Matemática – Aula 1

Decomposição Matricial

A decomposição matricial é uma fatoração[1] de uma matriz em alguma forma canônica ou padrão. Existem diversas decomposições matriciais que são úteis para determinadas classes de problemas, como, por exemplo, a fatoração LU para a resolução de sistemas de equações lineares e a decomposição em valores característicos para a análise de estabilidade de sistemas dinâmicos. Geralmente estas decomposições são utilizadas para simplificar a análise de sistemas ou para implementar de maneira mais eficiente os algoritmos numéricos que envolvem operações com matrizes e vetores.

Em muitas aplicações práticas, vários cálculos matriciais, como inversão, determinante e solução de sistemas lineares não são factíveis de serem realizados de forma explícita ótima. Desta forma, a conversão de problemas de cálculos matriciais difíceis em várias tarefas mais fáceis, como a solução de sistemas triangulares ou diagonais, é de grande valia. Além disto, matrizes de dados representando alguma observação numérica são geralmente de grande dimensão e de difícil análise e, portanto, a decomposição destas matrizes em formas canônicas pode revelar suas características e estruturas inerentes, ajudando na interpretação de seus significados de maneira mais rápida.

Nesta aula serão abordados os métodos de decomposição matricial que têm grande aplicação na engenharia química: decomposição em valores característicos, decomposição de Jordan, decomposição em valores singulares, fatoração LU e fatoração QR. Alguns exemplos típicos, tais como análise de estabilidade de sistemas dinâmicos, seleção de estruturas de controle e resolução de sistemas lineares, serão usados na aula prática para fixar a aplicação destes métodos. Na próxima seção serão apresentadas algumas definições e conceitos necessários para abordar os métodos de decomposição.

A tabela a seguir lista outros métodos de decomposição com os respectivos tipos de matrizes em que podem ser aplicados.

---

[1] Fatoração é a decomposição de um objeto em um produto de outros objetos, ou fatores, que multiplicados resultam no objeto original.

Decomposição	Tipo de Matriz	Fatoração	Notação
LU	m × n	$\mathbf{L} \cdot \mathbf{U}$	**L**: triangular inferior (m × m) **U**: triangular superior (m × n)
LUP ou LU com pivotamento parcial	m × n	$\mathbf{P}^{-1} \cdot \mathbf{L} \cdot \mathbf{U}$	**L**: triangular inferior (m × m) **U**: triangular superior (m × n) **P**: matriz de permutação (m × m)
LDU	quadrada	$\mathbf{L} \cdot \mathbf{D} \cdot \mathbf{U}$	**L**: triangular inferior **U**: triangular superior **D**: matriz diagonal
Cholesky	simétrica e positiva definida	$\mathbf{L} \cdot \mathbf{L}^H$	**L**: triangular inferior
LDL	simétrica	$\mathbf{L} \cdot \mathbf{D} \cdot \mathbf{L}^H$	**L**: triangular inferior **D**: matriz diagonal
SVD ou em valores singulares	m × n	$\mathbf{U} \cdot \sum \cdot \mathbf{V}^H$	**U**: matriz unitária (m × m) **V**: matriz unitária (n × n) ∑: diagonal (m × n) com os p valores singulares ≥ 0, p = min(m,n)
QR	m × n	$\mathbf{Q} \cdot \mathbf{R}$	**Q**: matriz ortogonal (m × m) **R**: triangular superior (m × n)
Espectral ou em valores característicos	quadrada e valores característicos distintos	$\mathbf{X} \cdot \mathbf{D} \cdot \mathbf{X}^{-1}$	**X**: matriz regular dos vetores característicos nas colunas **D**: matriz diagonal dos valores característicos
Jordan	quadrada	$\mathbf{T} \cdot \mathbf{J} \cdot \mathbf{T}^{-1}$	**T**: transformação similar **J**: forma canônica de Jordan
Schur	quadrada	$\mathbf{U} \cdot \mathbf{S} \cdot \mathbf{U}^H$	**U**: matriz unitária **S**: forma canônica de Schur, triangular superior com os valores característicos na diagonal
Hessenberg	quadrada	$\mathbf{Q} \cdot \mathbf{H} \cdot \mathbf{Q}^T$	**Q**: matriz ortogonal **H**: forma canônica de Hessenberg, elementos abaixo da sub-diagonal nulos. Se **A** for simétrica ou hermitiana então **H** é tridiagonal.

1.1 Conceitos básicos e definições

Transformação linear: Sejam U e V espaços vetoriais sobre $\Re$. Uma aplicação F: $U \to V$ é chamada transformação linear de U em V se, e somente se,

(a) $F(u_1 + u_2) = F(u_1) + F(u_2)$, $\forall\ u_1, u_2 \in U$, e

(b) $F(\alpha\, u) = \alpha\, F(u)$, $\forall \alpha \in \Re$ e $\forall\ u \in U$.

No caso em que $U = V$, uma transformação linear F: $U \to U$ é chamada também de **operador linear**.

<u>Exemplo</u>: Para verificar se a aplicação F: $\Re^3 \to \Re^2$ definida por $F(x, y, z) = (x, 2x - z)$, $\forall\ (x, y, z) \in \Re^3$, é uma transformação linear, sejam $\mathbf{u}_1 = (x_1, y_1, z_1)$ e $\mathbf{u}_2 = (x_2, y_2, z_2)$ em $\Re^3$ e:

(a) $F(\mathbf{u}_1 + \mathbf{u}_2) = F(x_1 + x_2, y_1 + y_2, z_1 + z_2) = (x_1 + x_2, 2x_1 + 2\,x_2 - z_1 - z_2) = (x_1, 2x_1 - z_1) + (x_2, 2x_2 - z_2) = F(\mathbf{u}_1) + F\,(\mathbf{u}_2)$

(b) $F(\alpha\, \mathbf{u}_1) = F(\alpha\, x_1, \alpha\, y_1, \alpha\, z_1) = (\alpha\, x_1, 2\alpha\, x_1 - \alpha\, z_1) = \alpha\, (x_1, 2x_1 - z_1) = \alpha\, F(\mathbf{u}_1)$, $\forall \alpha \in \Re$.

logo, a aplicação é uma transformação linear.

Verifica-se prontamente que a multiplicação de uma matriz $\mathbf{A} \in \Re^{m\times n}$ por um vetor $\mathbf{x} \in \Re^n$ é uma transformação linear, pois $\mathbf{A}\cdot(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}\cdot\mathbf{x}_1 + \mathbf{A}\cdot\mathbf{x}_2$ e $\mathbf{A}\cdot\alpha\,\mathbf{x} = \alpha\,\mathbf{A}\cdot\mathbf{x}$.

Núcleo ou espaço nulo: Sejam U e V espaços vetoriais sobre $\Re$ e F: $U \to V$ uma transformação linear. Indica-se por Ker(F) ou null(F) e denomina-se núcleo ou espaço nulo de F o seguinte subconjunto de U:

$$\mathrm{Ker}(F) = \mathrm{null}(F) = \{\mathrm{u} \in U \mid F(u) = 0\} \subseteq U$$

<u>Exemplo</u>: Seja F: $\Re^2 \to \Re^3$ a transformação linear dada por $F(x, y) = (0, x + y, 0)$, $\forall\ (x, y) \in \Re^2$, o núcleo de F é obtido fazendo $F(x, y) = (0, 0, 0) = (0, x + y, 0)$. Logo, $x = -y$ e Ker(F) = $\{(x, -x) \mid x \in \Re\}$, como ilustra a figura abaixo:

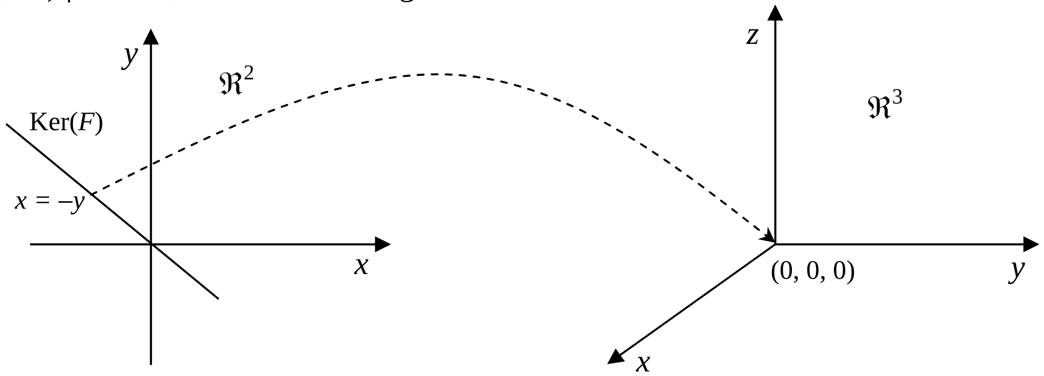


Figura 1: Núcleo de uma transformação linear

O espaço nulo de uma matriz $\mathbf{A} \in \Re^{m\times n}$ é o subespaço Ker(**A**) = null(**A**) = $\{\mathbf{x} \in \Re^n \mid \mathbf{A}\cdot\mathbf{x} = \mathbf{0}\} \subseteq \Re^n$, ou seja, é o conjunto de todas as soluções do sistema homogêneo $\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$. No caso de uma matriz regular (inversível), Ker(**A**) = {**0**}.

Imagem ou espaço gerado: Sejam U e V espaços vetoriais sobre $\Re$ e F: $U \to V$ uma transformação linear. Indica-se por Im(F) ou range(F) e denomina-se imagem ou espaço gerado de F o seguinte subconjunto de V:

$$\mathrm{Im}(F) = \mathrm{range}(F) = \{F(u) \mid u \in U\} \subseteq V$$

A imagem ou espaço gerado por uma matriz $\mathbf{A} \in \Re^{m\times n}$ é o subespaço $\mathrm{Im}(\mathbf{A}) \subseteq \Re^m$ gerado pela transformação linear $F(\mathbf{x}) = \mathbf{A}\cdot\mathbf{x}$, isto é,

$$\mathrm{Im}(\mathbf{A}) = \mathrm{range}(\mathbf{A}) = \{\mathbf{A}\cdot\mathbf{x} \mid \mathbf{x} \in \Re^n\} \subseteq \Re^m$$

Da mesma forma, a imagem ou espaço gerado pela transposta de uma matriz $\mathbf{A} \in \Re^{m\times n}$ é o subespaço $\mathrm{Im}(\mathbf{A}^T) \subseteq \Re^n$ gerado pela transformação linear $F(\mathbf{y}) = \mathbf{A}^T\cdot\mathbf{y}$, isto é,

$$\mathrm{Im}(\mathbf{A}^T) = \mathrm{range}(\mathbf{A}^T) = \{\mathbf{A}^T\cdot\mathbf{y} \mid \mathbf{y} \in \Re^m\} \subseteq \Re^n$$

Um resultado importante é sobre a dimensão destes subespaços. Sejam U e V espaços vetoriais de dimensão finita sobre $\Re$. Dada uma transformação linear F: $U \to V$, então

$$\dim U = \dim \mathrm{Ker}(F) + \dim \mathrm{Im}(F)$$

Conjunto gerador: para um conjunto de vetores $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$ o subespaço:

$$\mathrm{span}(S) = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_r\mathbf{v}_r$$

gerado por todas as combinações lineares dos vetores de S é chamado espaço gerado por S, denominado conjunto gerador.

Com o uso deste conceito, pode-se dizer que range(**A**) é o espaço gerado pelos vetores colunas de **A** e range($\mathbf{A}^T$) é o espaço gerado pelos vetores linhas de **A**.

Posto ou *rank*: o posto ou rank de uma matriz $\mathbf{A} \in \Re^{m\times n}$ é o número máximo de linhas ou colunas linearmente independentes. Observe que dim Im(**A**) = rank(**A**) e, portanto, se Ker(**A**) = {**0**}, então rank(**A**) = n, pois $\dim \Re^n = n = \dim \mathrm{Ker}(\mathbf{A}) + \dim \mathrm{Im}(\mathbf{A})$. Da mesma forma, se $\mathrm{Ker}(\mathbf{A}^T) = \{\mathbf{0}\}$, então rank(**A**) = m.

Matriz simétrica: $\mathbf{A} = \mathbf{A}^T$

Matriz hermitiana (ou auto-adjunta): $\mathbf{A} = \mathbf{A}^H$, ou seja, matriz complexa igual a sua transposta conjugada ($a_{ij} = \overline{a}_{ji}$).

Exemplo de matriz hermitiana: $A = \begin{bmatrix} 1 & 3+i \\ 3-i & 4 \end{bmatrix}$

Matriz ortogonal (**Q**): $\mathbf{Q}^{-1} = \mathbf{Q}^{T}$

Matriz unitária (**U**): $\mathbf{U}^{-1} = \mathbf{U}^{H}$

As matrizes hermitiana e unitária em $\mathbb{C}$ (campo complexo) estão, respectivamente, para as matrizes simétrica e ortogonal em $\Re$. Por isso, as descrições a seguir serão limitadas ao campo dos números reais, podendo ser estendidas para $\mathbb{C}$ substituindo as matrizes ortogonais por matrizes unitárias e as matrizes simétricas por matrizes hermitianas.

1.2 Decomposição em valores e vetores característicos

Dada uma matriz $\mathbf{A} \in \Re^{n \times n}$ pode-se determinar um escalar λ e um vetor $\mathbf{v}$ tal que a equação:

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

seja satisfeita, o escalar λ é chamado de <u>*valor característico*</u> ou <u>*autovalor*</u> da matriz **A** e **v** é chamado de <u>*vetor característico*</u> ou <u>*autovetor*</u> de **A.** A equação de definição do valor e vetor característico pode também ser escrita na forma:

$$(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} = \mathbf{0}$$

transformando-se assim em um sistema linear e homogêneo de equações que apresenta solução apenas se a matriz $\mathbf{A} - \lambda \cdot \mathbf{I}$ for singular, isto é:

$$\det(\mathbf{A} - \lambda \cdot \mathbf{I}) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} = p(\lambda) = 0$$

que é um polinômio de grau n em λ chamado de <u>*polinômio característico*</u> de **A** cujas n raízes são os <u>*valores característicos*</u> ou <u>*autovalores*</u> de **A**. Verifica-se, pela expansão deste determinante, que o único termo de grau n em λ é o correspondente ao produto da diagonal principal de $\mathbf{A} - \lambda \cdot \mathbf{I}$, isto é : $(a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdots (a_{nn} - \lambda)$ sendo todos os demais termos de grau inferior a n, além disto como $p(0) = \det(\mathbf{A})$, o termo independente de λ em $p(\lambda)$ é $\det(\mathbf{A})$, permitindo assim concluir que

$$p(\lambda) = (-\lambda)^n + (a_{11} + a_{22} + \cdots a_{nn})(-\lambda)^{n-1} + \cdots + \det(\mathbf{A}) = 0$$

multiplicando-se cada membro por $(-1)^n$, tem-se:

$$p(\lambda) = \lambda^n - (a_{11} + a_{22} + \cdots a_{nn})\lambda^{n-1} + \cdots + (-1)^n \det(\mathbf{A}) = 0$$

(note que apesar de ter-se multiplicado membro a membro da expressão por $(-1)^n$, manteve-se a notação $p(\lambda)$ para designar o polinômio característico, já que o mesmo está igualado a zero sendo assim irrelevante seu sinal). Pela expressão de $p(\lambda)$ deduz-se que:

(a) $\lambda_1 + \lambda_2 + \cdots \lambda_n = a_{11} + a_{22} + \cdots a_{nn}$, ou seja, $\sum_{i=1}^{n} \lambda_i = tr(\mathbf{A})$, (traço da matriz **A**);

(b) $\lambda_1 \cdot \lambda_2 \cdots \lambda_n = \det(\mathbf{A})$, ou seja, $\prod_{i=1}^{n} \lambda_i = \det(\mathbf{A})$;

(c) como $p(\lambda) = \det(\mathbf{A} - \lambda \cdot \mathbf{I}) = 0$, se **A** for singular tem-se $\det(\mathbf{A}) = 0$ desta forma $p(0) = \det(\mathbf{A}) = 0$, isto é, se **A** for singular $\lambda = 0$ é necessariamente um valor característico de **A**.

Após determinados os valores característicos de **A**, os vetores característicos são determinados através de:

$$(\mathbf{A} - \lambda_i \cdot \mathbf{I}) \cdot \mathbf{v}_i = \mathbf{0} \quad \text{para } i = 1, 2, ..., n,$$

porém, sabe-se que para qualquer matriz quadrada **M**:

$$\mathbf{M} \cdot \text{adj}(\mathbf{M}) = \det(\mathbf{M}) \cdot \mathbf{I}$$

que aplicado a $\mathbf{M} = \mathbf{A} - \lambda_i \cdot \mathbf{I}$ em vista de $\det(\mathbf{A} - \lambda_i \cdot \mathbf{I}) = 0$, tem-se:

$$(\mathbf{A} - \lambda_i \cdot \mathbf{I}) \cdot \text{adj}(\mathbf{A} - \lambda_i \cdot \mathbf{I}) = 0, \quad \text{ou ainda,}$$

$$(\mathbf{A} - \lambda_i \cdot \mathbf{I}) \cdot \left[C^{te} \cdot \text{adj}\left(\mathbf{A} - \lambda_i \cdot \mathbf{I}\right) \right] = \mathbf{0}$$

deste modo qualquer vetor coluna não nulo da matriz $\text{adj}(\mathbf{A} - \lambda_i \cdot \mathbf{I})$ multiplicado por qualquer constante real é um vetor característico de **A**. Como a matriz adjunta de uma matriz quadrada é obtida transpondo-se a matriz construída substituindo cada elemento por seu cofator, para calcular o vetor característico $\mathbf{v}_i$ basta calcular os cofatores não nulos de uma linha qualquer de $(\mathbf{A} - \lambda_i \cdot \mathbf{I})$ multiplicados por uma constante real conveniente.

Pré-multiplicando pela matriz **A** a equação de definição dos valores e vetores característicos tem-se: $\mathbf{A}^2 \cdot \mathbf{v} = \lambda \cdot (\mathbf{A} \cdot \mathbf{v})$, mas $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$, assim $\mathbf{A}^2 \cdot \mathbf{v} = \lambda^2 \cdot \mathbf{v}$, repetindo o procedimento a esta última equação tem-se $\mathbf{A}^3 \cdot \mathbf{v} = \lambda^3 \cdot \mathbf{v}$, e assim sucessivamente, permitindo escrever:

$$\mathbf{A}^m \cdot \mathbf{v} = \lambda^m \cdot \mathbf{v} \quad \text{para } m = 1, 2, ..,$$

isto é, os valores característicos de $\mathbf{A}^m$ são os valores característicos de **A** elevados a mesma potência m e os vetores característicos são os mesmos.

Se a matriz **A** é regular (admite inversa) então $\lambda = 0$ não é valor característico, pois $p(0) = \det(\mathbf{A}) \neq 0$, assim da definição de valores e vetores característicos:

$$\mathbf{A}^{-1} \cdot \mathbf{v} = \frac{1}{\lambda} \cdot \mathbf{v}$$

conclui-se que os valores característicos de $\mathbf{A}^{-1}$ são os recíprocos dos valores característicos de **A** e os vetores característicos são os mesmos. Desta forma, a afirmação de que os valores característicos de $\mathbf{A}^m$ são os valores característicos de **A** elevados a mesma potência m e os vetores característicos são os mesmos vale para todos os valores inteiros de m, isto é para $m = 0, \pm 1, \pm 2, ...$

Estas últimas propriedades podem também ser aplicadas a funções polinomiais de **A** do tipo:

$$q(\mathbf{A}) = \mathbf{A}^m + a_1\mathbf{A}^{m-1} + a_2\mathbf{A}^{m-2} + \cdots + a_{m-1}\mathbf{A} + a_m\mathbf{I}$$

assim $q(\mathbf{A})\cdot\mathbf{v} = \mathbf{A}^m\cdot\mathbf{v} + a_1\mathbf{A}^{m-1}\cdot\mathbf{v} + a_2\mathbf{A}^{m-2}\cdot\mathbf{v} + \cdots + a_{m-1}\mathbf{A}\cdot\mathbf{v} + a_m\mathbf{v}$, e se **v** é um vetor característico de **A,** tem-se:

$$q(\mathbf{A})\cdot\mathbf{v} = \lambda^m\cdot\mathbf{v} + a_1\lambda^{m-1}\cdot\mathbf{v} + a_2\lambda^{m-2}\cdot\mathbf{v} + \cdots + a_{m-1}\lambda\cdot\mathbf{v} + a_m\mathbf{v} = q(\lambda)\cdot\mathbf{v},$$

isto é se λ é um valor característico e **v** o correspondente vetor característico de **A**, então $q(\lambda)$ é valor característico e **v** o correspondente vetor característico de $q(\mathbf{A})$.

Pela propriedade acima e a propriedade de que $\sum_{i=1}^{n}\lambda_i = tr(\mathbf{A})$, tem-se que:

$$S_m = \sum_{i=1}^{n}\lambda_i^m = tr\left(\mathbf{A}^m\right)$$

isto é, a soma da *m*-ésima potência dos valores característicos de **A** é igual ao traço da matriz $\mathbf{A}^m$.

Considerando a seguinte expansão do polinômio característico de **A**:

$$p(\lambda) = \lambda^n + c_1\cdot\lambda^{n-1} + c_2\cdot\lambda^{n-2} + \cdots + c_{n-1}\cdot\lambda + c_n$$

que pode também ser expresso pelo produto dos monômios $(\lambda - \lambda_i)$, isto é:

$$p(\lambda) = (\lambda-\lambda_1)\cdot(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

além disto:

$$\frac{dp(\lambda)}{d\lambda} = (\lambda-\lambda_2)\cdots(\lambda-\lambda_n) + (\lambda-\lambda_1)\cdot(\lambda-\lambda_3)\cdots(\lambda-\lambda_n) + \cdots + (\lambda-\lambda_1)\cdot(\lambda-\lambda_2)\cdots(\lambda-\lambda_{n-1}) =$$

$$= \frac{p(\lambda)}{(\lambda-\lambda_1)} + \frac{p(\lambda)}{(\lambda-\lambda_2)} + \cdots + \frac{p(\lambda)}{(\lambda-\lambda_n)} = \sum_{i=1}^{n}\frac{p(\lambda)}{(\lambda-\lambda_i)}$$

e cada uma destas parcelas pode ser obtida pela divisão da forma original de $p(\lambda)$ por $(\lambda - \lambda_i)$:

$$\frac{p(\lambda)}{(\lambda-\lambda_i)} = \lambda^{n-1} + (c_1+\lambda_i)\cdot\lambda^{n-2} + \left(c_2 + c_1\cdot\lambda_i + \lambda_i^2\right)\cdot\lambda^{n-3} + \left(c_3 + c_2\cdot\lambda_i + c_1\cdot\lambda_i^2 + \lambda_i^3\right)\cdot\lambda^{n-4} +$$

$$+\cdots+\left(c_{n-1} + c_{n-2}\cdot\lambda_i + c_{n-3}\cdot\lambda_i^2 + \cdots + c_1\cdot\lambda_i^{n-2} + \lambda_i^{n-1}\right) \text{ para } i = 1, 2, ..., n$$

assim:

$$\frac{dp(\lambda)}{d\lambda} = \sum_{i=1}^{n}\frac{p(\lambda)}{(\lambda-\lambda_i)} = n\cdot\lambda^{n-1} + (nc_1+S_1)\cdot\lambda^{n-2} + \left(nc_2 + c_1\cdot S_1 + S_2\right)\cdot\lambda^{n-3} +$$

$$\left(nc_3 + c_2\cdot S_1 + c_1\cdot S_2 + S_3\right)\cdot\lambda^{n-4} + \cdots + \left(nc_{n-1} + c_{n-2}\cdot S_1 + c_{n-3}\cdot S_2 + \cdots + c_1\cdot S_{n-2} + S_{n-1}\right)$$

subtraindo desta expressão a expressão obtida derivando diretamente a forma original de $p(\lambda)$, isto é, $\frac{dp(\lambda)}{d\lambda} = n\lambda^{n-1} + (n-1)c_1\cdot\lambda^{n-2} + (n-2)c_2\cdot\lambda^{n-3} + \cdots + c_{n-1}$, tem-se:

$$\begin{cases} c_1 + S_1 = 0 \\ 2c_2 + c_1\cdot S_1 + S_2 = 0 \\ \vdots \\ (n-1)c_{n-1} + c_{n-2}\cdot S_1 + c_{n-3}\cdot S_2 + \cdots + c_1\cdot S_{n-2} + S_{n-1} = 0 \end{cases},$$

e como $p(\lambda_i) = \lambda_i^n + c_1 \cdot \lambda_i^{n-1} + c_2 \cdot \lambda_i^{n-2} + \cdots + c_{n-1} \cdot \lambda_i + c_n = 0$ para $i = 1, 2, ..., n$, tem-se também pelo $\sum_{i=1}^{n} p(\lambda_i) = 0$:

$nc_n + c_{n-1} \cdot S_1 + c_{n-2} \cdot S_2 + \cdots + c_1 \cdot S_{n-1} + S_n = 0$, resultando assim em um sistema linear triangular cuja solução pode ser expressa na forma recursiva:

$$\begin{cases} c_1 = -S_1 \\ c_2 = -\dfrac{1}{2}(c_1 \cdot S_1 + S_2) \\ \vdots \\ c_i = -\dfrac{1}{i}(c_{i-1} \cdot S_1 + c_{i-2} \cdot S_2 + \cdots + c_1 \cdot S_{i-1} + S_i) \\ \vdots \\ c_n = -\dfrac{1}{n}(c_{n-1} \cdot S_1 + c_{n-2} \cdot S_2 + \cdots + c_1 \cdot S_{n-1} + S_n) \end{cases}$$

Este método é chamado de *método de Leverrier*, que determina recursivamente os coeficientes de $p(\lambda)$ a partir do cálculo dos traços das sucessivas potências de **A** de 1 a *n*.

Exemplo: Aplique o método de Leverrier para determinar o polinômio característico, os valores característicos e os vetores característicos da matriz:

$\mathbf{A} = \begin{pmatrix} -2,50 & -1,00 & -1,25 \\ -0,50 & -4,00 & -1,75 \\ 1,00 & 2,00 & 0,50 \end{pmatrix}$, assim $\mathbf{A}^2 = \begin{pmatrix} 5,50 & 4,00 & 4,25 \\ 1,50 & 13,00 & 6,75 \\ -3,00 & -8,00 & -4,50 \end{pmatrix}$ e

$\mathbf{A}^3 = \begin{pmatrix} -11,50 & -13,00 & -11,75 \\ -3,50 & -40,00 & -21,25 \\ 7,00 & 26,00 & 15,50 \end{pmatrix}$, resultando em S_1 = tr(**A**) = –6; S_2 = tr($\mathbf{A}^2$) = 14 e S_3 = tr($\mathbf{A}^3$) = –36. E, recursivamente: $c_1 = -S_1 = 6$; $c_2 = -\frac{1}{2}\left[6 \cdot (-6) + 14\right] = 11$ e

$c_3 = -\frac{1}{3}\left[11 \cdot (-6) + 6 \cdot 14 - 36\right] = 6$, logo: $p(\lambda) = \lambda^3 + 6 \cdot \lambda^2 + 11 \cdot \lambda + 6$, por inspeção verifica-se que as três raízes características são: $\lambda_1 = -1$; $\lambda_2 = -2$ e $\lambda_3 = -3$. Para determinar os correspondentes vetores característicos assim procede-se:

1º Vetor Característico: $\lambda_1 = -1$,

$\mathbf{A} - \lambda_1 \cdot \mathbf{I} = \mathbf{A} + \mathbf{I} = \begin{pmatrix} -1,50 & -1,00 & -1,25 \\ -0,50 & -3,00 & -1,75 \\ 1,00 & 2,00 & 1,50 \end{pmatrix}$, cofatores da primeira linha: –1,00; –1,00 e 2,00,

multiplicando estes cofatores por –1, tem-se: $\mathbf{v_1} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$

2º Vetor Característico: $\lambda_2 = -2$,

$$\mathbf{A} - \lambda_2 \cdot \mathbf{I} = \mathbf{A} + 2 \cdot \mathbf{I} = \begin{pmatrix} -0,50 & -1,00 & -1,25 \\ -0,50 & -2,00 & -1,75 \\ 1,00 & 2,00 & 2,50 \end{pmatrix}$$, cofatores da primeira linha:

$-1{,}50$; $-0{,}50$ e $1{,}00$, multiplicando estes cofatores por -2, tem-se: $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$

<u>3º Vetor Característico</u>: $\lambda_3 = -3$, $\mathbf{A} - \lambda_3 \cdot \mathbf{I} = \mathbf{A} + 3 \cdot \mathbf{I} = \begin{pmatrix} 0,50 & -1,00 & -1,25 \\ -0,50 & -1,00 & -1,75 \\ 1,00 & 2,00 & 3,50 \end{pmatrix}$,

cofatores da segunda linha: 1,00; 3,00 e $-2{,}00$, mantendo estes cofatores como elementos do vetor, tem-se: $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$

Note que: $\mathbf{A}^3 + 6 \cdot \mathbf{A}^2 + 11 \cdot \mathbf{A} + 6 \cdot \mathbf{I} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ e que construindo uma matriz cujos vetores coluna são os vetores característicos de **A**, isto é: $\mathbf{X} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 3 \\ -2 & -2 & -2 \end{pmatrix}$, tem-se:

$$\mathbf{A} \cdot \mathbf{X} = \begin{pmatrix} -1 & -6 & -3 \\ -1 & -2 & -9 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 3 \\ -2 & -2 & -2 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \mathbf{X} \cdot \mathbf{D} \text{ ou } \mathbf{X}^{-1} \cdot \mathbf{A} \cdot \mathbf{X} = \mathbf{D}$$, onde

D=diag(-1, -2, -3).

Uma propriedade muito importante de valores característicos e vetores característicos diz respeito a matrizes simétricas, para isto associa-se um outro problema de valores e vetores característicos à mesma matriz **A**:

Determinar os valores de μ e de **w** que satisfazem a $\mathbf{w}^\mathbf{T} \cdot \mathbf{A} = \mu \mathbf{w}^\mathbf{T}$ ou, de forma análoga, $\mathbf{A}^T \cdot \mathbf{w} = \mu \mathbf{w}$ (na realidade está se estudando o problema dos valores e vetores característicos associado à matriz transposta de **A**) este problema pode também ser colocado na forma: $(\mathbf{A}^T - \mu \mathbf{I}) \cdot \mathbf{w} = \mathbf{0}$, que, de forma análoga ao caso anterior, só apresenta solução se $\det(\mathbf{A}^T - \mu \mathbf{I}) = \det(\mathbf{A} - \mu \mathbf{I}) = 0$, resultando assim no mesmo polinômio característico do problema anterior. Assim os valores característicos deste novo problema são os mesmos do problema anterior, entretanto os vetores característicos não são os mesmos (a não ser que a matriz **A** seja simétrica). Neste caso os vetores característicos são os vetores coluna não nulos de adj($\mathbf{A}^\mathrm{T} - \lambda\mathbf{I}$), que são na realidade os vetores linha não nulos da matriz adj($\mathbf{A} - \lambda\mathbf{I}$) (que são os cofatores não nulos de uma coluna da matriz $\mathbf{A} - \lambda\mathbf{I}$ multiplicados por constante real conveniente).

Sejam $\mathbf{v}_i$ e $\mathbf{w}_j$ dois vetores característicos de **A** e $\mathbf{A}^\mathrm{T}$, respectivamente, correspondentes a valores característicos distintos $\lambda_i \neq \lambda_j$, assim:

$$\begin{cases} \mathbf{A}\cdot\mathbf{v}_i = \lambda_i\cdot\mathbf{v}_i \\ \mathbf{w}_j^{\mathbf{T}}\cdot\mathbf{A} = \lambda_j\cdot\mathbf{w}_j^T \end{cases}$$

pré-multiplicando a primeira expressão por $\mathbf{w}_j^T$, tem-se: $\mathbf{w}_j^T\cdot\mathbf{A}\cdot\mathbf{v}_i = \lambda_i\mathbf{w}_j^T\cdot\mathbf{v}_i$, mas: $\mathbf{w}_j^{\mathbf{T}}\cdot\mathbf{A} = \lambda_j\cdot\mathbf{w}_j^T$, então $\lambda_j\mathbf{w}_j^T\cdot\mathbf{v}_i = \lambda_i\mathbf{w}_j^T\cdot\mathbf{v}_i \Rightarrow (\lambda_j-\lambda_i)\mathbf{w}_j^T\cdot\mathbf{v}_i = 0$ como $\lambda_i \neq \lambda_j$, esta última expressão só é nula se $\mathbf{w}_j^T\cdot\mathbf{v}_i = 0$,isto é os vetores $\mathbf{v}_i$ e $\mathbf{w}_j$ para $i \neq j$ são *ortogonais* entre si.

Exemplo: No exemplo anterior tinha-se:

1º Vetor Característico: $\lambda_1 = -1$,

$$\mathbf{A}-\lambda_{\mathbf{1}}\cdot\mathbf{I} = \mathbf{A}+\mathbf{I} = \begin{pmatrix} -1,50 & -1,00 & -1,25 \\ -0,50 & -3,00 & -1,75 \\ 1,00 & 2,00 & 1,50 \end{pmatrix}$$

, cofatores da primeira coluna: –1,00; –1,00 e –2,00, multiplicando estes cofatores por –1, tem-se: $\mathbf{w_1} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$

2º Vetor Característico: $\lambda_2 = -2$,

$$\mathbf{A}-\lambda_2\cdot\mathbf{I} = \mathbf{A}+2\cdot\mathbf{I} = \begin{pmatrix} -0,50 & -1,00 & -1,25 \\ -0,50 & -2,00 & -1,75 \\ 1,00 & 2,00 & 2,50 \end{pmatrix}$$

, cofatores da primeira coluna: –1,50; 0,00 e –0,75, multiplicando estes cofatores por –4/3, tem-se: $\mathbf{w}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$

3º Vetor Característico: $\lambda_3 = -3$,

$$\mathbf{A}-\lambda_3\cdot\mathbf{I} = \mathbf{A}+3\cdot\mathbf{I} = \begin{pmatrix} 0,50 & -1,00 & -1,25 \\ -0,50 & -1,00 & -1,75 \\ 1,00 & 2,00 & 3,50 \end{pmatrix}$$

, cofatores da segunda coluna: 0,00; 3,00 e 1,50, dividindo estes cofatores por 1,5, tem-se: $\mathbf{w}_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$

Note que: $\mathbf{w}_1^T\cdot\mathbf{v}_1 = -2$; $\mathbf{w}_1^T\cdot\mathbf{v}_2 = \mathbf{w}_1^T\cdot\mathbf{v}_3 = 0$

$\mathbf{w}_2^T\cdot\mathbf{v}_2 = 4$; $\mathbf{w}_2^T\cdot\mathbf{v}_1 = \mathbf{w}_2^T\cdot\mathbf{v}_3 = 0$

$\mathbf{w}_3^T\cdot\mathbf{v}_3 = 4$; $\mathbf{w}_3^T\cdot\mathbf{v}_1 = \mathbf{w}_3^T\cdot\mathbf{v}_2 = 0$

assim redefinindo os vetores $\mathbf{w}_1$, $\mathbf{w}_2$ e $\mathbf{w}_3$ como : $\mathbf{w}_{1,\text{novo}} = \mathbf{w}_1 / (-2)$; $\mathbf{w}_{2,\text{novo}} = \mathbf{w_2} / 4$ e $\mathbf{w}_{3,\text{novo}} = \mathbf{w_3} / 4$, tem-se:

$$\mathbf{w}_{\mathbf{1},novo} = \begin{pmatrix} -0,5 \\ -0,5 \\ -1 \end{pmatrix} ; \mathbf{w}_{2,novo} = \begin{pmatrix} 0,50 \\ 0,00 \\ 0,25 \end{pmatrix} \text{ e } \mathbf{w}_{3,novo} = \begin{pmatrix} 0,00 \\ 0,50 \\ 0,25 \end{pmatrix}$$

, vê-se que a matriz :

$$\mathbf{Q} = \begin{pmatrix} \mathbf{w}_{1,novo}^T \\ \mathbf{w}_{2,novo}^T \\ \mathbf{w}_{3,novo}^T \end{pmatrix} = \begin{pmatrix} -0,50 & -0,50 & -1,00 \\ 0,50 & 0,00 & 0,25 \\ 0,00 & 0,50 & 0,25 \end{pmatrix} = \mathbf{X}^{-1}, \text{ pois : } \mathbf{Q}\cdot\mathbf{X} = \mathbf{I}$$

No caso particular de **A** ser simétrica como $\mathbf{v}_i = \mathbf{w}_i$ para todo i tem-se os vetores característicos <u>*mutuamente ortogonais*</u> entre si, isto é $\mathbf{v}_j^T \cdot \mathbf{v}_i = 0$ para $i \neq j$, se além disto $\|\mathbf{v}_i\| = 1$ (vetores característicos de norma unitária) tem-se $\mathbf{v}_j^T \cdot \mathbf{v}_i = \delta_{ij}$ e a matriz composta pelos vetores característicos $\mathbf{X} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix}$ é uma <u>*matriz ortonormal*</u>.

Uma outra propriedade de matrizes simétricas é que só apresentam valores característicos reais, isto pode ser demonstrado considerando a hipótese oposta, isto é seja $\lambda_i = \alpha + \beta \cdot \boldsymbol{i}$ um valor característico de **A** correspondendo a um vetor característico (também complexo): $\mathbf{v}_i = \mathbf{a} + \mathbf{b} \cdot \boldsymbol{i}$. Como todos os elementos da matriz **A** são reais os coeficientes do polinômio característico serão também todos reais, desta forma se $\lambda_i = \alpha + \beta \cdot \boldsymbol{i}$ é raiz de $p(\lambda)$ o seu conjugado $\overline{\lambda}_i = \alpha - \beta \cdot \boldsymbol{i}$ também o será correspondendo a um vetor característico $\overline{\mathbf{v}}_i = \mathbf{a} - \mathbf{b} \cdot \boldsymbol{i}$. Demonstrou-se, entretanto, que em matrizes simétricas $\mathbf{v}_i = \mathbf{w}_i$ e que de forma geral $(\lambda_j - \lambda_i)\mathbf{w}_j^T \cdot \mathbf{v}_i = 0$ para $\lambda_i \neq \lambda_j$, que no caso de matrizes simétricas será $(\lambda_j - \lambda_i)\mathbf{v}_j^T \cdot \mathbf{v}_i = 0$, adotando nesta expressão $\lambda_j = \overline{\lambda}_i$ e $\mathbf{v}_j = \overline{\mathbf{v}}_i$, tem-se: $\lambda_j - \lambda_i = 2\beta \cdot \boldsymbol{i}$ e $\mathbf{v}_j^T \cdot \mathbf{v}_i = \mathbf{a}^T \cdot \mathbf{a} + \mathbf{b}^T \cdot \mathbf{b} = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2$, assim $(\lambda_j - \lambda_i)\mathbf{v}_j^T \cdot \mathbf{v}_i = 2\beta \cdot \left(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2\right) \cdot \boldsymbol{i} = 0$ e como $\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 > 0$ tem-se necessariamente $\beta \equiv 0$ o que contradiz a hipótese inicial da matriz admitir um valor característico complexo.

<u>Exemplo</u>: Determine o polinômio característico, os valores característicos e os vetores característicos da matriz simétrica:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 2 & 0 & -1 \end{pmatrix}, \text{ assim: } \mathbf{A}^2 = \begin{pmatrix} 6 & -3 & 0 \\ -3 & 5 & -2 \\ 0 & -2 & 5 \end{pmatrix} \text{ e } \mathbf{A}^3 = \begin{pmatrix} 9 & -12 & 12 \\ -12 & 13 & -4 \\ 12 & -4 & -5 \end{pmatrix}, \text{ resultando em}$$

$S_1 = \mathrm{tr}(\mathbf{A}) = 2$; $S_2 = \mathrm{tr}(\mathbf{A}^2) = 16$ e $S_3 = \mathrm{tr}(\mathbf{A}^3) = 17$. E, recursivamente: $c_1 = -S_1 = -2$; $c_2 = -\frac{1}{2}[-2\cdot 2 + 16] = -6$ e $c_3 = -\frac{1}{3}[-6\cdot 2 - 2\cdot 16 + 17] = 9$, logo: $p(\lambda) = \lambda^3 - 2\cdot\lambda^2 - 6\cdot\lambda + 9$, por inspeção verifica-se que: $\lambda_1 = 3$ é raiz de $p(\lambda)$, assim: $\frac{\lambda^3 - 2\cdot\lambda^2 - 6\cdot\lambda + 9}{\lambda - 3} = \lambda^2 + \lambda - 3$ cujas raízes são $\lambda_2 = \frac{\sqrt{13} - 1}{2}$ e $\lambda_3 = -\left(\frac{\sqrt{13} + 1}{2}\right)$. Para determinar os correspondentes vetores característicos assim procede-se:

<u>1º Vetor Característico</u>: $\lambda_1 = 3$,

$\mathbf{A}-\lambda_1\cdot\mathbf{I}=\mathbf{A}-3\mathbf{I}=\begin{pmatrix}-2 & -1 & 2\\ -1 & -1 & 0\\ 2 & 0 & -4\end{pmatrix}$, cofatores da primeira linha: 4; −4 e 2, dividindo estes cofatores por 2, tem-se: $\mathbf{v}_1=\begin{pmatrix}2\\-2\\1\end{pmatrix}$

<u>2º Vetor Característico</u>: $\lambda_2=\dfrac{\sqrt{13}-1}{2}$,

$\mathbf{A}-\lambda_2\cdot\mathbf{I}=\mathbf{A}-\left(\dfrac{\sqrt{13}-1}{2}\right)\cdot\mathbf{I}=\begin{pmatrix}\dfrac{3-\sqrt{13}}{2} & -1 & 2\\ -1 & \dfrac{5-\sqrt{13}}{2} & 0\\ 2 & 0 & -\left(\dfrac{1+\sqrt{13}}{2}\right)\end{pmatrix}$, cofatores da terceira linha: $\sqrt{13}-5$; -2 e $6-2\sqrt{13}$, mantendo estes cofatores como os elementos do vetor característico, tem-se: $\mathbf{v}_2=\begin{pmatrix}\sqrt{13}-5\\-2\\6-2\sqrt{13}\end{pmatrix}$

<u>3º Vetor Característico</u>: $\lambda_3=-\left(\dfrac{\sqrt{13}+1}{2}\right)$

$\mathbf{A}-\lambda_3\cdot\mathbf{I}=\mathbf{A}+\left(\dfrac{\sqrt{13}+1}{2}\right)\cdot\mathbf{I}=\begin{pmatrix}\dfrac{3+\sqrt{13}}{2} & -1 & 2\\ -1 & \dfrac{5+\sqrt{13}}{2} & 0\\ 2 & 0 & \dfrac{\sqrt{13}-1}{2}\end{pmatrix}$, cofatores da terceira linha: $-\left(5+\sqrt{13}\right)$; -2 e $6+2\sqrt{13}$, mantendo estes cofatores como os elementos do vetor característico, tem-se: $\mathbf{v}_3=\begin{pmatrix}-\left(5+\sqrt{13}\right)\\-2\\6+2\sqrt{13}\end{pmatrix}$

Note que $\mathbf{A}^3-2\cdot\mathbf{A}^2-6\cdot\mathbf{A}+9\cdot\mathbf{I}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$

e que $\mathbf{v}_1^T\cdot\mathbf{v}_1=9$; $\mathbf{v}_1^T\cdot\mathbf{v}_2=\mathbf{v}_1^T\cdot\mathbf{v}_3=0$

$\mathbf{v}_2^T\cdot\mathbf{v}_2=7,411257$; $\mathbf{v}_2^T\cdot\mathbf{v}_1=\mathbf{v}_2^T\cdot\mathbf{v}_3=0$

$\mathbf{v}_3^T \cdot \mathbf{v}_3 = 252,588743;\ \mathbf{v}_3^T \cdot \mathbf{v}_1 = \mathbf{v}_3^T \cdot \mathbf{v}_2 = 0$

assim redefinindo os vetores $\mathbf{v}_1$, $\mathbf{v}_2$ e $\mathbf{v}_3$ como:

$\mathbf{v}_{1,novo} = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{pmatrix} 0,667 \\ -0,667 \\ 0,333 \end{pmatrix}$;$\mathbf{v}_{2,novo} = -\dfrac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{pmatrix} 0,512 \\ 0,735 \\ 0,445 \end{pmatrix}$ e $\mathbf{v}_{3,novo} = \dfrac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \begin{pmatrix} -0,542 \\ -0,126 \\ 0,831 \end{pmatrix}$, vê-se que a matriz :

$\mathbf{X} = \begin{pmatrix} \mathbf{v}_{1,novo} & \mathbf{v}_{2,novo} & \mathbf{v}_{3,novo} \end{pmatrix} = \begin{pmatrix} 0,667 & 0,512 & -0,542 \\ -0,667 & 0,735 & -0,126 \\ 0,333 & 0,445 & 0,831 \end{pmatrix}$, é ortonormal pois

$\mathbf{X}^T \cdot \mathbf{X} = \mathbf{X} \cdot \mathbf{X}^T = \mathbf{I}$.

Uma outra propriedade importante relativa a valores característicos de matrizes diz respeito à sua <u>*invariância*</u> à seguinte transformação da matriz **A**: $\mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}$ onde **P** é uma matriz não-singular, assim os valores característicos de $\mathbf{B} = \mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}$ são as raízes de

$$p(\lambda) = \det(\mathbf{B} - \lambda\mathbf{I}) = \det\left(\mathbf{P}^{-1} \cdot \mathbf{A} \cdot \mathbf{P} - \lambda\mathbf{I}\right) = \det\left[\mathbf{P}^{-1} \cdot (\mathbf{A} - \lambda\mathbf{I}) \cdot \mathbf{P}\right] =$$

$$= \det\left(\mathbf{P}^{-1}\right) \cdot \det(\mathbf{A} - \lambda\mathbf{I}) \cdot \det(\mathbf{P}) = \frac{1}{\det(\mathbf{P})} \cdot \det(\mathbf{A} - \lambda\mathbf{I}) \cdot \det(\mathbf{P}) = \det(\mathbf{A} - \lambda\mathbf{I})$$

mantendo-se assim idêntico ao polinômio característico de **A**, desta forma os valores característicos de **A** e de $\mathbf{B} = \mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}$ são os mesmos. Entretanto os vetores característicos de **A** e **B** são distintos, assim sejam $\mathbf{v}_i$ e $\mathbf{u}_i$ os vetores característicos de **A** e **B**, respectivamente, correspondentes ao mesmo valor característico λ_i, isto é: $\mathbf{A} \cdot \mathbf{v}_i = \lambda_i \cdot \mathbf{v}_i$ e $\mathbf{B} \cdot \mathbf{u}_i = \lambda_i \cdot \mathbf{u}_i$ ou seja: $\mathbf{P}^{-1} \cdot \mathbf{A} \cdot \mathbf{P} \cdot \mathbf{u}_i = \lambda_i \cdot \mathbf{u}_i$, pré-multiplicando membro a membro desta última expressão por **P**, tem-se: $\mathbf{A} \cdot (\mathbf{P} \cdot \mathbf{u}_i) = \lambda_i \cdot (\mathbf{P} \cdot \mathbf{u}_i)$ que confrontado com a definição de vetor característico de **A** permite concluir que $\mathbf{v}_i = \mathbf{P} \cdot \mathbf{u}_i$ ou $\mathbf{u}_i = \mathbf{P}^{-1} \cdot \mathbf{v}_i$, isto pode ser interpretando como uma mudança de coordenadas, assim os elementos do vetor $\mathbf{u}_i$ nada mais são que os componentes do vetor $\mathbf{v}_i$ na base composta pelos vetores coluna da matriz **P**.

<u>Exemplo</u>: Em exemplo anterior determinaram-se os valores e vetores característicos da matriz: $\mathbf{A} = \begin{pmatrix} -2,50 & -1,00 & -1,25 \\ -0,50 & -4,00 & -1,75 \\ 1,00 & 2,00 & 0,50 \end{pmatrix}$, mostre que a transformação:

$\mathbf{B} = \mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}$ com $\mathbf{P} = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -4 & -3 \\ -1 & 1 & 3 \end{pmatrix}$ não modifica os valores característicos de **A** e os novos vetores característicos são $\mathbf{u}_i = \mathbf{P}^{-1} \cdot \mathbf{v}_i$.

Invertendo a matriz **P**, tem-se: $\mathbf{P}^{-1} = \begin{pmatrix} 4,5 & 2,5 & 1,0 \\ -3,0 & -2,0 & -1,0 \\ 2,5 & 1,5 & 1,0 \end{pmatrix}$, assim:

$\mathbf{B} = \mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P} = \begin{pmatrix} 10,50 & 17,50 & -2,50 \\ -8,25 & -14,25 & 0,75 \\ 5,75 & 8,75 & -2,25 \end{pmatrix}$; $\mathbf{B}^2 = \begin{pmatrix} -48,50 & -87,50 & -7,50 \\ 35,25 & 65,25 & 8,25 \\ -24,75 & -43,75 & -2,75 \end{pmatrix}$ e

$\mathbf{B}^3 = \begin{pmatrix} 169,50 & 332,50 & 72,50 \\ -120,75 & -240,75 & -57,75 \\ 85,25 & 166,25 & 35,25 \end{pmatrix}$ resultando em S_1 = tr(**B**) = –6; S_2 = tr($\mathbf{B}^2$) = 14 e

S_3 = tr($\mathbf{B}^3$) = –36. E, recursivamente: $c_1 = -S_1 = 6$; $c_2 = -\frac{1}{2}\left[6\cdot(-6)+14\right] = 11$ e

$c_3 = -\frac{1}{3}\left[11\cdot(-6)+6\cdot 14-36\right] = 6$, logo: $p(\lambda) = \lambda^3 + 6\cdot\lambda^2 + 11\cdot\lambda + 6$, idêntico ao polinômio característico de **A**, desta forma os valores característicos mantém-se inalterados. Para determinar os correspondentes vetores característicos assim procede-se:

1º Vetor Característico: $\lambda_1 = -1$,

$\mathbf{B} - \lambda_1\cdot\mathbf{I} = \mathbf{B} + \mathbf{I} = \begin{pmatrix} 11,50 & 17,50 & -2,50 \\ -8,25 & -13,25 & 0,75 \\ 5,75 & 8,75 & -1,25 \end{pmatrix}$, cofatores da primeira linha: 10; –6 e 4,00, dividindo estes cofatores por 2, tem-se: $\mathbf{u_1} = \begin{pmatrix} 5 \\ -3 \\ 2 \end{pmatrix}$ logo: $\mathbf{P}\cdot\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \mathbf{v}_1$

2º Vetor Característico: $\lambda_2 = -2$,

$\mathbf{B} - \lambda_2\cdot\mathbf{I} = \mathbf{B} + 2\cdot\mathbf{I} = \begin{pmatrix} 12,50 & 17,50 & -2,50 \\ -8,25 & -12,25 & 0,75 \\ 5,75 & 8,75 & -0,25 \end{pmatrix}$, cofatores da primeira linha –3,5; 2,25 e –1,75, multiplicando estes cofatores por –4, tem-se: $\mathbf{u}_2 = \begin{pmatrix} 14 \\ -9 \\ 7 \end{pmatrix}$ logo: $\mathbf{P}\cdot\mathbf{u}_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \mathbf{v}_2$

3º Vetor Característico: $\lambda_3 = -3$,

$\mathbf{B} - \lambda_3\cdot\mathbf{I} = \mathbf{B} + 3\cdot\mathbf{I} = \begin{pmatrix} 13,50 & 17,50 & -2,50 \\ -8,25 & -11,25 & 0,75 \\ 5,75 & 8,75 & 0,75 \end{pmatrix}$, cofatores da primeira linha: –15; 10,5 e –7,5, multiplicando estes cofatores por –2/3, tem-se: $\mathbf{u}_3 = \begin{pmatrix} 10 \\ -7 \\ 5 \end{pmatrix}$ logo: $\mathbf{P}\cdot\mathbf{u}_3 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = \mathbf{v}_3$

1.3 Decomposição de Jordan

À toda matriz quadrada **A** associa-se a *transformação linear* **u** = **A**·**v**, onde $\mathbf{A} \in \Re^{n\times n}$, $\mathbf{u} \in \Re^n$ e $\mathbf{v} \in \Re^n$. Expressando os vetores **v** e **u** em uma base composta por *n* vetores linearmente independentes: $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$, tem-se: $\mathbf{v} = \mathbf{P}\cdot\hat{\mathbf{v}}$ e $\mathbf{u} = \mathbf{P}\cdot\hat{\mathbf{u}}$, onde: $\mathbf{P} = \left(\mathbf{p}_1\ \mathbf{p}_2 \cdots\ \mathbf{p}_n\right)$, $\hat{\mathbf{v}}$ e $\hat{\mathbf{u}}$ são os componentes de **u** e **v** na nova base, assim: $\mathbf{P}\cdot\hat{\mathbf{u}} = \mathbf{A}\cdot\mathbf{P}\cdot\hat{\mathbf{v}} \Rightarrow \hat{\mathbf{u}} = (\mathbf{P}^{-1}.\mathbf{A}\cdot\mathbf{P})\cdot\hat{\mathbf{v}}$, permitindo interpretar a matriz: $\hat{\mathbf{A}} = \mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}$ como os *componentes* de **A** na nova base $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$. Este procedimento consiste em uma *transformação de coordenadas* podendo assim ser resumido: um vetor $\mathbf{v} \in \Re^n$ tem como componentes em uma base formada pelos *n* vetores linearmente independentes $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$, os elementos de $\hat{\mathbf{v}}$ relacionados com os elementos de **v** através de $\mathbf{v} = \mathbf{P}\cdot\hat{\mathbf{v}}$ ou $\hat{\mathbf{v}} = \mathbf{P}^{-1}\cdot\mathbf{v}$, onde $\mathbf{P} = \left(\mathbf{p}_1\ \mathbf{p}_2 \cdots\ \mathbf{p}_n\right)$ e uma matriz $\mathbf{A} \in \Re^{n\times n}$ tem como componentes em uma base formada pelos *n* vetores linearmente independentes $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$, os elementos de $\hat{\mathbf{A}}$ relacionados com os elementos de **A** através de $\hat{\mathbf{A}} = \mathbf{P}^{-1}.\mathbf{A}\cdot\mathbf{P}$ ou $\mathbf{A} = \mathbf{P}.\hat{\mathbf{A}}\cdot\mathbf{P}^{-1}$.

Uma propriedade importante da transformação de coordenadas diz respeito a potências da matriz **A**, assim tem-se:

$$\hat{\mathbf{A}}^2 = \left(\mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}\right)\cdot\left(\mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P}\right) = \mathbf{P}^{-1}\cdot\mathbf{A}^2\cdot\mathbf{P}$$

$$\text{e}\quad \hat{\mathbf{A}}^3 = \hat{\mathbf{A}}\cdot\hat{\mathbf{A}}^2 = \left(\mathbf{P}^{-1}.\mathbf{A}\cdot\mathbf{P}\right)\cdot\left(\mathbf{P}^{-1}.\mathbf{A}^2\cdot\mathbf{P}\right) = \mathbf{P}^{-1}.\mathbf{A}^3\cdot\mathbf{P}\ ,$$

e assim sucessivamente, permitindo concluir que: $\hat{\mathbf{A}}^m = \mathbf{P}^{-1}.\mathbf{A}^m\cdot\mathbf{P}$ ou $\mathbf{A}^m = \mathbf{P}.\hat{\mathbf{A}}^m\cdot\mathbf{P}^{-1}$, para *m* = 0, 1, 2, ... Se a matriz **A** for regular tem-se:

$$\mathbf{A}^{-1} = (\mathbf{P}.\hat{\mathbf{A}}\cdot\mathbf{P}^{-1})^{-1} = (\mathbf{P}^{-1})^{-1}\cdot(\mathbf{P}.\hat{\mathbf{A}})^{-1}\mathbf{P}.\hat{\mathbf{A}}^{-1}\cdot\mathbf{P}^{-1} \quad \text{e} \quad \hat{\mathbf{A}}^{-1} = \mathbf{P}^{-1}.\mathbf{A}^{-1}\cdot\mathbf{P}\ ,$$

assim para **A** regular tem-se: $\hat{\mathbf{A}}^m = \mathbf{P}^{-1}.\mathbf{A}^m\cdot\mathbf{P}$ ou $\mathbf{A}^m = \mathbf{P}.\hat{\mathbf{A}}^m\cdot\mathbf{P}^{-1}$, para *m* = 0, ±1, ±2, ...

Esta mesma propriedade pode ser aplicada a funções polinomiais de **A** do tipo: $p(\mathbf{A}) = \mathbf{A}^m + a_1\mathbf{A}^{m-1} + a_2\mathbf{A}^{m-2} + \cdots + a_{m-1}\mathbf{A} + a_m\mathbf{I}$, pré-multiplicando esta expressão por $\mathbf{P}^{-1}$ e pós-multiplicando por **P**, tem-se:

$$\mathbf{P}^{-1}\cdot p(\mathbf{A})\cdot\mathbf{P} = \mathbf{P}^{-1}\cdot\left(\mathbf{A}^m + a_1\mathbf{A}^{m-1} + a_2\mathbf{A}^{m-2} + \cdots + a_{m-1}\mathbf{A} + a_m\mathbf{I}\right)\cdot\mathbf{P} = \mathbf{P}^{-1}\cdot\mathbf{A}^m\cdot\mathbf{P} +$$

$$+ a_1\mathbf{P}^{-1}\cdot\mathbf{A}^{m-1}\cdot\mathbf{P} + a_2\mathbf{P}^{-1}\cdot\mathbf{A}^{m-2}\cdot\mathbf{P} + \cdots + a_{m-1}\mathbf{P}^{-1}\cdot\mathbf{A}\cdot\mathbf{P} + a_m\mathbf{I} = \hat{\mathbf{A}}^m + a_1\hat{\mathbf{A}}^{m-1} +$$

$$+ a_2\hat{\mathbf{A}}^{m-2} + \cdots + a_{m-1}\hat{\mathbf{A}} + a_m\mathbf{I} = p\left(\hat{\mathbf{A}}\right) \Rightarrow \mathbf{P}^{-1}\cdot p(\mathbf{A})\cdot\mathbf{P} = p\left(\hat{\mathbf{A}}\right) \ \text{e}\ \mathbf{P}\cdot p(\hat{\mathbf{A}})\cdot\mathbf{P}^{-1} = p\left(\mathbf{A}\right)$$

A questão que se propõe agora é como selecionar a base composta por $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$ de tal forma que a matriz **A** assuma sua forma mais simples (*forma canônica*), duas situações se apresentam:

i) A matriz **A** apresenta *n* valores característicos distintos e, em conseqüência, *n* vetores característicos linearmente independentes, sejam estes vetores: $\mathbf{p}_1,\ \mathbf{p}_2,\cdots,\ \mathbf{p}_n$, como $\mathbf{A}\cdot\mathbf{p}_i = \lambda_i\cdot\mathbf{p}_i$ para *i* = 1, $\cdots,n$ tem-se: $\mathbf{A}\cdot\mathbf{P} = \left(\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n\right)$, mas:

$$\lambda_1\mathbf{p}_1 = \mathbf{P}\cdot\begin{pmatrix}\lambda_1\\0\\\vdots\\0\end{pmatrix};\lambda_2\mathbf{p}_2 = \mathbf{P}\cdot\begin{pmatrix}0\\\lambda_2\\\vdots\\0\end{pmatrix};\cdots;\lambda_n\mathbf{p}_n = \mathbf{P}\cdot\begin{pmatrix}0\\0\\\vdots\\\lambda_n\end{pmatrix} \Rightarrow \mathbf{A}\cdot\mathbf{P} = \left(\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n\right) = \mathbf{P}\cdot\mathbf{D}\text{, onde}$$

$$\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \text{ isto é:}$$

se **A** apresenta n valores característicos distintos então $\mathbf{P^{-1}} \cdot \mathbf{A} \cdot \mathbf{P} = \mathbf{D}$ ou $\mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} = \mathbf{A}$ sendo **P** uma matriz $n \times n$ cujos vetores colunas são os vetores característicos $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n$ de **A** correspondentes aos valores característicos λ_1, λ_2, ..., λ_n da matriz **D**, que é a forma mais simples que a matriz **A** pode assumir (no caso, a *forma diagonal* e o procedimento é chamado de *diagonalização*).

Exemplo: No exemplo anterior tinha-se: $\mathbf{A} = \begin{pmatrix} -2,50 & -1,00 & -1,25 \\ -0,50 & -4,00 & -1,75 \\ 1,00 & 2,00 & 0,50 \end{pmatrix}$, e a matriz dos vetores característicos $\mathbf{P} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 3 \\ -2 & -2 & -2 \end{pmatrix}$

Invertendo a matriz **P**, tem-se: $\mathbf{P}^{-1} = \begin{pmatrix} -0,50 & -0,50 & -1,00 \\ 0,50 & 0,00 & 0,25 \\ 0,00 & 0,50 & 0,25 \end{pmatrix}$, assim:

$$\mathbf{P^{-1}} \cdot \mathbf{A} \cdot \mathbf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \mathbf{D} \text{ e } \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} = \begin{pmatrix} -2,50 & -1,00 & -1,25 \\ -0,50 & -4,00 & -1,75 \\ 1,00 & 2,00 & 0,50 \end{pmatrix} = \mathbf{A}$$

ii) A matriz **A** apresenta valores característicos múltiplos e, neste caso, não é possível determinar n vetores característicos linearmente independentes. Seja o primeiro valor característico λ_1 um valor característico de multiplicidade m sendo os $(n-m)$ restantes distintos entre si e diferentes de λ_1. Ao valor característico λ_1 associa-se um vetor característico $\mathbf{p}_1$ tal que $\mathbf{A} \cdot \mathbf{p_1} = \lambda_1 \cdot \mathbf{p_1}$ e aos demais λ_k os vetores característicos $\mathbf{p}_k$ que: satisfazem a: $\mathbf{A} \cdot \mathbf{p}_k = \lambda_k \cdot \mathbf{p}_k$ para $k = m+1, m+2, ...,n$. Os vetores característicos $\mathbf{p}_1$, $\mathbf{p}_{m+1}$, $\mathbf{p}_{m+2}$, ..., $\mathbf{p}_n$ constituem a primeira coluna e as colunas $m+1$, $m+2$, ..., n da matriz **P**, para determinar as demais colunas desta matriz (colunas: 2, 3, ...,m) assim procede-se:

$$\mathbf{A} \cdot \mathbf{p}_j = \lambda_1 \cdot \mathbf{p}_j + \mathbf{p_{j-1}} \text{ para } j = 2, 3, ..., m.$$

Deste modo tem-se:

$$\mathbf{A} \cdot \mathbf{P} = \begin{pmatrix} \lambda_1\mathbf{p_1} & \lambda_1\mathbf{p}_2 + \mathbf{p_1} & \lambda_1\mathbf{p}_3 + \mathbf{p}_2 & \cdots & \lambda_1\mathbf{p}_m + \mathbf{p}_{m-1} & \lambda_{m+1}\mathbf{p}_{m+1} & \cdots & \lambda_n\mathbf{p}_n \end{pmatrix}, \text{ mas:}$$

$$\lambda_1 \mathbf{p_1} = \mathbf{P}\begin{pmatrix}\lambda_1\\0\\0\\\vdots\\0\\0\\\vdots\\0\end{pmatrix};\lambda_1\mathbf{p}_2+\mathbf{p}_1=\mathbf{P}\begin{pmatrix}1\\\lambda_1\\0\\\vdots\\0\\0\\\vdots\\0\end{pmatrix};\lambda_1\mathbf{p}_3+\mathbf{p}_2=\mathbf{P}\begin{pmatrix}0\\1\\\lambda_1\\\vdots\\0\\0\\\vdots\\0\end{pmatrix};\cdots;\lambda_1\mathbf{p}_m+\mathbf{p}_{m-1}=\mathbf{P}\begin{pmatrix}0\\0\\\vdots\\\vdots\\1\\\lambda_1\\\vdots\\0\end{pmatrix};\cdots;\lambda_n\mathbf{p}_n=\mathbf{P}\cdot\begin{pmatrix}0\\0\\\vdots\\\lambda_n\end{pmatrix}\Rightarrow$$

$$\mathbf{A}\cdot\mathbf{P}=\mathbf{P}\cdot\begin{pmatrix}\lambda_1 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ 0 & \lambda_1 & 1 & \cdots & 0 & 0 & \cdots & 0\\ 0 & 0 & \lambda_1 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0\\ 0 & 0 & 0 & \cdots & \lambda_1 & 0 & \cdots & 0\\ 0 & 0 & 0 & \cdots & 0 & \lambda_{m+1} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & \lambda_n\end{pmatrix}=\mathbf{P}\cdot\mathbf{J}$$

, onde $\mathbf{J}$ é forma mais simples da matriz **A** chamada de *forma canônica de Jordan* de **A**.

Exemplo: com: $\mathbf{A}=\begin{pmatrix}0 & 1 & -1\\ 0 & 2 & 1\\ 4 & -2 & 5\end{pmatrix}$, tem-se:

$p(\lambda)=\det\left[\begin{pmatrix}\lambda & -1 & 1\\ 0 & \lambda-2 & -1\\ -4 & 2 & \lambda-5\end{pmatrix}\right]=(\lambda-2)^2(\lambda-3)$ assim os valores característicos são $\lambda_1=\lambda_2=2$ e $\lambda_3=3$

1º Vetor Característico: $\lambda_1 = 2$,

$\mathbf{A}-\lambda_1\cdot\mathbf{I}=\mathbf{A}-2\mathbf{I}=\begin{pmatrix}-2 & 1 & -1\\ 0 & 0 & 1\\ 4 & -2 & 3\end{pmatrix}$, cofatores da primeira linha: 2; 4 e 0,00, dividindo estes cofatores por 2, tem-se: $\mathbf{p_1}=\begin{pmatrix}1\\2\\0\end{pmatrix}$

e $(\mathbf{A}-\lambda_1\cdot\mathbf{I})\mathbf{p_2}=\mathbf{p}_1=\begin{pmatrix}-2 & 1 & -1\\ 0 & 0 & 1\\ 4 & -2 & 3\end{pmatrix}\cdot\begin{pmatrix}p_{21}\\p_{22}\\p_{23}\end{pmatrix}=\begin{pmatrix}-2p_{21}+p_{22}-p_{23}\\ p_{23}\\ 4p_{21}-2p_{22}+3p_{23}\end{pmatrix}=\mathbf{p_1}=\begin{pmatrix}1\\2\\0\end{pmatrix}$

logo: $p_{23} = 2$ e $\begin{cases} -2p_{21} + p_{22} = 1 + p_{23} = 3 \\ 4p_{21} - 2p_{22} = -3p_{23} = -6 \end{cases}$ em que uma das soluções é: $p_{21} = 0$ e $p_{22} = 3$, logo:

$$\mathbf{p}_2 = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$$

<u>2º Vetor Característico</u>: $\lambda_3 = 3$,

$\mathbf{A} - \lambda_3 \cdot \mathbf{I} = \mathbf{A} - 3\mathbf{I} = \begin{pmatrix} -3 & 1 & -1 \\ 0 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix}$, cofatores da primeira linha: 0; 4 e 4, dividindo estes cofatores por 4, tem-se: $\mathbf{p}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, assim $\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 1 \\ 0 & 2 & 1 \end{pmatrix}$

Invertendo a matriz **P**, tem-se: $\mathbf{P}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}$, assim:

$$\mathbf{P}^{-1} \cdot \mathbf{A} \cdot \mathbf{P} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{J} \quad \text{e} \quad \mathbf{P} \cdot \mathbf{J} \cdot \mathbf{P}^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 2 & 1 \\ 4 & -2 & 5 \end{pmatrix} = \mathbf{A}$$

1.4 Decomposição em valores e vetores singulares (SVD)

Dada uma matriz $\mathbf{A} \in \Re^{m \times n}$ e lembrando dos conceitos básicos que o range(**A**) é o espaço gerado pelos vetores colunas de **A** e range($\mathbf{A}^T$) é o espaço gerado pelos vetores linhas de **A**, a decomposição SVD é capaz de obter simultaneamente as bases ortonormais destes subespaços.

Qualquer matriz $\mathbf{A} \in \Re^{m \times n}$ pode ser decomposta na forma:

$$\mathbf{A} = \mathbf{U} \cdot \Sigma \cdot \mathbf{V}^T$$

onde $\mathbf{U} \in \Re^{m \times m}$ é uma matriz ortogonal cujas colunas são os vetores característicos de $\mathbf{A} \cdot \mathbf{A}^T$, $\mathbf{V} \in \Re^{n \times n}$ é uma matriz ortogonal cujas colunas são os vetores característicos de $\mathbf{A}^T \cdot \mathbf{A}$, $\Sigma \in \Re^{m \times n}$ é uma matriz diagonal contento a raiz quadrada dos valores característicos de $\mathbf{A} \cdot \mathbf{A}^T$ (que são equivalentes aos valores característicos de $\mathbf{A}^T \cdot \mathbf{A}$), arranjados em ordem decrescente. Os vetores característicos de $\mathbf{A} \cdot \mathbf{A}^T$ e $\mathbf{A}^T \cdot \mathbf{A}$ estão arranjados nas colunas de **U** e **V**, respectivamente, na ordem de seus valores característicos na matriz Σ. Os elementos, σ_i, da diagonal de Σ são denominados de valores singulares de **A**, sendo todos não-negativos. Além disso, o número de valores singulares positivos é igual ao rank(**A**). Os vetores colunas de **U** são denominados de vetores singulares à esquerda de **A** e os vetores colunas de **V** são denominados de vetores singulares à direita de **A**, e as relações entre estes vetores são:

$$\mathbf{A} \cdot \mathbf{v}_i = \sigma_i\, \mathbf{u}_i \quad \text{e} \quad \mathbf{A}^T \cdot \mathbf{u}_i = \sigma_i\, \mathbf{v}_i$$

A decomposição SVD revela várias propriedades intrínsecas da matriz **A** e é numericamente estável para os cálculos. Algumas propriedades são listadas abaixo para uma matriz com r valores singulares positivos:

1) rank($\mathbf{A}$) = r

2) null($\mathbf{A}$) = span($\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, ..., \mathbf{v}_n$)

3) range($\mathbf{A}$) = span($\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r$)

4) range($\mathbf{A}^T$) = span($\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$)

5) $\mathbf{A} = \sum_{i=1}^{r} \sigma_i\, \mathbf{u}_i \cdot \mathbf{v}_i^T$

6) $\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$ (norma de Frobenius)

7) $\|\mathbf{A}\|_2 = \sigma_1$

A matriz:

$$\mathbf{A}^\dagger = \mathbf{V} \cdot \Sigma^\dagger \cdot \mathbf{U}^T = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T$$

é chamada de pseudo-inversa de **A**, onde os elementos da diagonal de $\Sigma^\dagger$ consistem no recíproco dos valores singulares positivos de Σ, na mesma ordem. A pseudo-inversa tem a propriedade $\mathbf{A}^\dagger \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^\dagger = \mathbf{I}$.

A solução do problema de valores singulares para o sistema linear, $\mathbf{y} = \mathbf{A}\cdot\mathbf{x}$, corresponde resolver o seguinte problema de otimização:

$$\max_{\mathbf{x}} \|\mathbf{y}\|_2^2$$

$$\text{sujeito a } \|\mathbf{x}\|_2^2 = 1$$

onde $\|\mathbf{y}\|_2^2 = \mathbf{y}^T\cdot\mathbf{y}$. Usando o conceito dos multiplicadores de Lagrange, o problema acima pode ser reescrito como:

$$\max_{\mathbf{x}} \left\{ S(\mathbf{x}) = \mathbf{x}^T \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} - \lambda\, (\mathbf{x}^T \cdot \mathbf{x} - 1) \right\}$$

cuja primeira condição de otimalidade é $\nabla S(\mathbf{x}) = \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} - \lambda\, \mathbf{x} = 0$, ou seja, a solução é equivalente ao problema de valor característico $\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \lambda\, \mathbf{x}$, cujos **x** ótimos locais correspondem aos vetores característicos de $\mathbf{A}^T \cdot \mathbf{A}$ ou os vetores singulares de **A** (também chamados de **componentes principais** de variação, pois indicam as direções de máxima variação de **y** em função das variações em **x** com mesma energia, isto é, $\|\mathbf{x}\|_2 = 1$) e os respectivos multiplicadores de Lagrange são os valores característicos de $\mathbf{A}^T \cdot \mathbf{A}$ ou o quadrado dos valores singulares de **A**.

Observe que para uma matriz de rank($\mathbf{A}$) = r, $\mathbf{A} = \mathbf{U} \cdot \Sigma \cdot \mathbf{V}^T = \sum_{i=1}^{r} \sigma_i\, \mathbf{u}_i \cdot \mathbf{v}_i^T$ e, portanto, $\mathbf{y} = \mathbf{A}\cdot\mathbf{x} = \mathbf{U} \cdot \Sigma \cdot \mathbf{V}^T \cdot \mathbf{x} = \sum_{i=1}^{r} \sigma_i\, \mathbf{u}_i \cdot \mathbf{v}_i^T \cdot \mathbf{x}$, indicando que a projeção do vetor **x** na direção do vetor $\mathbf{v}_i$ (ou seja, $\mathbf{v}_i^T \cdot \mathbf{x}$) é amplificada por σ_i na direção $\mathbf{u}_i$ do vetor **y**, sendo $i = 1$ a direção de maior amplificação e $i = r$ a direção de menor amplificação. Dependendo do valor de σ_i, uma pequena mudança em **x** pode causar uma grande mudança em **y**, mas isto vai depender do ângulo entre os vetores **x** e $\mathbf{v}_i$.

<u>Exemplo</u>: usando a função demonstrativa **eigshow(A)** do MATLAB, que apresenta de forma gráfica um vetor unitário **x** e sua transformação linear **A·x** para uma matriz de dimensão 2 × 2 e valores de **x**, com $\|\mathbf{x}\| = 1$, é possível pela movimentação do mouse sobre a figura localizar as direções características da matriz. As Figuras 2a, 2b e 2c ilustram esta movimentação para a matriz

$$\boldsymbol{A} = \begin{bmatrix} 1/4 & 3/4 \\ 1 & 1/2 \end{bmatrix}$$

onde os valores característicos, λ, e os vetores característicos, **x**, resultantes das soluções não-triviais do sistema de equações lineares

$$\mathbf{A}\cdot\mathbf{x} = \lambda\,\mathbf{x}$$

$$\lambda_1 = 5/4 \quad \text{e} \quad \mathbf{x}_1 = [3/5 \quad 4/5]^T \qquad ; \qquad \lambda_2 = -1/2 \quad \text{e} \quad \mathbf{x}_2 = [\sqrt{2}/2 \quad -\sqrt{2}/2\,]^T$$

são visualizados quando os dois vetores (**x** e **A·x**) estão na mesma direção. Mostrando que o operador **A**, na direção de **x**, corresponde a uma redução ou ampliação por um fator λ. Quando os sentidos dos dois vetores são opostos tem-se um valor característico negativo.

Pode-se observar que os dois vetores característicos não são os eixos maior e menor da elipse formada pelas transformações lineares. Seriam para o caso particular de matrizes simétricas. Matrizes 2 × 2 com um par de valores característicos complexos não possuem vetores característicos reais.

Movendo dois vetores unitários, **x** e **y**, perpendiculares e suas correspondentes transformações lineares, **A·x** e **A·y**, pode-se visualizar os valores e vetores singulares, resultantes das soluções não-triviais dos sistemas de equações lineares

$$\mathbf{A}^T\cdot\mathbf{A}\cdot\mathbf{v} = \sigma^2\,\mathbf{v} \qquad\qquad \mathbf{A}\cdot\mathbf{A}^T\cdot\mathbf{u} = \sigma^2\,\mathbf{u}$$

$$\sigma_1 = 1{,}2792, \quad \mathbf{x} = \mathbf{v}_1 = [0{,}7678 \quad 0{,}6407]^T \quad \text{e} \quad \mathbf{A}\cdot\mathbf{x} = \sigma_1\,\mathbf{u}_1 = [0{,}6725 \quad 1{,}0881]^T$$

$$\sigma_2 = 0{,}4886, \quad \mathbf{y} = \mathbf{v}_2 = [-0{,}6407 \quad 0{,}7678]^T \quad \text{e} \quad \mathbf{A}\cdot\mathbf{y} = \sigma_2\,\mathbf{u}_2 = [0{,}4156 \quad -0{,}2569]^T$$

onde σ são os valores singulares, **v** e **u** são os vetores singulares à direita e à esquerda, respectivamente. Eles surgem no momento em que as transformações são perpendiculares entre si, conforme mostra a Figura 2d. Observa-se que isto acontece quando os vetores das transformações são os eixos maior e menor da elipse, mostrando, por exemplo, as direções de máxima e mínima amplificação de sinais, respectivamente. Para o caso particular de uma matriz quadrada, simétrica e positiva definida, as decomposições em valores característicos e em valores singulares são equivalentes.

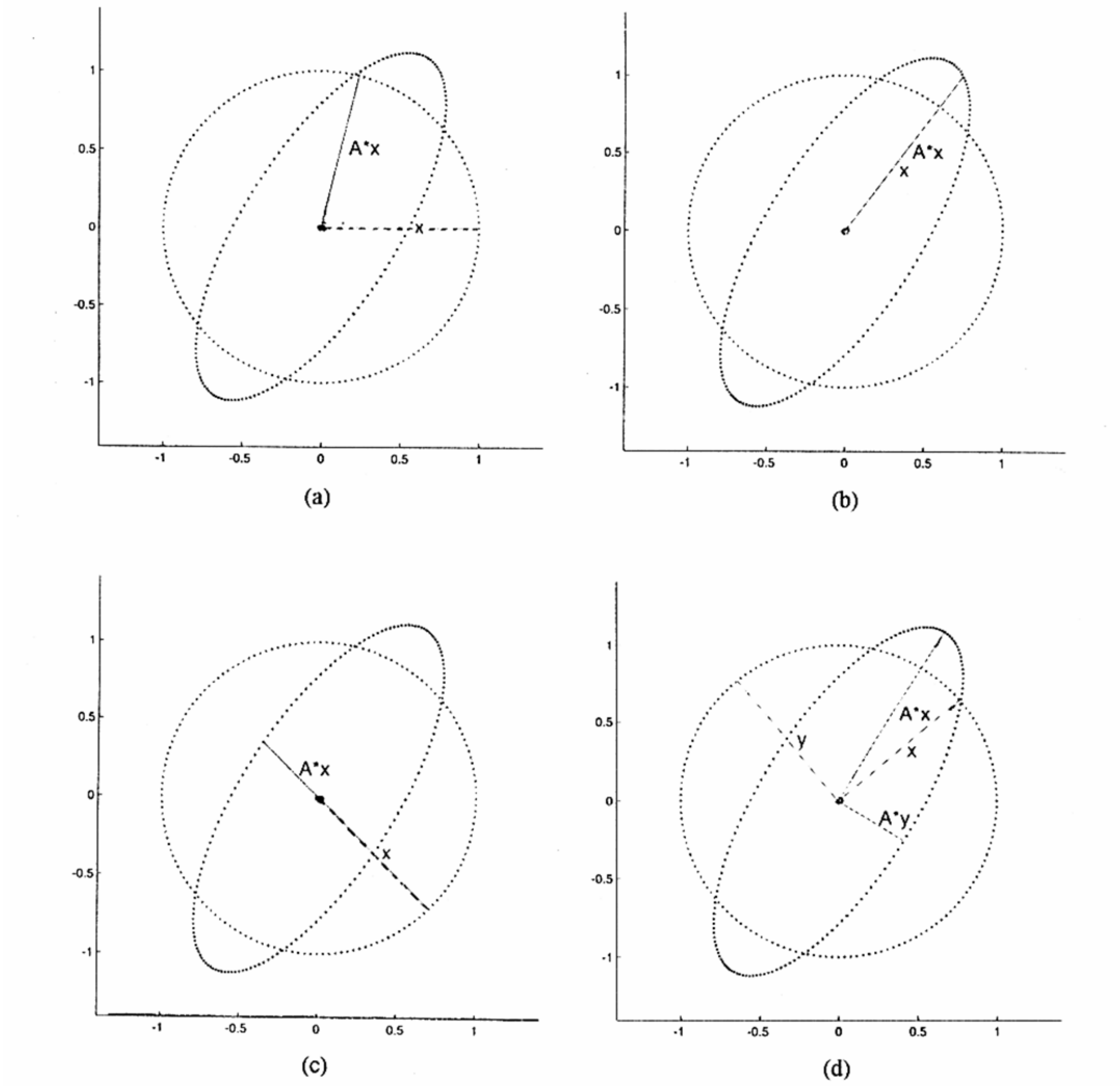


Figura 2: Direções características de uma matriz.

Para determinar se o sistema linear, $\mathbf{y} = \mathbf{A}\cdot\mathbf{x}$, está bem escalonado deve-se verificar o condicionamento da matriz **A**, que na norma 2 é dado por:

$$\gamma(\mathbf{A}) = \overline{\sigma}(\mathbf{A}) / \underline{\sigma}(\mathbf{A})$$

onde $\overline{\sigma}(\mathbf{A})$ é o maior valor singular de **A** e $\underline{\sigma}(\mathbf{A})$ é o menor valor singular não-nulo de **A**.

A melhor maneira de escalonar um sistema é atacando a origem do problema, ou seja um apropriado adimensionamento das variáveis dependentes e das equações do problema. Uma maneira numérica de determinar quais as variáveis devem ser re-escalonadas é através do cálculo do condicionamento mínimo, isto é, determinar as matrizes que pré- e pós-multiplicadas pela matriz **A** resultam em um γ mínimo (γ*), isto é:

$$\gamma^*(\mathbf{A}) = \min_{\mathbf{L},\mathbf{R}} \gamma(\mathbf{L}\cdot\mathbf{A}\cdot\mathbf{R})$$

Considerando **L** e **R** matrizes diagonais, então tem-se como resultado do problema de otimização acima quais as saídas e entradas devem ser re-escalonadas, respectivamente, pois:

$$\mathbf{y}_e = \mathbf{L} \cdot \mathbf{y} \qquad \text{e} \qquad \mathbf{x}_e = \mathbf{R}^{-1} \cdot \mathbf{x}$$

No MATLAB, o condicionamento de uma matriz **A** pode ser obtido através da função *cond*(**A**), e as matrizes de escalonamento podem ser calculadas pela função:

[*ub*, **D**] = *mu* (**H**, *ones* (2**n*, 2), 'C')

onde n é a ordem da matriz **A**, $\mathbf{H} = \begin{bmatrix} 0 & \mathbf{A}^{-1} \\ \mathbf{A} & 0 \end{bmatrix}$, $\gamma^*(\mathbf{A}) = \left(\dfrac{ub(1)}{ub(2)}\right)^2$,

L = *diag*(**D**(n+1:2*n)) e **R** = *inv*(*diag*(**D**(1:n))). Para verificar se os cálculos estão corretos, basta calcular o *cond*(**L*****A*****R**), que deve ser igual a $\gamma^*(\mathbf{A})$.

1.5 Fatoração LU

Eliminação Gaussiana: O propósito da eliminação Gaussiana é reduzir a matriz **A** a uma estrutura triangular (métodos de triangularização) ou diagonal (método de Gauss-Jordan) através de operações da álgebra elementar. Dentre os diversos algoritmos de eliminação Gaussiana temos os seguintes:

$$\left.\begin{array}{lll} a_{kj} \leftarrow \dfrac{a_{kj}}{a_{kk}} & , & j = k, k+1, \ldots, n \\ a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj} & , & \begin{array}{l} j = k, \ldots, n \\ i = 1, \ldots, n\, (\neq k) \end{array} \end{array}\right\} k = 1, \ldots, n \qquad \text{(Gauss-Jordan)}$$

$$\begin{array}{l} k = 1, \ldots, n-1 \\ j = k+1, \ldots, n+1 \\ i = k+1, \ldots, n \end{array} \left\{ \begin{array}{l} a_{kj} \leftarrow \dfrac{a_{kj}}{a_{kk}} \\ \\ a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj} \end{array} \right. \qquad \text{(triangularização SAXPY)}$$

onde a_{ij} são os elementos da matriz aumentada: $\tilde{\mathbf{A}}$ = [**A** **b**]. No caso do método de Gauss-Jordan, a solução é encontrada na (n+1)-ésima coluna da matriz aumentada, após as operações de eliminação Gaussiana. Nos métodos de triangularização é necessário ainda realizar operações de substituição (para matriz triangular inferior) ou retro-substituição (para matriz triangular superior), isto é,

$$x_1 = \frac{a_{1,n+1}}{a_{1,1}}, \quad x_i = \frac{1}{a_{i,i}}\left(a_{i,n+1} - \sum_{j=1}^{i-1} a_{i,j}\, x_j\right), \quad i = 2,\ldots,n \qquad \text{substituição}$$

$$x_n = \frac{a_{n,n+1}}{a_{n,n}}, \quad x_i = \frac{1}{a_{i,i}}\left(a_{i,n+1} - \sum_{j=i+1}^{n} a_{i,j}\, x_j\right), \quad i = n-1,\ldots,1 \qquad \text{retro-substituição}$$

De modo a evitar prováveis divisões por zero (dos elementos a_{kk}) e também garantir a estabilidade numérica do algoritmo (devido a problemas de arredondamento), faz-se necessário o uso de técnicas de **pivotamento**. Pivotamentos são operações de trocas de linhas e/ou colunas de modo a obter uma matriz tendo na diagonal elementos com maior valor absoluto. Quando são efetuadas somente trocas de linhas, diz-se um **pivotamento parcial**. No **pivotamento total** tem-se trocas de linhas e colunas. As operações de pivotamento podem ser representadas por matrizes de permutações **P** e **Q**:

$$\mathbf{P\,A\,x} = \mathbf{P\,B} \qquad \text{(pivotamento parcial)}$$

$$\mathbf{P\,A\,Q\,Q^{-1}\,x} = \mathbf{P\,B} \qquad \text{(pivotamento total)}$$

Exemplo: considere o sistema de equações algébricas lineares:

$$\begin{cases} 2\cdot x_1 - 7\cdot x_2 + 4\cdot x_3 = 9 \\ x_1 + 9\cdot x_2 - 6\cdot x_3 = 1 \\ -3\cdot x_1 + 8\cdot x_2 + 5\cdot x_3 = 6 \end{cases}$$

permitindo identificar: $\mathbf{A} = \begin{bmatrix} 2 & -7 & 4 \\ 1 & 9 & -6 \\ -3 & 8 & 5 \end{bmatrix}$ e $\mathbf{b} = \begin{bmatrix} 9 \\ 1 \\ 6 \end{bmatrix}$, assim a *matriz aumentada* é:

2	-7	4	9
1	9	-6	1
-3	8	5	6

1) Método de Eliminação por Triangularização da Matriz Aumentada

1ª Etapa) Reposicionamento das linhas (pivotamento parcial) de modo que a primeira linha contenha o maior elemento (em módulo) da primeira coluna (**pivô**):

2	-7	4	9
1	9	-6	1
-3	8	5	6

-3	8	5	6
1	9	-6	1
2	-7	4	9

2ª Etapa) Normalização dos elementos da primeira linha, dividindo-os pelo 1º elemento da mesma: $a_{kj} \leftarrow \dfrac{a_{kj}}{a_{kk}}$ ($k = 1$ e $j = k, ..., n+1$)

1	-8/3	-5/3	-2
1	9	-6	1
2	-7	4	9

3ª Etapa) Eliminação dos elementos da primeira coluna da segunda e terceira linhas: $a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj}$ ($k = 1$, $j = k, ..., n+1$ e $i = k+1, ..., n$)

1	-8/3	-5/3	-2
0	35/3	-13/3	3
2	-7	4	9

1	-8/3	-5/3	-2
0	35/3	-13/3	3
0	-5/3	22/3	13

Repete-se o procedimento para próximas linhas até a triangularização da matriz aumentada:

4ª Etapa) Reposicionamento das linhas de modo que a segunda linha contenha o maior elemento (em módulo) da segunda coluna [sem levar em consideração a primeira linha]:

1	-8/3	-5/3	-2
0	35/3	-13/3	3
0	-5/3	22/3	13

Não há necessidade do reposicionamento, pois o maior elemento (em módulo) da segunda linha, desconsiderando-se a primeira linha, já se encontra na segunda linha (pivô).

5ª Etapa) Normalização dos elementos da segunda linha, dividindo-os pelo 2º elemento da mesma:

1	-8/3	-5/3	-2
0	1	-13/35	9/35
0	-5/3	22/3	13

6ª Etapa) Eliminação dos elementos da segunda coluna da terceira linha:

1	-8/3	-5/3	-2
0	1	-13/35	9/35
0	0	141/21	94/7

7ª Etapa) Normalização dos elementos da terceira linha, dividindo-os pelo 3º elemento da mesma:

1	-8/3	-5/3	-2
0	1	-13/35	9/35
0	0	1	2

8ª Etapa) Determinação recursiva de x_1, x_2 e x_3 , iniciando com x_3. Esta última forma da matriz aumentada traduz o sistema linear:

$$x_n = \frac{a_{n,n+1}}{a_{n,n}}, \quad x_i = \frac{1}{a_{i,i}}\left(a_{i,n+1} - \sum_{j=i+1}^{n} a_{i,j}\, x_j\right), \; i = n\text{–}1,...,1 \;\; \text{(retro-substituição)}$$

$$\begin{cases} x_1 - \frac{8}{3}\cdot x_2 - \frac{5}{3}\cdot x_3 = -2 \\ \qquad x_2 - \frac{13}{35}\cdot x_3 = \frac{9}{35} \\ \qquad\qquad x_3 = 2 \end{cases} \Rightarrow \begin{cases} x_3 = 2 \\ x_2 = \frac{9}{35} + \frac{13}{35}\cdot x_3 \\ x_1 = -2 + \frac{8}{3}\cdot x_2 + \frac{5}{3}\cdot x_3 \end{cases}$$

assim:

$$\begin{cases} x_3 = 2 \\ x_2 = \frac{9}{35} + \frac{13}{35}\cdot 2 = \frac{35}{35} = 1 \\ x_1 = -2 + \frac{8}{3}\cdot 1 + \frac{5}{3}\cdot 2 = \frac{18}{3} - 2 = 6 - 2 = 4 \end{cases}$$

deste modo a solução é: $\boxed{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}}$

Nota: no exemplo acima a diagonal da matriz também foi dividida pelos pivôs durante a etapa de eliminação Gaussiana e, por isso, na etapa de retro-substituição não houve a necessidade da divisão pelos elementos da diagonal da matriz, pois estes eram unitários. O algoritmo da triangularização SAXPY não realiza esta divisão, deixando-a para a etapa de retro-substituição.

2) Método de Eliminação por Diagonalização da Matriz Aumentada

1ª Etapa) Reposicionamento das linhas (pivotamento parcial) de modo que a primeira linha contenha o maior elemento (em módulo) da primeira coluna (pivô):

2	-7	4	9
1	9	-6	1
-3	8	5	6

-3	8	5	6
1	9	-6	1
2	-7	4	9

2ª Etapa) Normalização dos elementos da primeira linha, dividindo-os pelo 1º elemento da linha: $a_{kj} \leftarrow \dfrac{a_{kj}}{a_{kk}}$ ($k = 1$ e $j = k, \ldots, n+1$)

1	-8/3	-5/3	-2
1	9	-6	1
2	-7	4	9

3ª Etapa) Eliminação dos elementos da primeira coluna da segunda e terceira linhas: $a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj}$ ($k = 1$, $j = k, \ldots, n+1$ e $i = 1, \ldots, n$ com $i \neq k$)

1	-8/3	-5/3	-2
0	35/3	-13/3	3
2	-7	4	9

1	-8/3	-5/3	-2
0	35/3	-13/3	3
0	-5/3	22/3	13

4ª Etapa) Reposicionamento das linhas de modo que a segunda linha contenha o maior elemento (em módulo) da segunda coluna [sem levar em consideração a primeira linha]:

1	-8/3	-5/3	-2
0	35/3	-13/3	3
0	-5/3	22/3	13

Não há necessidade do reposicionamento, pois o maior elemento (em módulo) da segunda linha, desconsiderando-se a primeira linha, já se encontra na segunda linha.

5ª Etapa) Normalização dos elementos da segunda linha, dividindo-os pelo 2º elemento da mesma:

1	-8/3	-5/3	-2
0	1	-13/35	9/35
0	-5/3	22/3	13

6ª Etapa) Eliminação dos elementos da segunda coluna da primeira e da terceira linhas:

1	0	-279/105	-46/35
0	1	-13/35	9/35
0	0	141/21	94/7

7ª Etapa) Normalização dos elementos da terceira linha, dividindo-os pelo 3º elemento da mesma:

1	0	-279/105	-46/35
0	1	-13/35	9/35
0	0	1	2

8ª Etapa) Eliminação dos elementos da terceira coluna da primeira e da segunda linhas:

1	0	0	4
0	1	0	1
0	0	1	2

Esta última forma da matriz aumentada traduz o sistema linear:

$$\begin{cases} x_1 = 4 \\ x_2 = 1 \\ x_3 = 2 \end{cases}$$

deste modo a solução é: $\boxed{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}}$

<u>Exemplo</u>: método de eliminação de Gauss para obtenção da matriz inversa. Seja a mesma matriz do exemplo ilustrativo do exemplo anterior:

$$\mathbf{A} = \begin{bmatrix} 2 & -7 & 4 \\ 1 & 9 & -6 \\ -3 & 8 & 5 \end{bmatrix},$$

neste caso deseja-se determinar a matriz: $\mathbf{B} = \mathbf{A}^{-1}$ tal que: $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}$, onde $\mathbf{I}$ é a matriz identidade com as mesmas dimensões da matriz $\mathbf{A}$.

Neste caso a *matriz aumentada* é:

2	-7	4	1	0	0
1	9	-6	0	1	0
-3	8	5	0	0	1

Aplicando-se procedimento de eliminação análogo ao anterior (diagonalização), ou seja, o método de eliminação por diagonalização da matriz aumentada:

1ª Etapa) Reposicionamento das linhas de modo que a primeira linha contenha o maior elemento (em módulo) da primeira coluna:

2	-7	4	1	0	0
1	9	-6	0	1	0
-3	8	5	0	0	1

-3	8	5	0	0	1
1	9	-6	0	1	0
2	-7	4	1	0	0

2ª Etapa) Normalização dos elementos da primeira linha, dividindo-os pelo 1º elemento da linha:

1	-8/3	-5/3	0	0	-1/3
1	9	-6	0	1	0
2	-7	4	1	0	0

3ª Etapa) Eliminação dos elementos da primeira coluna da segunda e terceira linhas:

1	-8/3	-5/3	0	0	-1/3
0	35/3	-13/3	0	1	1/3
2	-7	4	1	0	0

1	-8/3	-5/3	0	0	-1/3
0	35/3	-13/3	0	1	1/3
0	-5/3	22/3	1	0	2/3

4ª Etapa) Reposicionamento das linhas de modo que a segunda linha contenha o maior elemento (em módulo) da segunda coluna [sem levar em consideração a primeira linha]:

1	-8/3	-5/3	0	0	-1/3
0	35/3	-13/3	0	1	1/3
0	-5/3	22/3	1	0	2/3

Não há necessidade do reposicionamento, pois o maior elemento (em módulo) da segunda linha, desconsiderando-se a primeira linha, já se encontra na segunda linha.

5ª Etapa) Normalização dos elementos da segunda linha, dividindo-os pelo 2º elemento da mesma:

1	-8/3	-5/3	0	0	-1/3
0	1	-13/35	0	3/35	1/35
0	-5/3	22/3	1	0	2/3

6ª Etapa) Eliminação dos elementos da segunda coluna da primeira e da terceira linhas:

1	0	-93/35	0	8/35	-9/35
0	1	-13/35	0	3/35	1/35
0	0	47/7	1	1/7	5/7

7ª Etapa) Normalização dos elementos da terceira linha, dividindo-os pelo 3º elemento da mesma:

1	0	-93/35	0	8/35	-9/35
0	1	-13/35	0	3/35	1/35
0	0	1	7/47	1/47	5/47

8ª Etapa) Eliminação dos elementos da terceira coluna da primeira e da segunda linhas:

1	0	0	93/235	67/235	6/235
0	1	0	13/235	22/235	16/235
0	0	1	7/47	1/47	5/47

As três últimas colunas desta última forma da matriz é a inversa da matriz original, isto é:

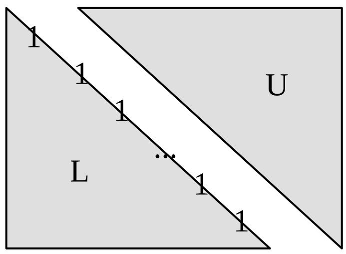
$$\mathbf{A}^{-1} = \begin{bmatrix} 93/235 & 67/235 & 6/235 \\ 13/235 & 22/235 & 16/235 \\ 7/47 & 1/47 & 5/47 \end{bmatrix}$$

para verificar se o valor da inversa é correto deve-se calcular: $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$

Fatoração LU: O processo de fatoração LU decompõe a matriz *A* em uma matriz triangular inferior, **L**, e outra triangular superior, **U**, com elementos unitários na diagonal principal da matriz *L* (método de Doolittle) ou da matriz **U** (método de Crout):

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$$

$$\begin{matrix} k = 1, \ldots, n-1 \\ i = k+1, \ldots, n \\ j = k+1, \ldots, n \end{matrix} \left\{ \begin{matrix} a_{ik} \leftarrow \dfrac{a_{ik}}{a_{kk}} \\ \\ a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj} \end{matrix} \right. \qquad \text{(Doolittle)}$$



com uma posterior substituição: $\mathbf{L} \cdot \mathbf{y} = \mathbf{b}$

e uma retro-substituição: $\mathbf{U} \cdot \mathbf{x} = \mathbf{y}$

As principais vantagens da fatoração em relação à eliminação Gaussiana é a redução do número de operações de $\frac{2}{3}n^3 + O(n^2)$ para $\frac{1}{3}n^3 + O(n^2)$, e a manutenção das operações básicas na matriz fatorada (matriz **L**, na fatoração LU), que pode ser aplicada para diferentes vetores **b**.

3) Método de Fatoração LU da Matriz Original

1ª Etapa) Reposicionamento das linhas (pivotamento parcial) de modo que a primeira linha contenha o maior elemento (em módulo) da primeira coluna (pivô). Somente nas etapas de pivotamento adiciona-se a matriz de pivotamento (inicialmente a matriz identidade) para armazenar as operações de trocas de linhas:

2	-7	4	1	0	0
1	9	-6	0	1	0
-3	8	5	0	0	1

-3	8	5	0	0	1
1	9	-6	0	1	0
2	-7	4	1	0	0

2ª Etapa) Normalização dos elementos da primeira coluna após a primeira linha, dividindo-os pelo 1º elemento da linha: $a_{ik} \leftarrow \dfrac{a_{ik}}{a_{kk}}$ $(k = 1 \text{ e } i = k+1, ..., n)$

-3	8	5
-1/3	9	-6
-2/3	-7	4

3ª Etapa) Fatoração dos elementos da segunda e terceira linhas após primeira coluna:

$a_{ij} \leftarrow a_{ij} - a_{ik}\, a_{kj}$ $(k = 1,\ \ i = k+1, ..., n \text{ e } j = k+1, ..., n)$

-3	8	5
-1/3	35/3	-13/3
-2/3	-7	4

-3	8	5
-1/3	35/3	-13/3
-2/3	-5/3	22/3

4ª Etapa) Reposicionamento das linhas de modo que a segunda linha contenha o maior elemento (em módulo) da segunda coluna [sem levar em consideração a primeira linha e a primeira coluna]:

-3	8	5	0	0	1
-1/3	35/3	-13/3	0	1	0
-2/3	-5/3	22/3	1	0	0

Não há necessidade do reposicionamento, pois o maior elemento (em módulo) da segunda linha, desconsiderando-se a primeira linha e primeira coluna, já se encontra na segunda linha.

5ª Etapa) Normalização dos elementos da segunda coluna após a segunda linha, dividindo-os pelo 2º elemento da mesma:

-3	8	5
-1/3	35/3	-13/3
-2/3	-1/7	22/3

6ª Etapa) Fatoração dos elementos da terceira linha após segunda coluna:

-3	8	5
-1/3	35/3	-13/3
-2/3	-1/7	47/7

7ª Etapa) Extraindo as matrizes **L** e **U**:

Matriz **L** (dos multiplicadores do processo de eliminação Gaussiana) com 1 na diagonal:

1	0	0
-1/3	1	0
-2/3	-1/7	1

Matriz **U**:

-3	8	5
0	35/3	-13/3
0	0	47/7

Para verificar se a fatoração está correta, o produto $\mathbf{P}^{-1} \cdot \mathbf{L} \cdot \mathbf{U}$ deve ser igual à matriz **A**.

8ª Etapa) Troca das linhas do vetor **b** de acordo com a matriz de permutação, isto é: $\mathbf{P} \cdot \mathbf{b}$.

9
1
6

matriz **P**:

0	0	1
0	1	0
1	0	0

6
1
9

9ª Etapa) Determinação recursiva de y_1, y_2 e y_3 , iniciando com y_1, do sistema $\mathbf{L} \cdot \mathbf{y} = \mathbf{b}$:

$$\begin{cases} y_1 = 6 \\ \frac{-1}{3}\cdot y_1 + y_2 = 1 \\ \frac{-2}{3}\cdot y_1 - \frac{5}{35}\cdot y_2 + y_3 = 9 \end{cases} \Rightarrow \begin{cases} y_1 = 6 \\ y_2 = 1 + \frac{1}{3}\cdot y_1 \\ y_3 = 9 + \frac{2}{3}\cdot y_1 + \frac{5}{35}\cdot y_2 \end{cases} \Rightarrow \begin{cases} y_1 = 6 \\ y_2 = 3 \\ y_3 = \frac{94}{7} \end{cases}$$

10ª Etapa) Determinação recursiva de x_1, x_2 e x_3 , iniciando com x_3, do sistema $\mathbf{U} \cdot \mathbf{x} = \mathbf{y}$:

$$\begin{cases} -3\cdot x_1 + 8\cdot x_2 + 5\cdot x_3 = 6 \\ \frac{35}{3}\cdot x_2 - \frac{13}{3}\cdot x_3 = 3 \\ \frac{47}{7}\cdot x_3 = \frac{94}{7} \end{cases} \Rightarrow \begin{cases} x_3 = \frac{94}{47} \\ x_2 = \frac{3}{35}\cdot\left(3 + \frac{13}{3}\cdot x_3\right) \\ x_3 = \frac{-1}{3}\cdot\left(6 - 5\cdot x_3 - 8\cdot x_2\right) \end{cases} \Rightarrow \begin{cases} x_1 = 4 \\ x_2 = 1 \\ x_3 = 2 \end{cases}$$

deste modo a solução é: $\boxed{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}}$

Caso desejássemos resolver outro sistema somente modificando o vetor **b**, bastaria repetir os passos 8 a 10, pois a matriz **A** já está fatorada. Do mesmo modo, para obter a inversa da matriz **A**, basta repetir estes três passos para os três vetores coluna da matriz identidade.

ANÁLISE DA SOLUÇÃO DE SISTEMAS ALGÉBRICOS LINEARES

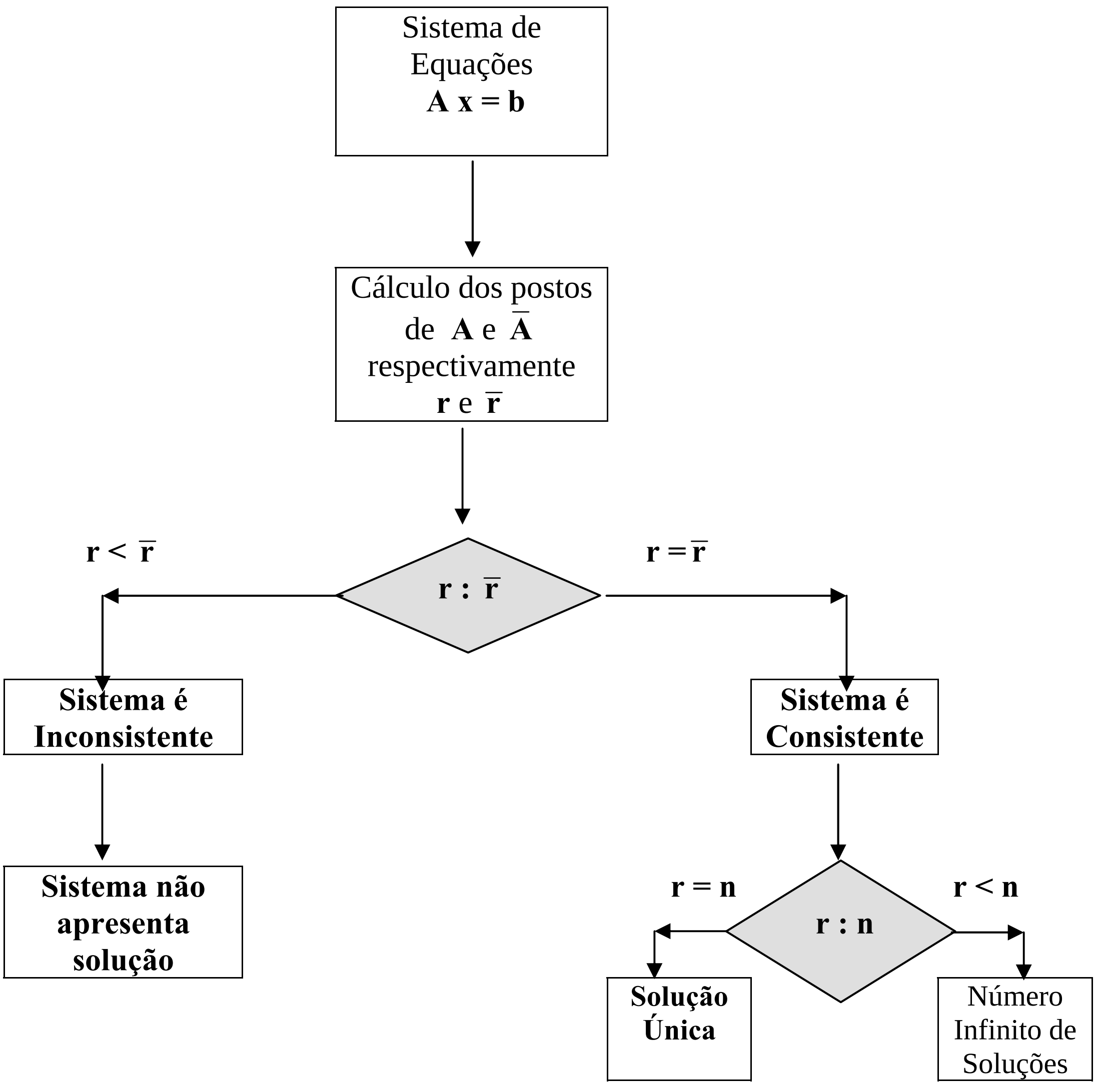
É o objetivo de esta seção fornecer os elementos para a análise de sistemas de equações algébricas lineares da forma: $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$, onde $\mathbf{A}$ é uma matriz quadrada (n,n) chamada de *<u>matriz dos coeficientes</u>* , $\mathbf{b}\in\Re^m$ chamado de *<u>vetor das constantes</u>* e $\mathbf{x}\in\Re^n$ chamado de *<u>vetor das incógnitas</u>* a solução deste sistema só existe se a matriz **A** for regular e pode ser expressa na forma: $\mathbf{x}=\mathbf{A}^{-1}\cdot\mathbf{b}$. Este procedimento já se encontra implementado, com grande eficiência, em inúmeros pacotes computacionais (**MATHCAD, MAPLE, MATLAB**, etc.) e dificilmente haverá a necessidade de reprogramá-lo. Entretanto dois aspectos de natureza qualitativa da estrutura do sistema devem ser analisados:

① Nem sempre o número de equações do sistema é igual ao número de incógnitas. Neste caso o sistema é descrito da mesma forma apresentada acima: $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$, mas a matriz **A** é retangular (m,n) o vetor $\mathbf{b}\in\Re^m$ e o vetor das incógnitas $\mathbf{x}\in\Re^n$, isto é m é o número de equações e n o número de incógnitas. O sistema de equações pode também ser rescrito na forma:

$$\mathbf{b}=\mathbf{A}\cdot\mathbf{x}=\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+\cdots+x_n\mathbf{a}_n,$$ onde $\mathbf{a}_k\in\Re^m$ para k = 1, 2, ..., n

são os vetores colunas da matriz **A**, desta forma os n elementos do vetor **x** podem ser interpretados como os componentes do vetor **b** na *base* formada pelos n vetores $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ sendo assim <u>obrigatoriamente</u> linearmente dependente do conjunto. Sendo r o número de vetores linearmente independentes no conjunto de n vetores $[r\le n]$ $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$, este deve ser igual ao número de vetores linearmente independentes do conjunto de n+1vetores $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$, **b**. Isto é o posto, r, da matriz $\mathbf{A}=\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{pmatrix}$ deve ser igual ao posto, $\overline{r}$, da matriz [chamada de *<u>matriz aumentada</u>*] $\overline{\mathbf{A}}=\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b}\end{pmatrix}$.

Quando $r=\text{posto}(\mathbf{A})=\overline{r}=\text{posto}(\overline{\mathbf{A}})$ o sistema é dito *<u>consistente</u>* e admite solução, entretanto só admite solução única de $r=\text{posto}(\mathbf{A})=n$. Se $m\ge n$ (número de equações ≥ número de incógnitas) o posto(**A**) é no <u>máximo</u> n, enquanto que se $m<n$ (número de equações < número de incógnitas) o posto(**A**) é no máximo $m<n$, portanto só há possibilidade do sistema apresentar solução única se $m\ge n$ (número de equações ≥ número de incógnitas). Caso $\mathbf{b}\equiv\mathbf{0}$ o sistema é dito *<u>homogêneo</u>* e como neste caso $\mathrm{r}=\text{posto}(\mathbf{A})$ será <u>sempre</u> igual a $\overline{r}=\text{posto}(\overline{\mathbf{A}})$ o sistema será sempre consistente e caso $r=n$ admite como solução única a *<u>solução trivial</u>* $\mathbf{x}\equiv\mathbf{0}$, deste modo para o sistema homogêneo de n equações e n incógnitas $\mathbf{A}\cdot\mathbf{x}=\mathbf{0}$ onde : $\mathbf{A}\in\Re^{\text{nxn}}$ e $\mathbf{x}\in\Re^{\text{n}}$, só apresenta solução não trivial se $r=\text{posto}(\mathbf{A})<n$, isto é a matriz **A** deve ser singular. O esquema de caracterização da *<u>consistência</u>* e da existência de solução de sistemas algébricos lineares é mostrado no diagrama abaixo:



Exemplos Ilustrativos:

(a) Analise a consistência do sistema linear de equações algébricas:

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 0 \\ x_1 - x_2 = -1 \end{cases} \Rightarrow \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 1 & -1 \end{pmatrix} \quad \mathrm{r} = 2 \text{ e a matriz aumentada: } \bar{\mathbf{A}} = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 2 & 0 \\ 1 & -1 & -1 \end{pmatrix} \quad \bar{\mathrm{r}} = 3,$$

sistema inconsistente. Haveria duas possibilidades de *corrigir* este sistema: (i) substituindo no lado direito da segunda equação 0 por 6, e neste caso a solução seria: $x_1 = 1$ e $x_2 = 2$; (ii) substituindo no lado direito da primeira equação 3 por 0, neste caso a solução seria x_1 =-0,5 e x_2 = +0,5.

(b) Analise a consistência e as possíveis soluções do sistema linear de equações algébricas:

$$\begin{cases} 2x_1 + 3x_2 = 0 \\ x_1 + x_2 = 0 \end{cases} \Rightarrow \mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \quad \mathrm{r} = 2 \text{ e a matriz aumentada: } \bar{\mathbf{A}} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 1 & 0 \end{pmatrix} \quad \bar{\mathrm{r}} = 2, \text{ assim o}$$

sistema é consistente e como $r = n = 2$ apresenta solução única que é a solução trivial $x_1 = x_2 = 0$ pois o sistema é homogêneo. [Este exemplo ilustra a observação feita anteriormente relativa a sistemas homogêneos].

(c) Analise a consistência e as possíveis soluções do sistema linear de equações algébricas:
$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 3x_1 + 6x_2 - x_3 = 8 \end{cases} \Rightarrow \mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 6 & -1 \end{pmatrix}$ $r = 2$ e a matriz aumentada: $\overline{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 3 & 6 & -1 & 8 \end{pmatrix}$ apresenta o posto $\overline{r} = 2$, assim o sistema é consistente (aliás se $m < n$ e se $r=m$ o sistema será sempre consistente, indicando que m vetores coluna de **A** constituem uma base de $\Re^m$ desta forma o vetor **b** necessariamente será linearmente dependente destes vetores e, em conseqüência, a matriz $\overline{\mathbf{A}}$ apresenta sempre posto igual ao de **A**, além disto como $r = m < n$ o sistema, neste caso, apresentará sempre um número infinito de soluções). Como $r = 2 < 3$ o sistema apresenta um número infinito de soluções, entretanto note que o sistema pode ser reescrito na forma: $\begin{cases} (x_1 + 2x_2) + x_3 = 4 \\ 3(x_1 + 2x_2) - x_3 = 8 \end{cases} \Rightarrow$ definindo: $z_1 = x_1 + 2\,x_2$, tem-se:

$$\begin{cases} z_1 + x_3 = 4 \\ 3z_1 - x_3 = 8 \end{cases} \Rightarrow \begin{cases} z_1 = (x_1 + 2x_2) = 3 \\ x_3 = 1 \end{cases}.$$

Este exemplo ilustra que em sistemas consistentes com menos equações do que incógnitas não se pode arbitrar indiscriminadamente (n-m) variáveis calculando as m restantes em função destas, neste processo de escolha de (m-n) entre as n incógnitas deve ser feita de modo que a matriz do sistema após esta escolha tenha posto = n. Assim se no exemplo o valor arbitrado de x_3 fosse diferente de <u>1</u> o sistema resultante seria inconsistente, pois com x_3=2, por exemplo, tem-se: $\begin{cases} x_1 + 2x_2 = 2 \\ 3x_1 + 6x_2 = 10 \end{cases} \Rightarrow \mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ $r = 1$ e a matriz aumentada:

$\overline{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 6 & 10 \end{pmatrix}$ tem o posto $\overline{\mathbf{r}} = 2$, sendo assim o sistema inconsistente. Entretanto se a variável x_1, por exemplo, tivesse um valor arbitrado α qualquer, ter-se-ia: $\begin{cases} 2x_2 + x_3 = 4 - \alpha \\ 6x_2 - x_3 = 8 - 3\alpha \end{cases} \Rightarrow \mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & -1 \end{pmatrix}$ $r = 2$ e a matriz aumentada: $\overline{\mathbf{A}} = \begin{pmatrix} 2 & 6 & 4\text{-}\alpha \\ 6 & -1 & 8\text{-}3\alpha \end{pmatrix}$ tem o posto $\overline{\mathbf{r}} = 2$ independente do valor de α, então neste caso o sistema é sempre consistente.

② Mesmo no caso em que m=n e em que **A** é regular a solução do sistema posta na forma: $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$ não assegura que a solução seja exata [uma prática recomendada é após o programa fornecer o vetor **x** calcular $\delta = \mathbf{A} \cdot \mathbf{x} - \mathbf{b}$ que é o chamado <u>*resíduo*</u> da solução. Quanto mais próximo δ estiver do vetor **0**, ou seja: $\|\delta\| \cong 0$, maior é a precisão do resultado] nem tão pouco que a obtenção da inversa da matriz **A** seja *fácil* [isto é especialmente verdadeiro se os elementos de **A** apresentarem ordens de grandeza muito distintas, neste caso a matriz é dita <u>*mal condicionada*</u>]. Estes dois fatos geralmente ocorrem devido ao <u>*mau condicionamento*</u> da matriz **A** que é medido pelos chamados <u>*números de condicionamento*</u>, valores elevados dos números de condicionamento é um forte indicativo de dificuldades numéricas na resolução do sistema e na inversão da matriz **A**. Os quatro números abaixo são usualmente considerados:

❶ M = n · M(**A**) · M($\mathbf{A}^{-1}$), onde M(**A**) = $\max\limits_{i,j} |a_{ij}|$ isto é, é o valor do módulo do elemento da matriz **A** que apresenta o maior valor absoluto;

❷ $N = N(\mathbf{A}) \cdot N(\mathbf{A}^{-1})$, onde N(**A**) é a norma de Frobenius de **A** definida por:

$N(\mathbf{A}) = \sqrt{tr\left(\mathbf{A} \cdot \mathbf{A}^T\right)}$;

❸ $P = \dfrac{|\lambda|}{|\mu|}$ onde $|\lambda|$ e $|\mu|$ são, respectivamente, os valores absolutos do maior e do menor valor característico de **A** em módulo (ou da parte real dos mesmos).

❹ $\kappa = \dfrac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}$, onde σ(**A**) são os valores singulares de A ou a raiz quadrada dos valores característicos de $\mathbf{A} \cdot \mathbf{A}^T$.

<u>Exemplo:</u> o sistema linear é o chamado problema de T. S. Wilson:

$$\begin{cases} 10x + 7y + 8z + 7w = 32 \\ 7x + 5y + 6z + 5w = 23 \\ 8x + 6y + 10z + 9w = 33 \\ 7x + 5y + 9z + 10w = 31 \end{cases}$$

a solução <u>exata</u> deste sistema é x = y = z = w = 1, entretanto adotando-se x = 6; y = -7,2; z = 2,9 e w = -0,1 os resultados de cada uma das equações são 32,1; 22,9; 32,9 e 31,1; e adotando x = 1,5; y = 0,18; z = 1,19 e w = 0,89 os correspondentes resultados são 32,01; 22,99; 32,99 e 31,01.

A matriz característica deste sistema é: $\mathbf{A} = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$ cuja inversa é:

$\mathbf{A}^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}$ e os valores característicos desta matriz são: 0,01015; 0,843107; 3,858057 e 30,288685, assim as duas normas destas matrizes são: M(**A**) = 10; M($\mathbf{A}^{-1}$)=68 , N(**A**)=30,5451 e N($\mathbf{A}^{-1}$) = 98,5292, então os números de condicionamento são:

❶ M = n · M(**A**) · M($\mathbf{A}^{-1}$) = 4 · 10 · 68 = 2720;

❷ $N = N(\mathbf{A}) \cdot N(\mathbf{A}^{-1})$ = 30,5451 · 98,5292 = 3009,58

❸ $P = \dfrac{|\lambda|}{|\mu|} = \dfrac{30,2887}{0,0102} = 2984,09$

❹ $\kappa = \dfrac{\sigma_{\max}}{\sigma_{\min}} = \dfrac{30,2887}{0,0102} = 2984,09$ (igual ao caso 3, pois a matriz **A** é simétrica)

1.6 Fatoração QR

Toda matriz **A** ∈ $\Re^{m\times n}$ com colunas linearmente independentes pode ser unicamente fatorada na forma **A** = **Q** · **R** com as colunas da matriz **Q** ∈ $\Re^{m\times n}$ formando uma base ortonormal para o range(**A**) e a matriz triangular superior **R** ∈ $\Re^{n\times n}$ tendo os elementos da diagonal positivos. Esta fatoração pode ser gerada pelo procedimento de ortogonalização de Gram-Schmidt, pois as colunas da matriz $\mathbf{Q} = (\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_n)$ são resultados da aplicação do processo de Gram-Schmidt nas colunas da matriz $\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n)$, e a matriz $\mathbf{R}$ é dada por:

$$\mathbf{R} = \begin{bmatrix} v_1 & \mathbf{q}_1^T \cdot \mathbf{a}_2 & \mathbf{q}_1^T \cdot \mathbf{a}_3 & \cdots & \mathbf{q}_1^T \cdot \mathbf{a}_n \\ 0 & v_2 & \mathbf{q}_2^T \cdot \mathbf{a}_3 & \cdots & \mathbf{q}_2^T \cdot \mathbf{a}_n \\ 0 & 0 & v_3 & \cdots & \mathbf{q}_3^T \cdot \mathbf{a}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & v_n \end{bmatrix}$$

onde $v_1 = \| \mathbf{a}_1 \|$ e $v_k = \left\| \mathbf{a}_k - \sum_{i=1}^{k-1} \mathbf{q}_i^T \cdot \mathbf{a}_k \cdot \mathbf{q}_i \right\|$ para $k = 2, 3, ..., n$.

Se a matriz $\mathbf{A} \in \Re^{n \times n}$ é regular, então $\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}$, pois as colunas da matriz Q são ortonormais entre si.

<u>Exemplo</u>: Obter a fatoração QR da matriz $\mathbf{A} = \begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}$ aplicando o processo de ortogonalização de Gram-Schmidt, ou seja:

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{v_1} \qquad \text{e} \qquad \mathbf{q}_k = \frac{\mathbf{a}_k - \sum_{i=1}^{k-1} \mathbf{q}_i^T \cdot \mathbf{a}_k \cdot \mathbf{q}_i}{v_k} \quad \text{para } k = 2, 3, ..., n$$

Então, para $k = 1$: $r_{11} = v_1 = \| \mathbf{a}_1 \| = 5$ e $\mathbf{q}_1 = \dfrac{\mathbf{a}_1}{v_1} = \begin{bmatrix} 0 \\ 3/5 \\ 4/5 \end{bmatrix}$

$k = 2$: $r_{12} = \mathbf{q}_1^T \cdot \mathbf{a}_2 = 25$; $\tilde{\mathbf{q}}_2 = \mathbf{a}_2 - r_{12} \cdot \mathbf{q}_1 = \begin{bmatrix} -20 \\ 12 \\ -9 \end{bmatrix}$; $r_{22} = v_2 = \|\tilde{\mathbf{q}}_2\| = 25$; $\mathbf{q}_2 = \dfrac{\tilde{\mathbf{q}}_2}{r_{22}} = \begin{bmatrix} -4/5 \\ 12/25 \\ -9/25 \end{bmatrix}$

k = 3: $r_{13} = \mathbf{q}_1^T \cdot \mathbf{a}_3 = -4$; $r_{23} = \mathbf{q}_2^T \cdot \mathbf{a}_3 = 10$; $\tilde{\mathbf{q}}_3 = \mathbf{a}_3 - r_{13} \cdot \mathbf{q}_1 - r_{23} \cdot \mathbf{q}_2 = \begin{bmatrix} -6 \\ -32/5 \\ 24/5 \end{bmatrix}$

$r_{33} = v_3 = \|\tilde{\mathbf{q}}_3\| = 10$; $\mathbf{q}_3 = \dfrac{\tilde{\mathbf{q}}_3}{r_{33}} = \begin{bmatrix} -3/5 \\ -16/25 \\ 12/25 \end{bmatrix}$ e, portanto,

$$\mathbf{Q} = \begin{bmatrix} 0 & -4/5 & -3/5 \\ 3/5 & 12/25 & -16/25 \\ 4/5 & -9/25 & 12/25 \end{bmatrix} \quad \text{e} \quad \mathbf{R} = \begin{bmatrix} 5 & 25 & -4 \\ 0 & 25 & 10 \\ 0 & 0 & 10 \end{bmatrix}$$

Verificando o resultado: $\mathbf{A} = \mathbf{Q} \cdot \mathbf{R} = \begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}$.

Lista de exercícios

1) Verificar se a aplicação F: $\Re^3 \to \Re^2$ definida por $F(x, y, z) = (z, x + y)$, $\forall\ (x, y, z) \in \Re^3$, é uma transformação linear.

2) Verificar se a aplicação F: $\Re \to \Re^2$ definida por $F(x) = (x, 2)$, $\forall\ x \in \Re$, é uma transformação linear.

3) Dada a matriz: $\mathbf{A} = \begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$

   a) compute seus valores característicos;
   b) compute o conjunto de vetores característicos de **A** e de $\mathbf{A}^\mathbf{T}$, verifique que estes dois conjuntos são mutuamente ortogonais;
   c) verifique que a matriz **A** satisfaz também sua equação característica.

4) Mostre que $\mathbf{A}^\dagger = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T$

Bibliografia

- Curso de Nivelamento em Matemática - E.C. Biscaia Jr., PEQ/COPPE-UFRJ, 2009.
- Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos - D. Sperandio, J.T. Mendes e L.H.M. Silva - Pearson-Prentice Hall, 2003.
- Matrix Analysis and Applied Linear Algebra - C.D. Meyer, SIAM, 2000.
- Matrix Computations - G. H. Golub e C. F. Van Loan - Johns Hopkins, 1996.
- Álgebra Linear a Aplicações - C.A. Callioli, H.H. Domingues e R.C.F. Costa - Atual Editora, 6ª ed., 1989.