Revisão de Álgebra Linear

Já discutimos anteriormente que neste curso iremos abordar apenas a "teoria de inversão discreta", ou seja, a parte da teoria que trabalha com parâmetros que são genuinamente discretos ou podem ser adequadamente aproximados como discretos.

Então neste curso usaremos alguns conceitos básicos da Álgebra Linear. Vamos agora apresentar de modo condensado alguns dos principais conceitos de Álgebra Linear que iremos empregar neste curso. Ressalto que o material apresentado a seguir não significa uma revisão completa de Álgebra Linear.

Espaço Vetorial Linear

Um dos conceitos mais importantes que iremos utilizar neste curso é o conceito de "Espaço Vetorial Linear". Em situações práticas nós estamos bem familiarizados com o conceito de Espaço vetorial linear porque já trabalhamos com conceitos concretos de vetores geométricos no espaço físico tridimensional.

Para discutir vetores geometricamente em um espaço 3D escolhemos em primeiro lugar um sistema de 3 eixos perpendiculares (x, y, z). Um vetor é representado por uma reta OP, em que O é um ponto na origem do sistema e o ponto P é especificado pelas coordenadas x_1, y_1, z_1. Um vetor geométrico $\bar{\mathbf{u}}$ pode ser representado como $\bar{\mathbf{u}}$ = (x_1, y_1, z_1) para indicar as componentes de $\bar{\mathbf{u}}$, ou seja, x_1, y_1, z_1 relativo ao sistema de coordenadas(x, y, z).

Inicialmente, os matemáticos introduziram a idéia de usar um par de números para localizar um ponto em um plano e uma trinca de números para localizar um ponto no espaço tridimensional. Muito mais tarde, fim do século VX, os matemáticos e físicos começaram a definir que quarto números poderiam localizar um ponto no espaço de 4 dimensões e cinco números poderiam localizar um ponto no espaço de 5 dimensões e assim por diante. Embora nossa visualização geométrica não se estenda além do espaço

3D é, no entanto, possível estender várias idéias além do espaço 3D trabalhando-se com propriedades analíticas ou numéricas dos pontos e vetores ao invés das propriedades geométricas que requerem uma visualização. Para tornar esta idéia mais precisa vamos a definição abaixo.

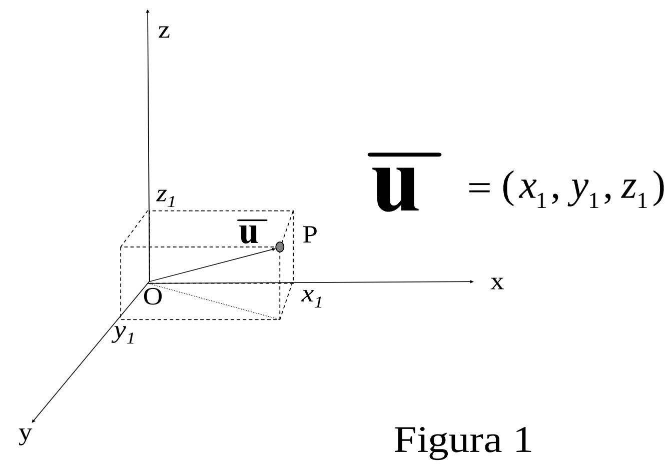


Figura 1

<u>Vetores no Espaço N:</u> Se N é um inteiro positivo, então N-uplas de coordenadas é uma seqüência de N números reais ($a_1, a_2, ..., a_N$). O conjunto de todas as N-uplas de coordenadas é chamado de espaço N e é denotado por R^N.

Já vimos anteriormente no estudo do espaço tridimensional que a tripla de números (x_1, y_1, z_1) na Figura 1 tem duas interpretações geométricas: i) pode ser interpretada como um ponto , em que (x_1, y_1, z_1) são as coordenadas deste ponto; ii) pode ser interpretada como um vetor, em que (x_1, y_1, z_1) são as componentes do vetor. Portanto, N-uplas de coordenadas ($a_1, a_2, ..., a_N$).pode tanto ser interpretada como um ponto em R^N como também um vetor em R^N. Óbvio que a visualização geométrica de um ponto ou vetor em R^N é impossível, no entanto, as propriedades aritméticas de adição de vetores e multiplicação de vetores por um escalar são propriedades padrão em R^N. Apresentaremos a seguir estas propriedades (axiomas) de operação no espaço N-dimensional.

Propriedades de operações Vetoriais no Espaço N:

Se $\mathbf{u} = (u_1, u_2, ... u_N)$, $\overline{\mathbf{v}} = (v_1, v_2, ... v_N)$ e $\overline{\mathbf{w}} = (w_1, w_2, ... w_N)$ são vetores no espaço R^N e k e β são escalares então

a) A adição é comutativa:

$$\overline{\mathbf{u}} + \overline{\mathbf{v}} = \overline{\mathbf{v}} + \overline{\mathbf{u}}$$

b) A adição é associativa:

$$\overline{\mathbf{u}} + (\overline{\mathbf{v}} + \overline{\mathbf{w}}) = (\overline{\mathbf{u}} + \overline{\mathbf{v}}) + \overline{\mathbf{w}}$$

c) Existe em R^N um único vetor $\overline{\mathbf{0}} = (0,0,...,0)$ tal que:

$$\overline{\mathbf{u}} + \overline{\mathbf{0}} = \overline{\mathbf{u}}$$

d) Para cada vetor $\overline{\mathbf{u}} \in R^N$, existe um único vetor $-\overline{\mathbf{u}} \in R^N$, tal que:

$$\overline{\mathbf{u}} + (\overline{\mathbf{u}}-) = \overline{\mathbf{0}}$$

e) A multiplicação de um vetor $\overline{\mathbf{u}} \in R^N$, por escalares é uma operação associativa

$$\kappa(\beta\ \overline{\mathbf{u}}) = (\kappa\beta)\ \overline{\mathbf{u}}$$

f) A multiplicação de um vetor $\overline{\mathbf{u}} \in R^N$, por escalares é uma operação distributiva relativamente à adição de escalares

$$(\kappa + \beta)\ \overline{\mathbf{u}} = \kappa\,\overline{\mathbf{u}} + \beta\,\overline{\mathbf{u}}$$

g) A multiplicação por um escalar é uma operação distributiva relativamente à adição vetorial

$$\kappa\,(\overline{\mathbf{u}} + \overline{\mathbf{v}}) = \kappa\,\overline{\mathbf{u}} + \kappa\,\overline{\mathbf{v}}$$

h) Existe em R^N um único escalar 1 tal que:

$$1\ \overline{\mathbf{u}} = \overline{\mathbf{u}}$$

Estes axiomas acima nos permite manipular com vetores $\in R^N$ sem a necessidade de expressa-los em termos de componentes, ou seja, sem a necessidade de uma visualização geométrica.

Espaço Vetorial Linear Abstrato

Primeiro descrevemos vetores e espaços vetoriais como entidades geométricas. Em seguida, estendemos a idéia de vetores e espaços vetoriais além dos espaço 3D e apresentamos axiomas de operações vetoriais no espaço N-dimensional. Tal axioma permitiu-nos trabalhar com vetores no espaço R^N sem requerer o uso de uma visualização geométrica.

No entanto, o conceito de espaço vetorial é muito mais amplo e definido usando-se conceitos abstratos da entidade "vetor". A este espaço vetorial abstrato foi estabelecido um conjunto de axiomas e se tais axiomas são obedecidos por uma classe de objetos então chamaremos este objetos de vetores que pertencem a um espaço vetorial linear. Vale ressaltar que o conceito de vetores como objetos de um espaço vetorial abstrato é uma espécie de generalização do conceito de um vetor até aqui realizado. Estes novos axiomas que serão apresentados a seguir são abstrações das propriedades importantes dos vetores em R^N, como conseqüência, vetores em R^N, automaticamente satisfazem estes axiomas.

Então, nosso novo conceito de vetores como objetos de um espaço vetorial irá incluir o velho conceito de vetor e vários outros novos tipos de vetores, dentre eles vários tipos de matrizes e funções.

Portanto, a definição de espaço vetorial linear abstrato de vetores é suficientemente genérica sendo, portanto, um poderoso instrumento para estender o nosso limitado poder de visualização geométrica a uma ampla variedade de problemas matemáticos. Vejamos agora os axiomas.

Axiomas do Espaço Vetorial Linear Abstrato:

Seja V um conjunto arbitrário e não vazio de objetos em que duas operações são definidas adição e multiplicação por escalares (conjunto de números) que formam um corpo F. Entenda a operação de adição como uma regra que associa a cada par de objetos $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$ em V um objeto $\bar{\mathbf{u}}+\bar{\mathbf{v}}$, chamado soma de $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$. Entenda a operação de multiplicação como uma regra que associa a cada escalar k em F e a cada objeto $\bar{\mathbf{u}}$ em V um objeto $k\ \bar{\mathbf{u}}$, chamado de produto escalar ou multiplicação escalar. Se os axiomas a seguir são satisfeitos por todos os objetos $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ e $\bar{\mathbf{w}}$ em V e todos os escalares k e β em F, então chamaremos V de espaço vetorial e chamaremos todos os objetos ($\bar{\mathbf{u}}, \bar{\mathbf{v}}$ e $\bar{\mathbf{w}}$) em V de vetores.

a. Se $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$ são objetos em V, então $\bar{\mathbf{u}}+\bar{\mathbf{v}}$ e
b. A adição é comutativa: $\bar{\mathbf{u}}+\bar{\mathbf{v}}=\bar{\mathbf{v}}+\bar{\mathbf{u}}$
c. A adição é associativa: $\bar{\mathbf{u}}+(\bar{\mathbf{v}}+\bar{\mathbf{w}})=(\bar{\mathbf{u}}+\bar{\mathbf{v}})+\bar{\mathbf{w}}$
d. Existe em V um único objeto $\bar{\mathbf{0}}$ chamado de vetor zero para V tal que: $\bar{\mathbf{u}}+\bar{\mathbf{0}}=\bar{\mathbf{u}}$
e. Para cada objeto $\bar{\mathbf{u}}$ em V, existe um objeto $-\bar{\mathbf{u}}$, em V tal que: $\bar{\mathbf{u}}+(\bar{\mathbf{u}}-)=\bar{\mathbf{0}}$
f. Se k é um escalar e $\bar{\mathbf{u}}$ um objeto em V, então $\kappa\,\bar{\mathbf{u}}$ está em V
g. A multiplicação de um objeto $\bar{\mathbf{u}}$ em V por escalares k e β em F, é uma operação associativa $\kappa(\beta\ \bar{\mathbf{u}})=(\kappa\beta)\ \bar{\mathbf{u}}$
h. A multiplicação de um objeto $\bar{\mathbf{u}}$ em V por escalares é uma operação distributiva relativamente à adição de escalares: $(\kappa+\beta)\ \bar{\mathbf{u}}=\kappa\,\bar{\mathbf{u}}+\beta\,\bar{\mathbf{u}}$
i. A multiplicação por um escalar k em F é uma operação distributiva relativamente à adição de objetos $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$ em V: $\kappa\,(\bar{\mathbf{u}}+\bar{\mathbf{v}})=\kappa\,\bar{\mathbf{u}}+\kappa\,\bar{\mathbf{v}}$
j. Existe em F um único escalar 1, chamado de identidade de F, tal que: $1\bar{\mathbf{u}}=\bar{\mathbf{u}}$

Dependendo da aplicação, os escalares podem ser números reais ou complexos. Os espaços vetoriais em que os escalares são números complexos são chamados de “Espaço Vetorial Complexo” e aqueles em que os escalares são números reais são chamados de “Espaço Vetorial Real”. Temos ainda um “Espaço Vetorial Zero” que é definido por um único objeto em V denotado por $\bar{\mathbf{0}}$, e definido para todos os escalares do corpo F. Se o objeto $\bar{\mathbf{0}}$ em V satisfaz os axiomas do espaço vetorial abstrato, então chamamos de “Espaço Vetorial Zero”

Espaço Vetorial X Geofísica:

Qual é a relação entre o conceito de espaço vetorial linear e a geofísica ?

Neste curso iremos apenas abordar a teoria de inversão discreta em que os parâmetros e os dados podem ser adequadamente aproximados como discretos. Portanto, as medidas observacionais geofísicas consistirão de um conjunto discreto de informações numéricas que iremos encapsular em um vetor. Além disso, os parâmetros também serão tratados como vetores.

Mais adiante veremos que há uma ampla classe de problemas inversos que podem ser escritos como um sistema de equações lineares

$$\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}$$

em que $\overline{\mathbf{y}}$ é o vetor N-dimensional contendo os dados geofísicos observados , $\overline{\mathbf{p}}$ é o vetor M-dimensional dos parâmetros e $\overline{\overline{\mathbf{A}}}$ é uma matriz N x M (operador linear). Então assumiremos que N medidas foram realizadas em algum experimento geofísico e trataremos estas N medidas como elementos de um vetor $\overline{\mathbf{y}} \in R^N$. Similarmente, os parâmetros podem ser representados como elementos do vetor $\overline{\mathbf{p}} \in R^M$. Há portanto um operador matricial $\overline{\overline{\mathbf{A}}}$ que transforma os parâmetros $\overline{\mathbf{p}}$ nos dados geofísicos $\overline{\mathbf{y}}$, i.e., mapea o espaço dos parâmetros P no espaço das medidas Y. A solução do problema inverso consistirá em determinar o vetor de parâmetros $\overline{\mathbf{p}}$ que produziu os dados geofísicos $\overline{\mathbf{y}}$. Nós discutiremos mais adiante, que não queremos ajustar (explicar) perfeitamente (exatamente) os dados observados porque eles sempre contém ruído. Portanto, procuraremos determinar o vetor de parâmetros $\overline{\mathbf{p}}$ que produza um ajuste que esteja o mais próximo possível dos dados observados medidos. Mas o que significa “próximo suficiente” ? Como podemos medir a proximidade entre dois conjuntos de dados? A resposta é muito simples: através da distância entre dois conjuntos de dados. Em outras palavras, tanto os nossos conjuntos de dados como de parâmetros devem estar definidos em um espaço métrico.

Matrizes

Definição: Uma matriz é um array retangular de números reais. Especificamente uma matriz consiste em NxM números reais dispostos em N linhas e M colunas fornecendo o seguinte array retangular N x M

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}} = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{pmatrix} & N \\ M & \end{array}$$

em que a_{ij} representa o elemento da i-ésima linha e j-ésima coluna da matriz $\overline{\overline{\mathbf{A}}}$. O conjunto de todas as NxM matrizes com elementos escalares é um espaço vetorial denominado R^{NxM}, com as regras usuais de adição e multiplicação por escalares.

Tamanho da Matriz (size): É dado pela especificação do número de linhas e número de colunas, ou seja, N x M.

Operações matriciais:

Se $\overline{\overline{\mathbf{A}}}$, $\overline{\overline{\mathbf{B}}}$ e $\overline{\overline{\mathbf{C}}}$ são matrizes N x M

(1) Adição de matrizes:

A operação $\overline{\overline{\mathbf{A}}}+\overline{\overline{\mathbf{B}}}$ é uma adição matricial que é realizada somando-se os correspondentes elementos de $\overline{\overline{\mathbf{A}}}$ e $\overline{\overline{\mathbf{B}}}$.

$$\overline{\overline{\mathbf{A}}}+\overline{\overline{\mathbf{B}}}=\overline{\overline{\mathbf{C}}}$$

$$[\,a_{ij}\,] + [\,b_{ij}\,] = [\,c_{ij}\,]$$

Propriedades da Adição Matricial

(1.1) A adição matricial é comutativa

$$\overline{\overline{\mathbf{A}}}+\overline{\overline{\mathbf{B}}}=\overline{\overline{\mathbf{B}}}+\overline{\overline{\mathbf{A}}}$$

(1.2) A adição matricial é <u>associativa</u>

$$\bar{\bar{\mathbf{A}}} + \left(\bar{\bar{\mathbf{B}}} + \bar{\bar{\mathbf{C}}}\right) = \left(\bar{\bar{\mathbf{A}}} + \bar{\bar{\mathbf{B}}}\right) + \bar{\bar{\mathbf{C}}}$$

(2) <u>Multiplicação por um escalar</u>

Se $\bar{\bar{\mathbf{A}}}$ é qualquer matriz N x M e k é um escalar, então $k\bar{\bar{\mathbf{A}}}$ é a matriz obtida pela multiplicação de cada elemento de $\bar{\bar{\mathbf{A}}}$ por k.

$$k\bar{\bar{\mathbf{A}}} \equiv \left[k a_{ij}\right]$$

<u>Propriedades</u>:

(2.1) A multiplicação de uma matriz $\bar{\bar{\mathbf{A}}}$ por escalares k e β é uma operação associativa

$$k\left(\beta\,\bar{\bar{\mathbf{A}}}\right) = (k\beta)\bar{\bar{\mathbf{A}}} \equiv \left[k\beta\, a_{ij}\right]$$

(2.2) A multiplicação de uma matriz $\bar{\bar{\mathbf{A}}}$ por escalares k e β é uma operação distributiva relativamente à adição de escalares

$$(k+\beta)\bar{\bar{\mathbf{A}}} = k\bar{\bar{\mathbf{A}}} + \beta\bar{\bar{\mathbf{A}}} \equiv \left[k\, a_{ij} + \beta\, a_{ij}\right]$$

$$(k-\beta)\bar{\bar{\mathbf{A}}} = k\bar{\bar{\mathbf{A}}} - \beta\bar{\bar{\mathbf{A}}} \equiv \left[k\, a_{ij} - \beta\, a_{ij}\right]$$

(2.3) A multiplicação de uma matriz $\bar{\bar{\mathbf{A}}}$ por um escalar k é uma operação distributiva relativamente à adição de matrizes

$$k\left(\bar{\bar{\mathbf{A}}} + \bar{\bar{\mathbf{B}}}\right) = k\bar{\bar{\mathbf{A}}} + k\bar{\bar{\mathbf{B}}} \equiv \left[k a_{ij} + k b_{ij}\right]$$

$$k\left(\bar{\bar{\mathbf{A}}} - \bar{\bar{\mathbf{B}}}\right) = k\bar{\bar{\mathbf{A}}} - k\bar{\bar{\mathbf{B}}} \equiv \left[k a_{ij} - k b_{ij}\right]$$

(2.4) A multiplicação de uma matriz $\bar{\bar{\mathbf{A}}}$ por um escalar -1 é

$$-1\cdot\bar{\bar{\mathbf{A}}} = -\bar{\bar{\mathbf{A}}} \equiv \left[-a_{ij}\right]$$

(3) <u>Multiplicação de matrizes</u>:

Se $\bar{\bar{\mathbf{A}}}$ é uma matriz N x L e $\bar{\bar{\mathbf{B}}}$ uma matriz L x M, então o produto $\bar{\bar{\mathbf{A}}}\,\bar{\bar{\mathbf{B}}}$ é uma matriz NxM. O (i,k)-ésimo elemento do produto $\bar{\bar{\mathbf{A}}}\,\bar{\bar{\mathbf{B}}}$ é obtido multiplicando-se a i-ésma linha de $\bar{\bar{\mathbf{A}}}$ pela k-ésima coluna de $\bar{\bar{\mathbf{B}}}$

$$\underset{(NxL)}{\bar{\bar{\mathbf{A}}}}\ \underset{(LxM)}{\bar{\bar{\mathbf{B}}}} = \underset{(NxM)}{\bar{\bar{\mathbf{C}}}}$$

O elemento c_{ik} é

$$c_{ik} = \sum_{j=1}^{L} a_{ij} b_{jk} \quad .$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1L} \\ a_{21} & a_{22} & \cdots & a_{2L} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NL} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1M} \\ b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & & \vdots \\ b_{L1} & b_{L2} & \cdots & b_{LM} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1M} \\ c_{21} & c_{22} & \cdots & c_{2M} \\ \vdots & \vdots & & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NM} \end{pmatrix}$$

$$\begin{pmatrix} [a_{11}b_{11} + a_{12}b_{21} + ... + a_{1L}b_{L1}] & [a_{11}b_{12} + a_{12}b_{22} + ... + a_{1L}b_{L2}] & \cdots & [a_{11}b_{1M} + a_{12}b_{2M} + ... + a_{1L}b_{LM}] \\ [a_{21}b_{11} + a_{22}b_{21} + ... + a_{2L}b_{L1}] & [a_{21}b_{12} + a_{22}b_{22} + ... + a_{2L}b_{L2}] & \cdots & [a_{21}b_{1M} + a_{22}b_{2M} + ... + a_{2L}b_{LM}] \\ \vdots & \vdots & & \vdots \\ [a_{N1}b_{11} + a_{N2}b_{21} + ... + a_{NL}b_{L1}] & [a_{N1}b_{12} + a_{N2}b_{22} + ... + a_{NL}b_{L2}] & \cdots & [a_{N1}b_{1M} + a_{N2}b_{2M} + ... + a_{NL}b_{LM}] \end{pmatrix}$$

Veja ainda que o produto $\bar{\bar{\mathbf{A}}}\bar{\bar{\mathbf{B}}}$ pode ser expresso como

$$\underset{(NxL)}{\bar{\bar{\mathbf{A}}}} \; \underset{(LxM)}{\bar{\bar{\mathbf{B}}}} = \underset{(NxM)}{\bar{\bar{\mathbf{C}}}}$$

$$\underset{(NxL)}{[\bar{\mathbf{a}}_1 \quad \bar{\mathbf{a}}_2 \quad \cdots \bar{\mathbf{a}}_L]} \; \underset{(LxM)}{\bar{\bar{\mathbf{B}}}} = \underset{(NxM)}{[\bar{\mathbf{c}}_1 \quad \bar{\mathbf{c}}_2 \quad \cdots \bar{\mathbf{c}}_M]}$$

Em que $\bar{\mathbf{a}}_j$ é o vetor formado pela j-ésima coluna da matriz $\bar{\bar{\mathbf{A}}}$ e $\bar{\mathbf{c}}_k$ é o vetor formado pela k-ésima coluna da matriz $\bar{\bar{\mathbf{C}}}$, sendo expresso como

$$\bar{\mathbf{c}}_k = \sum_{j=1}^{L} \bar{\mathbf{a}}_j b_{jk} \,, \qquad\qquad k = 1,2,...,M$$

<u>Caso particular:</u>

<u>Multiplicação de uma matriz por vetor</u>

Se $\bar{\bar{\mathbf{A}}}$ é uma matriz N x M e $\bar{\mathbf{p}}$ um vetor M x 1, então o produto $\bar{\bar{\mathbf{A}}}\bar{\mathbf{p}}$ é um vetor N x 1.

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}}\ \underset{(M\times 1)}{\overline{\mathbf{p}}} = \underset{(N\times 1)}{\overline{\mathbf{y}}}$$

cujo i-ésimo elemento do vetor $\overline{\mathbf{y}}$ é expresso como

$$y_i = \sum_{j=1}^{M} a_{ij}\, p_j \qquad\qquad i = 1,2,\ldots,N$$

Considerando $\overline{\mathbf{a}}_j$ o vetor N-dimensional formado pela j-ésima coluna da matriz $\overline{\overline{\mathbf{A}}}$ então o vetor $\overline{\mathbf{y}}$ pode ser expresso como

$$\underset{N\times 1}{\overline{\mathbf{y}}} = \sum_{j=1}^{M} \overline{\mathbf{a}}_j\, p_j$$

<u>Propriedades da Multiplicação de Matrizes:</u>

(3.1) A multiplicação de matrizes obedece a lei distributiva

$$\overline{\overline{\mathbf{A}}}\left(\overline{\overline{\mathbf{B}}} + \overline{\overline{\mathbf{C}}}\right) = \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}} + \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{C}}}$$

$$\overline{\overline{\mathbf{A}}}\left(\overline{\overline{\mathbf{B}}} - \overline{\overline{\mathbf{C}}}\right) = \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}} - \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{C}}}$$

$$\left(\overline{\overline{\mathbf{B}}} + \overline{\overline{\mathbf{C}}}\right)\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{C}}}\,\overline{\overline{\mathbf{A}}}$$

$$\left(\overline{\overline{\mathbf{B}}} - \overline{\overline{\mathbf{C}}}\right)\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}} - \overline{\overline{\mathbf{C}}}\,\overline{\overline{\mathbf{A}}}$$

(3.2) A multiplicação de matrizes obedece a lei associativa

$$\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right)\overline{\overline{\mathbf{C}}} = \overline{\overline{\mathbf{A}}}\left(\overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{C}}}\right)$$

<u>Observações práticas:</u>

- Em geral a lei comutativa da multiplicação de matrizes em que NÃO é válido

  $$\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right) \neq \overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}}$$

- Na manipulação de matrizes, os colchetes podem ser removido e as potências podem ser combinadas, por exemplo

  $$\left(\overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}}^{2}\right)\overline{\overline{\mathbf{A}}}\left(\overline{\overline{\mathbf{A}}}^{-2}\overline{\overline{\mathbf{B}}}\right) = \overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}$$

- A lei do cancelamento em geral não é valida na multiplicação de matrizes. Então se $\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{0}}}$ onde $\overline{\overline{\mathbf{0}}}$ é uma matriz nula isto <u>NÃO</u> implica $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{0}}}$ ou $\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{0}}}$
- Se $\overline{\overline{\mathbf{A}}}$ é uma matriz N x L, $\overline{\overline{\mathbf{B}}}$ uma matriz L x M e k um escalar então
  $k\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right) = \left(k\overline{\overline{\mathbf{A}}}\right)\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{A}}}\left(k\overline{\overline{\mathbf{B}}}\right)$

TIPOS ESPECIAIS DE MATRIZES

(1) <u>Matriz Nula:</u> $\overline{\overline{\mathbf{0}}}$

Matriz com todos os elementos zeros. Se $\overline{\overline{\mathbf{A}}}$ é uma matriz N x M , uma matriz Nula N x M pode ser definida como

$$\overline{\overline{\mathbf{A}}} - \overline{\overline{\mathbf{A}}} \equiv \left[a_{ij} - a_{ij}\right] = \overline{\overline{\mathbf{0}}}$$

Propriedades da matriz Nula:

$$\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{0}}} = \overline{\overline{\mathbf{A}}}$$
$$\overline{\overline{\mathbf{0}}} - \overline{\overline{\mathbf{A}}} = -\overline{\overline{\mathbf{A}}}$$
$$\overline{\overline{\mathbf{A}}} - \overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{0}}}$$
$$\overline{\overline{\mathbf{0}}}\,\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{0}}} = \overline{\overline{\mathbf{0}}}$$

- A lei do cancelamento em geral não é valida na multiplicação de matrizes. Então se $\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{0}}}$ onde $\overline{\overline{\mathbf{0}}}$ é uma matriz nula isto <u>NÃO</u> implica $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{0}}}$ ou $\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{0}}}$
- se $\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{C}}}$ então $\overline{\overline{\mathbf{A}}}\left(\overline{\overline{\mathbf{B}}} - \overline{\overline{\mathbf{C}}}\right) = \overline{\overline{\mathbf{0}}}$ e isto NÃO implica $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{0}}}$ ou $\overline{\overline{\mathbf{B}}} = \overline{\overline{\mathbf{C}}}$

(2) <u>Matriz Transposta:</u> $\overline{\overline{\mathbf{A}}}^T$

Se $\overline{\overline{\mathbf{A}}}$ é uma matriz N x M então a transposta de $\overline{\overline{\mathbf{A}}}$ é uma matriz denominada de $\overline{\overline{\mathbf{A}}}^T$ de dimensão M X N cuja i-ésima coluna é a j-ésima linha de $\overline{\overline{\mathbf{A}}}$. Se

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}} \equiv \left[a_{i\,j}\right] \text{ então } \underset{(M\times N)}{\overline{\overline{\mathbf{A}}}^T} \equiv \left[a_{j\,i}\right]$$

Propriedades da matriz Transposta:

- A Transposta da soma de duas matrizes é a soma das matrizes transpostas

$$\left(\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{B}}}\right)^T = \overline{\overline{\mathbf{A}}}^T + \overline{\overline{\mathbf{B}}}^T$$

- A transposta da transposta de uma matriz é igual a matriz

$$\left(\overline{\overline{\mathbf{A}}}^{T}\right)^{T} = \overline{\overline{\mathbf{A}}}$$

- Se $\overline{\overline{\mathbf{S}}}$ é uma matriz diagonal [veja definição no item (4)] então a sua transposta é igual a matriz

$$\left(\overline{\overline{\mathbf{S}}}\right)^{T} = \overline{\overline{\mathbf{S}}}$$

- A transposta do produto de duas matrizes é o produto das transpostas na ordem inversa

$$\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right)^{T} = \overline{\overline{\mathbf{B}}}^{T}\overline{\overline{\mathbf{A}}}^{T}$$

- Se k é um escalar então

$$\left(k\overline{\overline{\mathbf{A}}}\right)^{T} = k\overline{\overline{\mathbf{A}}}^{T}$$

(3) Matriz Quadrada

É uma matriz cujo número de linhas é igual ao número de colunas.

$\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}} \rightarrow$ matriz quadrada de ordem N

$\underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \rightarrow$ matriz quadrada de ordem M

- Diagonal principal : Se $\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}}$ é uma matriz quadrada chamamos de diagonal principal os elementos a_{ii} , i=1,2,…,N.

- Traço: Se $\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}}$ é uma matriz quadrada, então o traço de $\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}}$ é definido como a soma dos elementos da diagonal principal

$$Tr\left(\overline{\overline{\mathbf{A}}}\right) = \sum_{i=1}^{N} a_{ii}$$

Propriedades do Traço:

$$Tr\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right) = Tr\left(\overline{\overline{\mathbf{B}}}\,\overline{\overline{\mathbf{A}}}\right)$$

(4) Matriz Diagonal

É uma matriz quadrada em que todos os elementos fora da diagonal principal são nulos.

$$a_{ij} = 0 \text{ se i} \neq \text{j}$$

$$a_{ij} \neq 0 \text{ se i} = \text{j}$$

(5) Matriz Identidade: $\overline{\overline{\mathbf{I}}}$

É uma matriz diagonal em que todos os elementos da diagonal principal são todos unitários.

$$\overline{\overline{\mathbf{I}}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

(6) Matriz Simétrica

É uma matriz quadrada tal que

$$\overline{\overline{\mathbf{A}}}^T = \overline{\overline{\mathbf{A}}}\,.$$

Em outras palavras, uma matriz é simétrica se os elementos da matriz estão simetricamente dispostos em relação a diagonal principal. Assim por exemplo se a matriz $\underset{(N \times N)}{\overline{\overline{\mathbf{A}}}}$ é simétrica temos que

$$\underset{(N \times N)}{\overline{\overline{\mathbf{A}}}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{12} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & \cdots & a_{NN} \end{pmatrix}$$

Neste caso se $\overline{\overline{\mathbf{A}}}$ é simétrica temos

$$\overline{\overline{\mathbf{A}}} \equiv \left[a_{ij}\right] = \left[a_{ji}\right]$$

Propriedades da matriz simétrica

- Se $\overline{\overline{\mathbf{A}}}$ é uma matriz N x M e $\overline{\overline{\mathbf{A}}}^T$ a sua transposta então

  $\underset{(M\times N)}{\overline{\overline{\mathbf{A}}}^T} \; \underset{(N\times M)}{\overline{\overline{\mathbf{A}}}}$ é uma matriz simétrica de ordem M e

  $\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}} \; \underset{(M\times N)}{\overline{\overline{\mathbf{A}}}^{\mathrm{T}}}$ é uma matriz simétrica de ordem N.

- Se $\overline{\overline{\mathbf{A}}}$ (N x N) é uma matriz simétrica de ordem N e $\overline{\overline{\mathbf{B}}}$ uma matriz qualquer N x M então

  $\overline{\overline{\mathbf{B}}}^T \, \overline{\overline{\mathbf{A}}} \, \overline{\overline{\mathbf{B}}}$ é uma matriz simétrica de ordem M

- O produto de duas matrizes simétricas em geral NÃO é simétrico. Então se $\overline{\overline{\mathbf{A}}}^T = \overline{\overline{\mathbf{A}}}$ e $\overline{\overline{\mathbf{B}}}^T = \overline{\overline{\mathbf{B}}}$

  $\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right)^T = \overline{\overline{\mathbf{B}}}^T \, \overline{\overline{\mathbf{A}}}^T$

(7) Matriz Anti-simétrica

É uma matriz quadrada tal que

$$\overline{\overline{\mathbf{A}}}^T = -\overline{\overline{\mathbf{A}}}.$$

Neste caso se $\overline{\overline{\mathbf{A}}}$ é anti-simétrica temos

$$\overline{\overline{\mathbf{A}}} \equiv \left[a_{ij}\right] = \left[-a_{ji}\right]\backslash$$

(8) Matriz Ortogonal

Uma matriz $\overline{\overline{\mathbf{A}}} \in R^{N\times N}$ é ortogonal se

$$\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}^T} \; \underset{(N\times N)}{\overline{\overline{\mathbf{A}}}} = \underset{(N\times N)}{\overline{\overline{\mathbf{A}}}} \; \underset{(N\times N)}{\overline{\overline{\mathbf{A}}}^T} = \overline{\overline{\mathbf{I}}}_N$$

De acordo com a definição de matriz inversa que iremos apresentar no tópico 13, se uma matriz $\overline{\overline{\mathbf{A}}} \in R^{N\times N}$ é ortogonal então

$$\underset{(N\times N)}{\overline{\overline{\mathbf{A}}}^{-1}} = \underset{(N\times N)}{\overline{\overline{\mathbf{A}}}^{T}}$$

Propriedades da Matriz Ortogonal:

- Cada coluna e cada linha da matriz é um vetor ortonormal,

$$\overline{\mathbf{a}}_i^T \cdot \overline{\mathbf{a}}_i = 1$$

(9) Matriz Real : Todos os elementos são reais

(10) Matriz Complexa: Tem elementos complexos

(11) Matriz Imaginária: Todos os elementos são imaginários ou nulos

(12) Matriz Hermitiana:

Uma matriz Hermitiana $\overline{\overline{\mathbf{A}}}^{H}$ é definida como a complexa conjugada $\overline{\overline{\mathbf{A}}}^{*}$ transposta.

$$\overline{\overline{\mathbf{A}}}^{H} = \overline{\overline{\mathbf{A}}}^{*T}$$

Veja que se $\overline{\overline{\mathbf{A}}}$ for uma matriz de número reais, então note que a Hermitiana é uma matriz transposta no caso Real.

(13) Matriz Inversa:

Se $\overline{\overline{\mathbf{A}}}$ é uma matriz quadrada de orem N, ($\overline{\overline{\mathbf{A}}} \in R^{N\times N}$) e se existe uma matriz $\overline{\overline{\mathbf{A}}}^{-1}$ tal que

$$\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{A}}}^{-1} = \overline{\overline{\mathbf{A}}}^{-1}\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{I}}}_N$$

então dizemos que:

- A matriz $\overline{\overline{\mathbf{A}}}$ é NÃO SINGULAR, ou seja, a matriz $\overline{\overline{\mathbf{A}}}$ possui inversa
- A inversa da matriz $\overline{\overline{\mathbf{A}}}$ é a matriz $\overline{\overline{\mathbf{A}}}^{-1}$

Chamamos de Matriz **NÃO SINGULAR** uma matriz que **possui INVERSA**
Chamamos de Matriz **SINGULAR** uma matriz que **NÃO possui INVERSA**

<u>Propriedades</u>

1. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz Não singular, então sua inversa é ÚNICA
2. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ e $\overline{\overline{\mathbf{B}}} \in R^{N \times N}$ são matrizes Não Singular então o produto $\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}$ é sempre Não singular cuja inversa é expressa como:
$$\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}\right)^{-1} = \overline{\overline{\mathbf{B}}}^{-1}\overline{\overline{\mathbf{A}}}^{-1}$$
3. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz Não singular então
$$\left(\overline{\overline{\mathbf{A}}}^{-1}\right)^{-1} = \overline{\overline{\mathbf{A}}}$$
4. Se $\overline{\overline{\mathbf{A}}}^{n} \in R^{N \times N}$, em que n=0,1,2,3, … , é uma matriz Não singular então
$$\left(\overline{\overline{\mathbf{A}}}^{n}\right)^{-1} = \left(\overline{\overline{\mathbf{A}}}^{-1}\right)^{n}$$ para todo n=0,1,2,…
5. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz Não singular e k é qualquer escalar diferente de zero então $k\overline{\overline{\mathbf{A}}}$ é também uma matriz Não singular
$$\left(k\overline{\overline{\mathbf{A}}}\right)^{-1} = \frac{1}{k}\overline{\overline{\mathbf{A}}}^{-1}$$
6. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz Não singular SIMÉTRICA (ou seja $\overline{\overline{\mathbf{A}}}^{T} = \overline{\overline{\mathbf{A}}}$), então sua inversa $\overline{\overline{\mathbf{A}}}^{-1}$ é também uma matriz SIMÉTRICA
7. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz Não singular ORTOGONAL (ou seja, $\underset{(N \times N)}{\overline{\overline{\mathbf{A}}}^{T}}\,\underset{(N \times N)}{\overline{\overline{\mathbf{A}}}} = \underset{(N \times N)}{\overline{\overline{\mathbf{A}}}}\,\underset{(N \times N)}{\overline{\overline{\mathbf{A}}}^{T}} = \overline{\overline{\mathbf{I}}}_N$), então sua inversa $\overline{\overline{\mathbf{A}}}^{-1} = \overline{\overline{\mathbf{A}}}^{T}$ é também uma matriz ORTOGONAL

8. Se $\overline{\overline{\mathbf{A}}} \in R^{N\times N}$ é uma matriz Não singular DIAGONAL, sua inversa $\overline{\overline{\mathbf{A}}}^{-1}$ é também uma matriz DIAGONAL cujo i-ésimo elemento da diagonal é expresso como

   $$\overline{\overline{\mathbf{A}}}^{-1} \equiv \left[\frac{1}{a_{ii}}\right]$$

9. Se $\overline{\overline{\mathbf{A}}} \in R^{N\times N}$ é uma matriz Não singular então $\left(\overline{\overline{\mathbf{A}}}^{-1}\right)^T = \left(\overline{\overline{\mathbf{A}}}^T\right)^{-1}$

10. Se $\overline{\overline{\mathbf{A}}} \in R^{N\times M}$ é uma matriz RETANGULAR então é uma matriz Singular (Não possui inversa).

A utilidade prática do estudo da matriz inversa é resolver um sistema de equações lineares que vamos ver mais adiante

<u>FUNÇÃO DETERMINANTE:</u>

Estamos todos familiarizados com funções do tipo f(x) =x^2 , que associa um número real f(x) com um número real da variável x. Como f(x) e x assumem apenas valores reais, tais funções são descritas como "funções de números reais de uma variável real"

A <u>função determinante</u> $\det(\overline{\overline{\mathbf{A}}})$ é uma função de número real de uma variável matricial. Em outras palavras, a função determinante associa um número real $\det(\overline{\overline{\mathbf{A}}})$ a uma matriz $\overline{\overline{\mathbf{A}}}$.

A utilidade prática do determinante é servir como uma ferramenta de avaliação de um sistema linear, vamos ver mais adiante

<u>Propriedades da Função Determinante:</u>

1. Se $\overline{\overline{\mathbf{A}}} \in R^{N\times N}$ e $\overline{\overline{\mathbf{A}}}^T$ é a sua transposta então

   $$\det(\overline{\overline{\mathbf{A}}}^T) = \det(\overline{\overline{\mathbf{A}}})$$

2. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz quadrada que contém uma linha ou coluna com elementos zero, então

$$\det(\overline{\overline{\mathbf{A}}}) = 0$$

3. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz quadrada que contém duas linhas (ou duas colunas) iguais então

$$\det(\overline{\overline{\mathbf{A}}}) = 0$$

4. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz quadrada cuja i-ésima linha (ou i-ésima coluna) é um múltiplo da j-ésima linha (ou j-ésima coluna) então

$$\det(\overline{\overline{\mathbf{A}}}) = 0$$

5. Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$, $\overline{\overline{\mathbf{B}}} \in R^{N \times N}$ e k qualquer escalar então temos que as seguintes propriedades

   - $\det(k\overline{\overline{\mathbf{A}}}) = k^{N} \det(\overline{\overline{\mathbf{A}}})$
   - $\det(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{B}}}) = \det(\overline{\overline{\mathbf{A}}}) \det(\overline{\overline{\mathbf{B}}})$
   - Em geral, $\det(\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{B}}}) \neq \det(\overline{\overline{\mathbf{A}}}) + \det(\overline{\overline{\mathbf{B}}})$

<u>**Importância do Determinante:**</u>

O determinante de uma matriz é uma ferramenta prática para testarmos se uma matriz possui inversa.

Lema: Uma matriz $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz **Não Singular** (possui inversa) se e somente se $\det(\overline{\overline{\mathbf{A}}}) \neq 0$

Logo uma matriz $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ **Singular** (NÃO possui inversa) se e somente se $\det(\overline{\overline{\mathbf{A}}}) = 0$

- Se $\overline{\overline{\mathbf{A}}} \in R^{N \times N}$ é uma matriz **Não Singular,** o determinante de sua inversa $\overline{\overline{\mathbf{A}}}^{-1}$ é:

$$\det(\overline{\overline{\mathbf{A}}}^{-1}) = \frac{1}{\det(\overline{\overline{\mathbf{A}}})}$$

REGRA DE CRAMER:

Se $\overline{\overline{\mathbf{A}}} \in R^{2 \times 2}$ é uma matriz **Não Singular,** a sua inversa $\overline{\overline{\mathbf{A}}}^{-1}$ é expressa como:

$$\overline{\overline{\mathbf{A}}}^{-1} = \frac{1}{\det(\overline{\overline{\mathbf{A}}})} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

<u>EQUAÇÕES LINEARES:</u>

Uma reta no plano x-y pode ser representada pela equação

$$a_1 x + a_2 y = b$$

Uma equação deste tipo é chamada de <u>equação linear</u> nas variáveis x e y.

Generalizando definimos uma equação linear em M variáveis $x_1, x_2, \ldots, x_M$ como sendo:

$$a_1 x_1 + a_2 x_2 + ,\ldots, + a_M x_M = b$$

Onde $a_1, a_2, \ldots, a_M$ e b são constantes reais e $x_1, x_2, \ldots, x_M$ são variáveis desconhecidas.

SISTEMA DE EQUAÇÕES LINEARES:

Um sistema de equações lineares (ou, simplesmente, sistema linear) é um conjunto finito de N equações lineares nas variáveis $x_1, x_2, ..., x_M$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1M}x_M = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2M}x_M = b_2 \\ \qquad \vdots \qquad\qquad\qquad\qquad = \vdots \\ a_{N1}x_1 + a_{N2}x_2 + ... + a_{NM}x_M = b_N \end{cases}$$

Este sistema linear apresenta N equações em M incógnitas.

Em notação matricial um sistema linear pode ser escrito por

$$\underset{(N \times M)}{\overline{\overline{\mathbf{A}}}} \; \underset{(M \times 1)}{\overline{\mathbf{x}}} = \underset{(N \times 1)}{\overline{\mathbf{b}}}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

Quantas soluções um sistema linear pode ter ?

Todo sistema de equações lineares pode:

1) Não ter solução
2) Ter exatamente uma solução
3) Ter infinitas soluções

Para responder a pergunta acima sobre quantas soluções um dado sistema pode ter, vamos introduzir um conceito chamado **Sistema Homogêneo de Equações Lineares.**

Sistema Homogêneo de Equações Lineares:

Um sistema linear homogêneo é expresso como

$$\underset{(N \times M)}{\overline{\overline{\mathbf{A}}}} \; \underset{(M \times 1)}{\overline{\mathbf{x}}} = \underset{(N \times 1)}{\overline{\mathbf{0}}}$$

ou seja,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

A solução do sistema homogêneo de equações ($\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{0}}$) é chamada de

- **SOLUÇÃO TRIVIAL** se $\overline{\mathbf{x}} = \overline{\mathbf{0}}$ e
- **SOLUÇÃO NÃO TRIVIAL** se $\overline{\mathbf{x}} \neq \overline{\mathbf{0}}$.

Agora responderemos a pergunta formulada anteriormente sobre o número soluções que um sistema linear não homogêneo pode ter.

Um sistema de equações lineares $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{b}}$ tem **EXATAMENTE UMA SOLUÇAO** se o sistema homogêneo $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{0}}$ tem apenas a **SOLUÇÃO TRIVIAL** ($\overline{\mathbf{x}} = \overline{\mathbf{0}}$).

Um sistema de equações lineares $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{b}}$ tem **INFINITAS SOLUÇÕES** se o sistema homogêneo $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{0}}$ tem **SOLUÇÃO NÃO TRIVIAL** ($\overline{\mathbf{x}} \neq \overline{\mathbf{0}}$).

Propriedades:

- Seja um sistema linear apresentando N equações lineares em M incógnitas definido por

$$\underset{(N \times M)}{\overline{\overline{\mathbf{A}}}}\;\underset{(M \times 1)}{\overline{\mathbf{x}}} = \underset{(N \times 1)}{\overline{\mathbf{b}}}$$

Se M > N o sistema linear terá INFINITAS SOLUÇÕES. Portanto, um sistema linear em que o número de incógnitas (M) é maior que o número de equações (N) tem infinitas soluções.

SISTEMA DE EQUAÇÕES LINEARES X GEOFÍSICA:

Vimos que na geofísica os dados coletados no campo (medidas observacionais de algum fenômeno físico) são simplesmente uma tabela de valores numéricos que são armazenados em um vetor. Assim dado um conjunto de N observações os dados medidos podem ser representados por

$$\overline{\mathbf{y}}^{\mathbf{o}} = \begin{bmatrix} y_1^0 \\ y_2^0 \\ \vdots \\ y_N^0 \end{bmatrix}$$

As medidas observacionais são portanto um vetor $\overline{\mathbf{y}}^{\mathbf{o}} \in R^N$.

Similarmente, os parâmetros (variável desconhecida do nosso problema) também serão um conjunto de M valores numéricos que são armazenados em um vetor

$$\overline{\mathbf{p}} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}$$

Já vimos que o princípio básico do problema inverso é que há alguma relação físico-matemática entre os dados $\overline{\mathbf{y}}^{\mathbf{o}}$ e os parâmetros $\overline{\mathbf{p}}$. Esta relação chamaremos de <u>modelo direto (forward operator)</u> ou <u>sistema</u> físico ou <u>relação funcional</u> que nada mais é que relacionam os dados coletados com os parâmetros. Assim por exemplo se medirmos a massa e o volume de um corpo temos dois elementos compondo o vetor $\overline{\mathbf{y}}^{\mathbf{o}}$, então podemos dizer que

$$\overline{\mathbf{y}}^{\mathbf{o}} = \begin{bmatrix} y_1^0 = \text{massa} \\ y_2^0 = \text{volune} \end{bmatrix}$$

Se desejamos estimar a densidade deste corpo temos um vetor de parâmetros $\overline{\mathbf{p}}$ com um único elemento

$$\overline{\mathbf{p}} = \left[p_1 = \text{densidade} \right]$$

O sistema físico ou modelo direto, neste caso é muito simples. Ele estabelece que a densidade multiplicada pelo volume é igual a massa

$$y_2^o \; p_1 = y_1^0 \; .$$

Em uma situação mais realista temos que os dados e os parâmetros estão relacionados de modo mais complicado. Neste caso simples acima temos apenas uma equação

$$y_2^o \; p_1 - y_1^0 = 0$$

ou seja,

$$f_1(\overline{\mathbf{y}}^{\mathbf{o}}, \overline{\mathbf{p}}) = 0$$

Em problema mais complexos temos N equações tal que

$$f_1(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = 0$$
$$f_2(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = 0$$
$$\vdots$$
$$f_N(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = 0$$

Podemos escrever este conjunto de N equações como uma equação vetorial

$$\bar{\mathbf{f}}\ (\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = \bar{\mathbf{0}}$$

ou simplesmente

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\mathbf{f}}\ (\bar{\mathbf{p}}).$$

Estas funções $f_i(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = 0, \quad i = 1, 2, \ldots, N$ que relacionam os parâmetros e os dados observados pode ser LINEAR ou NÃO LINEAR. Se as funções $f_i(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}}) = 0, \quad i = 1, 2, \ldots, N$ são funções lineares em relação aos M parâmetros, então

$$\frac{\partial f_i(\bar{\mathbf{y}}^{\mathbf{o}}, \bar{\mathbf{p}})}{\partial p_j}, \quad i = 1, 2, \ldots, N \text{ e } j = 1, 2, \ldots, M,$$

NÃO é função dos parâmetros (não dependem dos parâmetros). Neste caso, dizemos que a relação funcional (modelo direto) entre os parâmetros e os dados é uma relação linear que pode ser representada por um sistema de N equações lineares em M incógnitas escrito em notação matricial como

$$\underset{(N \times M)}{\bar{\bar{\mathbf{A}}}}\ \underset{(M \times 1)}{\bar{\mathbf{p}}} = \underset{(N \times 1)}{\bar{\mathbf{y}}^{\mathbf{o}}}$$

COMBINAÇÃO LINEAR:

Veja que a multiplicação de uma matriz $\overline{\overline{\mathbf{A}}}$ por um vetor de parâmetros $\overline{\mathbf{p}}$, resultando no vetor de dados observados $\overline{\mathbf{y}}^{\mathbf{o}}$,

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}}\ \underset{(M\times 1)}{\overline{\mathbf{p}}} = \underset{(N\times 1)}{\overline{\mathbf{y}}^{\mathbf{o}}},$$

pode ser rescrita como

$$\underset{N\times 1}{\overline{\mathbf{y}}^{o}} = \sum_{j=1}^{M} \overline{\mathbf{a}}_j\, p_j ,$$

em que $\overline{\mathbf{a}}_j$ é um vetor (N x 1) formado pela j-ésima coluna da matriz $\overline{\overline{\mathbf{A}}}$. Em outras palavras, podemos dizer que

$$\overline{\mathbf{y}}^{o} = \overline{\mathbf{a}}_1 p_1 + \overline{\mathbf{a}}_2 p_2 + \overline{\mathbf{a}}_3 p_3 + \ldots + \overline{\mathbf{a}}_M\, p_M$$

Então dizemos que o vetor dos dados observados $\overline{\mathbf{y}}^{\mathbf{o}}$ é uma <u>COMBINAÇÃO LINEAR</u> dos vetores $\overline{\mathbf{a}}_1, \mathbf{a}_2, \overline{\mathbf{a}}_3, \ldots, \overline{\mathbf{a}}_M$, que são os M vetores colunas da matriz $\overline{\overline{\mathbf{A}}}$, e $p_1, p_2, p_3, \ldots, p_M$ são escalares que formam o vetor de parâmetros $\overline{\mathbf{p}}$.

INDEPENDÊNCIA LINEAR:

Dizemos que um conjunto de vetores $\overline{\mathbf{a}}_1, \mathbf{a}_2, \overline{\mathbf{a}}_3, \ldots, \overline{\mathbf{a}}_M$ é LINEARMENTE INDEPENDENTE (LI) se existirem coeficientes $p_1, p_2, p_3, \ldots, p_M$ TODOS NULOS tal que

$$\overline{\mathbf{a}}_1 p_1 + \overline{\mathbf{a}}_2 p_2 + \overline{\mathbf{a}}_3 p_3 + \ldots + \overline{\mathbf{a}}_M\, p_M = \overline{\mathbf{0}}$$

Em outras palavras, os vetores $\bar{\mathbf{a}}_1, \mathbf{a}_2, \bar{\mathbf{a}}_3, \ldots, \bar{\mathbf{a}}_M$ forma um conjunto LI se o sistema homogêneo, $\bar{\bar{\mathbf{A}}}\bar{\mathbf{p}} = \bar{\mathbf{0}}$, admitir apenas a solução trivial ($\bar{\mathbf{p}} = \bar{\mathbf{0}}$); caso contrário, estes vetores formam um conjunto LINEARMENTE DEPENDENTE (LD)

<u>Propriedades:</u>

- Um conjunto de vetores é LI se e somente se NENHUM vetor deste conjunto de vetores pode ser expresso como uma combinação linear dos outros vetores.
- Um conjunto de vetores é LD se pelo menos um vetor deste conjunto de vetores pode ser expresso como uma combinação linear dos outros vetores.
- Se um dos vetores $\bar{\mathbf{a}}_1, \mathbf{a}_2, \bar{\mathbf{a}}_3, \ldots, \bar{\mathbf{a}}_M$ é nulo, o conjunto é LD
- Seja um conjunto de vetores $\bar{\mathbf{a}}_1, \mathbf{a}_2, \bar{\mathbf{a}}_3, \ldots, \bar{\mathbf{a}}_M \in R^N$. Se M > N, então este conjunto de vetores é LD

Estes conceitos de dependência e independência linear que acabamos de discutir são conceitos fundamentais da Álgebra Linear. Veremos que tais conceitos auxiliam no entendimento de um sistema linear e no estudo de espaço vetoriais.

<u>BASE E DIMENSÃO DE UM ESPAÇO VETORIAL:</u>

Geralmente, quando pensamos numa linha pensamos em um espaço de uma dimensão, quando pensamos em um plano pensamos no espaço de duas dimensões e quando pensamos no espaço ao nosso redor pensamos no espaço de três dimensões. Para entendermos o que é de fato a DIMENSÃO de um espaço vetorial precisamos primeiro definir a BASE deste espaço.

<u>BASE:</u> A base de um espaço vetorial é um conjunto LI de vetores que geram o espaço vetorial.

<u>Propriedades:</u>

- Um conjunto de vetores que constituem uma base de um espaço vetorial NÃO é único.

- O número de vetores em uma base é ÚNICO, então se um espaço vetorial tem uma base contendo N vetores qualquer outra base deste mesmo espaço vetorial terá N vetores

DIMENSÃO:

É o número de vetores em uma base de um espaço vetorial (ex.: o espaço vetorial R^2 tem uma base formada por dois vetores, logo a dimensão deste espaço é 2).

POSTO DE UMA MATRIZ:

Uma matriz $\overline{\overline{\mathbf{A}}}$ (N x M)

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{pmatrix}$$

É formada por N vetores linhas ($\overline{\mathbf{r}}_i \in R^M$, $i = 1,2,..., N$)

$$\begin{aligned} \overline{\mathbf{r}}_1 &= (a_{11} \quad a_{12} \quad \cdots \quad a_{1M}) \\ \overline{\mathbf{r}}_2 &= (a_{21} \quad a_{22} \quad \cdots \quad a_{2M}) \\ &\vdots \qquad\qquad \vdots \\ \overline{\mathbf{r}}_N &= (a_{N1} \quad a_{N2} \quad \cdots \quad a_{NM}) \end{aligned}$$

Ou formada por M vetores colunas $\overline{\mathbf{c}}_j \in R^N$, $j = 1,2,..., M$

$$\bar{\mathbf{c}}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{bmatrix} \quad \bar{\mathbf{c}}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{N2} \end{bmatrix} \quad \cdots \quad \bar{\mathbf{c}}_M = \begin{bmatrix} a_{1M} \\ a_{2M} \\ \vdots \\ a_{NM} \end{bmatrix}$$

Podemos então escrever a matriz $\overline{\overline{\mathbf{A}}}$ (N x M) de dois diferentes modos

$$\overline{\overline{\mathbf{A}}} = \begin{bmatrix} \bar{\mathbf{r}}_1 \\ \bar{\mathbf{r}}_2 \\ \vdots \\ \bar{\mathbf{r}}_N \end{bmatrix}$$

ou

$$\overline{\overline{\mathbf{A}}} = \begin{bmatrix} \bar{\mathbf{c}}_1 & \bar{\mathbf{c}}_2 & \cdots & \bar{\mathbf{c}}_M \end{bmatrix}$$

ESPAÇO LINHA DE UMA MATRIZ $\overline{\overline{\mathbf{A}}}$ (N x M)

A matriz $\overline{\overline{\mathbf{A}}}$ tem um total e N linhas, ou seja, N vetores linhas ($\bar{\mathbf{r}}_i \in R^M, i = 1,2,..., N$) em que cada vetor linha tem M elementos.

Então o espaço linha da matriz $\overline{\overline{\mathbf{A}}}$ (N x M) é aquele subespaço de R^M gerado pelo conjunto das linhas $\overline{\overline{\mathbf{A}}}$. Em outras palavras, são os vetores linha que são LI.

ESPAÇO COLUNA DE UMA MATRIZ $\overline{\overline{\mathbf{A}}}$ (N x M)

A matriz $\overline{\overline{\mathbf{A}}}$ tem um total e M colunas, ou seja, M vetores coluna $(\overline{\mathbf{c}}_j \in R^N,\ j = 1,2,\ldots, M\ )$ em que cada vetor coluna tem N elementos.

Então o espaço coluna da matriz $\overline{\overline{\mathbf{A}}}$ (N x M) é aquele subespaço de R^N gerado pelo conjunto das colunas de $\overline{\overline{\mathbf{A}}}$. Em outras palavras, são os vetores colunas que são LI.

Se $\overline{\overline{\mathbf{A}}}$ (N x M) é qualquer matriz, então o espaço linha e o espaço coluna tem a mesma dimensão.

POSTO (RANK) DA MATRIZ $\overline{\overline{\mathbf{A}}}$ (N X M) :

É a dimensão do espaço linha (ou do espaço coluna) da matriz $\overline{\overline{\mathbf{A}}}$. Em outras palavras, é o número de vetores linhas (ou vetores colunas) que formam um conjunto LI de vetores.

Propriedades:

- Posto de $\overline{\overline{\mathbf{A}}}$ = Posto de $\overline{\overline{\mathbf{A}}}^T$

POSTO X SISTEMA DE EQUAÇÕES LINEARES:

Vamos agora estabelecer a relação entre o posto da matriz $\overline{\overline{\mathbf{A}}}$ (N x M) e a solução de um sistema não homogêneo de equações lineares $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$.

Se $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$ é um sistema de N equações lineares em M incógnitas $p_1, p_2, p_3, \ldots, p_M$ e $\overline{\mathbf{a}}_1, \overline{\mathbf{a}}_2, \overline{\mathbf{a}}_3, \ldots, \overline{\mathbf{a}}_M \in \mathrm{R}^N$ são os vetores colunas da matriz $\overline{\overline{\mathbf{A}}}$. Então podemos dizer que o vetor de dados observados é uma combinação linear dos vetores colunas da matriz $\overline{\overline{\mathbf{A}}}$, ou seja,

$$\overline{\mathbf{y}}^o = \overline{\mathbf{a}}_1 p_1 + \overline{\mathbf{a}}_2 p_2 + \overline{\mathbf{a}}_3 p_3 + \ldots + \overline{\mathbf{a}}_M\ p_M$$

Já vimos que um sistema linear pode: 1) não ter solução (sistema inconsistente); 2) ter uma única solução; e 3) ter infinitas soluções.

- O sistema $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$ é **inconsistente** se e somente se o posto da matriz aumentada $\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right]$ é MAIOR que o posto da matriz $\overline{\overline{\mathbf{A}}}$.
  $r\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right] > r\left[\overline{\overline{\mathbf{A}}}\right]$
- O sistema $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$ tem **solução única** se o posto da matriz aumentada $\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right]$ é IGUAL ao posto (r) da matriz $\overline{\overline{\mathbf{A}}}$ e é igual ao número de incógnitas r = M.
  $r\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right] = r\left[\overline{\overline{\mathbf{A}}}\right] = M$
- O sistema $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$ tem **infinitas soluções** se o posto da matriz aumentada $\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right]$ é IGUAL ao posto (r) da matriz $\overline{\overline{\mathbf{A}}}$ e é MENOR que o número de incógnitas r < M.
  $r\left[\overline{\overline{\mathbf{A}}}, \overline{\mathbf{y}}^o\right] = r\left[\overline{\overline{\mathbf{A}}}\right] < M$

POSTO X SISTEMA DE EQUAÇÕES LINEARES X SISTEMA HOMOGÊNEO:

Qual é a relação entre a solução de um sistema não homogêneo, $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$, e o correspondente sistema homogêneo, $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{0}}$?

A relação entre estes dois sistemas é descrito pela **Alternativa de Fredholm** que possibilita responder a questão da existência e unicidade da solução do sistema linear arbitrário, $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^o$, através do simples estudo do sistema homogêneo correspondente, $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{0}}$.

Alternativa de Fredholm: Existe precisamente UMA ÚNICA solução de $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^{o}$ se e somente se a solução trivial, $\overline{\mathbf{p}} = \overline{\mathbf{0}}$, é a única solução do sistema homogêneo correspondente, $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{0}}$.

Como conseqüência temos a seguinte propriedade:

$\overline{\mathbf{p}} = \overline{\mathbf{0}}$ é a única solução do sistema homogêneo $\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{0}}$ se e somente se o posto de é igual ao número de incógnitas (M)

$$r\left\lfloor \overline{\overline{\mathbf{A}}} \right\rfloor = M$$

NORMA DE VETORES:

Vamos agora introduzir a noção de tamanho de um vetor. Em duas dimensões, medimos o comprimento geométrico de um vetor via teorema de Pitágoras como sendo $(x_1^2 + x_2^2)^{1/2}$. Antes de definirmos a norma de vetores vamos ao conceito de produto interno (ou produto escalar) Euclideano.

PRODUTO ESCALAR EUCLIDEANO.

Seja dois vetores $\overline{\mathbf{u}}$ e $\overline{\mathbf{v}}$, $\in \mathrm{R}^{\mathrm{N}}$, então o produto escalar Euclideano é definido como:

$$\left\langle \overline{\mathbf{u}}, \overline{\mathbf{v}} \right\rangle = u_1 v_1 + u_2 v_2 + \ldots + u_N v_N \,.$$

O produto escalar Euclideano pode também ser escrito como:

$$\left\langle \overline{\mathbf{u}}, \overline{\mathbf{v}} \right\rangle = \overline{\mathbf{u}}^{T}\overline{\mathbf{v}} = u_1 v_1 + u_2 v_2 + \ldots + u_N v_N$$

Propriedades:

Seja três vetores $\overline{\mathbf{u}}$, $\overline{\mathbf{v}}$ e $\overline{\mathbf{w}}$, $\in \mathrm{R}^{\mathrm{N}}$ e k um escalar qualquer então temos as seguintes propriedades:

a) $\langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \rangle = \langle \bar{\mathbf{v}}, \bar{\mathbf{u}} \rangle$

b) $\langle \bar{\mathbf{u}} + \bar{\mathbf{v}}, \bar{\mathbf{w}} \rangle = \langle \bar{\mathbf{u}}, \bar{\mathbf{w}} \rangle + \langle \bar{\mathbf{v}}, \bar{\mathbf{w}} \rangle$

c) $\langle k\, \bar{\mathbf{u}}, \bar{\mathbf{v}} \rangle = k \langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \rangle$

d) $\langle \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle \geq 0$

e) Se $\langle \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle = 0$, então $\bar{\mathbf{v}} = \bar{\mathbf{0}}$

<u>NORMA EUCLIDEANA DE UM VETOR</u> $\in \mathrm{R}^{\mathrm{N}}$

Se o vetor $\bar{\mathbf{u}} \in \mathrm{R}^{\mathrm{N}}$, então o comprimento (tamanho) deste vetor é expresso como:

$$\|\bar{\mathbf{u}}\|_2 = \langle \bar{\mathbf{u}}, \bar{\mathbf{u}} \rangle^{1/2} = (\bar{\mathbf{u}}^T \bar{\mathbf{u}})^{1/2} = \left(u_1{}^2 + u_2{}^2 + \ldots + u_N{}^2 \right)^{1/2}$$

sendo $\| \cdot \|_2$ a norma Euclideana de um vetor ou simplesmente a norma 2 de um vetor. Portanto $\|\bar{\mathbf{u}}\|_2$ fornece o tamanho do vetor $\bar{\mathbf{u}}$.

PROPRIEDADES DA NORMA EUCLIDEANA

- $\|\bar{\mathbf{u}}\|_2 \geq 0$
- $\|\bar{\mathbf{u}}\|_2 = 0 \ \text{se}\, \bar{\mathbf{u}} = \bar{\mathbf{0}}$
- $\|k\bar{\mathbf{u}}\|_2 = k \|\bar{\mathbf{u}}\|_2$
- $\|\bar{\mathbf{u}} + \bar{\mathbf{v}}\|_2 \leq \|\bar{\mathbf{u}}\|_2 + \|\bar{\mathbf{v}}\|_2$

DISTÂNCIA EUCLIDEANA ENTRE DOIS VETORES $\in \mathrm{R}^{\mathrm{N}}$

Seja dois vetores $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$, $\in \mathrm{R}^{\mathrm{N}}$, então a distância Euclideana entre $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$ é definida como

$$d\left(\bar{\mathbf{u}}, \bar{\mathbf{v}}\right) = \left\| \bar{\mathbf{u}} - \bar{\mathbf{v}} \right\|_2 = \left[(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_N - v_N)^2 \right]^{1/2}$$

Rescrevendo a equação acima temos

$$d\left(\bar{\mathbf{u}}, \bar{\mathbf{v}}\right) = \left\| \bar{\mathbf{u}} - \bar{\mathbf{v}} \right\|_2 = \left[\left(\bar{\mathbf{u}} - \bar{\mathbf{v}}\right)^T \left(\bar{\mathbf{u}} - \bar{\mathbf{v}}\right) \right]^{1/2}$$

$$d\left(\bar{\mathbf{u}}, \bar{\mathbf{v}}\right) = \left\| \bar{\mathbf{u}} - \bar{\mathbf{v}} \right\|_2 = \left[\bar{\mathbf{u}}^T \bar{\mathbf{u}} - 2\bar{\mathbf{u}}^T \bar{\mathbf{v}} + \bar{\mathbf{v}}^T \bar{\mathbf{v}} \right]^{1/2}$$

$$d\left(\bar{\mathbf{u}}, \bar{\mathbf{v}}\right)^2 = \left\| \bar{\mathbf{u}} - \bar{\mathbf{v}} \right\|_2^2 = \left[\bar{\mathbf{u}}^T \bar{\mathbf{u}} - 2\bar{\mathbf{u}}^T \bar{\mathbf{v}} + \bar{\mathbf{v}}^T \bar{\mathbf{v}} \right]$$

$$d\left(\bar{\mathbf{u}}, \bar{\mathbf{v}}\right)^2 = \left\| \bar{\mathbf{u}} - \bar{\mathbf{v}} \right\|_2^2 = \left\| \bar{\mathbf{u}} \right\|_2^2 - 2\left\langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \right\rangle + \left\| \bar{\mathbf{v}} \right\|_2^2$$

ÂNGULO ENTRE DOIS VETORES

Seja dois vetores $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$, $\in \mathrm{R}^{\mathrm{N}}$, então o ângulo entre $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$ é definido como

$$\left\langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \right\rangle = \left\| \bar{\mathbf{u}} \right\|_2 \left\| \bar{\mathbf{v}} \right\|_2 . \cos\theta$$

$$\cos\theta = \frac{\left\langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \right\rangle}{\left\| \bar{\mathbf{u}} \right\|_2 \left\| \bar{\mathbf{v}} \right\|_2}$$

VETORES ORTOGONAIS

Se dois vetores $\bar{\mathbf{u}}$ e $\bar{\mathbf{v}}$, $\in \mathrm{R}^{\mathrm{N}}$, são chamados **vetores ortogonais** então

$$\langle \overline{\mathbf{u}}, \overline{\mathbf{v}} \rangle = 0$$

$$\overline{\mathbf{u}}^T \overline{\mathbf{v}} = 0$$

$$u_1 v_1 + u_2 v_2 + \ldots + u_N v_N = 0$$

<u>CONJUNTO DE VETORES ORTOGONAIS EM UM ESPAÇO VETORIAL</u>

Um conjunto S de vetores $S = \{\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_M\}$, em que $\overline{\mathbf{v}}_i \in R^N$, é chamado de conjunto ortogonal se todos os pares de vetores é ortogonal, ou seja,

$$\overline{\mathbf{v}}_i^T \overline{\mathbf{v}}_j = 0 \quad \text{para i} \neq \text{j em que } i = 1,2,\ldots,M \text{ e } j = 1,2,\ldots,M$$

<u>CONJUNTO DE VETORES ORTONORMAIS EM UM ESPAÇO VETORIAL</u>

Um conjunto S de vetores $S = \{\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_M\}$, em que $\overline{\mathbf{v}}_i \in R^N$, é chamado de conjunto ortonormal se este conjunto de vetores é ortogonal e cada vetor tem norma 1, ou seja,

$$\overline{\mathbf{v}}_i^T \overline{\mathbf{v}}_j = 0 \quad \text{para i} \neq \text{j em que } i = 1,2,\ldots,M \text{ e } j = 1,2,\ldots,M$$

e

$$\overline{\mathbf{v}}_i^T \overline{\mathbf{v}}_i = 1 \quad i = 1,2,\ldots,M$$

<u>BASE ORTOGONAL:</u> É uma base contendo vetores ORTOGONAIS

<u>BASE ORTONORMAL:</u> É uma base contendo vetores ORTONORMAIS

VETOR GRADIENTE

Dado um vetor de parâmetros $\bar{\mathbf{p}} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}$ (M x 1), definiremos o OPERADOR GRADIENTE DOS PARÂMETROS como um vetor $\overline{\nabla}_{\bar{\mathbf{p}}} \in R^M$ que é definido como

$$\overline{\nabla}_{\bar{\mathbf{p}}} = \begin{bmatrix} \dfrac{\partial}{\partial p_1} \\ \dfrac{\partial}{\partial p_2} \\ \vdots \\ \dfrac{\partial}{\partial p_M} \end{bmatrix}$$

Então $\overline{\nabla}_{\bar{\mathbf{p}}} \in R^M$ é um vetor (M X 1) cujo o j-ésimo elemento é a derivada parcial com relação ao j-ésimo parâmetro p_j

$$\overline{\nabla}_{\bar{\mathbf{p}}} \equiv \left[\frac{\partial}{\partial p_j} \right] \quad j = 1,2,\ldots,M$$

Veja que $\overline{\nabla}_{\bar{\mathbf{p}}} \in R^M$ é um vetor operador que deve ser aplicado a um escalar ou a transposta de um vetor coluna.

Vetor Gradiente Aplicado a um Escalar

Seja Q um escalar então se aplicarmos o vetor gradiente $\overline{\nabla}_{\bar{\mathbf{p}}} \in R^M$ ao escalar Q obtemos um vetor (M x 1) expresso como

$$\overline{\nabla}_{\overline{\mathbf{p}}}\{Q\} = \begin{bmatrix} \dfrac{\partial Q}{\partial p_1} \\ \dfrac{\partial Q}{\partial p_2} \\ \vdots \\ \dfrac{\partial Q}{\partial p_M} \end{bmatrix}, \quad \in \mathrm{R}^{\mathrm{M}}$$

Vetor Gradiente Aplicado a um Vetor

Seja um vetor N-dimensional $\overline{\mathbf{x}} \in R^N$, então aplicando o operador gradiente $\overline{\nabla}_{\overline{\mathbf{p}}} \in R^M$ ao vetor $\overline{\mathbf{x}} \in R^N$ obteremos a primeira derivada do vetor $\overline{\mathbf{x}}$ em relação aos parâmetros $\overline{\mathbf{p}}$, que é definida calculando

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \underset{(1\times N)}{\{\overline{\mathbf{x}}^T\}} = \begin{bmatrix} \dfrac{\partial}{\partial p_1} \\ \dfrac{\partial}{\partial p_2} \\ \vdots \\ \dfrac{\partial}{\partial p_M} \end{bmatrix}_{(M\times 1)} \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}_{(N\times 1)}$$

O resultado é uma matriz

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \underset{(1\times N)}{\{\overline{\mathbf{x}}^T\}} = \begin{pmatrix} \dfrac{\partial x_1}{\partial p_1} & \dfrac{\partial x_2}{\partial p_1} & \cdots & \dfrac{\partial x_N}{\partial p_1} \\ \dfrac{\partial x_1}{\partial p_2} & \dfrac{\partial x_2}{\partial p_2} & \cdots & \dfrac{\partial x_N}{\partial p_2} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial x_1}{\partial p_M} & \dfrac{\partial x_2}{\partial p_M} & \cdots & \dfrac{\partial x_N}{\partial p_M} \end{pmatrix}_{(M\times N)}$$

Então o vetor operador gradiente aplicado a um vetor gera uma matriz.

<u>Caso Especial</u> $\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}}\ \underset{(1\times M)}{\left\{\overline{\mathbf{p}}^T\right\}}$:

Veja que se $\overline{\mathbf{p}} \in R^M$ é um vetor de parâmetros, então aplicando o operador gradiente $\overline{\nabla}_{\overline{\mathbf{p}}} \in R^M$ ao vetor $\overline{\mathbf{p}} \in R^M$ obteremos a primeira derivada do vetor $\overline{\mathbf{p}}$ em relação aos parâmetros $\overline{\mathbf{p}}$, que é a matriz identidade

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}}\ \underset{(1\times M)}{\left\{\overline{\mathbf{p}}^T\right\}} = \begin{bmatrix} \dfrac{\partial}{\partial p_1} \\ \dfrac{\partial}{\partial p_2} \\ \vdots \\ \dfrac{\partial}{\partial p_M} \end{bmatrix}_{(M\times 1)} \begin{bmatrix} p_1 & p_2 & \cdots & p_M \end{bmatrix}_{(1\times M)}$$

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}}\ \underset{(1\times M)}{\left\{\overline{\mathbf{p}}^T\right\}} = \begin{pmatrix} \dfrac{\partial p_1}{\partial p_1} & \dfrac{\partial p_2}{\partial p_1} & \cdots & \dfrac{\partial p_M}{\partial p_1} \\ \dfrac{\partial p_1}{\partial p_2} & \dfrac{\partial p_2}{\partial p_2} & \cdots & \dfrac{\partial p_M}{\partial p_2} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial p_1}{\partial p_M} & \dfrac{\partial p_2}{\partial p_M} & \cdots & \dfrac{\partial p_M}{\partial p_M} \end{pmatrix}_{(M\times M)} = \underset{(M\times M)}{\overline{\overline{\mathbf{I}}}}$$

<u>Vetor Gradiente Aplicado a um Vetor que Multiplica uma Matriz</u>

Seja o produto $\overline{\mathbf{x}}^T\ \overline{\overline{\mathbf{A}}}$ em que o vetor $\overline{\mathbf{x}} \in R^N$ é função de $\overline{\mathbf{p}}$ e a matriz $\overline{\overline{\mathbf{A}}} \in R^{N\times M}$ não é função de $\overline{\mathbf{p}}$. Então aplicando o operador gradiente $\overline{\nabla}_{\overline{\mathbf{p}}} \in R^M$ ao vetor $\bar{\mathbf{x}}^T \; \bar{\bar{\mathbf{A}}}$ obteremos a primeira derivada do vetor $\bar{\mathbf{x}}^T \; \bar{\bar{\mathbf{A}}}$ em relação aos parâmetros $\bar{\mathbf{p}}$, que é definida como

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times N)}{\bar{\mathbf{x}}^T} \; \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}} \right\} = \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times N)}{\bar{\mathbf{x}}^T} \right\} \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}}$$

No caso especial em que $\bar{\mathbf{x}} = \bar{\mathbf{p}} \in R^M$ temos que

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{p}}^T} \; \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}} \right\} = \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{p}}^T} \right\} \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}}$$

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{p}}^T} \; \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}} \right\} = \underset{(N\times M)}{\bar{\bar{\mathbf{A}}}}$$

<u>FORMAS QUADRÁTICAS</u>

Até agora mostramos equações lineares, ou seja, equações do tipo

$$a_1 p_1 + a_2 p_2 + ,..., + a_M p_M = y_1^o$$

Se temos um sistema de equações lineares com N equações temos

$$\begin{cases} a_{11} p_1 + a_{12} p_2 + \ldots + a_{1M} p_M & = & y_1^0 \\ a_{21} p_1 + a_{22} p_2 + \ldots + a_{2M} p_M & = & y_2^0 \\ \vdots & = & \vdots \\ a_{N1} p_1 + a_{N2} p_2 + \ldots + a_{NM} p_M & = & y_N^0 \end{cases}$$

Este sistema de N equações lineares em M incógnitas pode ser escrito em notação matricial como

$$\underset{(N\times M)}{\overline{\overline{\mathbf{A}}}}\;\underset{(M\times 1)}{\overline{\mathbf{p}}} = \underset{(N\times 1)}{\overline{\mathbf{y}}^{\mathbf{o}}}$$

Note que na forma linear temos funções lineares, ou seja, as M variáveis ou incógnitas ($p_1, p_2, p_3, ..., p_M$) são todas potência de 1 e não há o produto entre variáveis. Ao contrário, na forma quadrática estudamos funções em que as variáveis são potências de 2 ou produto entre variáveis.

<u>A Forma Quadrática em Duas Variáveis</u>

A forma quadrática em duas variáveis (p_1 e p_2) é definida como

$$ap_1^2 + 2bp_1p_2 + cp_2^2$$

Em notação matricial temos

$$\begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

Note que neste caso simples em que temos duas variáveis p_1 e p_2 a matriz (2 x 2) é simétrica em que os elementos da diagonal são os coeficientes dos termos quadráticos e os elementos fora da diagonal são ½ dos coeficiente dos termos que são os produtos entre variáveis.

<u>A Forma Quadrática em M Variáveis</u>

A forma quadrática em M variáveis ($p_1, p_2, p_3, ..., p_M$) pode ser escrita em forma matricial como

$$\begin{bmatrix} p_1 & p_2 & \cdots & p_M \end{bmatrix} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}$$

em que $\overline{\overline{\mathbf{A}}}$ é uma matriz simétrica (M X M) que em forma compacta temos

$$Q = \bar{\mathbf{p}}^T \bar{\bar{\mathbf{A}}} \; \bar{\mathbf{p}}$$

Usando a propriedade que $\bar{\bar{\mathbf{A}}}$ é uma matriz simétrica então $\bar{\bar{\mathbf{A}}}^T = \bar{\bar{\mathbf{A}}}$ logo temos

$$Q = \bar{\mathbf{p}}^T \bar{\bar{\mathbf{A}}} \; \bar{\mathbf{p}} = \bar{\mathbf{p}}^T \bar{\bar{\mathbf{A}}}^T \; \bar{\mathbf{p}} = \left(\bar{\bar{\mathbf{A}}}\bar{\mathbf{p}}\right)^T \bar{\mathbf{p}}$$

<u>Derivadas de uma Forma Quadrática</u>

Vamos definir uma forma quadrática como

$$Q = \bar{\mathbf{x}}^T \bar{\bar{\mathbf{A}}} \; \bar{\mathbf{x}}$$

em que o vetor $\bar{\mathbf{x}} \in R^M$ é função de $\bar{\mathbf{p}}$ e $\bar{\bar{\mathbf{A}}} \in R^{M \times M}$ é uma matriz simétrica que não é função de $\bar{\mathbf{p}}$. Então aplicando o operador gradiente $\bar{\nabla}_{\bar{\mathbf{p}}} \in R^M$ a forma quadrática Q temos

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \underset{(M\times M)}{\bar{\bar{\mathbf{A}}}} \underset{(M\times 1)}{\bar{\mathbf{x}}} \right\} = \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \right\} \underset{(M\times M)}{\bar{\bar{\mathbf{A}}}} \underset{(M\times 1)}{\bar{\mathbf{x}}} + \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \underset{(M\times M)}{\bar{\bar{\mathbf{A}}}^T} \right\} \underset{(M\times 1)}{\bar{\mathbf{x}}}$$

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \underset{(M\times M)}{\bar{\bar{\mathbf{A}}}} \underset{(M\times 1)}{\bar{\mathbf{x}}} \right\} = 2\, \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \right\} \underset{(M\times M)}{\bar{\bar{\mathbf{A}}}} \underset{(M\times 1)}{\bar{\mathbf{x}}}$$

Casos Especiais:

(1) No caso especial em que $\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{I}}}$ temos a forma quadrática

$$Q = \bar{\mathbf{x}}^T \; \bar{\mathbf{x}}$$

Então aplicando o operador gradiente $\bar{\nabla}_{\bar{\mathbf{p}}} \in R^M$ a forma quadrática acima temos

$$\underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \underset{(M\times 1)}{\bar{\mathbf{x}}} \right\} = 2\, \underset{(M\times 1)}{\bar{\nabla}_{\bar{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\bar{\mathbf{x}}^T} \right\} \underset{(M\times 1)}{\bar{\mathbf{x}}}$$

(2) No caso especial em que $\overline{\overline{\mathbf{A}}} \neq \overline{\overline{\mathbf{I}}}$ é uma matriz simétrica M x M e $\overline{\mathbf{x}} = \overline{\mathbf{p}} \in R^M$ temos a seguinte forma quadrática

$$Q = \overline{\mathbf{p}}^T \overline{\overline{\mathbf{A}}} \overline{\mathbf{p}}$$

Então aplicando o operador gradiente $\overline{\nabla}_{\overline{\mathbf{p}}} \in R^M$ a forma quadrática acima temos

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \underset{(M\times 1)}{\overline{\mathbf{p}}} \right\} = 2 \underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \right\} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \underset{(M\times 1)}{\overline{\mathbf{p}}}$$

Usando a propriedade que $\overline{\overline{\mathbf{A}}}$ é uma matriz simétrica , $\overline{\overline{\mathbf{A}}}^T = \overline{\overline{\mathbf{A}}}$, então temos

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \underset{(M\times 1)}{\overline{\mathbf{p}}} \right\} = 2 \underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \right\} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}^T} \underset{(M\times 1)}{\overline{\mathbf{p}}}$$

Como $\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \right\} = \overline{\overline{\mathbf{I}}}$ temos

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}} \underset{(M\times 1)}{\overline{\mathbf{p}}} \right\} = 2 \underset{(M\times M)}{\overline{\overline{\mathbf{A}}}^T} \underset{(M\times 1)}{\overline{\mathbf{p}}}$$

(3) No caso especial em que $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{I}}}$ temos a forma quadrática a seguinte forma quadrática

$$Q = \overline{\mathbf{p}}^T \overline{\mathbf{p}}$$

Então aplicando o operador gradiente $\overline{\nabla}_{\overline{\mathbf{p}}} \in R^M$ à forma quadrática temos

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \underset{(M\times 1)}{\overline{\mathbf{p}}} \right\} = 2 \underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \right\} \underset{(M\times 1)}{\overline{\mathbf{p}}}$$

$$\underset{(M\times 1)}{\overline{\nabla}_{\overline{\mathbf{p}}}} \left\{ \underset{(1\times M)}{\overline{\mathbf{p}}^T} \underset{(M\times 1)}{\overline{\mathbf{p}}} \right\} = 2 \underset{(M\times 1)}{\overline{\mathbf{p}}} \, .$$

Exercício Teórico 1:

1) Se $\overline{\overline{\mathbf{A}}} \in R^{N \times M}$ pode ser rescrita como $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{U}}}\,\overline{\overline{\mathbf{S}}}\,\overline{\overline{\mathbf{V}}}^{\mathbf{T}}$, em que $\overline{\overline{\mathbf{U}}}$ (N x N) e $\overline{\overline{\mathbf{V}}}$ (M x M) são matrizes ortogonais e $\overline{\overline{\mathbf{S}}}$ é uma matriz diagonal (N X M).

1.1. Ache a transposta de $\overline{\overline{\mathbf{A}}}$

1.2. Ache as matrizes $\overline{\overline{\mathbf{A}}}^T\overline{\overline{\mathbf{A}}}$ e $\overline{\overline{\mathbf{A}\mathbf{A}}}^T$ e prove que estas matrizes são simétricas

1.3. Ache as matrizes $\left(\overline{\overline{\mathbf{A}}}^T\overline{\overline{\mathbf{A}}}\right)^{-1}$ e $\left(\overline{\overline{\mathbf{A}\mathbf{A}}}^T\right)^{-1}$

1.4. Mostre que o (i,j)-ésimo elemento da matriz $\left(\overline{\overline{\mathbf{A}}}^T\overline{\overline{\mathbf{A}}}\right)^{-1}$ é expresso por

$$\sum_{j=1}^{M} \frac{v_{k\,j} v_{j\,k}}{s_j^2} \qquad k = 1,2...,M$$

em que v_{kj} é o elemento (k,j) da matriz $\overline{\overline{\mathbf{V}}}$ e s_j é o j-ésimo elemento da matriz $\overline{\overline{\mathbf{S}}}$

1.5. Ache a matriz $\overline{\overline{\mathbf{A}}}^{+} = \left(\overline{\overline{\mathbf{A}}}^T\overline{\overline{\mathbf{A}}}\right)^{-1}\overline{\overline{\mathbf{A}}}^T$

1.6. Ache a matriz $\overline{\overline{\mathbf{A}}}^{+} = \overline{\overline{\mathbf{A}}}^T\left(\overline{\overline{\mathbf{A}}}\,\overline{\overline{\mathbf{A}}}^T\right)^{-1}$,

1.7. Considere que $\overline{\overline{\mathbf{S}}}$ é uma matriz diagonal (N X M) contendo r valores não nulos em que $r < M = N$

$$\overline{\overline{\mathbf{S}}} = \begin{bmatrix} s_1 & & & & & & \\ & s_2 & & & & & \\ & & \ddots & & & & \\ & & & s_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

Esta matriz acima pode ser particionada como $\bar{\bar{\mathbf{S}}} = \begin{bmatrix} \bar{\bar{\mathbf{S}}}_r & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} \end{bmatrix}$ em que $\bar{\bar{\mathbf{S}}}_r$ é uma matriz diagonal (r X r). Mostre que a matriz $\bar{\bar{\mathbf{A}}} \in R^{N \times M}$ pode ser rescrita como $\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{U}}}_\mathbf{r} \; \bar{\bar{\mathbf{S}}}_\mathbf{r} \; \bar{\bar{\mathbf{V}}}_\mathbf{r}^\mathbf{T}$

em que a matriz $\bar{\bar{\mathbf{U}}}_r = \left[\bar{\mathbf{u}}_1 \, \bar{\mathbf{u}}_2 .. \bar{\mathbf{u}}_r \right]$ é uma matriz (N x r) e a matriz $\bar{\bar{\mathbf{V}}}_r = \left[\bar{\mathbf{v}}_1 \, \bar{\mathbf{v}}_2 \ldots \bar{\mathbf{v}}_r \right]$ é uma matriz (M x r).

1.8. Considere que $\bar{\bar{\mathbf{S}}}$ é uma matriz diagonal (N X M) contendo r valores não nulos e M-r valores muito pequenos em que $r < M \le N$. Portanto, esta matriz pode ser particionada como

$$\bar{\bar{\mathbf{S}}} = \begin{bmatrix} \bar{\bar{\mathbf{S}}}_r & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{S}}}_{M-r} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} \end{bmatrix}$$

em que $\bar{\bar{\mathbf{S}}}_r$ é uma matriz diagonal de dimensão r x r composta dos r elementos Não Nulos e $\bar{\bar{\mathbf{S}}}_{M-r}$ é uma matriz diagonal de dimensão M-r x M-r composta dos M-r elementos pequenos (próximos de zero).

a) Como a matriz $\bar{\bar{\mathbf{A}}} \in R^{N \times M}$ pode ser rescrita?

b) No caso em que $\bar{\bar{\mathbf{S}}}_{M-r} = \bar{\bar{0}} \in \mathrm{R}^{\mathrm{M-r}}$ mostre que $\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{U}}}_\mathbf{r} \; \bar{\bar{\mathbf{S}}}_\mathbf{r} \; \bar{\bar{\mathbf{V}}}_\mathbf{r}^\mathbf{T}$

2) Se $\bar{\bar{\mathbf{A}}} \in R^{N \times M}$ é uma matriz expressa como $\bar{\bar{\mathbf{A}}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ (N=2 e M=3) e que pode ser rescrita como $\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{U}}}\,\bar{\bar{\mathbf{S}}}\,\bar{\bar{\mathbf{V}}}^\mathbf{T}$ em que $\bar{\bar{\mathbf{U}}}$ e $\bar{\bar{\mathbf{V}}}$ são matrizes ortogonais expressas respectivamente como,

$$\overline{\overline{\mathbf{U}}} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}, \quad \overline{\overline{\mathbf{V}}} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e $\overline{\overline{\mathbf{S}}}$ é uma matriz diagonal contendo apenas um valor não nulo (r=1) $\overline{\overline{\mathbf{S}}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

2.1. Mostre que a matriz $\overline{\overline{\mathbf{A}}}$ pode ser rescrita como $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{U}}}_{\mathbf{r}} \, \overline{\overline{\mathbf{S}}}_{\mathbf{r}} \, \overline{\overline{\mathbf{V}}}_{\mathbf{r}}^{\mathbf{T}}$

2.2 Mostre que $\overline{\overline{\mathbf{U}}}^{\mathbf{T}} \, \overline{\overline{\mathbf{U}}} = \overline{\overline{\mathbf{U}}} \, \overline{\overline{\mathbf{U}}}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}$

2.3 Mostre que $\overline{\overline{\mathbf{V}}}^{\mathbf{T}} \, \overline{\overline{\mathbf{V}}} = \overline{\overline{\mathbf{V}}} \, \overline{\overline{\mathbf{V}}}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}$

2.4 Mostre que $\overline{\overline{\mathbf{U}}}_r{}^{\mathbf{T}} \, \overline{\overline{\mathbf{U}}}_r = \overline{\overline{\mathbf{I}}}$ e $\overline{\overline{\mathbf{U}}}_{N-r}{}^{\mathbf{T}} \, \overline{\overline{\mathbf{U}}}_{N-r} = \overline{\overline{\mathbf{I}}}$

2.5 Mostre que $\overline{\overline{\mathbf{V}}}^{\mathbf{T}}{}_r \, \overline{\overline{\mathbf{V}}}_r = \overline{\overline{\mathbf{I}}}$ e $\overline{\overline{\mathbf{V}}}^{\mathbf{T}}{}_{M-r} \, \overline{\overline{\mathbf{V}}}_{M-r} = \overline{\overline{\mathbf{I}}}$

2.6 Mostre que $\overline{\overline{\mathbf{U}}}_r \, \overline{\overline{\mathbf{U}}}_r^T \neq \overline{\overline{\mathbf{I}}}$ e $\overline{\overline{\mathbf{U}}}_{N-r}{}^{\mathbf{T}} \, \overline{\overline{\mathbf{U}}}_{N-r}^T \neq \overline{\overline{\mathbf{I}}}$

2.7 Mostre que $\overline{\overline{\mathbf{V}}}_r \, \overline{\overline{\mathbf{V}}}_r^T \neq \overline{\overline{\mathbf{I}}}$ e $\overline{\overline{\mathbf{V}}}_{M-r} \, \overline{\overline{\mathbf{V}}}_{M-r}^T \neq \overline{\overline{\mathbf{I}}}$

2.8 Mostre que $\overline{\overline{\mathbf{U}}}_r \, \overline{\overline{\mathbf{U}}}_r^{\mathbf{T}} + \overline{\overline{\mathbf{U}}}_{N-r} \, \overline{\overline{\mathbf{U}}}_{N-r}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}_{(N \times N)}$

2.9 Mostre que $\overline{\overline{\mathbf{V}}}_r \, \overline{\overline{\mathbf{V}}}_{\mathrm{r}}^{\mathrm{T}} + \overline{\overline{\mathbf{V}}}_{M-r} \, \overline{\overline{\mathbf{V}}}_{\mathrm{M-r}}^{\mathrm{T}} = \overline{\overline{\mathbf{I}}}_{(M \times M)}$

2.10 Mostre que $\overline{\overline{\mathbf{V}}}_{\mathrm{r}}^{\mathrm{T}} \, \overline{\overline{\mathbf{V}}}_{M-r} = \overline{\overline{0}}_{(M \times M)}$ e $\overline{\overline{\mathbf{U}}}_r \, \overline{\overline{\mathbf{U}}}_{N-r}^{\mathbf{T}} = \overline{\overline{\mathbf{0}}}_{(N \times N)}$

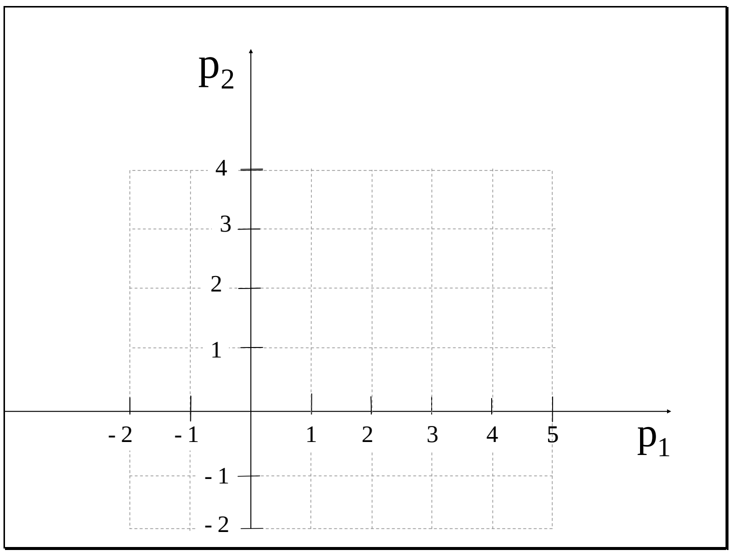
3) Se o vetor $\overline{\boldsymbol{\varepsilon}}$ (N x 1) definido como $\overline{\boldsymbol{\varepsilon}} = \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}$ mostre que

3.1 $\| \overline{\boldsymbol{\varepsilon}} \|_2^{\,2} = \overline{\mathbf{y}}^{\mathbf{o}\,\mathrm{T}} \overline{\mathbf{y}}^{\mathbf{o}} - 2\,\overline{\mathbf{p}}^T \, \overline{\overline{\mathbf{A}}}^T \, \overline{\mathbf{y}}^{\mathbf{o}} + \overline{\mathbf{p}}^T \, \overline{\overline{\mathbf{A}}}^T \, \overline{\overline{\mathbf{A}}} \, \overline{\mathbf{p}}$

3.2 $\overline{\nabla}_{\overline{\mathbf{p}}} \left\{ \| \overline{\boldsymbol{\varepsilon}} \|_2^2 \right\} = -\,2\,\overline{\overline{\mathbf{A}}}^{\mathrm{T}} \, \overline{\mathbf{y}}^{\mathbf{o}} + \overline{\overline{\mathbf{A}}}^{\mathrm{T}} \, \overline{\overline{\mathbf{A}}} \, \overline{\mathbf{p}}$

4) Considere que a matriz $\overline{\overline{\mathbf{A}}} \in \mathrm{R}^{1\times 2}$, o vetor $\overline{\mathbf{y}}^{\mathbf{o}} \in R^1$ e o vetor $\overline{\mathbf{p}} \in R^2$ são respectivamente iguais a $\overline{\overline{\mathbf{A}}}_{(1\times 2)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, $\overline{\mathbf{y}}^{\mathbf{o}}_{(1\times 1)} = 2$ e $\overline{\mathbf{p}}_{(2\times 1)} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$. Considere a função $Q = \| \overline{\boldsymbol{\varepsilon}} \|_2^2$ em que $\overline{\boldsymbol{\varepsilon}}$ é um vetor de dimensão 1 definido como $\overline{\boldsymbol{\varepsilon}} = \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}$.

4.1 Represente graficamente esta função *Q* no espaço de parâmetros (p_1-p_2)



4.2 O mínimo da função *Q* produzido na questão 4.1 não está representado por ponto e sim por uma curva de isovalor. Assinale no gráfico esta curva que representa o mínimo de *Q*.

4.3 Mostre que qualquer par de parâmetros p_1 e p_2 que caem sobre o mínimo da função *Q* satisfaz exatamente a equação $\overline{\overline{\mathbf{A}}}\overline{\mathbf{p}} = \overline{\mathbf{y}}^{\mathbf{o}}$. Faça esta demonstração usando pelo menos 4 diferentes pares de parâmetros.

5) Considerando uma matriz $\overline{\overline{\mathbf{R}}} \in \mathrm{R}^{L\times M}$ que não é função de $\overline{\mathbf{p}}$, calcule o vetor gradiente: $\overline{\nabla}_{\overline{\mathbf{p}}} \left\{ \left\| \overline{\overline{\mathbf{R}}}\,\overline{\mathbf{p}} \right\|_2^2 \right\}$