Departamento de Ciências Florestais
ESALQ - USP

Elementos de Álgebra de Matrizes Importantes para Regressão Linear

João L. F. Batista

Piracicaba
– 2004 –

Sumário

1 Introdução

- A importância da álgebra de matrizes na regressão linear múltipla se relaciona principalmente a dois aspectos dos problemas de regressão múltipla:
  - envolvem sistemas de equações extensos; e
  - envolvem grande conjuntos de dados.
- A álgebra de matrizes na regressão múltipla permite:
  - representação compacta dos sistemas;
  - operações eficientes independentemente do tamanho dos sistemas;
  - notação matemática clara e concisa.

2 Definção de Matrizes

Matrix é um conjunto retangular de elementos arranjados em *linhas* e *colunas*:

$$\begin{array}{l c} & \begin{array}{rr} \text{Coluna 1} & \text{Coluna 2} \end{array} \\ \begin{array}{l} \text{Linha 1} \\ \text{Linha 2} \\ \text{Linha 3} \end{array} & \left[\begin{array}{rr} 34.090 & 2 \\ 125.001 & 9 \\ 5.671 & 5 \end{array}\right] \end{array}$$

A matrix acima é designada de matrix de *ordem* 3×2 por que possui 3 linhas por duas colunas.

Outro exemplos de matrizes são:

$$\text{Matrix } 2 \times 2 : \begin{bmatrix} 0 & 3 \\ 8 & 2 \end{bmatrix} \quad \text{Matrix } 2 \times 3 : \begin{bmatrix} 8 & 9 & 4 \\ 7 & 2 & 1 \end{bmatrix}$$

$$\text{Matrix } 4 \times 1 : \begin{bmatrix} 12 \\ 34 \\ 45 \\ 76 \end{bmatrix} \quad \text{Matrix } 1 \times 3 : \begin{bmatrix} 0.002 & 0.201 & 1.345 \end{bmatrix}$$

Utilizamos índices para representar os elementos de uma matrix:

$$\begin{array}{l c} & \begin{array}{ccc} j=1 & j=2 & j=3 \end{array} \\ \begin{array}{l} i=1 \\ i=2 \end{array} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \end{array}$$

sendo que o índice i representa as *linhas* e o índice j representa as *colunas*.

Para diferenciar matrizes de outros entes matemáticos, o nome das matrizes são geralmente indicados por letras em **negrito**, como $\mathbf{A}$, $\mathbf{X}$, $\mathbf{Z}$. Assim a matrix acima pode ser definida como:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

ou então pela expressão:

$$\mathbf{A} = [a_{ij}] \qquad i = 1, 2; \quad j = 1, 2, 3.$$

Note que em ambas notações fica claro que $\mathbf{A}$ é uma matrix 2×3.

Para representar resumidamente uma grande matrix com ordem $n \times p$, frequentemente utilizamos a expressão:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{bmatrix}$$

ou na forma abreviada:

$$\mathbf{A} = [a_{ij}] \qquad i = 1, \dots, n; \quad j = 1, \dots, p.$$

2.1 Algumas Matrizes Especiais

Matrix Quadrada é uma matrix com igual número de linhas e colunas, por exemplo:

$$\text{Matrix } 1 \times 1 : \begin{bmatrix} 0.002 \end{bmatrix} \qquad \text{Matrix } 2 \times 2 : \begin{bmatrix} 0 & 3 \\ 8 & 2 \end{bmatrix}$$

$$\text{Matrix } 3 \times 3 : \begin{bmatrix} 8 & 9 & 4 \\ 7 & 2 & 1 \\ 10 & -3 & -5 \end{bmatrix} \quad \text{Matrix } 4 \times 4 : \begin{bmatrix} 12 & 34 & 45 & 76 \\ 2 & 6 & -1 & 0 \\ -3 & 7 & 10 & 8 \\ 1 & -9 & 0 & 34 \end{bmatrix}$$

Vetores: uma matrix é chamada de *vetor coluna* quando ela possui apenas uma coluna:

$$\mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Uma matrix é dita *vetor linha* quando formada por apenas uma linha:

$$\mathbf{b} = \begin{bmatrix} 3 & 8 & -2 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}$$

Assumiremos que o formato padrão dos vetores é a forma de **vetor coluna**, assim, quando nos referirmos genericamente a um *vetor*, entenda-se *vetor coluna*.

Matrix Escalar ou (simplismente) Escalar é uma matrix composta de apenas uma linha e uma coluna:

$$\mathbf{A} = [7].$$

Matrix Simétrica é uma *matrix quadrada* em que a permutação do número da linha pela coluna resulta em elementos de mesmo valor, i.e.,

$$a_{ij} = a_{ji}, \quad \text{para todo } i \text{ e } j.$$

Por exemplo, na matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 7 & 3 \\ 2 & 3 & 6 \end{bmatrix},$$

temos que:

$$\begin{aligned} a_{12} &= a_{21} = 0 \\ a_{13} &= a_{31} = 2 \\ a_{23} &= a_{32} = 3 \end{aligned}$$

Matrix Diagonal é uma *matrix quadrada* cujos únicos elementos não nulos estão na diagonal principal. Exemplos:

$$\mathbf{A} = \begin{bmatrix} 7 & 0 \\ 0 & 9 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{bmatrix}$$

Note que toda matrix diagonal é necessariamente simétrica.

Matrix Identidade é uma *matrix diagonal* cujos os elementos da diagonal principal são todos unitários:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Vetores e Matrizes Unitários são aqueles em que *todos os elementos* são unitários. Particularmente úteis à análise de regressão são os vetores colunas unitários:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

e as matrizes quadradas unitárias:

$$\mathbf{J}_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \mathbf{J}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Matrix Triangular é uma matrix cujos elementos acima ou abaixo da diagonal principal são nulos.

Matrix Triangular Superior elementos *abaixo* da diagonal principal são *nulos*:

$$\mathbf{L} = \begin{bmatrix} 6 & 3 & 5 & 7 \\ 0 & 9 & 2 & 1 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Matrix Triangular Inferior elementos *acima* da diagonal principal são *nulos*:

$$\mathbf{U} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 5 & 7 & 0 & 0 \\ 2 & 9 & 8 & 0 \\ 7 & 4 & 1 & 8 \end{bmatrix}$$

3 Manipulação Algébrica de Matrizes

3.1 Igualdade de Matrizes

Duas matrizes $\mathbf{A}$ e $\mathbf{B}$ são consideradas iguais quando *todos os seus elementos correspondentes são iguais,* i.e., quando

$$a_{ij} = b_{ij} \quad \text{para todo } i \text{ e } j.$$

A primeira exigência da igualdade é que $\mathbf{A}$ e $\mathbf{B}$ tenham a mesma ordem. Por exemplo, se $\mathbf{A}$ e $\mathbf{B}$ são:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -4 & 0 \\ 0 & 5 \\ 3 & 4 \end{bmatrix}$$

a expressão $\mathbf{A} = \mathbf{B}$ implica que:

$$a_{11} = -4, \quad a_{12} = 0, \quad a_{21} = 0, \quad a_{22} = 5, \quad a_{31} = 3, \quad a_{32} = 4.$$

3.2 Transposição de Matrizes

A matrix (ou vetor) transposta de uma matrix $\mathbf{A}$, denotada por $\mathbf{A}'$, é obtida criando-se uma matrix cuja a $k^{\text{ésima}}$ linha é a $k^{\text{ésima}}$ coluna da matrix original.

Ou seja, se a matrix original é

$$\mathbf{A} = [a_{ij}]$$

então a sua transposta será

$$\mathbf{B} = \mathbf{A}' \Leftrightarrow b_{ij} = a_{ji} \quad \text{para todo } i \text{ e } j.$$

Se a matrix $\mathbf{A}$ tem ordem $n \times p$, a sua transposta $\mathbf{A}'$ terá ordem $p \times n$.

Exemplo numérico:

$$\mathbf{A} = \begin{bmatrix} 34 & 12 \\ 21 & 43 \\ 63 & -30 \end{bmatrix} \quad \Rightarrow \quad \mathbf{A}' = \begin{bmatrix} 34 & 21 & 63 \\ 12 & 43 & -30 \end{bmatrix}.$$

3.2.1 Implicações da Transposição

- Se a matrix $\mathbf{A}$ é *simétrica* então $\mathbf{A}' = \mathbf{A}$.
- Para qualquer matrix $\mathbf{A}$, temos que $(\mathbf{A}')' = \mathbf{A}$.
- A transposta de um *vetor coluna* é um *vetor linha*:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \Rightarrow \quad \mathbf{a}' = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

  Notação:

  - utilizaremos a notação de **letra minúscula em negrito** para vetor coluna: $\mathbf{a}$;
  - os vetores linhas serão **sempre** representados como vetores transpostos: $\mathbf{a}'$.

3.3 Adição e Subtração

Adicionar ou **subtrair** duas matrizes requer que ambas tenham a mesma ordem. A adição/subtração é realizando somando-se ou adicionando-se os *elementos correspondentes*:

$$\mathbf{A} = \begin{bmatrix} 34 & 12 \\ 21 & 43 \\ 63 & -30 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -40 & 3 \\ 0 & 10 \\ 2 & 50 \end{bmatrix}$$

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} 34+(-40) & 12+3 \\ 21+0 & 43+10 \\ 63+2 & -30+50 \end{bmatrix} = \begin{bmatrix} -6 & 15 \\ 21 & 53 \\ 65 & 20 \end{bmatrix}$$

$$\mathbf{A}-\mathbf{B} = \begin{bmatrix} 34-(-40) & 12-3 \\ 21-0 & 43-10 \\ 63-2 & -30-50 \end{bmatrix} = \begin{bmatrix} 74 & 9 \\ 21 & 33 \\ 61 & -80 \end{bmatrix}$$

Portanto, dado as matrizes:

$$\mathbf{A} = [a_{ij}] \quad \text{e} \quad \mathbf{B} = [b_{ij}] \quad i = 1, \ldots, n; \quad j = 1, \ldots, p,$$

a adição/subtração corresponde a seguinte operação

$$\mathbf{A}+\mathbf{B} = [a_{ij} + b_{ij}] \quad \text{e} \quad \mathbf{A}-\mathbf{B} = [a_{ij} - b_{ij}]$$

A **matrix nula** na operação de adição/subtração é a matrix de zeros:

$$\mathbf{0} = [x_{ij} = 0] \qquad i = 1, \ldots, n; \quad j = 1, \ldots, p,$$

$$\mathbf{A} + \mathbf{0} \;=\; [a_{ij} + 0] = [a_{ij}] = \mathbf{A}$$

3.3.1 Propriedades da Adição/Subtração

Comutativa: $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$.

Associativa: $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$.

Matrix Nula: $\mathbf{A}+(-\mathbf{A}) = [a_{ij} - a_{ij}] = [0] = \mathbf{0}$.

Transposição: $(\mathbf{A}+\mathbf{B})' = \mathbf{A}'+\mathbf{B}'$.

3.4 Multiplicação

3.4.1 Multiplicação de Matrix por Escalar

A multiplicação de uma matrix por um escalar é realizada multiplicando-se cada elemento da matrix pelo escalar:

$$\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 9 & 5 \end{bmatrix} \qquad 4\mathbf{A} = \begin{bmatrix} 4(3) & 4(-2) \\ 4(9) & 4(5) \end{bmatrix} = \begin{bmatrix} 12 & -8 \\ 36 & 20 \end{bmatrix}$$

Ou seja, se $\mathbf{A}$ é uma matrix e k um escalar, então:

$$k\mathbf{A} \;=\; \mathbf{A}k = [k \times a_{ij}]$$

Analogamente, a divisão de $\mathbf{A}$ por k deve ser interpretada como a multiplicação de $\mathbf{A}$ pelo inverso de k:

$$(1/k)\mathbf{A} \;=\; \mathbf{A}(1/k) = \left[\frac{a_{ij}}{k}\right]$$

Propriedades:

$$\begin{aligned}
(p+q)\mathbf{A} &= p\mathbf{A} + q\mathbf{A} \\
p(\mathbf{A}+\mathbf{B}) &= p\mathbf{A} + p\mathbf{B} \\
p(q\mathbf{A}) &= (pq)\mathbf{A} \\
(-1)\mathbf{A} &= -\mathbf{A} \\
k\mathbf{0} &= \mathbf{0} \\
0\mathbf{A} &= \mathbf{0}
\end{aligned}$$

3.4.2 Produto Interno

O **produto interno** é gerado pela multiplicação de um vetor linha e um vetor coluna (*nessa ordem*). Para que o produto exista ambos vetores devem ter o mesmo número de elementos e o resultado será um *escalar*:

$$\begin{bmatrix} 2 & 7 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 9 \end{bmatrix} = (2)(3) + (7)(4) + (-1)(9) = 25$$

Note que, a multiplicação dos vetores inicia-se com a multiplicação de cada elemento do vetor linha pelo *respectivo* elemento do vetor coluna, e temina com a soma dos produtos obtidos.

Seja $\mathbf{a}'$ um vetor $1 \times n$ e $\mathbf{b}$ um vetor $n \times 1$, então:

$$\mathbf{a}'\,\mathbf{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Propriedade: uma vez que $a_i b_i$ é igual a $b_i a_i$, temos que $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$.

3.4.3 Produto Externo

O **produto externo** é gerado pela multiplicação de um vetor coluna e um vetor linha (*nessa ordem*). Para que o produto exista ambos vetores devem ter o mesmo número de elementos e o resultado será uma *matrix quadrada*:

$$\begin{bmatrix} 3 \\ 4 \\ 9 \end{bmatrix} \begin{bmatrix} 2 & 7 & -1 \end{bmatrix} = \begin{bmatrix} 3(2) & 3(7) & 3(-1) \\ 4(2) & 4(7) & 4(-1) \\ 9(2) & 9(7) & 9(-1) \end{bmatrix} = \begin{bmatrix} 6 & 21 & -3 \\ 8 & 28 & -4 \\ 18 & 63 & -9 \end{bmatrix}$$

Note que, a multiplicação dos vetores é realizada multiplicando-se *cada linha* do vetor **coluna** por *cada coluna* do vetor **linha**. Seja $\mathbf{a}$ um vetor $m \times 1$ e $\mathbf{b}'$ um vetor $1 \times n$, então:

$$\mathbf{a}\,\mathbf{b}' = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & a_2 b_3 & \cdots & a_2 b_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & a_n b_3 & \cdots & a_n b_n \end{bmatrix}.$$

Note que a matrix resultante tem ordem $n \times n$.

Propriedade: uma vez que $a_i b_j$ é igual a $b_j a_i$, temos que $\mathbf{a}\mathbf{b}' = (\mathbf{b}\mathbf{a}')'$.

3.4.4 Multiplicação de Matrix por Vetor

Uma **matrix** pode ser multiplicada por um **vetor coluna** (*nessa ordem*) se o número de *elementos* no vetor for igual ao número de *colunas* da matrix.

A multiplicação é obtida realizando o **produto interno** de cada linha da matrix com o vetor, sendo que o resultado será um vetor com o mesmo número de linhas da matrix.

Por exemplo, tomemos uma matrix $\mathbf{A}$ e identifiquemos as suas linhas:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 5 \\ 2 & 3 & -4 \\ 6 & -3 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \mathbf{a}'_3 \end{bmatrix} \quad \text{onde} \quad \begin{matrix} \mathbf{a}'_1 = \begin{bmatrix} 4 & 1 & 5 \end{bmatrix} \\ \mathbf{a}'_2 = \begin{bmatrix} 2 & 3 & -4 \end{bmatrix} \\ \mathbf{a}'_3 = \begin{bmatrix} 6 & -3 & 0 \end{bmatrix} \end{matrix}$$

Agora tomemos um vetor:

$$\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}.$$

Calculemos o produto interno de cada linha de $\mathbf{A}$ por $\mathbf{b}$:

$$\begin{aligned}
\mathbf{a'}_1\mathbf{b} &= \begin{bmatrix} 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 4(1) + 1(3) + 5(2) = 17 \\
\mathbf{a'}_2\mathbf{b} &= \begin{bmatrix} 2 & 3 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 2(1) + 3(3) + (-4)(2) = 3 \\
\mathbf{a'}_3\mathbf{b} &= \begin{bmatrix} 6 & -3 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 6(1) + (-3)(3) + 0(2) = -3
\end{aligned}$$

Os escalares resultantes compõe o **vetor coluna** que é o resultado da multiplicação:

$$\mathbf{c} = \mathbf{A}\mathbf{b} \quad \Rightarrow \quad \mathbf{c} = \begin{bmatrix} 17 \\ 3 \\ -3 \end{bmatrix}.$$

Para que a multiplicação de matrix por vetor seja possível há a necessidade de compatibilidade de ordem entre a matrix e o vetor:

$$\underbrace{(n \times \overbrace{p) \times (p} \times 1)}_{(n \times 1)},$$

e o resultado será sempre um vetor com ordem $n \times 1$.

Utilizando uma notação algébrica mais detalhada, a multiplicação matrix-vetor pode ser apresentada na forma:

$$\mathbf{c} = \mathbf{A}\mathbf{b},$$

se temos

$$\mathbf{A} = [a_{ij}], \quad i = 1, \ldots, n;\ j = 1, \ldots, p.$$

e também

$$\mathbf{b} = [b_j], \quad j = 1, \ldots, p;$$

Então o produto é

$$\mathbf{c} = [c_i] \quad \text{onde} \quad c_i = \sum_{j=1}^{p} a_{ij} b_j \quad \text{e} \quad i = 1, 2, \ldots, n.$$

3.4.5 Combinação Linear

A multiplicação matrix-vetor pode ser vista como uma representação de um sistema de equações de primeiro grau (sistema de equações lineares) onde o resultado é uma **combinação linear** das colunas da matrix.

Consideremos novamente o produto matrix-vetor:

$$\mathbf{c} = \mathbf{A}\mathbf{b},$$

mas tomemos o vetor $\mathbf{b}$ como sendo composto de incógnitas

$$\begin{bmatrix} 7 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 5 \\ 2 & 3 & 4 \\ 6 & 3 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Essa multiplicação representa o sistema de equações

$$\begin{aligned} 7 &= 4b_1 + 1b_2 + 5b_3 \\ 3 &= 2b_1 + 3b_2 + 4b_3 \\ 2 &= 6b_1 + 3b_2 + 1b_3 \end{aligned}$$

o qual pode ser escrito na forma

$$\begin{bmatrix} 7 \\ 3 \\ 2 \end{bmatrix} = b_1 \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} + b_3 \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$$

mostrando que o resultado à esquerda do sinal de igualdade ($\mathbf{c}$) é uma **combinação linear** das colunas da matrix ($\mathbf{A}$), sendo os coeficientes os elementos do vetor ($\mathbf{b}$). O caso geral pode ser expresso como

$$\begin{aligned} \mathbf{c} &= \mathbf{A}\mathbf{b} \\ &= b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_p\mathbf{a}_p \end{aligned}$$

3.4.6 Multiplicação Matrix por Matrix

Considere uma matrix $\mathbf{A}$ com ordem $n \times p$ e uma matrix $\mathbf{B}$ com ordem $p \times k$.

A **pré**-multiplicação da matrix $\mathbf{B}$ pela matrix $\mathbf{A}$, resultando no produto $\mathbf{AB}$, é obtida com os seguintes passos:

1. Decompondo a matrix $\mathbf{A}$ em *linhas*:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_n \end{bmatrix}.$$

2. Decompondo a matrix $\mathbf{B}$ em *colunas*:

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{bmatrix}.$$

3. Realizando os produtos internos de cada linha de $\mathbf{A}$ por cada coluna de $\mathbf{B}$:

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a'}_1 \\ \mathbf{a'}_2 \\ \vdots \\ \mathbf{a'}_n \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{bmatrix} = \begin{bmatrix} \mathbf{a'}_1\mathbf{b}_1 & \mathbf{a'}_1\mathbf{b}_2 & \cdots & \mathbf{a'}_1\mathbf{b}_k \\ \mathbf{a'}_2\mathbf{b}_1 & \mathbf{a'}_2\mathbf{b}_2 & \cdots & \mathbf{a'}_2\mathbf{b}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a'}_n\mathbf{b}_1 & \mathbf{a'}_n\mathbf{b}_2 & \cdots & \mathbf{a'}_n\mathbf{b}_k \end{bmatrix}$$

Num exemplo numérico:

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} 34 & 12 \\ 21 & 43 \\ 63 & -30 \end{bmatrix} = \begin{bmatrix} \mathbf{a'}_1 \\ \mathbf{a'}_2 \\ \mathbf{a'}_3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 34 & 12 \end{bmatrix} \\ \begin{bmatrix} 21 & 43 \end{bmatrix} \\ \begin{bmatrix} 63 & -30 \end{bmatrix} \end{bmatrix} \\
\mathbf{B} &= \begin{bmatrix} 0 & 3 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 \\ 4 \end{bmatrix} & \begin{bmatrix} 3 \\ -2 \end{bmatrix} \end{bmatrix} \\
\mathbf{AB} &= \begin{bmatrix} \begin{bmatrix} 34 & 12 \end{bmatrix} \\ \begin{bmatrix} 21 & 43 \end{bmatrix} \\ \begin{bmatrix} 63 & -30 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 0 \\ 4 \end{bmatrix} & \begin{bmatrix} 3 \\ -2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 34(0) + 12(4) & 34(3) + 12(-2) \\ 21(0) + 43(4) & 21(3) + 43(-2) \\ 63(0) + (-30)(4) & 63(3) + (-30)(-2) \end{bmatrix} \\
&= \begin{bmatrix} 48 & 78 \\ 172 & -23 \\ -120 & 249 \end{bmatrix}
\end{aligned}$$

Note os seguintes aspectos importantes:

- Há necessidade de **compatibilidade** de ordem entre as matrizes para que a multiplicação seja possível:

$$\underbrace{(n \times \overbrace{p) \times (p}\times k)}_{(n \times k)} \implies \underbrace{(3 \times \overbrace{2) \times (2}\times 2)}_{(3 \times 2)}$$

- Nesse caso o produto $\mathbf{A}\,\mathbf{B}$ (**pré**-multiplicação de $\mathbf{B}$ por $\mathbf{A}$) é possível, mas o produto $\mathbf{B}\,\mathbf{A}$ (**pós**-multiplicação de $\mathbf{B}$ por $\mathbf{A}$) não existe, pois não há compatibilidade de ordem.

$$\underbrace{(2 \times \overbrace{2) \times (3}^{\neq} \times 2)}_{?}$$

Apresentando a multiplicação de matrix com uma notação algébrica detalhada temos: o produto

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

onde as matrizes são:

$$\begin{aligned}
\mathbf{A} &= [a_{il}] \quad i = 1, 2, \ldots, n; \quad l = 1, 2, \ldots, k; \\
\mathbf{B} &= [b_{lj}] \quad l = 1, 2, \ldots, k; \quad j = 1, 2, \ldots, p \\
\mathbf{C} &= [c_{ij}] \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, p;
\end{aligned}$$

o produto resulta na matrix $\mathbf{C}$ com elementos

$$c_{ij} = \sum_{l=1}^{k} a_{ik}\, b_{kj}.$$

3.4.7 Propriedades da Multiplicação de Matrizes

Algumas propriedades da multiplicação de escalares não são válidas no caso da multiplicação de matrizes:

Comutativa: não é válida pois geralmente

$$\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}$$

mesmo que haja compatibilidade de ordem para a pré e pós-multiplicação.

Lei do Cancelamento: não é válida pois

$$\mathbf{A}\mathbf{B} = \mathbf{0} \not\Rightarrow \mathbf{A} = \mathbf{0} \quad \text{ou} \quad \mathbf{B} = 0$$

e ainda temos que

$$\mathbf{A}\mathbf{B} = \mathbf{A}\mathbf{C} \not\Rightarrow \mathbf{B} = \mathbf{C}.$$

As *propriedade válidas* para multiplicação de matrizes são:

Lei Associativa:

$$(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C})$$

Lei Distribuitiva:

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

Transposta do Produto:

$$(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$$

Matrix Unitária na Multiplicação: seja $\mathbf{A}$ uma matrix quadrada e $\mathbf{I}$ uma matrix identidade, então

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A}$$

3.5 Representação Matricial do Somatório de Valores

- Matrizes também são convenientes para representar a somatório de valores:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n = \mathbf{i}'\mathbf{x}.$$

- No caso do vetor ser composto de uma mesma constante: $\mathbf{x} = a\mathbf{i}$, temos:

$$\sum_{i=1}^{n} x_i = \mathbf{i}'\mathbf{x} = \mathbf{i}'(a\mathbf{i}) = a(\mathbf{i}\mathbf{i}') = na.$$

- Se tivermos uma constante a multiplicando cada elemento x_i temos:

$$\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i = a\mathbf{i}'\mathbf{x}.$$

- No caso de calcularmos a média amostra de n observações temos $a = 1/n$, logo:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\mathbf{i}'\mathbf{x}.$$

- O somatório do quadrado das observações e o somatório do produto de observações pareadas também pode ser obtida com o produto interno:

$$\begin{aligned}
\sum_{i=1}^{n} x_i^2 = &\quad x_1(x_1) + x_2(x_2) + \cdots + x_n(x_n) &= \mathbf{x}'\mathbf{x}; \\
\sum_{i=1}^{n} x_i y_i = &\quad x_1(y_1) + x_2(y_2) + \cdots + x_n(y_n) &= \mathbf{x}'\mathbf{y}.
\end{aligned}$$

- As matrizes também podem ser utilizadas para representar os desvios das observações em relação à média amostral e a soma do quadrado desses desvios (*soma de quadrados*).

  Primeiro representamos um vetor com a média amostral:

$$\begin{bmatrix} \overline{x} \\ \overline{x} \\ \vdots \\ \overline{x} \end{bmatrix} = \mathbf{i}\overline{x} = \mathbf{i}(\frac{1}{n}\mathbf{i}'\mathbf{x}) = \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}$$

  A matrix $(1/n)\mathbf{i}\mathbf{i}'$ é uma matrix $n \times n$, onde todos os elementos são iguais a $1/n$.

  Agora podemos representar o vetor de desvios em relação à média amostral:

$$\begin{bmatrix} x_1 - \overline{x} \\ x_2 - \overline{x} \\ \vdots \\ x_n - \overline{x} \end{bmatrix} = [\mathbf{x} - \mathbf{i}\overline{x}] = \left[\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right]$$

  como temos $\mathbf{x} = \mathbf{I}\mathbf{x}$, então

$$\left[\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right] = \left[\mathbf{I}\mathbf{x} - \frac{1}{n}\mathbf{i}\mathbf{i}'\mathbf{x}\right] = \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]\mathbf{x} = \mathrm{M}^0\mathbf{x}$$

- A matrix M^0 tem os elementos da diagonal todos igual a $(1 - 1/n)$ e os demais elementos $1/n$ e será sempre utilizada para gerar os desvios em relação à média amostral. Ela tem propriedades que facilitam a computação das somas de quadrados e de produtos:

  Seu produto com vetor unitário é nula:

$$\mathrm{M}^0\mathbf{i} = \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]\mathbf{i} = \mathbf{i} - \frac{1}{n}\mathbf{i}(\mathbf{i}'\mathbf{i}) = \mathbf{0}.$$

  Isso também implica em $\mathbf{i}'\mathrm{M}^0 = \mathbf{0}$.

  Matrix Simétrica:

$$\mathrm{M}^{0\prime} = \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]' = \left[\mathbf{I}' - \frac{1}{n}(\mathbf{i}')'\mathbf{i}'\right] = \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] = \mathrm{M}^0$$

  Matrix Idempotente: é uma matrix que é igual ao seu quadrado:

$$\begin{aligned}(\mathrm{M}^0)^2 = \mathrm{M}^0\mathrm{M}^0 &= \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]\left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] \\ &= \mathbf{I}\left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] - \frac{1}{n}\mathbf{i}\mathbf{i}'\left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right]\end{aligned}$$

$$
\begin{aligned}
&= \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] - \left[\mathbf{I}(\frac{1}{n}\mathbf{i}\mathbf{i}') - (\frac{1}{n}\mathbf{i}\mathbf{i}')(\frac{1}{n}\mathbf{i}\mathbf{i}')\right] \\
&= \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] - \left[(\frac{1}{n}\mathbf{i}\mathbf{i}') - (\frac{1}{n^2}\mathbf{i}(\mathbf{i}'\mathbf{i})\mathbf{i}')\right] \\
&= \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] - \left[(\frac{1}{n}\mathbf{i}\mathbf{i}') - (\frac{1}{n}\mathbf{i}\mathbf{i}')\right] \\
&= \left[\mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}'\right] - \mathbf{0} \\
&= \mathbf{M}^0
\end{aligned}
$$

- A matrix $\mathbf{M}^0$ pode ser utilizada para obter os somatórios associados aos desvios em relação à média amostral:

  Soma dos desvios em relação à média:

  $$\sum_{i=1}^{n}(x_i - \overline{x}) = \mathbf{i}'[\mathbf{M}^0\mathbf{x}] = \mathbf{0}'\mathbf{x} = 0.$$

  Soma de quadrados dos desvios em relação à média:

  $$
  \begin{aligned}
  \sum_{i=1}^{n}(x_i - \overline{x})^2 &= (\mathbf{x} - \overline{x}\mathbf{i})'(\mathbf{x} - \overline{x}\mathbf{i}) \\
  &= (\mathbf{M}^0\mathbf{x})'(\mathbf{M}^0\mathbf{x}) \\
  &= \mathbf{x}'\mathbf{M}^{0\prime}\mathbf{M}^0\mathbf{x} \\
  &= \mathbf{x}'\mathbf{M}^0\mathbf{x}
  \end{aligned}
  $$

  Soma de produtos dos desvios em relação às médias:

  $$
  \begin{aligned}
  \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) &= (\mathbf{x} - \overline{x}\mathbf{i})'(\mathbf{y} - \overline{y}\mathbf{i}) \\
  &= (\mathbf{M}^0\mathbf{x})'(\mathbf{M}^0\mathbf{y}) \\
  &= \mathbf{x}'\mathbf{M}^{0\prime}\mathbf{M}^0\mathbf{y} \\
  &= \mathbf{x}'\mathbf{M}^0\mathbf{y}
  \end{aligned}
  $$

4 Geometria Matricial

Os vetores apresentados, sejam eles vetores colunas, sejam as colunas de uma matriz, podem ser intrepretados geometricamente como grandezas vetoriais que definem espaços geométricos.

Essa interpretação é útil pois permite a compreensão mais aprofundada da álgebra, por meio da geometria, mas também nos permitirá a interpretação geométrica do modelo linear múltiplo e de seu ajuste.

4.1 Espaços Vetoriais

- Os n elementos de um vetor coluna

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

podem ser vistos como definindo um ponto num *espaço com* p *dimensões*.

- A figura 1 ilustra os seguintes vetores num espaço *bi-dimensional*:

$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

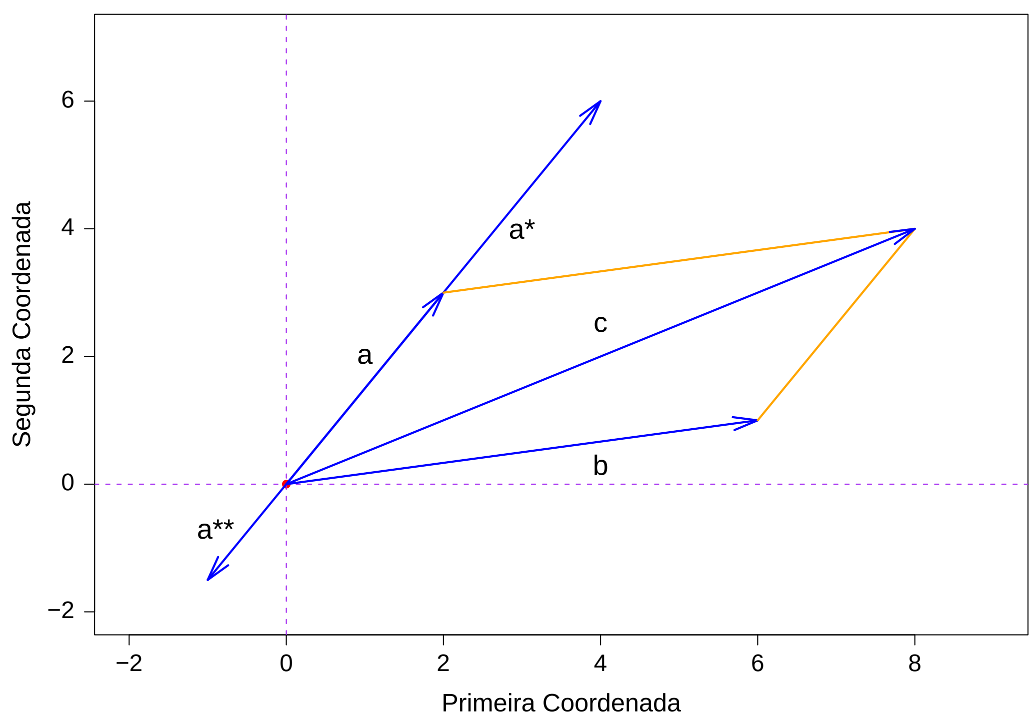

Figura 1: Figura ilustrativa do espaço vetorial *bi-dimensional* definido por *multiplicação por escalar* e *adição de vetores*.

- Duas operações básicas são definidas num espaço vetorial:

Multiplicação por escalar: que altera o *comprimento* ou o *sentido* do vetor, mas nunca altera a sua *direção*:

$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \mathbf{a}^* = 2\mathbf{a} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \quad \mathbf{a}^{**} = -\frac{1}{2}\mathbf{a} = \begin{bmatrix} -1 \\ -3/2 \end{bmatrix}$$

Adição de vetores: que pode gerar um terceiro vetor a partir da soma de dois vetores:

$$\mathbf{a} + \mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

DEFINIÇÃO: ESPAÇO VETORIAL

Espaço vetorial é o conjunto de todos os vetores que é fechado sob multiplicação por escalar e adição de vetores.

- Para vetores bi-dimensionais, o conjunto $\Re^2$ é um espaço vetorial.
- Para vetores p-dimensionais, o conjutno $\Re^p$ é um espaço vetorial.

4.2 Base de Espaços Vetoriais

- Considere as matrizes

$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

e as seguintes operações vetoriais:

$$\mathbf{a}^* = 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \qquad \mathbf{b}^* = 0\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1.5\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 1.5 \end{bmatrix}$$

$$\mathbf{c} = 1\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \qquad \mathbf{d} = 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 7 \end{bmatrix}$$

$$\mathbf{e} = 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1.5\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 7.5 \end{bmatrix} \qquad \mathbf{f} = -1\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

- Todas essa operações vetoriais geram **combinações lineares** dos vetores $\mathbf{a}$ e $\mathbf{b}$.
- A figura 2 ilustra essas combinações lineares e mostra que *qualquer vetor* em $\Re^2$ pode ser expresso como uma combinação linear dos vetores $\mathbf{a}$ e $\mathbf{b}$.

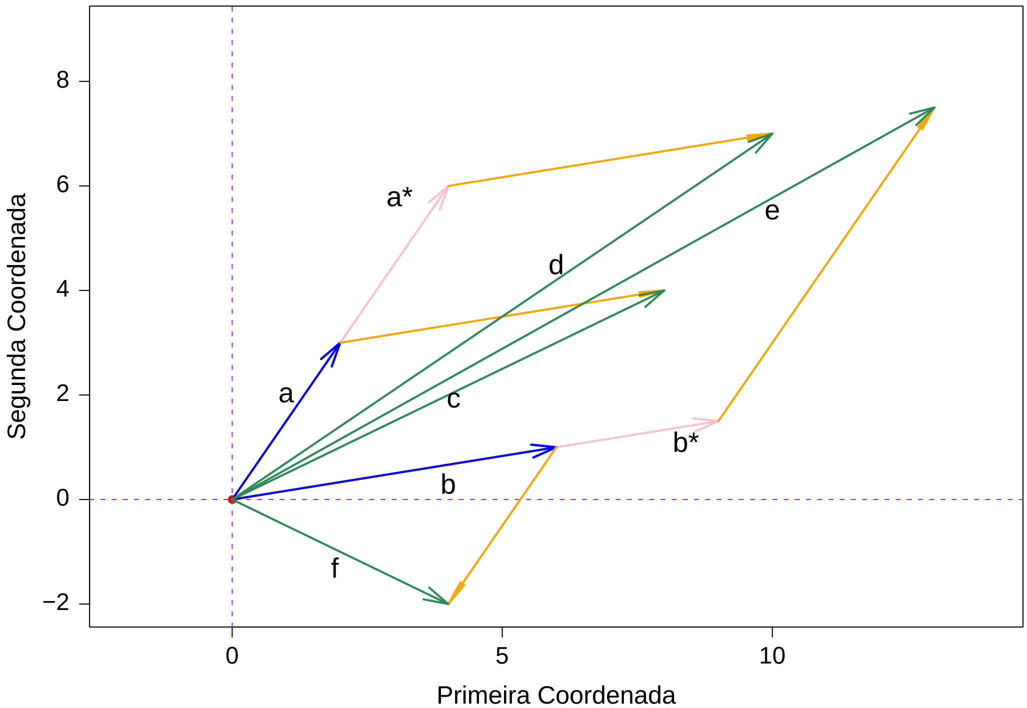

Figura 2: Figura ilustrativa das combinações lineraes de dois vetores num espaço vetorial bidimensional.

- Analisemos com mais detalhe as combinações lineares possíveis em $\Re^2$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

- Qualquer vetor $\mathbf{c}$ em $\Re^2$ pode ser expresso como combinação linear dos vetores $\mathbf{a}$ e $\mathbf{b}$

$$\mathbf{c} = \alpha_1 \mathbf{a} + \alpha_2 \mathbf{b}.$$

- A combinação linear implica num sistema de equações

$$\begin{aligned} c_1 &= \alpha_1 a_1 + \alpha_2 b_1 \\ c_2 &= \alpha_1 a_2 + \alpha_2 b_2. \end{aligned}$$

- A solução do sistema é dado por

$$\alpha_1 = \frac{b_2 c_1 - b_1 c_2}{a_1 b_2 - b_1 a_2} \qquad \alpha_2 = \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - b_1 a_2}$$

- Essa solução só será **única** se $(a_1 b_2 - b_1 a_2) \neq 0$.
- A situação de igualdade implica em

$$a_1 b_2 - b_1 a_2 = 0 \quad \Rightarrow \quad \frac{a_1}{a_2} = \frac{b_1}{b_2}$$

  ou seja, se o vetor $\mathbf{b}$ for múltiplo do vetor $\mathbf{a}$ o sistema **não terá** solução única.

- **Conclusão:** para que os vetores $\mathbf{a}$ e $\mathbf{b}$ sejam vetores-base para $\Re^2$ é necessário que eles tenham **direções diferentes** (apontem para direções diferentes).

  Nesse caso, o denominador acima não será nulo e o sistema de equações terá solução única, o que implica que *qualquer vetor* $\mathbf{c}$ em $\Re^2$ pode ser expresso como uma *combinação linear única* de $\mathbf{a}$ e $\mathbf{b}$.

- A base de um espaço vetorial não é única pois ela pode ser representada por qualquer conjunto de vetores que satisfaçam a definição.

  Entretanto, *para uma dada base*, apenas uma combinação linear dos vetores base produzirá um vetor particular no espaço vetorial.

4.3 Dependência Linear

- Os resultados da secção anterior *podem* dar a impressão de que qualquer conjunto de p vetores é uma base para $\Re^p$.
- Mas observando a figura 2 vemos que os vetores $\mathbf{a}$ e $\mathbf{b}$ são base para $\Re^2$, mas os vetores $\mathbf{a}$ e $\mathbf{a}^*$ não são.
- A diferença está no fato de que os vetores $\mathbf{a}$ e $\mathbf{b}$ são **linearmente independentes** (apontam para direções diferentes), enquanto que os vetores $\mathbf{a}$ e $\mathbf{a}^*$ são **linearmente dependentes** (apontam para a mesma direção).

DEFINIÇÃO: DEPENDÊNCIA LINEAR

Um conjunto de vetores é dito *linearmente dependente* se qualquer vetor no conjunto é uma combinação linear dos demais vetores do conjunto.

Exemplo de conjunto de vetores linearmente dependentes: considere os vetores

$$\mathbf{a} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Eles formam um conjunto linearmente dependente pois

$$-2\,\mathbf{a} + 7\,\mathbf{b} - \mathbf{c} = \mathbf{0}.$$

DEFINIÇÃO: INDEPENDÊNCIA LINEAR

Um conjunto de vetores é *linearmente independente* se, e somente se, a solução para qualquer combinação linear deles:

$$\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \cdots + \alpha_p \mathbf{a}_p$$

são escalares nulos:

$$\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0.$$

Essa definição nos permite re-definir a *base* de um espaço vetorial:

DEFINIÇÃO: BASE DE UM ESPAÇO VETORIAL

A *base* de um espaço vetorial p-dimensional é qualquer conjunto de p vetores linearmente independentes naquele espaço vetorial.

4.4 Subespaços Vetoriais

DEFINIÇÃO: SUBESPAÇO VETORIAL

O conjunto de todas as combinações lineares de um conjunto de vetores é o *subespaço gerado* por esse conjunto de vetores.

- Pela definição acima, o *subespaço* gerado por uma base de $\Re^p$ é o próprio espaço vetorial $\Re^p$.
- Nos exemplos acima os vetores $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ formam uma base que gera o espaço vetorial $\Re^2$.

  Logicamente que $\mathbf{c}$ é desnecessário, pois os vetores $\mathbf{a}$ e $\mathbf{b}$ já são uma base para $\Re^2$.

- A conseqüência é que um dado vetor $\mathbf{d}$ poderá ser expresso por *infinitas* combinações lineares de $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$, mas por apenas *uma única* combinação linear de $\mathbf{a}$ e $\mathbf{b}$.

  A título de exemplo, considere os vetores

$$\mathbf{a} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \qquad \mathbf{d} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}.$$

  O vetor $\mathbf{d}$ é expresso por uma única combinação linear dos vetores $\mathbf{a}$ e $\mathbf{b}$

$$\mathbf{d} = -4\,\mathbf{a} + 14\,\mathbf{b}.$$

  Entretanto, como temos $\mathbf{d} = -2\,\mathbf{c}$, as combinações lineares de $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ para expressar $\mathbf{d}$ são

$$\mathbf{d} = k\,(-4\,\mathbf{a} + 14\,\mathbf{b}) + (k-1)\,(-2\,\mathbf{c})\,, \qquad \text{para} \quad k = 1, 2, 3, \ldots$$

- Consideremos agora os seguintes vetores tridimensionais

$$\mathbf{a}' = \begin{bmatrix} a_1 & a_2 & 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{b}' = \begin{bmatrix} b_1 & b_2 & 0 \end{bmatrix}.$$

  Esses vetores não geram o espaço tridimensional de $\Re^3$, pois, por exemplo, o vetor $\mathbf{c}' = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ não pode ser expresso como uma combinação linear deles.

  Entretanto, esses vetores podem representar qualquer vetor tridimensional *cujo terceiro elemento seja zero*, por isso, eles geram um **subespaço** de $\Re^3$.

- O subespaço gerado pelos vetores tridimensionais $\mathbf{a}$ e $\mathbf{b}$ acima **não é** o espaço $\Re^2$, mas apenas **um plano** no espaço $\Re^3$: o plano que contem todos os vetores em $\Re^3$ cujo terceiro elemento é nulo.

- Qualquer plano em $\Re^3$, independentemente de sua orientação, forma um subespaço bidimensional em $\Re^3$ que pode ser gerado por qualquer par de vetores independentes que estejam nesse subespaço.

- O espaço gerado por um conjunto de vetores em $\Re^p$ tem **no máximo** p dimensões.

  Se o espaço gerado tem **menos** que p dimensões ele é um *subespaço*, designado de *hiperplano*.

- A idéia central é que *qualquer conjunto de vetores* gera *algum espaço vetorial*, o qual pode ser o espaço completo onde os vetores residem ou um subespaço dele.

4.5 Posto de uma Matrix

- Uma matrix deve ser vista como um conjunto de vetores coluna.
  - O número de colunas na matrix é igual o número de vetores.
  - O número de linhas na matrix é igual o número de coordenadas em cada vetor.

DEFINIÇÃO: ESPAÇO COLUNA

O *espaço coluna* de uma matrix é o espaço vetorial gerado por seus vetores coluna.

- Considere a matrix

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

  ela possui três vetores de $\Re^3$, mas o terceiro vetor é a soma dos dois primeiros, por isso o espaço coluna dessa matrix não tem três dimensões, mas apenas duas, pois o seus os seus vetores colunas geram um subespaço bidimensional em $\Re^3$.

- Assim a matrix $\mathbf{A}$ tem *posto coluna* igual a 2.

DEFINIÇÃO: POSTO COLUNA

O *posto coluna* de uma matrix é a dimensão do espaço vetorial gerado por seus vetores coluna.

- Considere a matrix

$$\mathbf{B} = \begin{bmatrix} 7 & 2 & 5 \\ 2 & 3 & 9 \\ 8 & 5 & 2 \\ 1 & 4 & 6 \end{bmatrix}.$$

  Essa matrix tem posto coluna igual a 3 e, portanto, seu espaço coluna é um subespaço em $\Re^4$.

- Considere agora a matrix transposta de $\mathbf{B}$

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

A matrix $\mathbf{C}$ é composta de quatro vetores coluna em $\Re^3$, conseqüentemente, seu espaço coluna é no máximo $\Re^3$ e seus vetores coluna devem ser linearmente dependentes.

O espaço coluna de $\mathbf{C}$ é $\Re^3$ e embora ele não seja o mesmo espaço coluna de $\mathbf{B}$ eles tem a mesma dimensão (3).

- O posto coluna de $\mathbf{C}$ é igual ao posto coluna de $\mathbf{B}$, mas as colunas de $\mathbf{C}$ são as *linhas* de $\mathbf{B}$, logo o posto coluna de $\mathbf{C}$ é igual o **posto linha** de $\mathbf{B}$.

IGUALDADE DE POSTO COLUNA E POSTO LINHA

O *posto coluna* e o *posto linha* de uma matrix são iguais, i.e., o *espaço linha* e o *espaço coluna* de uma matrix têm a mesma dimensão.

- A igualdade de posto coluna e posto linha nos permite abandonar a distinção e trabalhar sempre com **posto** de uma matrix:

$$\text{posto}\,(\mathbf{A}) = \text{posto}\,(\mathbf{A}') \leq \min(\text{número de linhas, número de colunas})$$

DEFINIÇÃO: POSTO COMPLETO

Se o posto de uma matrix é igual ao número de colunas dela, a matrix tem *posto completo*.

4.5.1 Propriedades do Posto de uma Matrix Importantes para Regressão

- Considere o produto matricial $\mathbf{C} = \mathbf{AB}$.
  - As colunas da matrix $\mathbf{C}$ são combinações lineares das colunas da matrix $\mathbf{A}$, portanto, cada coluna de $\mathbf{C}$ está no espaço coluna da matrix $\mathbf{A}$, portanto
  $$\text{posto}\,(\mathbf{C}) \leq \text{posto}\,(\mathbf{A})\,.$$
  - As linhas da matrix $\mathbf{C}$ são combinações lineares das linhas da matrix $\mathbf{B}$ e, consequentemente,
  $$\text{posto}\,(\mathbf{C}) \leq \text{posto}\,(\mathbf{B})\,.$$
  - Uma vez que o posto coluna e posto linha de uma matrix é sempre igual, a conclusão é
  $$\text{posto}\,(\mathbf{AB}) \leq \min\left(\text{posto}\,(\mathbf{A})\,,\text{posto}\,(\mathbf{B})\right).$$

- Uma propriedade que tem um papel central na aplicação da algebra de matrizes na regressão linear é:

$$\text{posto}\,(\mathbf{A}) = \text{posto}\,(\mathbf{A}'\mathbf{A}) = \text{posto}\,(\mathbf{A}\mathbf{A}')\,.$$

4.6 Determinante de uma Matrix

- Considere a matrix

$$\begin{aligned} \mathbf{A} &= [\mathbf{ab}] \\ &= \begin{bmatrix} 2 & 6 \\ 3 & 1 \end{bmatrix} \end{aligned}$$

que é mostrada na figura 3.

- A área do paralelogramo formada pelas colunas de $\mathbf{A}$ é o valor absoluto (sem sinal) do **determinante** de $\mathbf{A}$. Nesse exemplo, o valor absoluto do determinante é obtido por

$$|a_1b_2 - b_1a_2| = |2(1) - 6(3)| = |-16| = 16.$$

- O determinante só é definido para *matrizes quadradas*.

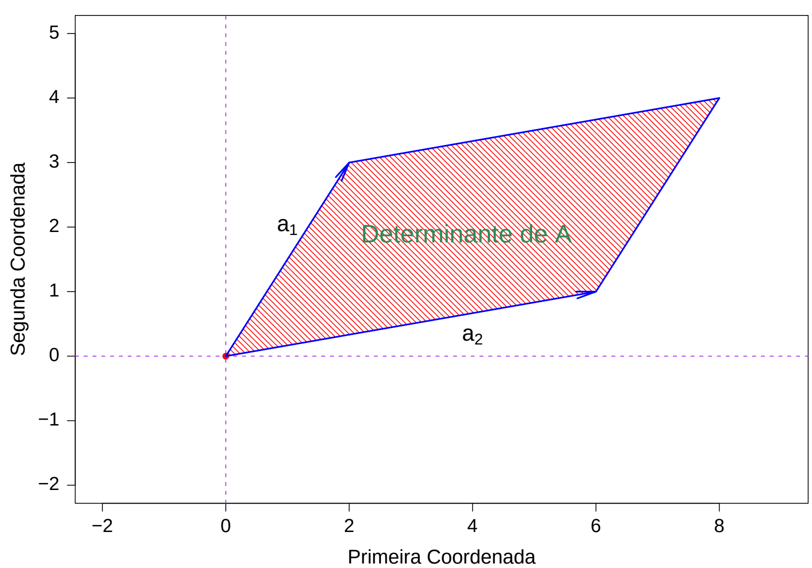

Figura 3: Figura ilustrativa do *determinante* de uma matrix a partir do *paralelogramo* formado pelo seus vetores coluna.

- Se as colunas de $\mathbf{A}$ fossem linearmente dependentes, i.e., uma coluna fosse um múltiplo escalar da outra, então ambos vetores estariam sobre a mesma linha e o paralelogramo colapsaria para uma linha e teria área nula.

- Esse conceito implica que *se as colunas de uma matrix são linearmente dependentes, então seu determinante é zero*.

- Já foi apresentado que para um par de vetores num espaço bidimensional, a dependência linear entre os vetores implica em:

$$|a_1b_2 - b_1a_2| = 0$$

(o determinante de $\mathbf{A}$ é $a_1b_2 - b_1a_2$).

- O mesmo conceito pode ser expandido para espaços de maiores dimensão. Considerando $\Re^3$, os vetores de uma matrix 3×3 definirão um sólido e o valor absoluto do determinante será o *volume* desse sólido.

- Se as colunas dessa matrix forem linearmente dependentes, então o sólido colapsará para uma figura plana cujo volume é nulo, e o determinante será zero.

- Generalizando o conceito para p dimensões temos a seguinte proposição:

PROPOSIÇÃO

O determinante de uma matrix é **não-nulo** se, e somente se, a matrix tiver **posto completo**.

4.6.1 Algums Aspectos Úteis do Determinante de uma Matrix

- Utiliziramentos a notação $|\mathbf{X}|$ para designar o determinante de uma matrix $\mathbf{X}$ (e não seu valor absoluto).

- O determinante de uma matrix 2×2 pode ser facilmente encontrado por:

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$$

- No caso de uma matrix $n \times n$, o determinante é encontrado usando a **expansão por cofatores**:

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{A}_{ij}|, \qquad i = 1, 2, \ldots, n$$

onde $\mathbf{A}_{ij}$ é a matrix obtida a partir de $\mathbf{A}$ elminando-se a linha i e a coluna j.

Exemplo com uma matrix 3×3:

$$\begin{aligned}
\mathbf{B} &= \begin{bmatrix} 2 & 0 & 1 \\ 3 & 5 & 2 \\ 1 & 4 & 6 \end{bmatrix} \\
|\mathbf{B}| &= 2(-1)^{1+1} \begin{vmatrix} 5 & 2 \\ 4 & 6 \end{vmatrix} + 0(-1)^{1+2} \begin{vmatrix} 3 & 2 \\ 1 & 6 \end{vmatrix} + 1(-1)^{1+3} \begin{vmatrix} 3 & 5 \\ 1 & 4 \end{vmatrix} \\
&= 2(22) + 0 + 1(7) \\
|\mathbf{B}| &= 51
\end{aligned}$$

- No caso de matrix transposta, temos: $|\mathbf{A}| = |\mathbf{A}'|$.
- No caso de matrix diagonal, temos

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \qquad \Rightarrow \qquad |\mathbf{D}| = d_1\, d_2\, \cdots\, d_n = \Pi_{i=1}^{n} d_i.$$

- Sendo a matrix identidade uma matrix diagonal especial temos que $|\mathbf{I}_n| = 1$.
- Determinante de matrix diagonal multiplicada por escalar: $|c\mathbf{D}| = c^n |\mathbf{D}|$.
- Determinante do produto de duas matrizes diagonais: $|\mathbf{DC}| = |\mathbf{D}| \cdot |\mathbf{C}|$.

4.7 Norma e Ortogonalidade de Vetores

DEFINIÇÃO: COMPRIMENTO DE VETOR

O comprimento, ou **norma**, de um vetor é a raiz quadrada do produto interno do vetor por ele mesmo:

$$||\mathbf{x}|| = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

DEFINIÇÃO: VETORES ORTOGONAIS

Dois vetores $\mathbf{a}$ e $\mathbf{b}$ são **ortogonais**, denotando por $\mathbf{a} \perp \mathbf{b}$, se, e somente se, o produto interno deles for nulo:

$$\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} = \mathbf{0}.$$

5 Solução de Sistemas Lineares

- A expressão matricial:

$$\mathbf{Ax} = \mathbf{b}$$

representa um **Sistema de Equações Lineares**, pois o seu desdobramento gera

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + & \cdots & + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + & \cdots & + a_{2p}x_p = b_2 \\ & \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + & \cdots & + a_{np}x_p = b_n \end{array}$$

- Para aplicação em regressão linear, consideraremos apenas sistemas *quadrados*, i.e., aqueles em que a matrix $\mathbf{A}$ é $p \times p$.

5.1 Sistemas de Equações Lineares

Há dois tipos de sistemas de equações lineares:

Sistemas Homogêneos: são aqueles que têm a forma

$$\mathbf{Ax} = \mathbf{0}.$$

Por definição, uma solução não nula para o sistema ($\mathbf{x} \neq \mathbf{0}$) existe se, e somente se, as colunas de $\mathbf{A}$ forem *linearmente dependentes*.

Isso implica que o determinante de $\mathbf{A}$ é nulo: $|\mathbf{A}| = 0$.

Sistemas Não-Homogêneos: são aqueles que têm a forma

$$\mathbf{Ax} = \mathbf{b}$$

onde $\mathbf{b} \neq \mathbf{0}$, i.e., $\mathbf{b}$ é num vetor não-nulo.

O vetor $\mathbf{b}$ é escolhido arbitrariamente e expressa as combinações lineares das colunas da matrix $\mathbf{A}$.

Se $\mathbf{b}$ tem p elementos, isso só será possível se as colunas de $\mathbf{A}$ gerarem todo o espaço p-dimensional $\Re^p$.

Isso é equivalente a dizer que as colunas de $\mathbf{A}$ são *linearmente independentes* e que o seu determinante não é nulo: $|\mathbf{A}| \neq 0$.

5.2 Matrix Inversa

- Para encontramos a solução do sistema (quadrado) $\mathbf{Ax} = \mathbf{b}$, necessitamos de uma operação matricial análoga à divisão entre escalares.
- Suponha que pudéssemos encontrar uma matrix $\mathbf{B}$ tal que $\mathbf{BA} = \mathbf{I}$.

  Nesse caso, a solução do sistema seria pré-multiplicando o sistema por ela:

$$\begin{aligned} \mathbf{BAx} &= \mathbf{Bb} \\ \mathbf{Ix} &= \mathbf{Bb} \\ \mathbf{x} &= \mathbf{Bb}. \end{aligned}$$

- Se a matrix $\mathbf{B}$ *existe*, então ela é a **matrix inversa** de $\mathbf{A}$:

$$\mathbf{B} = \mathbf{A}^{-1}.$$

- Pela definição temos $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.
- Mas a relação também é válida para pós-multiplicação: $\mathbf{AA}^{-1} = \mathbf{I}$.
- Se a matrix inversa existe, então ela é única:

  Suponha que existam duas matrizes tal que $\mathbf{B} = \mathbf{A}^{-1}$ e $\mathbf{C} = \mathbf{A}^{-1}$, então temos que

$$\begin{aligned} \mathbf{CAB} &= \mathbf{CAB} \\ (\mathbf{CA})\mathbf{B} &= \mathbf{C}(\mathbf{AB}) \\ \mathbf{IB} &= \mathbf{CI} \\ \mathbf{B} &= \mathbf{C} \end{aligned}$$

DEFINIÇÃO: MATRIX NÃO-SINGULAR

Se a inversa de uma matrix existe, essa matrix é dita **não-singular**.

5.2.1 Algumas Propriedades de Matrizes Inversas

- Matriz inversa de uma matrix 2×2:

$$\begin{aligned}
\mathbf{AB} &= I \\
&\Rightarrow \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&\Rightarrow \quad \begin{cases} a_{11}b_{11} + a_{12}b_{21} = 1 \\ a_{11}b_{12} + a_{12}b_{12} = 0 \\ a_{21}b_{11} + a_{22}b_{21} = 0 \\ a_{21}b_{12} + a_{22}b_{12} = 1 \end{cases} \\
&\Rightarrow \quad \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \\
&\Rightarrow \quad \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}
\end{aligned}$$

- Matriz inversa de uma matrix $n \times n$: utilizando a notação a^{ij} para indicar o $(ij)^{\text{ésimo}}$ elemento de $\mathbf{A}^{-1}$, então

$$a^{ij} = \frac{|\mathbf{C}_{ij}|}{|\mathbf{A}|}$$

onde $|\mathbf{C}_{ij}|$ é o *(j i)* **ésimo** **co-fator** de $\mathbf{A}$:

$$|\mathbf{C}_{ij}| = (-1)^{i+j} |\mathbf{A}_{ij}|$$

Exemplo numérico:

$$\mathbf{B} = \begin{bmatrix} 2 & 0 & 1 \\ 3 & 5 & 2 \\ 1 & 4 & 6 \end{bmatrix} \qquad |\mathbf{B}| = 51$$

$$a^{11} = \tfrac{(-1)^{1+1}}{51} \begin{vmatrix} 5 & 2 \\ 4 & 6 \end{vmatrix} = \tfrac{22}{51} \quad a^{21} = \tfrac{(-1)^{1+2}}{51} \begin{vmatrix} 3 & 2 \\ 1 & 6 \end{vmatrix} = \tfrac{-16}{51} \quad a^{31} = \tfrac{(-1)^{1+3}}{51} \begin{vmatrix} 3 & 5 \\ 1 & 4 \end{vmatrix} = \tfrac{7}{51}$$

$$a^{12} = \tfrac{(-1)^{2+1}}{51} \begin{vmatrix} 0 & 1 \\ 4 & 6 \end{vmatrix} = \tfrac{4}{51} \quad a^{22} = \tfrac{(-1)^{2+2}}{51} \begin{vmatrix} 2 & 1 \\ 1 & 6 \end{vmatrix} = \tfrac{11}{51} \quad a^{32} = \tfrac{(-1)^{2+3}}{51} \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} = \tfrac{-8}{51}$$

$$a^{13} = \tfrac{(-1)^{3+1}}{51} \begin{vmatrix} 0 & 1 \\ 5 & 2 \end{vmatrix} = \tfrac{-5}{51} \quad a^{23} = \tfrac{(-1)^{3+2}}{51} \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = \tfrac{-1}{51} \quad a^{33} = \tfrac{(-1)^{3+3}}{51} \begin{vmatrix} 2 & 0 \\ 3 & 5 \end{vmatrix} = \tfrac{10}{51}$$

$$\mathbf{B}^{-1} = \frac{1}{51}\begin{bmatrix} 22 & 4 & -5 \\ -16 & 11 & -1 \\ 7 & -8 & 10 \end{bmatrix}$$

- Matrix inversa de matrix diagonal:

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \qquad \Rightarrow \qquad \mathbf{D}^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}$$

- Sendo a matrix identidade uma matrix diagonal especial, temos: $\mathbf{I}^{-1} = \mathbf{I}$.
- Se a matrix $\mathbf{A}$ é simétrica, então $\mathbf{A}^{-1}$ também é simétrica.
- Alguns resultados computacionais:

$$\begin{aligned} |\mathbf{A}^{-1}| &= \frac{1}{|\mathbf{A}|} \\ (\mathbf{A}^{-1})^{-1} &= \mathbf{A} \\ (\mathbf{A}^{-1})' &= (\mathbf{A}')^{-1} \end{aligned}$$

- No caso de matrizes $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$, todas elas quadradas e não-singulares:

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{C}^{-1}$$

- Note que é possível que exista a matrix $(\mathbf{AB})^{-1}$, mesmo no caso em que as matrizes $\mathbf{A}$ e $\mathbf{B}$ não sejam nem mesmo quadradas.

  Por exemplo, se $\mathbf{A}$ matrix de ordem $(n \times p)$, é possível que exista a matrix $(\mathbf{A}'\mathbf{A})^{-1}$ que é simétrica $(p \times p)$.

5.3 Solução de Sistemas Lineares Não-Homogêneos

Dado um sistema linear não-homogêneo

$$\mathbf{Ax} = \mathbf{b},$$

se a matrix $\mathbf{A}$ é não-singular, a **solução única** para o sistema é

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

6 Auto-Valores e Auto-Vetores

- Para se analisar uma matrix quadrada $\mathbf{A}$ é útil utilizarmos a solução para o seguinte sistema de equações:
$$\mathbf{A}\mathbf{c} = \lambda\mathbf{c},$$
onde

  $\mathbf{c}$ é chamado de **auto-vetor** e

  λ é o **auto-valor** associado a ele.

- Entretanto, há uma indeterminação nesse sistema, pois se $\mathbf{c}$ é uma solução para o sistema, então o vetor $k\mathbf{c}$ também será será uma solução.

- Consideraremos então os **vetores normalizados**, i.e., aqueles que possuem norma unitária
$$\mathbf{c}'\mathbf{c} = 1,$$
como solução do sistema.

- Assim a solução, consiste em λ e os $n-1$ elementos de $\mathbf{c}$.

6.1 Solução do Sistema

- Apresentaremos aqui apenas a idéia da solução do sistema.

- Transformando o sistema original para um sistema homogêneo temos
$$\mathbf{A}\mathbf{c} = \lambda\mathbf{I}\mathbf{c} \quad \Rightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{c} = \mathbf{0}.$$

- Esse sistema só tem solução se a matrix $(\mathbf{A} - \lambda\mathbf{I})$ for singular, i.e., tiver determinante nulo.

- Logo, para encontrarmos os auto-valores devemos solucionar
$$|\mathbf{A} - \lambda\mathbf{I})| = 0,$$
que gera uma **equação polinômial em** λ chamada de equação característica.

- Por exemplo, considere a matrix (2×2):
$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 8 & 7 \end{bmatrix} \quad \Rightarrow \quad \begin{aligned} |\mathbf{A} - \lambda\mathbf{I}| &= \begin{vmatrix} 3-\lambda & 4 \\ 8 & 7-\lambda \end{vmatrix} \\ &= (3-\lambda)(7-\lambda) - 4(8) \\ &= \lambda^2 - 10\lambda - 11 \end{aligned}$$
As raízes dessa equação característica são $\lambda = 11$ e $\lambda = -1$.

- Aplicando-se os auto-valores ao sistema homogêneo obtem-se os auto-vetores.
- No exemplo acima teremos

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{I})\mathbf{c} &= 0\\ \begin{bmatrix} 3-\lambda & 4 \\ 8 & 7-\lambda \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.\end{aligned}$$

Para $\lambda=-1$ temos:

$$4c_1+4c_2=0 \quad \text{e} \quad 8c_1+8c_2=0 \quad \Rightarrow \quad c_1=c_2.$$

Para $\lambda=11$ temos:

$$-8c_1+4c_2=0 \quad \text{e} \quad 8c_1-4c_2=0 \quad \Rightarrow \quad c_1=\frac{1}{2}c_2.$$

Para obtermos a solução devemos considerar que $\mathbf{c}'\mathbf{c}=1$, assim se obtem:

$$\text{Para } \lambda=11 \quad \Rightarrow \quad \mathbf{c}=\begin{bmatrix} -1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

$$\text{Para } \lambda=-1 \quad \Rightarrow \quad \mathbf{c}=\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

6.2 Resultados Úteis para Matrizes Simétricas

- Se $\mathbf{A}$ é uma matrix $p\times p$ simétrica, então
  1. $\mathbf{A}$ tem p auto-vetores *distintos*

$$\mathbf{c}_1,\ \mathbf{c}_2,\ \mathbf{c}_3,\ \ldots,\ \mathbf{c}_p.$$

  2. $\mathbf{A}$ tem p auto-valores *reais* não necessariamente distintos

$$\lambda_1,\ \lambda_2,\ \lambda_3,\ \ldots,\ \lambda_p.$$

- Os auto-vetores de $\mathbf{A}$ são **orto-normais**:

$$\begin{aligned}\mathbf{c}_i{}'\mathbf{c}_j &= 1 \qquad \text{se } i=j;\\ \mathbf{c}_i{}'\mathbf{c}_j &= 0 \qquad \text{se } i\neq j.\end{aligned}$$

- Os auto-vetores de $\mathbf{A}$ podem ser organizados numa matrix orto-normal

$$\mathbf{C} = [\mathbf{c}_1 \, \mathbf{c}_2 \, \mathbf{c}_3 \, \ldots \, \mathbf{c}_p]$$

enquanto que os auto-valores podem ser organizados numa matrix diagonal

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p \end{bmatrix}$$

- Utilizando essas matrizes, o todas as equações do sistema homogêneo

$$\mathbf{A}\mathbf{c}_i = \lambda_i \mathbf{c}_i$$

pode ser representado numa única expressão

$$\mathbf{AC} = \mathbf{C\Lambda}$$

- Como a matrix de auto-vetores é orto-normal, ela tem as seguintes propriedades:

$$\begin{aligned} \mathbf{C}'\mathbf{C} &= \mathbf{C}\mathbf{C}' = \mathbf{I} \\ \mathbf{C}' &= \mathbf{C}^{-1} \end{aligned}$$

DEFINIÇÃO: DIAGONALIZAÇÃO DE UMA MATRIX SIMÉTRICA

A *diagonalização* de uma matrix simétrica $\mathbf{A}$ é

$$\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{C}'\mathbf{C}\mathbf{\Lambda} = \mathbf{I}\mathbf{\Lambda} = \mathbf{\Lambda}.$$

DEFINIÇÃO: DECOMPOSIÇÃO ESPECTRAL DE UMA MATRIX SIMÉTRICA

A *decomposição espectral* de uma matrix simétrica $\mathbf{A}$ é

$$\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}' = \sum_{i=1}^{p} \lambda_i \mathbf{c}_i \mathbf{c}_i{}'.$$

Uma vez que o produto externo $\mathbf{c}_i \mathbf{c}_i{}'$ gera uma matrix $(p \times p)$, na decomposição espectral, a matrix $\mathbf{A}$ é expressa como o somatório de p matrizes de posto 1.

6.3 Posto de Matrizes

- Para qualquer matrix $\mathbf{A}$ e duas matrizes *não-singulares* $\mathbf{B}$ e $\mathbf{C}$, o posto do produto $\mathbf{BAC}$ é igual ao posto de $\mathbf{A}$.

$$\begin{aligned}
\text{posto}\,(\mathbf{BCA}) &= \text{posto}\,((\mathbf{BA})\mathbf{C}) = \min(\text{posto}\,(\mathbf{BA})\,, \text{posto}\,(\mathbf{C})) \\
&= \text{posto}\,(\mathbf{BA}) \qquad [\text{pois } \mathbf{C} \text{ tem posto completo}] \\
\\
\text{posto}\,(\mathbf{BA}) &= \text{posto}\,(\mathbf{A}'\mathbf{B}') = \min(\text{posto}\,(\mathbf{A}')\,, \text{posto}\,(\mathbf{B}')) \\
&= \text{posto}\,(\mathbf{A}') \qquad [\text{pois } \mathbf{B}' \text{ tem posto completo}] \\
\\
\text{posto}\,(\mathbf{A}') &= \text{posto}\,(\mathbf{A}) \\
\\
\text{posto}\,(\mathbf{BCA}) &= \text{posto}\,(\mathbf{A})
\end{aligned}$$

- No caso de uma matrix *simétrica* $\mathbf{A}$, usando a decomposição espectral temos

$$\text{posto}\,(\mathbf{A}) = \text{posto}\,(\mathbf{C}\boldsymbol{\Lambda}\mathbf{C}') = \text{posto}\,(\boldsymbol{\Lambda})$$

  O posto de $\boldsymbol{\Lambda}$ é simplismente o número de elementos não nulos na sua diagonal principal, pois $\boldsymbol{\Lambda}$ é uma matrix diagonal.

TEOREMA: POSTO DE MATRIX SIMÉTRICA

O posto de uma matrix simétrica é o número de auto-valores não nulos que ela contem.

- Esse teorema pode ser generalizado se lembrarmos que toda matrix pode gerar uma matrix simétrica se for pré-multiplicada ou pela sua transposta e que

$$\text{posto}\,(\mathbf{A}) = \text{posto}\,(\mathbf{A}'\mathbf{A}) = \text{posto}\,(\mathbf{A}\mathbf{A}')\,.$$

TEOREMA: POSTO DE UMA MATRIX

O posto de *qualquer* matrix $\mathbf{A}$ é igual ao número de auto-valores não nulos presentes na matrix $\mathbf{A}'\mathbf{A}$.

TEOREMA: AUTO-VALORES DO PRODUTO EXTERNO DE UMA MATRIX

Os auto-valores não-nulos da matrix $\mathbf{AA}'$ são os mesmo da matrix $\mathbf{A}'\mathbf{A}$.

6.4 Número Condicional de uma Matriz

- Como foi visto, o posto de uma matrix é um número discreto e, portanto, o fato de uma matrix ter ou não ter posto completo também é.
- Entretanto, freqüentemente na regressão linear lidados com matrizes que têm posto completo, pois todos os seus auto-valores são não-nulos, mas elas estão muito próximo de serem singulares, pois algumas colunas estão muito próximas de serem uma combinação linear das demais.
- O **número condicional** de uma matrix é uma forma de detectarmos esse problema:

$$\gamma = \left[\frac{\max(\text{auto-valores})}{\min(\text{auto-valores})}\right]^{1/2}$$

- Para uma matrix *não-quadrada* $\mathbf{X}$, utilizaremos a matrix transformada $\mathbf{X}'\mathbf{X}$ e como os auto-valores dependem da escala das observações, devemos *re-escalonar* a matrix $\mathbf{X}$ para que suas colunas tenham comprimento 1.

  Isso é obtido, dividindo os elementos da coluna por sua norma.

6.5 Determinante por Decomposição

- Para uma matrix simétrica $\mathbf{A}$, a diagonalização dela implica em

$$\mathbf{C}'\mathbf{AC} = \mathbf{\Lambda} \quad \Rightarrow \quad |\mathbf{C}'\mathbf{AC}| = |\mathbf{\Lambda}|$$

- Desenvolvendo o determinante da diagonalização temos:

$$\begin{aligned} |\mathbf{C}'\mathbf{AC}| &= |\mathbf{C}'| \cdot |\mathbf{A}| \cdot |\mathbf{C}| \\ &= |\mathbf{C}'| \cdot |\mathbf{C}| \cdot |\mathbf{A}| \\ &= |\mathbf{C}'\mathbf{C}| \cdot |\mathbf{A}| \\ &= |\mathbf{I}| \cdot |\mathbf{A}| \\ &= 1 \cdot |\mathbf{A}| \\ &= |\mathbf{A}| \\ &= |\mathbf{\Lambda}| \end{aligned}$$

- Como $\mathbf{\Lambda}$ é uma matrix diagonal, o seu determinante é igual ao produto dos elementos da diagonal principal.

TEOREMA: DETERMINANTE DE MATRIX SIMÉTRICA

O determinante de uma matrix simétrica é igual ao produto de seus auto-valores.

6.6 Potências de uma Matrix

TEOREMA: POTÊNCIA DE MATRIX SIMÉTRICA

Para toda matrix simétrica *não-singular* temos

$$\mathbf{A}^K = \mathbf{C}'\mathbf{\Lambda}^K\mathbf{C},$$

para $K = \ldots, -2, -1, 0, 1, 2, \ldots$.

Alguns exemplos:

- Para $K = 2$:

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{C\Lambda C'}^2 = (\mathbf{C'\Lambda C})(\mathbf{C\Lambda C'}) \\ &= \mathbf{C\Lambda C'C\Lambda C'} \\ &= \mathbf{C\Lambda I\Lambda C'} \\ &= \mathbf{C\Lambda}^2\mathbf{C'} \end{aligned}$$

- Para $K = -1$:

$$\begin{aligned} \mathbf{A}^{-1} &= \mathbf{C\Lambda C'}^{-1} \\ &= (\mathbf{C}')^{-1}(\mathbf{\Lambda})^{-1}(\mathbf{C})^{-1} \\ &= \mathbf{C}'\mathbf{\Lambda}^{-1}\mathbf{C} \end{aligned}$$

DEFINIÇÃO: MATRIX DEFINIDA POSITIVA

Uma matrix é dita **definida positiva** quando todos os seus auto-valores são positivos.

TEOREMA: POTÊNCIA DE MATRIX DEFINIDA POSITIVA

Para toda matrix definida positiva temos

$$\mathbf{A}^r = \mathbf{C}'\mathbf{\Lambda}^r\mathbf{C},$$

para qualquer número real r.

- Por exemplo para $r = 1/2$ temos:

$$\begin{aligned}\mathbf{A}^{1/2} &= \mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{C}' \\ &= \mathbf{C} \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{\lambda_n} \end{bmatrix} \mathbf{C}'\end{aligned}$$

- Essa equação satisfaz a exigência de raiz quadrada, pois

$$\begin{aligned}\mathbf{A}^{1/2}\mathbf{A}^{1/2} &= (\mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{C}')(\mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{C}') \\ &= \mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{C}'\mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{C}' \\ &= \mathbf{C}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{C}' \\ &= \mathbf{C}\mathbf{\Lambda}\mathbf{C}' \\ &= \mathbf{A}\end{aligned}$$

7 Vetores e Matrizes Aleatórios

A única diferença de vetor ou matrix aleatórios das matrizes comuns é que os seus elementos são *variáveis aleatórias*. Por exemplo, o vetor

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

é um vetor aleatório se os elementos $Y_1, Y_2, \ldots, Y_n$ forem todas variáveis aleatórias.

7.1 Valor Esperado

O valor esperado de um vetor aleatório e o vetor dos valores esperados de cada elemento:

$$\boldsymbol{E}\{\mathbf{Y}\} = \begin{bmatrix} \boldsymbol{E}\{Y_1\} \\ \boldsymbol{E}\{Y_2\} \\ \vdots \\ \boldsymbol{E}\{Y_n\} \end{bmatrix}$$

Na regressão linear, geralmente temos a pressuposição para o termo dos os erros que $\varepsilon_i \sim N(0, \sigma^2)$ Se o termo do erro for representado como um vetor aleatório temos:

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \Rightarrow \boldsymbol{E}\{\varepsilon\} = \begin{bmatrix} \boldsymbol{E}\{\varepsilon_1\} \\ \boldsymbol{E}\{\varepsilon_2\} \\ \vdots \\ \boldsymbol{E}\{\varepsilon_n\} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}$$

7.2 Matrix de Variância-Covariância de um Vetor Aleatório

Se $\mathbf{Y}$ é um vetor aleatório $(n \times 1)$, a sua matrix de variância-covariância é definida como:

$$\boldsymbol{Var}\{\mathbf{Y}\}_{(n\times n)} = \begin{bmatrix} \boldsymbol{Var}\{Y_1\} & \boldsymbol{Cov}\{Y_1, Y_2\} & \dots & \boldsymbol{Cov}\{Y_1, Y_n\} \\ \boldsymbol{Cov}\{Y_2, Y_1\} & \boldsymbol{Var}\{Y_2\} & \dots & \boldsymbol{Cov}\{Y_2, Y_n\} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{Cov}\{Y_n, Y_1\} & \boldsymbol{Cov}\{Y_n, Y_2\} & \dots & \boldsymbol{Var}\{Y_n\} \end{bmatrix}$$

Algumas característica desta matrix:

- Os elementos da diagonal principal ($i = j$) são a variância de cada variável aleatória: $\boldsymbol{Var}\{Y_1\}$, $\boldsymbol{Var}\{Y_2\}, \dots,$ $\boldsymbol{Var}\{Y_n\}$.
- Os elementos fora da diagonal principal ($i \neq j$) são as covariância entre as variáveis aleatórias: $\boldsymbol{Cov}\{Y_1, Y_2\}, \boldsymbol{Cov}\{Y_1, Y_n\}, \dots \boldsymbol{Cov}\{Y_i, Y_j\} \dots$.
- Como $\boldsymbol{Cov}\{Y_i, Y_j\} = \boldsymbol{Cov}\{Y_j, Y_i\}$ a matrix é *simétrica*.
- A definição tradicional de variância-covariância pode ser generalizada para um vetor de variáveis aleatórias, portanto:

$$\boldsymbol{Var}\{\mathbf{Y}\} = \boldsymbol{E}\{[\mathbf{Y} - \boldsymbol{E}\{\mathbf{Y}\}]\,[\mathbf{Y} - \boldsymbol{E}\{\mathbf{Y}\}]'\}$$

Uma das pressuposições da regressão linear é que os erros se comportam como $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. Neste caso, a matrix de variância-covariância dos erros é:

$$\boldsymbol{Var}\{\varepsilon\}_{(n\times n)} = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \sigma^2 \mathbf{I}_n$$

Portanto, em notação matricial, esta pressuposião fica:

$$\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n).$$

7.3 Algumas Propriedades envolvendo Vetores Aleatório

Já foi visto que uma variável aleatória W pode ser definida como uma combinação linear de uma série de variáveis aleatórias Y_i ($i = 1, \dots, n$), na forma:

$$W = a_1 Y_1 + a_2 Y_2 + \dots + a_n Y_n = \sum_{i=1}^{n} a_i Y_i,$$

onde a_i ($i = 1, \dots, n$) são constantes.

Generalizando a idéia de *função linear de variáveis aleatórias*, podemos dizer que um vetor aleatório $\mathbf{W}$ é uma combinação linear do vetor aleatório $\mathbf{Y}$ na forma

$$\mathbf{W} = \mathbf{A}\mathbf{Y}$$

onde $\mathbf{A}$ é uma matrix de constantes (não aleatória).

Neste caso, são válidas as seguintes propriedades:

$$\begin{aligned} \boldsymbol{E}\{\mathbf{A}\} &= \mathbf{A} \\ \boldsymbol{E}\{\mathbf{W}\} &= \boldsymbol{E}\{\mathbf{AY}\} = \mathbf{A}\boldsymbol{E}\{\mathbf{Y}\} \\ \boldsymbol{Var}\{\mathbf{W}\} &= \boldsymbol{Var}\{\mathbf{AY}\} = \mathbf{A}\,\boldsymbol{Var}\{\mathbf{Y}\}\mathbf{A}' \end{aligned}$$

Nota: em algebra de matrizes o equivalente a $(ab)^2 = a^2 b^2$ é a forma quadrática: $(\mathbf{AB})^2 = \mathbf{ABB}'\mathbf{A}'$.

8 Exercícios

1. Para as matrizes

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 3 \\ -4 & 8 & 2 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 7 & 2 \\ 4 & -1 \\ 2 & -3 \end{bmatrix}$$

calcule: $\mathbf{AB}$, $\mathbf{A}'\mathbf{B}'$ e $\mathbf{BA}$.

2. Expanda o produto matricial

$$\mathbf{X} = \{[\mathbf{AB} + (\mathbf{CD})'] \, [(\mathbf{EF})^{-1} + \mathbf{GH}]\}'$$

tomando todas as matrizes como quadradas e $\mathbf{E}$ e $\mathbf{F}$ como não-singulares.

3. Seja $\mathbf{A}$ uma matrix quadrada cujas colunas são $[\mathbf{a}_1 \, \mathbf{a}_2 \, \ldots \, \mathbf{a}_m]$ e seja $\mathbf{B}$ qualquer *re-arranjamento das colunas* da matrix identidade $m \times m$.

   Qual operação é realizada ao se calcular o produto matricial $\mathbf{AB}$? E o produto matricial $\mathbf{BA}$?

4. Considere a matrix
$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 7 & -5 & 1 \end{bmatrix}.$$
   (a) Compute $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.
   (b) Compute $\mathbf{H} = \mathbf{I} - \mathbf{P}$.
   (c) Verifique que $\mathbf{PH} = \mathbf{0}$.
   (d) Demonstre que $\mathbf{P}$ e $\mathbf{H}$ são idempotentes.

5. Uma matrix $\mathbf{A}$ é chamada de *nilpotente* se

$$\lim_{K \to \infty} \mathbf{A}^K = \mathbf{0}.$$

   Demonstre que a condição necessária e suficiente para uma matrix simétrica ser nilpotente é que o valor absoluto de todos seus auto-valores é menor que um.