrow$ Veja o exemplo anterior para notar que o mesmo <u>não vale</u> para os <u>autovetores</u> da matriz.