ÁLGEBRA MATRICIAL

Frederico Machado Almeida
frederico.almeida@unb.br

Departamento de Estatística
Instituto de Exatas
Universidade de Brasília

Algébra Matricial

Antes de inciar com o conteúdo da disciplina de `GLM`, vamos rever ou ver algumas propriedades matriciais importantes para o curso.

- **Espaço Vetorial:** O espaço vetorial é um conjunto

$$\mathcal{V}_n = \{v_i = (v_{i1}, \cdots, v_{in}, v_{ij} \in \mathcal{R}, \quad i = 1, \cdots, I)\},$$

  no qual é fechado sobre a adição, multiplicação por escalares, e contém $\mathbf{0}$.

Definição 1 (Matriz): é uma forma retangular de números. Ou seja, uma matriz $A_{m\times n}$ com m-linhas e n-colunas, sendo n e m inteiros positivos é dada por:

$$A_{m\times n} = \left(a_{ij}\right), \text{ com } i = 1, \cdots, m \text{ e } j = 1, \cdots, n.$$

- Existem diferentes formas de definir uma matriz. Como por exemplo, $A_{m\times n} = [a_{ij}]$ ou $A_{m\times n} = \{a_{ij}\}$.
- Cada linha ou coluna de A é chamado de vetor. Assim, $\mathbf{a}_{.j}$ e $\mathbf{a}_{i.}$ são vetores linha e coluna, respectivamente.

Algébra Matricial

Leis da Álgebra (Regra de Adição):

Considere A, B e C matrizes $m \times n$ e a, b, c escalares. Então, valem as seguintes propriedades:

1. $(A + B) + C = A + (B + C)$
2. $A + B = B + A$
3. $A + (-A) = \mathbf{0}$
4. $A + \mathbf{0} = \mathbf{0} + A$
5. $c\,(A + B) = cA + cB$
6. $(a + b)\,C = aC + bC$
7. $abC = a\,(bC) = b\,(aC)$
8. $\mathbf{0}A = \mathbf{0}$
9. $\mathbf{I}_m A = A$.

Leis da Álgebra (Regra da Multiplicação):

Seja A uma matriz $m \times n$, e B e C matrizes com a mesma dimensão. Considere igualmente a um escalar. Então, valem as seguintes propriedades:

1. $(AB)\,C = A\,(BC)$
2. $A\,(B + C) = AB + AC$
3. $(A + B)\,C = AC + BC$
4. $a\,(BC) = (aB)\,C$
5. $\mathbf{I}_m A = A\mathbf{I}_n = A$
6. $\mathbf{0}_m A = A\mathbf{0}_n = \mathbf{0}$.

Obs.1: no caso do produto matricial, é possível que $AB \neq BA$ em geral, em alguns casos podemos ter AB definido e BA não.

Definição 2 (`Matriz quadrada`): Uma matriz $A_{m\times n}$ é dita ser quadrada se ela tiver o mesmo número de linhas e colunas. Isto é, se $m = n$, consequentemente, $A_{n\times n} = (a_{ij})$, com $i, j = 1, 2, \cdots, n$.

Definição 3 (`Matriz diagonal`): Uma matriz quadrada $n \times n$ é dita ser diagonal, quando $a_{ij} = 0$ para $\forall_{i\neq j}$. Isto é,

$$A_{n\times n} = \mathrm{diag}\{a_{11}, \cdots, a_{nn}\}.$$

Algébra Matricial

Definição 4 (`Soma de matrizes`): Duas matrizes A e B, podem ser somadas elemento a elemento, se elas tiverem a mesma dimensão, digamos $m \times n$. Assim, se $A_{m\times n} = (a_{ij})$ e $B_{m\times n} = (b_{ij})$ segue que,

$$A + B = (a_{ij} + b_{ij}), \text{ com } i = 1, \cdots, m \text{ e } j = 1, \cdots, n.$$

Definição 5 (`Produto de matrizes`): Para multiplicar a matriz A pela matriz B e formar o produto AB, é necessário que a dimensão da coluna de A seja igual ao número de linhas de B. Assim se A é uma matriz $m \times n$, então, B deverá ser uma matriz com n linhas. Logo,

$$AB = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Definição 6 (Transposição de matrizes): Seja $A = (a_{ij})$ uma matriz $m \times n$. A transposta de A denotada por $A^\top$ (também chamada de A-prime) é uma matriz $n \times m$, obtida pela troca de linhas e colunas de A. De tal forma que $A^\top_{n\times m} = (a_{ji})$.

Algumas propriedades: Assuma que B e C sejam matrizes cujas dimensões permitem a validade das seguintes propriedades.

1. $\left(A^\top\right)^\top = A$
2. $(aA)^\top = aA^\top$
3. $(A+B)^\top = A^\top + B^\top$
4. $(aA+bB)^\top = aA^\top + bB^\top$
5. $(AB)^\top = B^\top A^\top$
6. $A^\top = B^\top$ se e somente se $A = B$

Algébra Matricial

7. $\mathbf{X}^\top\mathbf{X} = \sum_{i=1}^{n} X_i^2$, sendo $\mathbf{X}$ um vetor $n \times 1$.

Obs.2: Se $\mathbf{X}$ for uma matriz $n \times p$ então, $\left(\mathbf{X}^\top\mathbf{X}\right)_{p\times p}$ será sempre definida e simétrica. O resultado segue imediatamente das `propriedades 1 e 5`.

Definição 7 (`Matriz simétrica`): uma matriz quadrada A é dita ser simétrica se, e somente se $A^\top = A$.

Algébra Matricial

Definição 8 (`Matriz de 1's`): Seguindo a definição de vetores, denote por $\mathbf{1}_m = (1, \cdots, 1)^\top$. Então, segue que

$$\mathbf{J}_{m\times n} = \mathbf{1}_m \times \mathbf{1}_n^\top = (a_{ij}), \text{ com } a_{ij} = 1 \text{ para todo } i \text{ e } j.$$

Para todo $i = 1, \cdots, m$ e $j = 1, \cdots, n$ então:

1. $\mathbf{J}_{m\times n}\mathbf{1}_n = n\mathbf{1}_m$
2. $\mathbf{J}_{n\times m}\mathbf{1}_m = m\mathbf{1}_n$
3. $\mathbf{1}_m^\top \mathbf{J}_{m\times n} = m\mathbf{1}_n$
4. $\mathbf{J}_{m\times n}\mathbf{J}_{n\times m} = n\mathbf{J}_{m\times m}$.

Vetores linearmente dependentes e independentes

Definição 9 (`Vetores dependentes:`) Seja $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ um conjunto de vetores colunas n-dimensionais pertencentes a $\mathbf{X}$. Esses m vetores são ditos serem `linearmente dependentes` **se e somente se** existir um vetor $C = (c_1, \cdots, c_m)^\top \in \mathcal{R}^m$, `com pelo menos um diferente de zero`, tal que $\mathbf{X}C = \sum_{i=1}^{m} c_i \mathbf{x}_i = 0$.

- A afirmação de que $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ é linearmente dependente, significa que, pelo menos um vetor pode ser escrito como uma combinação linear dos outros.
- Por outro lado, se $c_1 = 0 \cdots, c_m = 0$ então, $\sum_{i=1}^{m} c_i \mathbf{x}_i = 0$ para quaisquer coleção de $\mathbf{x}_i$'s então, $\{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$ são ditos serem **linearmente independentes**.

Definição 10 (Posto de uma matriz): Seja A uma matriz $n \times m$. O posto de A (ou rank de A), sinalizado por $r(A)$, é o número máximo de linhas ou colunas linearmente independentes de A.

(i) Se A é uma matriz $n \times m$, e $r(A) = m$, então, A tem posto de coluna completo.

(ii) $r(A)$ é um inteiro positivo, excepto o fato de que $r(A) = 0$ se e somente se A for uma matriz nula.

(iii) $r(A_{n\times m}) \leq n$, e $\leq m$. Ou seja, o posto de uma matriz A, deve ser no máximo, igual ao número de linhas ou de colunas.

(iv) Se A é uma matriz quadrada $n \times n$, então, $r(A) \leq n$.

Obs.3: Métodos para calcular o posto de uma matriz estão disponíveis na literatura.

Algébra Matricial

Propriedades de rank:

1. $A_{m\times n}$ tem $r(A) = r$ se a maior sub-matriz não singular de A possui tamanho r.
2. Para $A_{m\times n}$, $r(A) \leq \min(m, n)$
3. $r(A + B) \leq r(A) + r(B)$
4. $r(AB) \leq \min\{r(A), r(B)\}$
5. Para matrizes `não singulares` A, B e uma matriz qualquer C, temos que,

$$r(C) = r(AC) = r(CB) = r(ACB).$$

6. $r(A) = r(A^\top) = r(A^\top A) = r(AA^\top)$
7. $r(A, \mathbf{b}) \geq r(A)$, isso é, adicionar coluna em A nunca reduz o seu rank

Definição 11 (`Traço da matriz`): O traço de uma matriz $A_{n\times n} = (a_{ij})$, com $i, j \in \{1, \cdots, \}$ (`notação:` $tr(A)$) é dada pela soma dos elementos da diagonal. Isto é,

$$tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + \cdots + a_{nn}$$

Obs.4: Quando A não é uma a matriz quadrada, o $tr(A)$ está bem definido, i.e, não existe.

Algébra Matricial

Propriedades do traço:

1. $tr(I_n) = n$
2. $tr(A) = tr(A^\top)$
3. $tr(A+B) = tr(B) + tr(A)$
4. $tr(aA) = \mathtt{a}tr(A)$, para $\forall a \in \mathcal{R}$
5. $tr(AB) = tr(BA)$
6. $tr(ABC) = tr(CAB) = tr(BCA)$
7. $tr(aA + bB) = \mathtt{a}tr(A) + \mathtt{b}tr(B)$
8. $tr(A^\top A) = tr(AA^\top) = \sum_{i=1}^{\ell}\sum_{j=1}^{c} a_{ij}^2 = \sum_{i,j=1}^{n} a_{ij}^2.$

Algébra Matricial

Definição 12 (Determinante de matriz): O determinante de uma matriz quadrada $A_{n\times n}$ é um escalar dado porque,

$$|A| = \sum_{i=1}^{n} a_{ij}\,(-1)^{i+j}\,|M_{ij}|,\ \text{para qualquer } i \text{ fixado},$$

onde M_{ij} é a matriz obtida removendo-se a i-ésima linha e j-ésima coluna de A.

Propriedades do determinante:

1. $|A| = |A^\top|$
2. $|cA| = c^n|A|$
3. $|AB| = |A||B|$
4. Se A é diagonal ou triangular inferior (ou superior), então, $|A| = \prod_{i=1}^{n} a_{ii}$
5. Se duas linhas ou colunas de A são iguais, então $|A| = 0$

Algébra Matricial

Propriedades do determinante (continuação):

- Se A tem linhas ou colunas iguais a 0, então $|A| = 0$.
- Se as linhas ou colunas de A são linearmente dependentes, então $|A| = 0$
- Seja B uma matriz obtida multiplicando uma linha ou coluna de A por um escalar c, então $|B| = c|A|$.
- Se B é uma matriz obtida trocando de posição uma coluna ou linha de A, etão $|B| = -|A|$
- Seja $A_{m\times n}$ e $B_{n\times m}$, então $|\mathbf{I}_m + AB| = |\mathbf{I}_n + BA|$.

Definição 13 (Inversa de uma matriz): Seja $A = (a_{ij})$, uma matriz quadrada e não singular. Isto é, $|A| \neq 0$. A inversa de A, denotada por A^{-1}, é uma matriz quadrada que satisfaz a condição $AA^{-1} = \mathbf{I}_n$.

Exemplo 5: Assuma que A é uma matriz 2×2 dada por:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ então, segue que } A^{-1} = \frac{1}{|A|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

onde, $|A| = a_{11}a_{22} - a_{21}a_{12}$. Assim, segue que,

Quando a dimensão de A é maior ou igual a 3 então,

$$[A_{n \times n} | \mathbf{I}_n] \sim [\mathbf{I}_n | A_{n \times n}^{-1}].$$

Algébra Matricial

Propriedades do inverso:

1. A^{-1} é único
2. $\left(A^{-1}\right)^{-1} = A$
3. $(AB)^{-1} = B^{-1}A^{-1}$
4. $(A^{\top})^{-1} = (A^{-1})^{\top}$
5. $|A^{-1}| = |A|^{-1}$
6. $(aA)^{-1} = a^{-1}A^{-1} = \left(\frac{1}{a}\right)A^{-1}$
7. $r(A^{-1}) = n$, se A for uma matriz $n \times n$
8. $(A_1A_2A_3\cdots A_p)^{-1} = A_p^{-1}A_{p-1}^{-1}\cdots A_2^{-1}A_1^{-1}$
9. Se $|A| \neq 0$, então $A^{\top}$ e A^{-1} são não-singulares.

Definição 14 (`Matriz positiva definida e semi-definida`): (i) Uma matriz simétrica A é `positiva definida` (`p.d`) se $\mathbf{x}_i^\top A\mathbf{x}_i > 0$, para todo vetor $\mathbf{x}_i$ de dimensão $n \times 1$, exceto $\mathbf{x}_i = \mathbf{0}$.

(ii) Uma matriz simétrica A é dita ser positiva `semi-definida` (`p.s.d`) se para todo para todo vetor $\mathbf{x}_i$ obtivermos, $\mathbf{x}_i^\top A\mathbf{x}_i \geq 0$.

`Exemplo 3:` Considere a matriz A de tamanho 3×3.

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Mostre que, para qualquer vetor $\mathbf{x} = (x_1, x_2, x_3)^\top$ não nulo, A será sempre positiva definida.

Exemplo 4: Considere a matriz A de tamanho 3×3.

$$A = \begin{bmatrix} 37 & -2 & -24 \\ -2 & 13 & -3 \\ -24 & -3 & 17 \end{bmatrix}.$$

1. Mostre que, para qualquer vetor $\mathbf{x} = (x_1, x_2, x_3)^\top$ não nulo, A será sempre positiva definida.
2. Apresente algum vetor $\mathbf{x}$ de reais, tal que $\mathbf{x}_i^\top A \mathbf{x}_i = 0$.

Algébra Matricial

Algumas propriedades:

1. Se A é $\texttt{p.d}$, então, A^{-1} existe e, também é $\texttt{p.d}$
2. Se $\mathbf{X}$ é uma matriz $n \times p$, então, $\mathbf{X}^\top\mathbf{X}$ e $\mathbf{X}\mathbf{X}^\top$ são matrizes $\texttt{p.s.d}$ e,
3. Se $\mathbf{X}$ é uma matriz $n \times p$ e $r(X) = p$ então, $\mathbf{X}^\top\mathbf{X}$ é $\texttt{p.d}$ e portanto, não singular.

Definição 15 (`Matriz idempotente`): Seja A uma matriz simétrica, com dimensão $n \times n$. Assim, A é dita ser `idempotente` se e somente se $A^2 = A$ (Em latim `idempotente significa: idem=mesma, potente=potência, i.e., mesma potência`).

`Exemplo 6:` Considere a matriz $A_{3\times 3}$, tal que,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Mostre que A é uma matriz idempotente.

Algébra Matricial

Propriedades de uma matriz idempotente: Defina A como uma matriz idempotente (A é $n \times n$). Então,

1. $r(A) = tr(A)$
2. $A_{n\times n}$ é positiva semi-definida, i.e, $a_{ii} \geq 0$, para $\forall i \in \{1, \cdots, n\}$
3. No decorrer do curso, vamos construir matrizes idempotentes muito usadas na estatística. Assim, se $\mathbf{X}$ denota uma matriz $n \times p$, com $r(\mathbf{X}) = p$. Defina $H = \mathbf{X}\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top$ e $M = \mathbf{I}_n - \mathbf{X}\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top$. É possível mostrar que H e M são matrizes simétricas e idempotentes (`verifique!`).
4. Posteriormente será mostrado que, $r(H) = p$ e $r(M) = n - p$.
5. Se $A^2 = A \Longrightarrow A^2 - A = A(A - \mathbf{I}_n)$. Porém, apesar de A ser idempotente, $(A - \mathbf{I}_n)$ não é idempotente (`verifique!`). No entanto, $(\mathbf{I}_n - A)$ é idempotente.
6. $\mathbf{I}_n - \bar{\mathbf{J}}_n = \mathbf{I}_n - \frac{1}{n}\mathbf{J}_n$ é idempotente.

Algébra Matricial

- Espaço das colunas de uma matriz: Seja $A_{m\times n}$ uma matriz na qual as colunas m-dimensional são $\mathbf{a}_1, \cdots, \mathbf{a}_n$. O espaço vetorial gerado pelas n colunas de A é chamado de espaço das colunas de A (notação $\mathcal{C}(A)$).

- A dimensão do espaço das colunas de A, isto é, $\mathtt{dim}(\mathcal{C}(A))$ é o número de colunas linearmente independentes de A.

- Espaço nulo de uma matriz: O espaço nulo (notação: $\mathcal{N}(A)$), de uma matriz $A_{m\times n}$ consiste de todos os vetores n-dimensionais $\mathbf{x}$ tal que, $A\mathbf{x} = \mathbf{0}$. Ou seja,

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathcal{R}^n, \text{ tal que, } A\mathbf{x} = \mathbf{0}\}.$$

Algébra Matricial

Propriedades do espaço das colunas:

1. $\texttt{dim}[\mathcal{C}(A)] = n - \texttt{dim}[\mathcal{N}(A)]$
2. $\mathcal{C}(A) = \{\mathcal{C}(A)\}^\top$
3. $\mathcal{C}(A^\top A) = \mathcal{C}(A^\top)$
4. Para qualquer A e B, $\mathcal{C}(AB) \subseteq \mathcal{C}(A)$.