MATRIZES E VETORES

1) CONCEITOS BÁSICOS

Os cálculos/operações assim como conceitos envolvendo matrizes e vetores constituem a base dos métodos numéricos que tratam da solução de sistemas lineares e não lineares de equações algébricas ou diferenciais. A representação destes sistemas em termos matriciais/vetoriais é extremamente mais compacta e é corrente na literatura técnica. Como visa-se neste curso apresentar os conceitos básicos deste assunto especialmente relacionados com aplicações em Engenharia Química, os elementos de matrizes e vetores serão em princípio números ou variáveis reais a não ser quando explicitamente especificados como complexos.

Uma *matriz* é um arranjo retangular de números em *m* linhas e *n* colunas, *m* x *n*, sendo representada como **A** (letras maiúsculas em negrito) pertencente a $\Re^{m \times n}$, isto é: $\mathbf{A} \in \Re^{\mathbf{mxn}}$. O elemento da linha *i* e coluna *j* de **A** é representado por a_{ij} (correspondente letra minúscula com o sub-índice *ij*) ou $(\mathbf{A})_{ij}$. A matriz *completa* é geralmente escrita na forma: $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ ou,

em forma mais compacta,por: $\mathbf{A} = (a_{ij})$ com *i*= 1, ..., *m* e *j*=1, ...*n*. Se duas matrizes **A** e **B** apresentam o mesmo número de linhas e o mesmo número de colunas são ditas do mesmo *tipo*.

Se $\mathbf{A} = (a_{ij})$ é tal que $a_{ij} = 0$ para todo *i* e *j* então a matriz **A** é dita nula e é representada por **0**.

Se *n=m* a matriz **A** é dita *quadrada*.

Se *n=m* e $a_{ij} = a_{ji}$ para *i*,*j* = 1, ... *n* a matriz quadrada **A** é dita *simétrica*.

Se *n=1* tem-se um *vetor coluna* ou simplesmente *vetor* designado por **v** (letra minúscula em negrito) e representado por: $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} \in \Re^m$

Se *m=1* tem-se um *vetor linha* designado por $\mathbf{v}^T$ (letra minúscula em negrito com o sobre-índice T de transposto) e representada por: $\mathbf{v}^T = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \in \Re^{1xn}$

Se *m=n=1* tem-se um *escalar* (real) α (letra minúscula grega), ou seja: $\alpha \in \Re$.

A matriz $\mathbf{A} \in \Re^{\mathbf{mxn}}$ pode ser *particionada* por:

a) Colunas na forma:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a_1} & \mathbf{a}_2 & \vdots & \mathbf{a_n} \end{pmatrix} \text{ onde } \mathbf{a_j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \Re^m$$ para $j = 1, \ldots, n$ são os n <u>vetores colunas</u> da matriz **A**;

b) <u>Linhas</u> na forma:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix} \text{ onde } \mathbf{a}_i^T = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \in \Re^{1xn}$$ para $i = 1, \ldots, m$ são os m <u>vetores linhas</u> da matriz **A**.

<u>2) OPERAÇÕES ENTRE MATRIZES</u>

As operações de adição ou subtração são definidas apenas para matrizes do mesmo tipo, assim se **A** e **B** são matrizes (m x n) então a matriz **C** , também (m x n), <u>*soma*</u> ou <u>*subtração*</u> de **A** com **B,** representada por **C = A ± B**, tem como termo geral :
$c_{ij} = a_{ij} \pm b_{ij}$ para $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Se <u>α</u> é um escalar qualquer, a matriz <u>α**A**</u> é uma matriz cujo termo geral é αa_{ij}.

A operação de multiplicação de matrizes está intimamente relacionada a transformações de coordenadas. Assim sejam as seguintes *transformações lineares*:

$$z_i = \sum_{j=1}^{n} a_{ij} y_j \text{ para } i = 1, \ldots, m \text{ e } y_j = \sum_{k=1}^{p} b_{jk} x_k \text{ para } j = 1, \ldots, m.,$$

expressando z_i em temos de x_k , por substituição tem-se:

$$z_i = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{p} b_{jk} x_k \right) = \sum_{k=1}^{p} \left(\sum_{j=1}^{n} a_{ij} \cdot b_{jk} \right) \cdot x_k$$

definindo: $c_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk}$ tem-se: $z_i = \sum_{k=1}^{m} c_{ik} \cdot x_k$, o que induz à definição da matriz: $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ onde **A** é (m,n) , **B** é (n,p) e **C** é (m,p) que apresenta como termo geral: $c_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk}$ para $i = 1, \ldots, m$ e $k = 1, \ldots, p$. Verificando-se assim que a operação $\mathbf{A} \cdot \mathbf{B}$ só é definida se o número de colunas de **A** (primeira parcela do produto) for igual ao número de linhas de **B** (segunda parcela do produto). É importante ressaltar que a lei de comutatividade não é satisfeita pelo produto entre matrizes, mesmo que $\mathbf{B} \cdot \mathbf{A}$ seja definida, isto é m=p e mesmo que $\mathbf{B} \cdot \mathbf{A}$ seja do mesmo tipo que $\mathbf{A} \cdot \mathbf{B}$, o que só ocorrerá se m=p=n (isto é ambas as matrizes são quadradas e de mesma dimensão), assim de uma forma geral tem-se: $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$.

Se a primeira parcela do produto é um vetor linha $\mathbf{u}^{\mathrm{T}}$ $(1,n)$ e a segunda parcela é um vetor coluna $\mathbf{v}$ $(n,1)$ então o produto $\mathbf{u^T}\cdot\mathbf{v}$ é um escalar: $\mathbf{u^T}\cdot\mathbf{v}=\sum_{j=1}^{n}u_j\cdot v_j$ que é comutável, isto é $\mathbf{u^T}\cdot\mathbf{v}=\mathbf{v^T}\cdot\mathbf{u}$. Este produto é chamado de <u>*produto escalar*</u> de dois vetores.

Se $\mathbf{A}$ é uma matriz (m,n) e $\mathbf{v}$ um vetor $(n,1)$ então o produto $\mathbf{A}\cdot\mathbf{v}$ é um vetor $\mathbf{u}$ $(m,1)$ cujo termo geral é: $u_i=\sum_{j=1}^{n}a_{ij}v_j$ para $i = 1, ..., m$. Este produto pode ser efetuado de duas formas distintas:

(a) por linhas (método **ij**) considerando a partição por linhas da matriz **A**, isto é: $\mathbf{A}=\begin{pmatrix}\mathbf{a}_1^T\\ \mathbf{a}_2^T\\ \vdots\\ \mathbf{a}_m^T\end{pmatrix}$,

então: $\mathbf{A}\cdot\mathbf{v}=\begin{pmatrix}\mathbf{a}_1^T\cdot\mathbf{v}\\ \mathbf{a}_2^T\cdot\mathbf{v}\\ \vdots\\ \mathbf{a}_m^T\cdot\mathbf{v}\end{pmatrix}$, isto é o elemento i de $\mathbf{u}$ é dado por $u_i=\mathbf{a}_i^T\cdot\mathbf{v}$ para $i = 1, ..., m$, que é o produto escalar do vetor composto pelos elementos da linha i da matriz **A** com o vetor **u**.

(b) por colunas (método **ji**): considerando a partição por colunas de **A**, isto é: $\mathbf{A}=\begin{pmatrix}\mathbf{a_1} & \mathbf{a}_2 & \vdots & \mathbf{a_n}\end{pmatrix}$, então:

$$\mathbf{u}=\mathbf{A}\cdot\mathbf{v}=\begin{pmatrix}\mathbf{a_1} & \mathbf{a}_2 & \vdots & \mathbf{a_n}\end{pmatrix}\cdot\begin{pmatrix}v_1\\ v_2\\ \vdots\\ v_n\end{pmatrix}=v_1\cdot\mathbf{a_1}+v_2\cdot\mathbf{a}_2+\cdots+v_n\cdot\mathbf{a}_n=\sum_{i=1}^{n}v_i\cdot\mathbf{a}_i$$

, isto é o vetor **u** é uma <u>*combinação linear*</u> dos vetores coluna de **A** sendo os coeficientes desta combinação os elementos do vetor **v.**

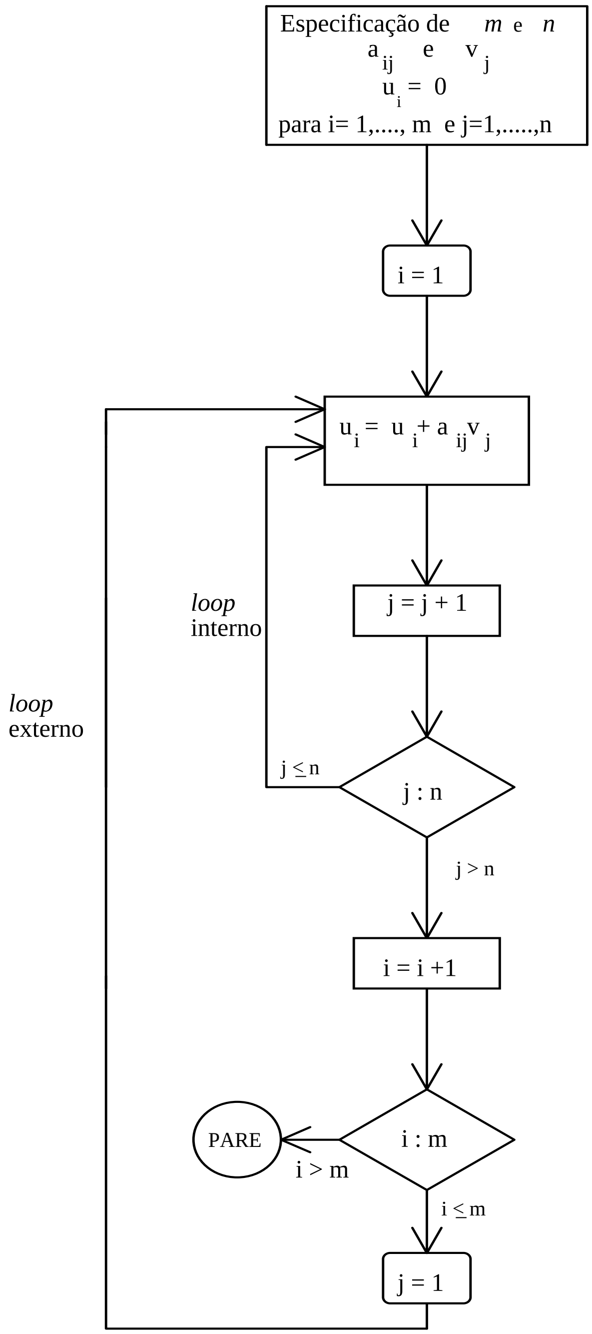
<u>Exemplo Ilustrativo</u> $\mathbf{A}=\begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix}$, $\mathbf{v}=\begin{pmatrix}7\\ 8\end{pmatrix}$

(a)método **ij**: $u_1=\begin{pmatrix}1 & 2\end{pmatrix}\cdot\begin{pmatrix}7\\ 8\end{pmatrix}=1\cdot7+2\cdot8=23$; $u_2=\begin{pmatrix}3 & 4\end{pmatrix}\cdot\begin{pmatrix}7\\ 8\end{pmatrix}=3\cdot7+4\cdot8=53$ e

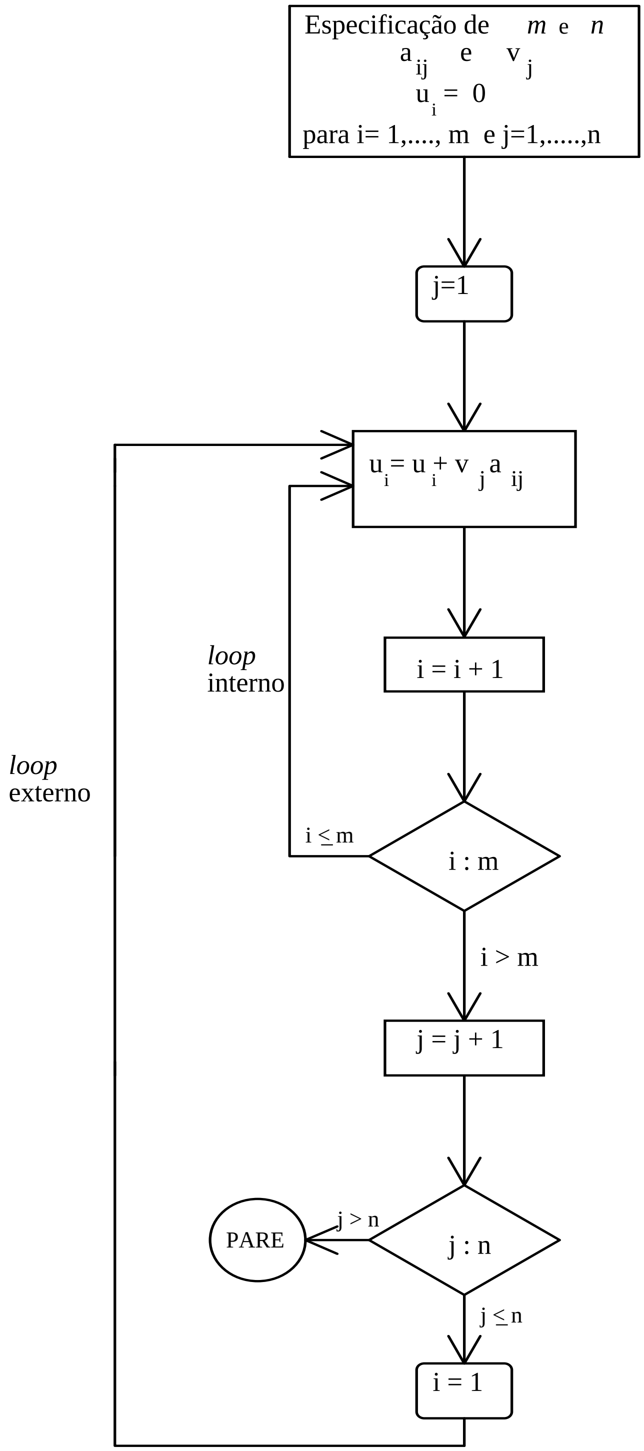
$u_3=\begin{pmatrix}5 & 6\end{pmatrix}\cdot\begin{pmatrix}7\\ 8\end{pmatrix}=5\cdot7+6\cdot8=83$, logo: $\mathbf{u}=\begin{pmatrix}23\\ 53\\ 83\end{pmatrix}$

(b) método **ji**: $\mathbf{u} = 7 \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 8 \cdot \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 23 \\ 53 \\ 83 \end{pmatrix}$.

A designação dos métodos como **ij** e como **ji** deve-se à forma como os *loops* de programação são efetuados, assim no primeiro método tem-se o seguinte fluxograma:



Note que neste caso o *loop* externo é em **i** (linha)e o *loop* interno é em **j** (coluna).
O segundo método é descrito pelo fluxograma:



Note que neste caso o *loop* externo é em **j** (coluna) e o *loop* interno é em **i** (linha).

Estes métodos podem também ser ilustrados acompanhando passo a passo o Exemplo Ilustrativo anterior segundo cada um dos algoritmos.

Método **ij**

i	1	1	2	2	3	3
j	1	2	1	2	1	2
$\mathbf{u_1}$	7	23	23	23	23	23
$\mathbf{u_2}$	0	0	21	53	53	53
$\mathbf{u_3}$	0	0	0	0	35	83

Método **ji**

j	1	1	1	2	2	2
i	1	2	3	1	2	3
$\mathbf{u_1}$	7	7	7	23	23	23
$\mathbf{u_2}$	0	21	21	21	53	53
$\mathbf{u_3}$	0	0	35	35	35	83

A operação de *transposição* de uma matriz **A** (*m*,*n*) consiste em trocar as linhas pelas colunas de **A**, esta nova matriz é chamada de matriz *transposta* de **A** , representada por $\mathbf{A}^T$, e é uma matriz (*n*,*m*) cujo termo da linha *j* e coluna *i* é $a_{ji}^T = a_{ij}$ para *j* = 1, ... , *n* e *i* = 1, ... , *m*. Se a matriz **A** é simétrica então: $\mathbf{A} = \mathbf{A}^T$.

As propriedades que serão descritas a seguir aplicam-se **exclusivamente** a matriz quadradas (*n*,*n*) e a vetores coluna (*n*,1) e a vetores linha (1,*n*).

Define-se como <u>*matriz identidade*</u> a matriz **I** cujo elemento geral é: $(\mathbf{I})_{ij} = \delta_{ij} = \begin{cases} 1 \text{ apenas se i=j} \\ 0 \text{ sempre que i} \neq \text{j} \end{cases}$, onde δ_{ij} é chamado de *delta de Kronecker*, deste modo a matriz identidade é uma matriz diagonal cujos termos da diagonal são todos unitários, assim: $\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$, entendendo-se como <u>*matriz diagonal*</u> uma matriz quadrada em que apenas os elementos da diagonal (também chamada de *diagonal principal*) são não nulos, geralmente uma matriz diagonal $\mathbf{D} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$ é representada na forma mais compacta: $\mathbf{D} = diag\begin{pmatrix} d_1 & d_2 & \cdots & d_n \end{pmatrix}$.
Note que toda matriz diagonal é simétrica.

Uma propriedade muito importante da matriz identidade é: $\mathbf{I} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I} = \mathbf{A}$, isto é, a matriz identidade pré-multiplicada ou pós-multiplicada por qualquer matriz quadrada de mesma dimensão não altera o valor de elemento algum desta matriz.

Uma matriz diagonal é um caso particular de matrizes dita *esparsas*, que são matrizes que apresentam um grande número de elementos nulos, sendo os elementos não nulos mais a exceção do que a regra. Algumas destas matrizes são apresentadas abaixo:
1) matrizes *tridiagonais* são matrizes que apresentam apenas os elementos da diagonal, os elementos sobre a diagonal e os elementos sob a diagonal não nulos, sendo os demais nulos, assim se **A** é uma matriz tridiagonal então:

$$a_{ij} = \begin{cases} \neq 0 \text{ se } i = j \text{ - diagonal} \\ \neq 0 \text{ se } i = j+1 \text{ (para i= 2,...,n)-sob a diagonal} \\ \neq 0 \text{ se } i = j - 1 \text{ (para i=1,...,n-1)- sobre a diagonal} \\ =0 \text{ em qualquer outro caso} \end{cases}$$

2) matrizes *bidiagonais* são matrizes que apresentam apenas os elementos da diagonal e os elementos sobre a diagonal ou sob a diagonal não nulos, no primeiro caso diz-se que a matriz é *bidiagonal superior* e no segundo caso *bidiagonal inferior.*
3) matrizes *triangulares* são matrizes que apresentam todos os elementos sob (ou sobre) a diagonal nulos, sendo neste caso chamada de matriz *triangular superior* ou matriz **U** (ou *triangular inferior* ou matriz **L**), assim:
$(\mathbf{U})_{ij} = 0$ se i > j e $(\mathbf{L})_{ij} = 0$ se j > i .

Algumas vezes para evitar ambigüidades representa-se a matriz identidade de dimensão n por $\mathbf{I}_n$.

O *traço* de uma matriz quadrada **A** é a soma dos elementos de sua diagonal, isto é:

$$tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \text{ .}$$

Uma matriz quadrada **A** é dita *positiva definida* se $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} > 0$ para todo vetor $\mathbf{x} \neq \mathbf{0}$ (isto é não nulo), caso $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} \geq 0$ a matriz **A** é dita *positiva semi-definida* e se $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} \geq 0$ para alguns vetores $\mathbf{x} \neq \mathbf{0}$ e se $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} < 0$ para algum vetor $\mathbf{x} \neq \mathbf{0}$ a matriz **A** é dita *não-definida*. Além disto, **A** é dita *negativa definida* se $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} < 0$ para todo vetor $\mathbf{x} \neq \mathbf{0}$ e é dita *negativa semi-definida* caso $\mathbf{x}^\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{x} \leq 0$.

O *determinante* de uma matriz **A** é um escalar obtido através da soma de todos os produtos possíveis envolvendo um elemento de cada linha e cada coluna da matriz, com o sinal positivo ou negativo conforme o número de permutações dos índices seja par ou ímpar. Sua obtenção e sua representação, apesar de ser um dos conceitos mais preliminares envolvendo matrizes, não são tarefas triviais e o conceito de determinante será utilizado nestas notas apenas como base de outras propriedades de matrizes quadradas. Assim, o determinante de **A** designado por det(**A**) pode ser representado por:
$\det(\mathbf{A}) = \sum \pm\, \mathrm{a}_{1,\mathrm{i}_1} \cdot a_{2,i_2} \cdot \cdots \cdot a_{n,i_n}$, ou então através do conceito de *cofator* do elemento *ij* da matriz **A** (representado por A_{ij})que é o determinante da matriz obtida cancelando a linha *i* e a coluna *j* da matriz **A** com o sinal mais ou menos conforme *i+j* seja par ou ímpar, assim:
$A_{ij} = (-1)^{i+j} \cdot \det(\Lambda_{ij})$ onde Λ_{ij} é matriz quadrada (n-1,n-1) obtida pela eliminação da linha *i* e a coluna *j* de **A.**. Tem-se então:

$$\det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij} \cdot A_{ij}$$

(expansão do determinante pela linha *i*),

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij} \cdot A_{ij}$$

(expansão do determinante pela coluna *j*).

Além disto: $\sum_{j=1}^{n} a_{ij} \cdot A_{kj} \equiv 0$ se $k \neq i$ (pois equivaleria a dizer que a matriz **A** apresenta duas linhas iguais, no caso as linhas *i* e *k*); e $\sum_{i=1}^{n} a_{ij} \cdot A_{ik} \equiv 0$ se $k \neq j$ (pois equivaleria a dizer que a matriz **A** apresenta duas colunas iguais, no caso as colunas *j* e *k*).

Na prática, entretanto, é praticamente impossível calcular o determinante de matrizes através destas regras gerais por envolver um número muito grande de termos [na realidade n!, assim mesmo com matrizes relativamente pequenas como com n=10 tem-se 3 milhões de termos]. Felizmente, para os nossos propósitos, apenas as regras a seguir serão suficientes:

① O determinante de uma matriz **A** mantém-se inalterado se somarem-se a todos os elementos de qualquer linha (ou coluna) os correspondentes elementos de uma outra linha (ou coluna) multiplicados pela mesma constante α;

②se a_{ij} é o único elemento não nulo da linha *i* ou da coluna *j* então: $\det(\mathbf{A}) = a_{ij} \cdot A_{ij}$;

③se $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ então : $\det(\mathbf{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \cdot d - b \cdot c$.

Da regra ① verifica-se que se det(**A**) = 0 então **A** apresenta duas linhas (ou colunas) proporcionais entre si, ou ainda, de uma forma mais geral, pode-se afirmar que uma linha (ou coluna) de **A** pode ser escrita como combinação linear de alguma ou algumas linhas (ou colunas) da mesma matriz. Da regra ② demonstra-se que se **A** for uma matriz triangular então det(**A**) é simplesmente o produto dos elementos de sua diagonal (note que o mesmo vale para matrizes bidiagonais que são também matrizes triangulares).

Se det(**A**) = 0 diz-se que a matriz **A** é <u>*singular*</u>, e caso det(**A**) ≠ 0 então **A** é dita <u>*regular*</u>.

Se $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ então det(**C**) = det(**A**) **.** det (**B**).

Se **B** = $\mathbf{A}^T$ então det(**B**) = det(**A**), isto é det($\mathbf{A}^T$) = det(**A**)

A <u>*matriz adjunta*</u> de uma matriz **A** é a matriz transposta da matriz obtida substituindo cada elemento da matriz **A** pelo seu correspondente cofator, isto é se **Ã** é a matriz adjunta de **A** então o elemento da linha *i* e coluna *j* de **Ã** é A_{ji}. A propriedade mais importante da matriz adjunta diz respeito aos produtos: **P=A Ã** e **Q=Ã A** o primeiro produto tem com termo geral: $p_{ij} = \sum_{k=1}^{n} a_{ik} \cdot \tilde{a}_{kj} = \sum_{k=1}^{n} a_{ik} \cdot A_{jk} = \det(\mathrm{A})\delta_{ij}$

e o segundo produto: $q_{ij} = \sum_{k=1}^{n} \tilde{a}_{ik} \cdot a_{kj} = \sum_{k=1}^{n} A_{ki} \cdot a_{kj} = \det(\mathrm{A})\delta_{ij}$, assim:

$\mathbf{A} \cdot \tilde{\mathbf{A}} = \tilde{\mathbf{A}} \cdot \mathbf{A} = \det(\mathbf{A})\mathbf{I}$. Deste modo se $\det(\mathbf{A}) \neq 0$ (**A** é regular) define-se:

$\mathbf{A}^{-1} = \dfrac{1}{\det(\mathbf{A})} \cdot \tilde{\mathbf{A}}$ a chamada *inversa* de **A** que tem como propriedade:

$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$ que existe apenas se $\det(\mathbf{A}) \neq 0$. Note que $\det\left(\mathbf{A}^{-1}\right) = \dfrac{1}{\det\left(\mathbf{A}\right)}$

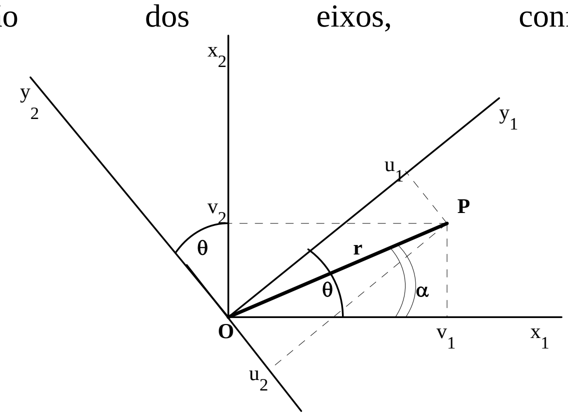
Exemplo Ilustrativo: Considere a seguinte matriz (2x2):

$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, assim, seus cofatores são: $\begin{cases} A_{11} = d \ ; \mathrm{A}_{12} = -c \\ A_{21} = -b \ ; \mathrm{A}_{22} = a \end{cases}$, permitindo determinar a matriz adjunta: $\tilde{\mathbf{A}} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, note que:

$\mathbf{A} \cdot \tilde{\mathbf{A}} = \tilde{\mathbf{A}} \cdot \mathbf{A} = (a \cdot d - b \cdot c)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det(\mathbf{A}) \cdot \mathbf{I} \Rightarrow \mathbf{A}^{-1} = \dfrac{1}{(a \cdot d - b \cdot c)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, isto é, para determinar a inversa de uma matriz (2x2) basta trocar os elementos da diagonal principal, trocar o sinal dos elementos da diagonal secundária e dividir a matriz resultante pelo determinante da matriz original.

Se $\mathbf{A}^{-1} = \mathbf{A}^{T}$, isto é a inversa da matriz é igual a sua transposta, então a matriz **A** é chamada de *matriz ortogonal*., e neste caso o det(**A**) = +1 ou -1 .

Exemplo Ilustrativo - Considere a mudança de coordenadas em $\Re^2$ resultante da simples rotação dos eixos, conforme mostrado abaixo:



vê-se da figura acima que no sistema original (x_1 , x_2) : $v_1 = r\cos(\alpha)$ e : $v_2 = r\,\mathrm{sen}(\alpha)$, o vetor **OP** faz um ângulo igual a **θ - α** com o eixo y_1 e projeta-se na porção negativa do eixo y_2, assim: $\begin{cases} u_1 = r \cdot \cos(\theta - \alpha) = r \cdot \cos\theta \cdot \cos\alpha + r \cdot sen\theta \cdot sen\alpha \\ u_2 = -r \cdot sen(\theta - \alpha) = -r \cdot sen\theta \cdot \cos\alpha + r \cdot \cos\theta \cdot sen\alpha \end{cases}$ ou seja:

$\begin{cases} u_1 = \cos\theta \cdot v_1 + sen\theta \cdot v_2 \\ u_2 = -sen\theta \cdot v_1 + \cos\theta \cdot v_2 \end{cases}$ ou, em termos matriciais: $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & sen\theta \\ -sen\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$,

identificando a matriz da transformação : $\mathbf{T} = \begin{pmatrix} \cos\theta & sen\theta \\ -sen\theta & \cos\theta \end{pmatrix} \Rightarrow \mathbf{T}^T = \begin{pmatrix} \cos\theta & -sen\theta \\ sen\theta & \cos\theta \end{pmatrix}$,

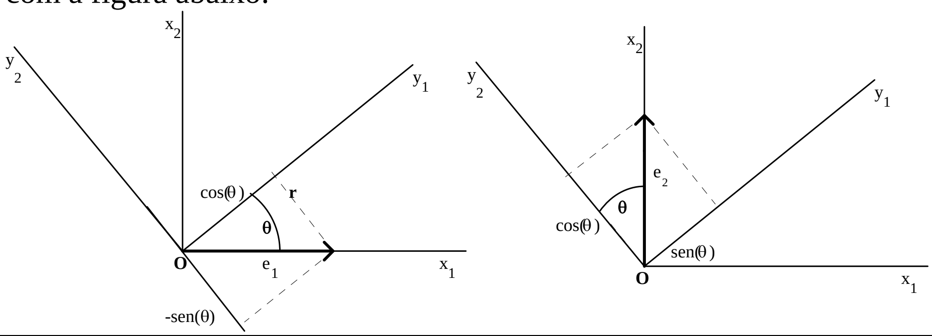
tem-se:

$$\mathbf{T} \cdot \mathbf{T}^T = \begin{pmatrix} \cos^2\theta + sen^2\theta & -\cos\theta sen\theta + sen\theta\cos\theta \\ -sen\theta\cos\theta + \cos\theta sen\theta & sen^2\theta + \cos^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{e}$$

$$\mathbf{T}^T \cdot \mathbf{T} = \begin{pmatrix} \cos^2\theta + sen^2\theta & \cos\theta sen\theta - sen\theta\cos\theta \\ sen\theta\cos\theta - \cos\theta sen\theta & sen^2\theta + \cos^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Verificando-se assim que a matriz **T** é uma matriz ortogonal.

É interessante verificar que os vetores coluna da matriz **T** são exatamente os componentes dos vetores $\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ no novo sistema de coordenadas, em acordo com a figura abaixo:



3) ALGUMAS PROPRIEDADES FUNDAMENTAIS DE OPERAÇÕES ENTRE MATRIZES

As leis de <u>associação</u> e de <u>comutação</u> são válidas para as operações de adição/subtração, assim: **(A+B)+C = A+(B+C)** e **A+B = B+A**.

São válidas também as leis de <u>associação</u> e de <u>distribuição</u> para a multiplicação, assim: **(AB)C = A(BC) ; A(B+C) = AB + AC** e **(A+B)C = AC + BC**

Para a matriz transposta tem-se as seguintes propriedades: $\mathbf{(A+B)^T = A^T + B^T}$ e $\mathbf{(AB)^T = B^T A^T}$

e para a matriz inversa: $\mathbf{(AB)^{-1} = B^{-1} A^{-1}}$ e $\mathbf{(A^{-1})^T = (A^T)^{-1}}$ Um <u>*menor de ordem p*</u> de uma matriz **A** (*n*,*n*) é o valor do determinante da matriz obtida eliminando-se *n-p* linhas e *n-p* colunas da matriz **A.** Se uma matriz **A** apresenta a propriedade de <u>todos</u> os menores de ordem (*r* + 1) serem nulos e de pelo menos um menor de ordem *r* ser não nulo então diz-se que a matriz **A** é de <u>*posto (rank) r*</u> . Note que todo matriz quadrade (*n,n*) <u>*regular*</u> (ou <u>*não singular*</u>) apresenta o posto igual a <u>*n*</u>.

Um conjunto de *n* vetores $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_n$ com *n* elementos é dito <u>*linearmente independente*</u> se os únicos valores de c_1 , c_2 , ...c_n tais que: $c_1\,\mathbf{u}_1+c_2\,\mathbf{u}_2+ ....+c_n\,\mathbf{u}_n= \mathbf{0}$ são:$c_1 =c_2 = ...=c_n = 0$. Neste caso os vetores $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_n$ formam uma <u>*base*</u> de $\Re^n$ e <u>todo</u> vetor deste espaço de <u>*dimensão n*</u> (que é o numero máximo de vetores linearmente independentes que pode existir neste espaço, que também é igual ao número de elementos destes vetores) pode ser expresso como uma combinação linear dos vetores da base, os coeficientes desta combinação linear são os <u>*componentes*</u> do vetor nesta base. Os <u>*componentes*</u> de um vetor qualquer do $\Re^n$ apenas confundem-se com seus <u>*elementos*</u> quando adota-se a <u>*base canônica*</u> do $\Re^n$, que é a base composta pelos vetores unitários $\mathbf{e}_i$ cujo único elemento não nulo é o i'ésimo, isto é : $e_{ij} = \delta_{ij}$, desta forma os vetores coluna ou os vetores linha da matriz identidade **I** são os vetores da base canônica do $\Re^n$.

Em uma matriz de <u>*posto r*</u> todos seus vetores linha (ou coluna) podem ser escritos como uma combinação linear de *r* vetores linha (ou coluna), desta forma o <u>*posto*</u> de uma matriz é também o número máximo de vetores linha (ou coluna) linearmente independentes.

Uma forma de determinar o posto de uma matriz é através do <u>*processo de ortogonalização de Gram-Schmidt*</u> aplicado aos vetores linha ou aos vetores coluna da matriz, este processo pode ser resumido na forma, sejam: $\mathbf{v}_1$, $\mathbf{v}_2$, ... , $\mathbf{v}_n$ os vetores coluna (ou linha) de **A**, então adota-se:

$\mathbf{u}_1 = \mathbf{v}_1$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v}_2^{\mathbf{T}} \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right)\cdot \mathbf{u}_1$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3^{\mathbf{T}} \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right)\cdot \mathbf{u}_1 - \left(\frac{\mathbf{v}_3^{\mathbf{T}} \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right)\cdot \mathbf{u}_2$$

....

$$\mathbf{u}_j = \mathbf{v}_j - \sum_{k=1}^{j-1}\left[\left(\frac{\mathbf{v}_j^{\mathbf{T}} \cdot \mathbf{u}_k}{\|\mathbf{u}_k\|^2}\right)\cdot \mathbf{u}_k\right] \text{ para j = 2, ..., n com } \mathbf{u}_1 = \mathbf{v}_1$$

onde $\|\mathbf{p}\| = \sqrt{p_1^2 + p_2^2 + \cdots + p_n^2}$ (módulo de **p**)

Encontrando-se durante este processo algum vetor $\mathbf{u}_k$ com módulo nulo (ou menor que um valor pequeno preestabelecido) abandona-se este vetor e prossegue-se o procedimento renumerando-se os vetores subseqüentes, ao final do processo o número de vetores $\mathbf{u}_k$ não nulos é igual ao posto da matriz. Este procedimento pode ser também aplicado a matrizes não-quadradas.

<u>Exemplos Ilustrativos</u> :Calcular através do processo de ortogonalização de Garm-Schmidt o posto de cada uma das matrizes abaixo:

(a) $\begin{pmatrix} 2 & -3 & 7 \\ 4 & 6 & 2 \\ -4 & 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} -1 & 2 & 3 \\ -2 & 4 & -1 \\ -1 & 2 & -4 \\ -5 & 10 & -6 \end{pmatrix}$; (c) $\begin{pmatrix} -1 & 2 & 3 & 1 \\ -2 & 4 & -1 & -3 \\ -1 & 2 & -4 & -4 \\ -5 & 10 & -6 & -10 \end{pmatrix}$.

(a) utilizando os vetores coluna da matriz, isto é: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix}$; $\mathbf{v}_2 = \begin{pmatrix} -3 \\ 6 \\ 0 \end{pmatrix}$ e $\mathbf{v}_3 = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}$,

tem-se: $\mathbf{u_1} = \mathbf{v}_1 = \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix} \Rightarrow \|\mathbf{u_1}\| = 6$; $\mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v_2} = \mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v}_3 = 18$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v_2^T} \cdot \mathbf{u_1}}{\|\mathbf{u_1}\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} -3 \\ 6 \\ 0 \end{pmatrix} - \frac{18}{6^2}\begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix} = \begin{pmatrix} -4 \\ 4 \\ 2 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 6 \text{ ; } \mathbf{u}_2^{\mathrm{T}} \cdot \mathbf{v}_3 = -18$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v_3^T} \cdot \mathbf{u_1}}{\|\mathbf{u_1}\|^2}\right) \cdot \mathbf{u}_1 - \left(\frac{\mathbf{v_3^T} \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right) \cdot \mathbf{u}_2 = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} - \frac{18}{6^2}\begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix} + \frac{18}{6^2}\begin{pmatrix} -4 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix} \Rightarrow \|\mathbf{u_3}\| = 6$$

como os 3 vetores $\mathbf{u}_1$, $\mathbf{u}_2$ e $\mathbf{u}_3$ são não nulos o posto da matriz é igual a 3.

utilizando os vetores linha da matriz, isto é: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -3 \\ 7 \end{pmatrix}$; $\mathbf{v}_2 = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}$ e $\mathbf{v}_3 = \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix}$,

tem-se: $\mathbf{u_1} = \mathbf{v}_1 = \begin{pmatrix} 2 \\ -3 \\ 7 \end{pmatrix} \Rightarrow \|\mathbf{u_1}\| = \sqrt{62}$; $\mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v_2} = 4; \mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v}_3 = -1$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v_2^T} \cdot \mathbf{u_1}}{\|\mathbf{u_1}\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix} - \frac{4}{\sqrt{62}}\begin{pmatrix} 2 \\ -3 \\ 7 \end{pmatrix} = \begin{pmatrix} 3{,}871 \\ 6{,}194 \\ 1{,}548 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 7{,}466 \text{ ; } \mathbf{u}_2^{\mathrm{T}} \cdot \mathbf{v}_3 = -13{,}935$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v_3^T} \cdot \mathbf{u_1}}{\|\mathbf{u_1}\|^2}\right) \cdot \mathbf{u}_1 - \left(\frac{\mathbf{v_3^T} \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right) \cdot \mathbf{u}_2 = \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{62}}\begin{pmatrix} 2 \\ -3 \\ 7 \end{pmatrix} + \frac{13{,}935}{7{,}466^2}\begin{pmatrix} 3{,}871 \\ 6{,}194 \\ 1{,}548 \end{pmatrix} =$$

$= \begin{pmatrix} -3 \\ 1{,}5 \\ 1{,}5 \end{pmatrix} \Rightarrow \|\mathbf{u_3}\| = \sqrt{13{,}5}$, novamente tem-se os 3 vetores $\mathbf{u}_1$, $\mathbf{u}_2$ e $\mathbf{u}_3$ não nulos e o posto da matriz é igual a 3.

(b) utilizando os vetores coluna da matriz, isto é: $\mathbf{v}_1 = \begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix}$; $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 10 \end{pmatrix}$ e $\mathbf{v}_3 = \begin{pmatrix} 3 \\ -1 \\ -4 \\ -6 \end{pmatrix}$,

$$\mathbf{u_1} = \mathbf{v}_1 = \begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} \Rightarrow \|\mathbf{u_1}\| = \sqrt{31} \text{ ; } \mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v_2} = -62; \mathbf{u}_1^{\mathrm{T}} \cdot \mathbf{v}_3 = 33$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v}_2^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 10 \end{pmatrix} + \frac{62}{31}\begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 0$$

novo: $\mathbf{u}_2 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} 3 \\ -1 \\ -4 \\ -6 \end{pmatrix} + \frac{33}{31}\begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} = \begin{pmatrix} 4,065 \\ 1,129 \\ -2,935 \\ -0,677 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 5{,}184$, como há apenas 2 vetores coluna linearmente independente o posto desta matriz é igual a 2;

utilizando os vetores linha da matriz, isto é :

$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$; $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}$; $\mathbf{v}_3 = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix}$ e $\mathbf{v}_4 = \begin{pmatrix} -5 \\ 10 \\ -6 \end{pmatrix}$, tem-se:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \Rightarrow \|\mathbf{u}_1\| = \sqrt{14};\ \mathbf{u}_1^T \cdot \mathbf{v}_2 = \mathbf{u}_1^T \cdot \mathbf{v}_4 = 7 \text{ e } \mathbf{u}_1^T \cdot \mathbf{v}_3 = -7$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v}_2^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} -2 \\ 4 \\ 1 \end{pmatrix} - \frac{7}{14}\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1,5 \\ 3 \\ -2,5 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = \sqrt{17{,}5};\ \mathbf{u}_2^T \cdot \mathbf{v}_3 = 17{,}5 \text{ e}$$

$$\mathbf{u}_2^T \cdot \mathbf{v}_4 = 52{,}5$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 - \left(\frac{\mathbf{v}_3^T \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right) \cdot \mathbf{u}_2 = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix} + \frac{7}{14}\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} - \frac{17{,}5}{17{,}5}\begin{pmatrix} -1,5 \\ 3 \\ -2,5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \|\mathbf{u}_3\| = 0$$ n

ovo: $\mathbf{u}_3$:

$$\mathbf{u}_3 = \mathbf{v}_4 - \left(\frac{\mathbf{v}_4^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 - \left(\frac{\mathbf{v}_4^T \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right) \cdot \mathbf{u}_2 = \begin{pmatrix} -5 \\ 10 \\ -6 \end{pmatrix} - \frac{7}{14}\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} - \frac{52{,}5}{17{,}5}\begin{pmatrix} -1,5 \\ 3 \\ -2,5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \|\mathbf{u}_3\| = 0$$

desta forma a matriz apresenta apenas 2 vetores linha linearmente independentes reconfirmando que a matriz tem posto = 2;

(c) utilizando os vetores coluna da matriz, isto é:

$\mathbf{v}_1 = \begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix}$; $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 10 \end{pmatrix}$; $\mathbf{v}_3 = \begin{pmatrix} 3 \\ -1 \\ -4 \\ -6 \end{pmatrix}$ e $\mathbf{v}_4 = \begin{pmatrix} 1 \\ -3 \\ -4 \\ -10 \end{pmatrix}$,

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} \Rightarrow \|\mathbf{u}_1\| = \sqrt{31};\ \mathbf{u}_1^T \cdot \mathbf{v}_2 = -62; \mathbf{u}_1^T \cdot \mathbf{v}_3 = 33 \text{ e } \mathbf{u}_1^T \cdot \mathbf{v}_4 = 59$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \left(\frac{\mathbf{v}_2^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 10 \end{pmatrix} + \frac{62}{31}\begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 0$$

novo: $\mathbf{u}_2 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 = \begin{pmatrix} 3 \\ -1 \\ -4 \\ -6 \end{pmatrix} + \frac{33}{31}\begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} = \begin{pmatrix} 4{,}065 \\ 1{,}129 \\ -2{,}935 \\ -0{,}677 \end{pmatrix} \Rightarrow \|\mathbf{u}_2\| = 5{,}184$,

$$\mathbf{u}_2^T \cdot \mathbf{v}_4 = 19{,}194$$

$$\mathbf{u}_3 = \mathbf{v}_4 - \left(\frac{\mathbf{v}_4^T \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right) \cdot \mathbf{u}_1 - \left(\frac{\mathbf{v}_4^T \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right) \cdot \mathbf{u}_2 = \begin{pmatrix} 1 \\ -3 \\ -4 \\ -10 \end{pmatrix} - \frac{59}{31}\begin{pmatrix} -1 \\ -2 \\ -1 \\ -5 \end{pmatrix} - \frac{19{,}194}{5{,}184^2}\begin{pmatrix} 4{,}065 \\ 1{,}129 \\ -2{,}935 \\ -0{,}677 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Rightarrow \|\mathbf{u}_3\| = 0$;como há apenas 2 vetores coluna linearmente independente o posto desta matriz é igual a 2, isto pode ser reconfirmado com os vetores linha da matriz.

4) FUNÇÕES DE MATRIZES

De forma análoga a funções analíticas de variáveis escalares que podem, em um certo domínio, ser expandidas em séries de potências da forma:

$f(x) = \sum_{i=0}^{\infty} c_i \cdot x^i$ onde : $c_i = \frac{1}{i!}\left[\frac{d^i f(x)}{dx^i}\right]_{x=0}$ tem-se as funções de matrizes que é um matriz da forma: $f(\mathbf{A}) = \sum_{i=0}^{\infty} c_i \cdot \mathbf{A}^i$. Como exemplo tem-se a função exponencial de uma matriz **A** definida , em analogia à função $e^x = \sum_{i=0}^{\infty} \frac{1}{i!} \cdot x^i$, pela série:

$\exp(\mathbf{A}) = e^{\mathbf{A}} = \sum_{i=0}^{\infty} \frac{1}{i!} \cdot \mathbf{A}^i$, note que esta função apresenta as propriedades:

i-) exp(**0**) = **I** onde **0** é a matriz nula;

ii-) $\exp(\mathbf{A}t) = e^{\mathbf{A}t} = \sum_{i=0}^{\infty} \frac{t^i}{i!} \cdot \mathbf{A}^i$ onde <u>t</u> é um escalar, assim:

$\frac{d\exp(\mathbf{A}t)}{dt} = \sum_{i=0}^{\infty} \frac{it^{i-1}}{i!} \cdot \mathbf{A}^i = \mathbf{A}\sum_{i=0}^{\infty} \frac{t^i}{i!} \cdot \mathbf{A}^i = \mathbf{A}\exp(\mathbf{A}t)$ ou seja se $\Phi(t) = \exp(\mathbf{A}t)$, tem-se:

$\Phi(0) = \mathbf{I}$ e $\frac{d\Phi(t)}{dt} = \mathbf{A} \cdot \Phi(t)$, ou seja a matriz $\Phi(t)$ é solução da equação diferencial ordinária <u>matricial</u> $\frac{d\Phi(t)}{dt} = \mathbf{A} \cdot \Phi(t)$, sujeita à condição inicial $\Phi(0) = \mathbf{I}$.

Uma forma mais simples para determinar funções de matrizes pode ser desenvolvida através da aplicação do <u>*Teorema de Cayley-Hamilton*</u> que estabelece que todo a matriz quadrada **A** é raiz de seu polinômio característico, isto é se

$p(\lambda) = \lambda^n + c_1 \cdot \lambda^{n-1} + c_2 \cdot \lambda^{n-2} + \cdots + c_{n-1} \cdot \lambda + c_n$ é o polinômio característico de **A**, então:

$p(\mathbf{A}) = \mathbf{A}^n + c_1 \cdot \mathbf{A}^{n-1} + c_2 \cdot \mathbf{A}^{n-2} + \cdots + c_{n-1} \cdot \mathbf{A} + c_n \cdot \mathbf{I} = \mathbf{0}$. A demonstração deste teorema pode ser feita definindo-se a matriz adjunta da matriz $\lambda\mathbf{I}$- $\mathbf{A}$, isto é: $\mathbf{C}$ = adj($\lambda\mathbf{I}$- $\mathbf{A}$) que pode ser expressa na forma: $\mathbf{C} = \mathbf{C}_1\lambda^{n-1} + \mathbf{C}_2\lambda^{n-2} + \cdots + \mathbf{C}_{n-1}\lambda + \mathbf{C}_n$, onde $\mathbf{C}_k$ k= 1, 2, ...,n são matrizes do mesmo tipo de $\mathbf{A}$, mas : ($\lambda\mathbf{I}$- $\mathbf{A}$)[adj($\lambda\mathbf{I}$- $\mathbf{A}$)]=det($\lambda\mathbf{I}$- $\mathbf{A}$)$\mathbf{I}$ = p(λ)$\mathbf{I}$, ou seja:

$$(\lambda\mathbf{I} - \mathbf{A}) \cdot \left(\mathbf{C}_1\lambda^{n-1} + \mathbf{C}_2\lambda^{n-2} + \cdots + \mathbf{C}_{n-1}\lambda + \mathbf{C}_n\right) = \left(\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n\right) \cdot \mathbf{I}$$

igualando os termos eqüipotentes de λ, tem-se:

$$\begin{cases} \mathbf{C}_1 = \mathbf{I} \rightarrow \mathbf{A}^n \cdot \Rightarrow \mathbf{A}^n \cdot \mathbf{C}_1 = \mathbf{A}^n \\ \mathbf{C}_2 - \mathbf{A} \cdot \mathbf{C}_1 = c_1\mathbf{I} \rightarrow \mathbf{A}^{n-1} \cdot \Rightarrow \mathbf{A}^{n-1} \cdot \mathbf{C}_2 - \mathbf{A}^n \cdot \mathbf{C}_1 = c_1\mathbf{A}^{n-1} \\ \mathbf{C}_3 - \mathbf{A} \cdot \mathbf{C}_2 = c_2\mathbf{I} \rightarrow \mathbf{A}^{n-2} \cdot \Rightarrow \mathbf{A}^{n-2} \cdot \mathbf{C}_3 - \mathbf{A}^{n-1} \cdot \mathbf{C}_2 = c_2\mathbf{A}^{n-2} \\ \vdots \\ \mathbf{C}_n - \mathbf{A} \cdot \mathbf{C}_{n-1} = c_{n-1}\mathbf{I} \rightarrow \mathbf{A} \cdot \Rightarrow \mathbf{A} \cdot \mathbf{C}_n - \mathbf{A}^2 \cdot \mathbf{C}_{n-1} = c_{n-1}\mathbf{A} \\ -\mathbf{A} \cdot \mathbf{C}_n = c_n\mathbf{I} \rightarrow \mathbf{A}^0 = \mathbf{I} \cdot \Rightarrow -\mathbf{A} \cdot \mathbf{C}_n = c_n\mathbf{I} \end{cases}$$ somando todos os termos

após $\Rightarrow$ tem-se: $\mathbf{A}^n + c_1 \cdot \mathbf{A}^{n-1} + c_2 \cdot \mathbf{A}^{n-2} + \cdots + c_{n-1} \cdot \mathbf{A} + c_n \cdot \mathbf{I} = p(\mathbf{A}) = \mathbf{0}$.

Uma conseqüência do teorema de Cayley-Hamilton é que:

$\mathbf{A}^n = -c_1 \cdot \mathbf{A}^{n-1} - c_2 \cdot \mathbf{A}^{n-2} - \cdots - c_{n-1} \cdot \mathbf{A} - c_n \cdot \mathbf{I}$, multiplicando membro a membro por **A:**

$\mathbf{A}^{n+1} = -c_1 \cdot \mathbf{A}^n - c_2 \cdot \mathbf{A}^{n-1} - \cdots - c_{n-1} \cdot \mathbf{A}^2 - c_n \cdot \mathbf{A}$ substituindo a expressão de $\mathbf{A}^n$, tem-se:

$\mathbf{A}^{n+1} = \left(c_1^2 - c_2\right) \cdot \mathbf{A}^{n-1} + \left(c_1c_2 - c_3\right) \cdot \mathbf{A}^{n-2} + \cdots + \left(c_1c_{n-1} - c_n\right) \cdot \mathbf{A} + c_1c_n \cdot \mathbf{I}$, e assim sucessivamente, o que permite concluir que :

$\mathbf{A}^m = d_1 \cdot \mathbf{A}^{n-1} + d_2 \cdot \mathbf{A}^{n-2} + \cdots + d_{n-1} \cdot \mathbf{A} + d_n \cdot \mathbf{I}$ para m = 0, 1, 2,..... Além disto se $\mathbf{A}$ é regular, multiplica-se membro a membro de p($\mathbf{A}$) por $\mathbf{A}^{-1}$, resultando em :

$\mathbf{A}^{n-1} + c_1 \cdot \mathbf{A}^{n-2} + c_2 \cdot \mathbf{A}^{n-3} + \cdots + c_{n-1} \cdot \mathbf{I} + c_n \cdot \mathbf{A}^{-1} = \mathbf{0}$ ou seja:

$\mathbf{A}^{-1} = -\dfrac{1}{c_n}\left(\mathbf{A}^{n-1} + c_1 \cdot \mathbf{A}^{n-2} + c_2 \cdot \mathbf{A}^{n-3} + \cdots + c_{n-1} \cdot \mathbf{I}\right)$ { note que c_n =$(-1)^n$det($\mathbf{A}$) $\neq$ 0 pois $\mathbf{A}$ é regular ou não-singular), assim sendo se $\mathbf{A}$ é regular:

$\mathbf{A}^m = d_1 \cdot \mathbf{A}^{n-1} + d_2 \cdot \mathbf{A}^{n-2} + \cdots + d_{n-1} \cdot \mathbf{A} + d_n \cdot \mathbf{I}$ para m = 0, ±1, ±2,......

A aplicação do Teorema de Cayley-Hamilton à série de potências $f(\mathbf{A}) = \sum_{i=0}^{\infty} c_i \cdot \mathbf{A}^i$ permite reescrevê-la na forma: $f(\mathbf{A}) = \sum_{i=0}^{n-1} \alpha_i \cdot \mathbf{A}^i$ pois potências superiores à (n-1) da matriz $\mathbf{A}$ pode, pelo teorema de Cayley-Hamilton, serem expressas em termos das (n-1) primeiras potências da matriz $\mathbf{A}$, além disto de acordo com a propriedade anteriormente apresentada de que se λ é um valor característico e $\mathbf{v}$ o correspondente vetor característico de $\mathbf{A}$, então q(λ) é valor característico e $\mathbf{v}$ o correspondente vetor característico de $q(\mathbf{A}) = \mathbf{A}^m + a_1\mathbf{A}^{m-1} + a_2\mathbf{A}^{m-2} + \cdots + a_{m-1}\mathbf{A} + a_m\mathbf{I}$ tem-se que os valores característicos de f($\mathbf{A}$) satisfazem a: $f(\lambda) = \sum_{i=0}^{n-1} \alpha_i \cdot \lambda^i$, então para determinar os coeficientes α_i , assim procede-se:

(i) se os valores característicos de **A** são todos distintos, resolve-se o sistema linear de equações: $\sum_{i=0}^{n-1} \alpha_i \cdot \lambda_k^i = f(\lambda_k)$ para $k = 1, 2, ...n$;

Se **A** é uma matriz (2,2), α_0 , α_1 é solução de :

$$\begin{cases} \alpha_0 + \alpha_1\lambda_1 = f(\lambda_1) \\ \alpha_0 + \alpha_1\lambda_2 = f(\lambda_2) \end{cases} \Rightarrow \begin{cases} \alpha_0 = \dfrac{\lambda_2 f(\lambda_1) - \lambda_1 f(\lambda_2)}{\lambda_2 - \lambda_1} \\ \alpha_1 = \dfrac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1} \end{cases}$$

caso $\lambda_1 = \lambda_2$ tem-se: $\begin{cases} \alpha_0 = \lim\limits_{\lambda_2 \to \lambda_1} \left[\dfrac{\lambda_2 f(\lambda_1) - \lambda_1 f(\lambda_2)}{\lambda_2 - \lambda_1} \right] = f(\lambda_1) - \lambda_1 f'(\lambda_1) \\ \alpha_1 = \lim\limits_{\lambda_2 \to \lambda_1} \left[\dfrac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1} \right] = f'(\lambda_1) \end{cases}$

o mesmo resultado poderia ser obtido derivando-se a segunda equação do sistema em relação a λ_2 e, em seguida, fazer $\lambda_2 \to \lambda_1$, assim:

$$\begin{cases} \alpha_0 + \alpha_1\lambda_1 = f(\lambda_1) \\ \alpha_1 = f'(\lambda_1) \end{cases} \Rightarrow \begin{cases} \alpha_0 = f(\lambda_1) - \lambda_1 f'(\lambda_1) \\ \alpha_1 = f'(\lambda_1) \end{cases}$$

<u>Exemplos Ilustrativos:</u> (a) para $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$ calcule $\mathbf{A}^{-3}$ e $\ln(\mathbf{A})$;

(b) para $\mathbf{A} = \begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix}$ calcule $\sqrt{\mathbf{A}}$.

(a) $p(\lambda) = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5) \Rightarrow \lambda_1 = 2$ e $\lambda_2 = 5$, assim:

$\begin{cases} \alpha_0 = \dfrac{5f(2) - 2f(5)}{3} \\ \alpha_1 = \dfrac{f(5) - f(2)}{3} \end{cases}$; para $f(x)=x^{-3}$ tem-se: $\begin{cases} \alpha_0 = \dfrac{5/8 - 2/125}{3} = 0{,}203 \\ \alpha_1 = \dfrac{1/125 - 1/8}{3} = -0{,}039 \end{cases}$ logo:

$\mathbf{A}^{-3} = 0{,}203\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 0{,}039\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0{,}047 & -0078 \\ -0{,}039 & 0{,}086 \end{pmatrix}$ e

para f(x)=ln(x), tem-se $\begin{cases} \alpha_0 = \dfrac{5\ln(2) - 2\ln(5)}{3} = 0{,}082287 \\ \alpha_1 = \dfrac{\ln(5) - \ln(2)}{3} = 0{,}30543 \end{cases}$ então:

$\mathbf{B} = \ln(\mathbf{A}) = 0{,}082287\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 0{,}305430\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1{,}304008 & 0{,}610860 \\ 0{,}30543 & 0{,}998577 \end{pmatrix}$, esta última função matricial está correta se a função inversa também é verdadeira, isto é : **A** = exp(**B**), para isto deve-se inicialmente determinar os valores característicos de **B** que são

η_1=0,693147 e η_2=1,609438 determina-se a seguir: $\begin{cases} \beta_0 = \dfrac{2\eta_2 - 5\eta_1}{\eta_2 - \eta_1} = -0{,}269412 \\ \beta_1 = \dfrac{5-2}{\eta_2 - \eta_1} = 3{,}27407 \end{cases}$ e então:

$$\exp(\mathbf{B}) = -0{,}269412\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 3{,}27407\begin{pmatrix} 1{,}304008 & 0{,}610860 \\ 0{,}30543 & 0{,}998577 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} = \mathbf{A}$$

(b) $p(\lambda) = \lambda^2 - 3\lambda - 4 = (\lambda+1)(\lambda-4) \Rightarrow \lambda_1 = -1$ e $\lambda_2 = 4$ então como $f(x) = \sqrt{x}$, tem-se: $f(\lambda_1) = \sqrt{-1} = \mathbf{i}$ e $f(\lambda_2) = \sqrt{4} = 2$ logo: $\begin{cases} \alpha_0 = \dfrac{4\mathbf{i}+2}{5} = 0{,}4 + 0{,}8\mathbf{i} \\ \alpha_1 = \dfrac{2-\mathbf{i}}{5} = 0{,}4 - 0{,}2\mathbf{i} \end{cases}$ e

$$\sqrt{\mathbf{A}} = (0{,}4+0{,}8\mathbf{i})\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (0{,}4-0{,}2\mathbf{i})\begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 2{,}4-0{,}2\mathbf{i} & -1{,}2+0{,}6\mathbf{i} \\ 0{,}8-0{,}4\mathbf{i} & -0{,}4+1{,}2\mathbf{i} \end{pmatrix}$$ note que:

$$\left(\sqrt{\mathbf{A}}\right)^2 = \begin{pmatrix} 2{,}4-0{,}2\mathbf{i} & -1{,}2+0{,}6\mathbf{i} \\ 0{,}8-0{,}4\mathbf{i} & -0{,}4+1{,}2\mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} 2{,}4-0{,}2\mathbf{i} & -1{,}2+0{,}6\mathbf{i} \\ 0{,}8-0{,}4\mathbf{i} & -0{,}4+1{,}2\mathbf{i} \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix} = \mathbf{A}$$

(ii) se a matriz **A** apresenta valores característicos múltiplos, por exemplo, $\lambda_1 = \lambda_2 = \cdots = \lambda_m \neq \lambda_{m+1} \neq \cdots \neq \lambda_n$, neste caso para levantar a indeterminação no cálculo dos coeficientes α_i, deriva-se em relação a λ_1 *m* vezes a equação correspondente a λ_1, assim: $\sum_{i=0}^{n-1} \alpha_i \cdot \lambda_1^i = f(\lambda_1)$; $\sum_{i=1}^{n-1} i\alpha_i \cdot \lambda_1^{i-1} = f'(\lambda_1)$; $\sum_{i=2}^{n-1} i(i-1)\alpha_i \cdot \lambda_1^{i-2} = f''(\lambda_1)$; ....;

$\sum_{i=m}^{n-1} i(i-1)\cdots(i-m+1)\alpha_i \cdot \lambda_1^{i-m} = \left.\dfrac{d^m f(\lambda)}{d\lambda^m}\right|_{\lambda_1}$ sendo as demais equações:

$\sum_{i=0}^{n-1} \alpha_i \cdot \lambda_k^i = f(\lambda_k)$ para $k = m+1, m+2, ...n$.

Exemplo Ilustrativo: para $\mathbf{A} = \begin{pmatrix} -1{,}750 & -2{,}000 & -0{,}250 \\ -0{,}125 & -2{,}000 & -0{,}375 \\ 0{,}250 & 2{,}000 & -0{,}250 \end{pmatrix}$ calcule exp(**A**);

$$p(\lambda) = det\begin{pmatrix} \lambda+1{,}750 & 2{,}000 & 0{,}250 \\ 0{,}125 & \lambda+2{,}000 & 0{,}375 \\ -0{,}250 & -2{,}000 & \lambda+0{,}250 \end{pmatrix} = \lambda^3 + 4\lambda^2 + 5\lambda + 2 = 0 \Rightarrow \lambda_1 = \lambda_2 = -1$$

e $\lambda_3 = -2$, assim:

$\begin{cases} \alpha_0 - \alpha_1 + \alpha_2 = \exp(-1) \\ \alpha_1 - 2\alpha_2 = \exp(-1) \\ \alpha_0 - 2\alpha_1 + 4\alpha_2 = \exp(-2) \end{cases}$ tem-se assim: $\begin{cases} \alpha_0 = 0{,}871094 \\ \alpha_1 = 0{,}638550 \\ \alpha_2 = 0{,}135335 \end{cases}$ logo:

$$\exp(\mathbf{A}) = 0{,}871094\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 0{,}638550\begin{pmatrix} -1{,}750 & -2{,}000 & -0{,}250 \\ -0{,}125 & -2{,}000 & -0{,}375 \\ 0{,}250 & 2{,}000 & -0{,}250 \end{pmatrix} + 0{,}135335\begin{pmatrix} -1{,}750 & -2{,}000 & -0{,}250 \\ -0{,}125 & -2{,}000 & -0{,}375 \\ 0{,}250 & 2{,}000 & -0{,}250 \end{pmatrix}^2 =$$

$$= \begin{pmatrix} 0{,}193471 & -1{,}065512 & -0{,}358348 \\ -0{,}029068 & 0{,}435547 & 0{,}062902 \\ 0{,}058136 & -0{,}135335 & -0{,}242076 \end{pmatrix}$$

Caso desejar-se determinar uma série de potências $f(\mathbf{A}t) = \sum_{i=0}^{\infty} c_i \cdot t^i \cdot \mathbf{A}^i = \boldsymbol{\Psi}(t)$ onde a variável t é uma variável escalar real, esta função pode ser rescrita na forma: $\boldsymbol{\Psi}(t) = \sum_{i=0}^{n-1} \left[\alpha_i(t)\right] \cdot \mathbf{A}^i$ onde $\alpha_i(t)$ é uma função escalar de t determinada através da solução de:

(i) $\sum_{i=0}^{n-1} \alpha_i(t) \cdot \lambda_k^i = f(\lambda_k t)$ para k = 1, 2, ...n se os valores característicos de **A** são todos distintos;

(ii) $$\sum_{i=0}^{n-1} \alpha_i(t) \cdot \lambda_1^i = f(\lambda_1 t) \;;$$

$$\sum_{i=1}^{n-1} i\alpha_i(t) \cdot \lambda_1^{i-1} = t\left[\frac{df(\xi)}{d\xi}\right]_{\xi=\lambda_1 t} \;;$$

$$\sum_{i=2}^{n-1} i(i-1)\alpha_i(t) \cdot \lambda_1^{i-2} = t^2\left[\frac{d^2 f(\xi)}{d\xi^2}\right]_{\xi=\lambda_1 t} \;;$$

$$\vdots\;\vdots\;\vdots$$

$$\sum_{i=m}^{n-1} i(i-1)\cdots(i-m+1)\alpha_i(t) \cdot \lambda_1^{i-m} = t^m\left[\frac{d^m f(\xi)}{d\xi^m}\right]_{\xi=\lambda_1 t}$$

sendo as demais equações:

$\sum_{i=0}^{n-1} \alpha_i(t) \cdot \lambda_k^i = f(\lambda_k t)$ para $k = m$+1, m+2, ...n. se

$\lambda_1 = \lambda_2 = \cdots = \lambda_m \neq \lambda_{m+1} \neq \cdots \neq \lambda_n$

<u>Exemplos Ilustrativos:</u> (a) para $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$ calcule exp(**At**);

(b) para $\mathbf{A} = \begin{pmatrix} -3 & -2 \\ 2 & 1 \end{pmatrix}$ calcule. exp(**At**).

(a) $p(\lambda) = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5) \Rightarrow \lambda_1 = 2$ e $\lambda_2 = 5$, assim:

$$\begin{cases} \alpha_0 = \dfrac{5e^{2t} - 2e^{5t}}{3} \\ \alpha_1 = \dfrac{e^{5t} - e^{2t}}{3} \end{cases} \text{; logo: } \exp(\mathbf{A}t) = \left(\frac{5e^{2t} - 2e^{5t}}{3}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{e^{5t} - e^{2t}}{3}\right)\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} \text{ então:}$$

$$\frac{d[\exp(\mathbf{A}t)]}{dt} = 10\left(\frac{e^{2t} - e^{5t}}{3}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{5e^{5t} - 2e^{2t}}{3}\right)\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} \text{ e}$$

$$\mathbf{A}[\exp(\mathbf{A}t)] = \left(\frac{5e^{2t} - 2e^{5t}}{3}\right)\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} + \left(\frac{e^{5t} - e^{2t}}{3}\right)\left\{7\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} - 10\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\} =$$

$$= 10\left(\frac{e^{2t} - e^{5t}}{3}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{5e^{5t} - 2e^{2t}}{3}\right)\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix} = \frac{d[\exp(\mathbf{A}t)]}{dt} \text{ e}$$

$\exp(\mathbf{A}0)=\mathbf{I}$, comprovando que esta matriz exponencial está correta.

(b) $p(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 \Rightarrow \lambda_1 = \lambda_2 = -1$ deve-se assim resolver o sistema:

$$\begin{cases} \alpha_0 + \alpha_1\lambda_1 = e^{\lambda_1 t} \\ \alpha_1 = te^{\lambda_1 t} \end{cases} \Rightarrow \begin{cases} \alpha_0 - \alpha_1 = e^{-t} \\ \alpha_1 = te^{-t} \end{cases} \Rightarrow \begin{cases} \alpha_0 = (1+t)e^{-t} \\ \alpha_1 = te^{-t} \end{cases} \Rightarrow \begin{cases} \dfrac{d\alpha_0}{dt} = -te^{-t} \\ \dfrac{d\alpha_1}{dt} = (1-t)e^{-t} \end{cases} \text{ assim:}$$

$$\exp(\mathbf{A}t) = (1+t)e^{-t}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + te^{-t}\begin{pmatrix} -3 & -2 \\ 2 & 3 \end{pmatrix};\ \exp(\mathbf{A}t)\Big|_{t=0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ e}$$

$$\frac{d\big[\exp(\mathbf{A}t)\big]}{dt} = -te^{-t}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-t)e^{-t}\begin{pmatrix} -3 & -2 \\ 2 & 3 \end{pmatrix} \text{ e}$$

$$\mathbf{A}[\exp(\mathbf{A}t)] = (1+t)e^{-t}\begin{pmatrix} -3 & -2 \\ 2 & 3 \end{pmatrix} - te^{-t}\left\{2\begin{pmatrix} -3 & -2 \\ 2 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\} = \frac{d[\exp(\mathbf{A}t)]}{dt},$$

comprovando que a matriz exponencial está correta.

5) FORMAS QUADRÁTICAS

Em $\Re^2$ a expressão geral das formas quadráticas é:

$$f(x_1, x_2) = c + b_1 \cdot x_1 + b_2 \cdot x_2 + \frac{a_{11}}{2} \cdot x_1^2 + a_{12} \cdot x_1 \cdot x_2 + \frac{a_{22}}{2} \cdot x_2^2,$$

cujas derivadas parciais são:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = b_1 + a_{11} \cdot x_1 + a_{12} \cdot x_2 \quad \text{e} \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = b_2 + a_{22} \cdot x_2 + a_{12} \cdot x_1$$

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} = a_{11};\quad \frac{\partial f(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\partial f(x_1, x_2)}{\partial x_2 \partial x_1} = a_{12} \quad \text{e} \quad \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} = a_{22}.$$

Esta forma quadrada pode ser rescrita em forma matricial, segundo:

$f(x_1, x_2) = c + (b_1 \quad b_2) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2} \cdot (x_1 \quad x_2) \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, ou seja, definindo:

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \Re^2$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \Re^2$ e $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \in \Re^{2x2}$ [matriz simétrica], tem-se:

$$f(\mathbf{x}) = c + \mathbf{b}^T \cdot \mathbf{x} + \frac{1}{2} \cdot \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x}$$

definindo o operador diferencial vetorial : $\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix}$ (operador *gradiente*) , tem-se:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(x_1, x_2)}{\partial x_1} \\ \frac{\partial f(x_1, x_2)}{\partial x_2} \end{pmatrix} = \mathbf{b} + \mathbf{A} \cdot \mathbf{x}$$ (vetor gradiente de uma função escalar f) e

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f(x_1, x_2)}{\partial x_1} \\ \frac{\partial f(x_1, x_2)}{\partial x_2} \end{pmatrix} = \frac{\partial^2 f(x_1, x_2)}{\partial^2 x_1} + \frac{\partial^2 f(x_1, x_2)}{\partial^2 x_2} = a_{11} + a_{22} = tr(\mathbf{A})$$

(Laplaciano de uma função escalar).

Define-se também a *matriz Hessiana* por:

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f(x_1, x_2)}{\partial x_1} & \frac{\partial f(x_1, x_2)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f(x_1, x_2)}{\partial^2 x_1} & \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x_1, x_2)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_1, x_2)}{\partial^2 x_2} \end{pmatrix} = \mathbf{A}.$$

Estas definições podem ser generalizadas para $\Re^n$, segundo:

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \Re^n$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \Re^n$ e $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \in \Re^{nxn}$ [matriz simétrica],

tem-se: $f(\mathbf{x}) = c + \mathbf{b}^T \cdot \mathbf{x} + \frac{1}{2} \cdot \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} = c + \sum_{i=1}^{n} b_i \cdot x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot x_i \cdot x_j$,

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} = \mathbf{b} + \mathbf{A} \cdot \mathbf{x},$$

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \cdots & \dfrac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} = a_{11} + a_{22} + \cdots + a_{nn} = \mathrm{tr}(\mathbf{A})$$

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \dfrac{\partial}{\partial x_2} \\ \vdots \\ \dfrac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} & \dfrac{\partial f(\mathbf{x})}{\partial x_2} & \cdots & \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial^2 x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial^2 x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial^2 x_n} \end{pmatrix} = \mathbf{A}$$

$H_{ij}(\mathbf{x}) = \dfrac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \dfrac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i} = H_{ji}(\mathbf{x})$ [matriz simétrica]. Note que caso a matriz **A** não seja simétrica redefinem-se seus elementos na forma:

$a_{ij,velha} = \left(\dfrac{a_{ij,velha} + a_{ji,velha}}{2} \right)$ ou, em termos matriciais, $\mathbf{A}_{nova} = \dfrac{1}{2} \cdot \left(\mathbf{A}_{velha} + \mathbf{A}_{velha}^T \right)$

A forma quadrática acima pode ser simplificada, através de um translação do eixo, tal que o termo $\mathbf{b}^T\mathbf{x}$ desapareçam, assim sejam as novas coordenadas (y_1 , y_2, ..., y_n) tais que: **x=y+d** assim: $\mathbf{b}^T \cdot \mathbf{x} = \mathbf{b}^T \cdot \mathbf{y} + \mathbf{b}^T \cdot \mathbf{d}$ e

$$\mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} = \left(\mathbf{y}^T + \mathbf{d}^T\right) \cdot (\mathbf{A} \cdot \mathbf{y} + \mathbf{A} \cdot \mathbf{d}) = \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y} + \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{d} + \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{y} + \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{d} =$$

$= \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y} + 2\mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{y} + \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{d}$ pois : $\mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{d} = \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{y}$ [**A** é simétrica], logo:

$$f(\mathbf{y}) = c + \mathbf{b}^T \cdot \mathbf{y} + \mathbf{b}^T \cdot \mathbf{d} + \frac{1}{2} \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y} + \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{y} + \frac{1}{2} \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{d}$$

identificando : $f(\mathbf{d}) = c + \mathbf{b}^T \cdot \mathbf{d} + \dfrac{1}{2} \cdot \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{d} = \hat{c}$ e definindo $\hat{\mathbf{b}} = \mathbf{b} + \mathbf{A} \cdot \mathbf{d}$

$f(\mathbf{y}) = \hat{c} + \hat{\mathbf{b}}^T \cdot \mathbf{y} + \frac{1}{2} \cdot \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y}$, adotando **d** tal que : $\hat{\mathbf{b}} = \mathbf{b} + \mathbf{A} \cdot \mathbf{d} = \mathbf{0} \Rightarrow \mathbf{d} = \mathbf{A}^{-1} \cdot \mathbf{b}$ o que só será possível se **A** for regular, assim chega-se a:

$f(\mathbf{y}) = \hat{c} + \frac{1}{2} \cdot \mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y}$ onde : **x=y+d** , $\mathbf{d} = \mathbf{A}^{-1} \cdot \mathbf{b}$ e $\hat{c} = f(\mathbf{d}) = c + \mathbf{b}^T \cdot \mathbf{d} + \frac{1}{2} \cdot \mathbf{d}^T \cdot \mathbf{A} \cdot \mathbf{d}$, neste novo sistema de coordenadas tem-se:

$$\nabla f(\mathbf{y}) = \begin{pmatrix} \frac{\partial f(\mathbf{y})}{\partial y_1} \\ \frac{\partial f(\mathbf{y})}{\partial y_2} \\ \vdots \\ \frac{\partial f(\mathbf{y})}{\partial y_n} \end{pmatrix} = \mathbf{A} \cdot \mathbf{y},$$

neste novo sistema de coordenadas o valor da variável independente **y** que anula o vetor gradiente é o valor nulo, isto é a origem : **y=0** e neste ponto o valor da função f(**y**) é igual a : $\hat{c}$. Esta condição, $\nabla f(\mathbf{y})$ =**0** é uma condição necessária para o ponto ser um extremo da função (máximo ou mínimo) e é chamado de *ponto crítico*, este ponto será um ponto de mínimo se para qualquer vizinhança de **y =0** , isto é : $\|\mathbf{y}\| \le \delta$, a função é f(**y**) > f(**0**) = $\hat{c}$, ou seja : $\mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y} > 0$ e, neste caso, a matriz **A** é chamada de *positiva definida* e caso em toda vizinhança de **y=0** a é f(**y**) < f(**0**) = $\hat{c}$, ou seja : $\mathbf{y}^T \cdot \mathbf{A} \cdot \mathbf{y} < 0$ e, neste caso, a matriz **A** é chamada de *negativa definida* e o ponto é um ponto de máximo. Em qualquer outra situação o ponto não é nem de máximo nem de mínimo, e no caso da matriz ser não definida tem-se o chamado *ponto de sela*.

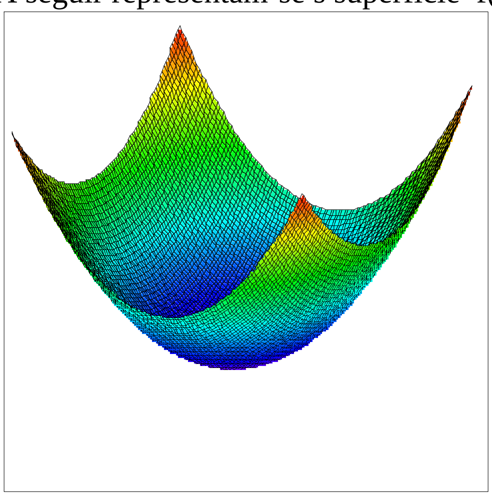
A forma quadrática pode também ser rescrita em sua *forma canônica*, de forma análoga à apresentada no processo de diagonalização de matrizes, assim considerando $\mathbf{y} = \mathbf{P} \cdot \mathbf{z}$, onde **P** é a matriz cujos vetores coluna são os vetores característicos normalizados de **A** (por enquanto considerados n vetores característicos linearmente independentes e ortogonais entre si, isto é os valores característicos são todos reais e distintos - matriz **A** é simétrica), tem-se assim:

$f(\mathbf{z}) = \hat{c} + \frac{1}{2} \mathbf{z}^T \cdot \left(\mathbf{P}^T \cdot \mathbf{A} \cdot \mathbf{P}\right) \cdot \mathbf{z} = \hat{c} + \frac{1}{2} \mathbf{z}^T \cdot \mathbf{D} \cdot \mathbf{z} = \hat{c} + \frac{1}{2} \cdot \sum_{i=1}^{n} \lambda_i \cdot z_i^2$, como à origem **y=0** correspondente também a **z=0**, tem-se **z=0** como ponto de mínimo se $\sum_{i=1}^{n} \lambda_i \cdot z_i^2 > 0$ para todo o domínio em que **z** ≠ **0** se $\lambda_i > 0$ para todo i = 1, $\cdots$,n , **z=0** é um ponto de máximo se $\sum_{i=1}^{n} \lambda_i \cdot z_i^2 < 0$ para todo o domínio em que **z** ≠ **0** se $\lambda_i < 0$ para todo i = 1, $\cdots$,n e **z=0** é um ponto de sela se não há vizinhança de **z=0** na qual $\sum_{i=1}^{n} \lambda_i \cdot z_i^2$ não muda de sinal o que ocorre se alguns $\lambda_i < 0$ e os demais são $\lambda_i > 0$. No $\Re^2$ a forma canônica assume a forma:

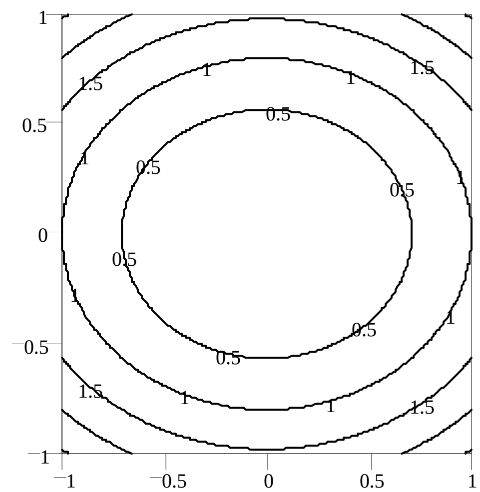
$f(z_1, z_2) = \hat{c} + \frac{1}{2}\left(\lambda_1 \cdot z_1^2 + \lambda_2 \cdot z_2^2\right)$, neste caso a forma das curvas de nível caracterizam as seguinte *cônicas*, de acordo com o sinal de Δ = det (**A**) = $a_{11} \cdot a_{22} - a_{12}^2 = \lambda_1 \cdot \lambda_2$, asssim com $\Delta > 0$: *elipse*; $\Delta < 0$: *hipérbole* e $\Delta = 0$: *parábola* .

(a) Elipse: neste caso os valores característicos têm o mesmo sinal, sendo **z=0** um ponto de mínimo se ambos forem positivos e um ponto de máximo se ambos forem positivos. O tamanho do eixo z_1 é $\frac{2}{\lambda_1}\left[K - \hat{c}\right]$ e do eixo z_2 é $\frac{2}{\lambda_2}\left[K - \hat{c}\right]$, onde $K=f(z_1,z_2)$ [verificando que se **z=0** é um ponto de mínimo $K > \hat{c}$, $\lambda_1 > 0$ e $\lambda_2 > 0$ e se **z=0** é um ponto de máximo $K < \hat{c}$, $\lambda_1 < 0$ e $\lambda_2 < 0$, deste modo em ambos os casos: $\frac{2}{\lambda_1}\left[K - \hat{c}\right] > 0$ e $\frac{2}{\lambda_2}\left[K - \hat{c}\right] > 0$].

A seguir representam-se s superfície $f(z_1,z_2)$ e as correspondentes curvas de contorno:



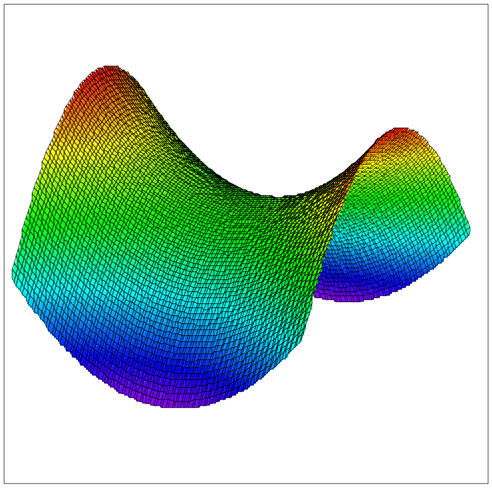
M



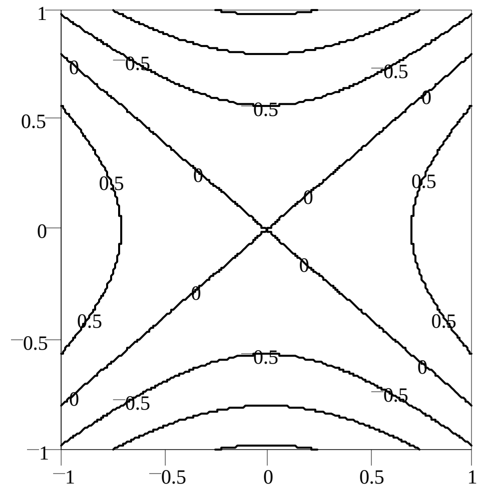
M

$$f(z_1, z_2) = z_1^2 + \left(5 \cdot z_2\right)^2$$

(b) Hipérbole: neste caso os valores característicos têm os sinais distintos, sendo **z=0** um ponto de sela Abaixo representam-se a superfície $f(z_1,z_2)$ e as correspondentes curvas de contorno



M



M

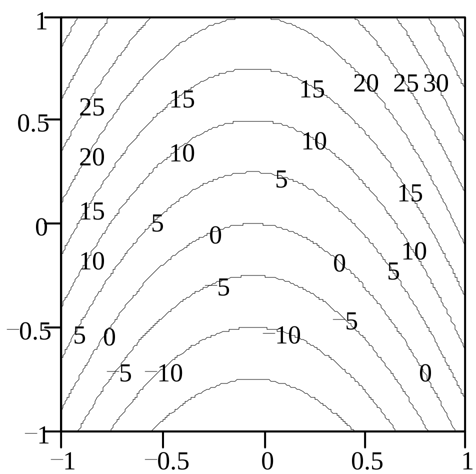
$$f(z_1, z_2) = z_1^2 - (5 \cdot z_2)^2$$

(c) Parábola: neste caso um dos valores característicos é nulo e portanto a matriz **A** é singular, desta forma não é possível fazer a translação de eixo que elimina o termo $\mathbf{b}^T\mathbf{x}$. Então neste caso a rotação dos eixos é aplicada diretamente às variáveis (x_1 , x_2), isto é : $\mathbf{x} = \mathbf{P} \cdot \mathbf{z}$, obtendo-se :

$$f(z_1, z_2) = c + \tilde{b}_1 \cdot z_1 + \tilde{b}_2 \cdot z_2 + \frac{\lambda_1}{2} \cdot z_1^2 \text{ se } \lambda_2 = 0 \text{ ou:}$$

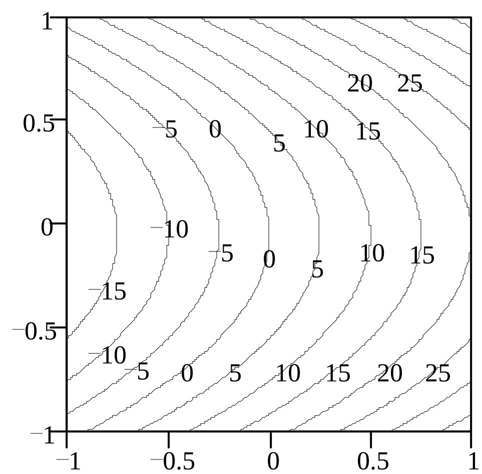
$f(z_1, z_2) = c + \tilde{b}_1 \cdot z_1 + \tilde{b}_2 \cdot z_2 + \frac{\lambda_2}{2} \cdot z_2^2$ se $\lambda_1 = 0$, onde $\mathbf{b} = \mathbf{P} \cdot \tilde{\mathbf{b}} \Leftrightarrow \tilde{\mathbf{b}} = \mathbf{P} \cdot \mathbf{b}$. Verifica-se assim que a condição necessária não é obtida em nenhum dos casos, pois no primeiro caso tem-se: $\nabla f(\mathbf{z}) = \begin{pmatrix} \frac{\partial f(z_1, z_2)}{\partial z_1} \\ \frac{\partial f(z_1, z_2)}{\partial z_2} \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 + \lambda_1 \cdot z_1 \\ \tilde{b}_2 \end{pmatrix}$ segundo componente não nulo,

e no segundo caso tem-se: $\nabla f(\mathbf{z}) = \begin{pmatrix} \frac{\partial f(z_1, z_2)}{\partial z_1} \\ \frac{\partial f(z_1, z_2)}{\partial z_2} \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 + \lambda_2 \cdot z_2 \end{pmatrix}$ primeiro componente não nulo. Deste modo em ambos os casos um dos componentes do vetor gradiente é constante não podendo ser anulado através da escolha de z_1 ou z_2, neste caso não se tem nem máximo nem mínimo. Abaixo, representa-se curvas de nível para cada um dos casos.



M
curvas de nível (um dos val. caract. =0)

$f(z_1, z_2) := z_1 + 2 \cdot z_2 + 5 \cdot (z_2)^2$



M
curvas de nível (um dos val. caract. =0)

$f(z_1, z_2) := z_2 + 2 \cdot z_1 + 5 \cdot (z_1)^2$

Lista de Exercícios

1) Mostre que todo matriz ortogonal apresenta o determinante +1 ou -1. *Sugestão*: parta dos princípios que det(**A**)=det($\mathbf{A}^{\mathbf{T}}$) e que det($\mathbf{A}^{\mathbf{-1}}$)=1/det(**A**).

2) Mostre que $(\mathbf{A.B})^{\mathbf{T}}=\mathbf{B}^{\mathbf{T}}.\mathbf{A}^{\mathbf{T}}$

3) Mostre que $(\mathbf{A.B})^{\mathbf{-1}}=\mathbf{B}^{\mathbf{-1}}.\mathbf{A}^{\mathbf{-1}}$

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