UMA ABORDAGEM GEOMÉTRICA DA TEORIA DE INVERSAS GENERALIZADAS

PAULO HENRIQUE SALES GUIMARÃES

2010

PAULO HENRIQUE SALES GUIMARÃES

UMA ABORDAGEM GEOMÉTRICA DA TEORIA DE INVERSAS GENERALIZADAS

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof.Dr. Lucas Monteiro Chaves

LAVRAS
MINAS GERAIS - BRASIL
2010

Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da Biblioteca Central da UFLA

Guimarães, Paulo Henrique Sales.
Uma abordagem geométrica da Teoria de Inversas Generalizadas / Paulo Henrique Sales Guimarães. – Lavras: UFLA, 2010.
67 p.: il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2010.
Orientador: Lucas Monteiro Chaves.
Bibliografia.

1. Álgebra linear. 2. Matriz particionada. 3. Modelos lineares. I. Universidade Federal de Lavras. II. Título.

CDD – 512.5

PAULO HENRIQUE SALES GUIMARÃES

UMA ABORDAGEM GEOMÉTRICA DA TEORIA DE INVERSAS GENERALIZADAS

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

APROVADA em 25 de fevereiro de 2010

Profa. Dra. Anliy Natsuyo Nashimoto Sargeant	UFOP
Prof. Dr. Augusto Ramalho de Morais	UFLA
Dr. Devanil Jaques de Souza	

Prof. Dr. Lucas Monteiro Chaves
UFLA
(Orientador)

LAVRAS
MINAS GERAIS – BRASIL

Ao mestre dos mestres, Jesus Cristo,
Aos meus amados pais Paulo e Niêza,
Aos meus queridos irmãos Renato e Lívia
Aos meus grandes amigos
Aos meus dedicados mestres,
Dedico.

AGRADECIMENTOS

A Deus, que se faz presente em minha vida, guiando meus passos e concedendo-me graças a cada dia.

À Universidade Federal de Lavras (UFLA), em especial ao Departamento de Ciências Exatas (DEX), pela realização deste trabalho.

À CNPQ, pelo suporte financeiro.

Aos professores do DEX, pelos valiosos ensinamentos e a professora Anliy, pela amizade e carinho.

Ao meu orientador, Lucas Monteiro Chaves, pela grande atenção, competência, dedicação e arte de ensinar.

A todos os funcionários do DEX, pelos serviços prestados.

A todos os meus colegas de curso, em especial as amigas Ana Paula e Deyse.

Ao colega Devanil que muito acrescentou em minha formação e que me ajudou muito na conclusão deste trabalho.

Ao meu amado pai que sempre me apoiou e ajudou nesta longa caminhada rumo ao conhecimento. À minha prezada mãe, pelo exemplo de vida, força, incentivo e dedicação.

Aos meus queridos irmãos Renato e Lívia.

Enfim, a todos que contribuíram, direta ou indiretamente, meus eternos agradecimentos.

SUMÁRIO

Página

LISTA DE TABELAS

LISTA DE FIGURAS

RESUMO

GUIMARÃES, Paulo Henrique Sales. **Uma abordagem geométrica da Teoria de Inversas Generalizadas.** 2010. 67 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG.[1]

Uma abordagem geométrica em termos de subespaços vetoriais e projetores lineares é utilizada para apresentar a teoria da inversa generalizada de Moore-Penrose. Suas principais propriedades são obtidas por este método. Algumas propriedades das equações normais também são demonstradas. Uma generalização desta interpretação geométrica da inversa de Moore-Penrose é aplicada para as inversas reflexivas em geral. Em particular a inversa de quadrados mínimos também é explicitada geometricamente.
.

[1] **Orientador:** Lucas Monteiro Chaves – UFLA

ABSTRACT

GUIMARÃES, Paulo Henrique Sales. A **geometrical approach of Generalized Inverse Theory**. 2010. 67 p. Dissertation (Master in Statistics and Agricultural Experimentation) - Federal University of Lavras, Lavras, MG.[1]

A geometrical approach regarding to vectorial subspaces and linear projectors is used to present the Moore-Penrose generalized inversion theory. Its main properties are acquired by this method. Some properties of the normal equations are demonstrated as well. A generalization from this Moore-Penrose inverse geometric interpretation is applied to the reflexive inverses overall. The minimum square inverse is particularly demonstrated geometrically.
.

1 **Adviser:** Lucas Monteiro Chaves – UFLA

1 INTRODUÇÃO

A álgebra linear é talvez a área da matemática mais acessível. Uma razão qual pode justificar tal fato é a sua natureza dupla, isto é, álgebra, com toda a sua abstração e elegância, e geometria, com todo o seu apelo intuitivo. Vetores, ângulos, subespaços vetoriais são conceitos geométricos e de fácil visualização. É também a área da matemática com maior aplicabilidade em estatística, modelos lineares e estatística multivariada que sob a hipótese de normalidade, são exemplos claros. Tem-se que o uso de transformações lineares ortogonais, diagonalização de matrizes simétricas e formas quadráticas são essenciais e de uso constante. Esses resultados de álgebra linear são encontrados em praticamente todos os livros utilizados nos cursos de graduação em matemática, engenharias e estatística. No entanto, o mesmo não acontece com a teoria de *inversas generalizadas*, muito importantes em estatística. Mesmo não exigindo nenhum pré-requisito a mais que outros resultados de álgebra linear acima citados o tópico sobre inversas generalizadas não é tratado, a não ser em alguns livros de álgebra de matrizes. Algumas referências básicas para a teoria é praticamente toda descrita por estatísticos. Desse modo a utilização de uma abordagem geométrica poderá dar uma contribuição ao entendimento de vários conceitos na estatística que são na sua maioria tratados puramente algébricos.

O fato interessante a ser observado é que a abordagem feita pelos autores estatísticos, em geral, é totalmente algébrica, extremamente analítica e abstrata. No entanto, as inversas generalizadas são utilizadas para se obterem estimadores e é essencial que as boas propriedades desses estimadores sejam explicitadas. Entretanto, o método puramente algébrico não deixa claro quais são essas boas propriedades. Nesse sentido, tem-se a impressão que os estatísticos não estão utilizando o método mais adequado a seus objetivos. No livro de Lima (2006) intitulado Álgebra Linear, uma abordagem bastante geométrica é desenvolvida para a obtenção da inversa generalizada de Moore-Penrose (denominada de pseudo-inversa), entretanto sem que nenhuma referência à estatística seja feita.

Estendendo a interpretação geométrica desenvolvida em Lima (2006) para a inversa generalizada de Moore-Penrose, uma construção semelhante é apresentada para inversas generalizadas reflexivas e de quadrados mínimos. Na bibliográfica básica consultada não se encontrou nenhuma referência que abordasse algo semelhante. Por último, o texto pretende ser, apesar da dificuldade de tal, uma referência didática em português, deste tópico que é muito utilizado pelos estatísticos, porém infelizmente pouco compreendido.

O objetivo deste trabalho é rever a teoria básica das inversas generalizadas sob o ponto de vista geométrico. Desta forma pretende-se explicitar propriedades que certamente ajudarão os estatísticos a compreender melhor a propriedade dos estimadores obtidos por métodos exclusivamente algébricos.

2 REFERENCIAL TEÓRICO

A álgebra linear é muito estudada na graduação de praticamente todos os cursos na área de ciências exatas. Tem-se que os conceitos básicos, definições e aplicações podem ser encontradas em excelentes livros tais como Lima (2006) e Hoffmann & Kunze (1971). Em razão disso, neste texto, as definições e conceitos gerais serão restritos ao mínimo necessário, a notação também será a usual, uma vez que se pressupõe que um futuro leitor desse texto tenha em mãos um livro básico de álgebra linear.

Todo espaço vetorial real, neste texto, será considerado como sendo o espaço Euclidiano $\mathbb{R}^n$ com a base canônica $e = \{e_1, e_2, ..., e_n\}$ e produto interno $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$. Assim, toda transformação linear $A : \mathbb{R}^n \to \mathbb{R}^m$ identifica-se com uma matriz $A_{m \times n}$. Tal identificação simplifica a exposição sem perda de generalidade dos resultados. Se $A_{m \times n} : \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear, então a matriz transposta $A_{n \times m}^t$ define a transformação linear $A_{n \times m}^t : \mathbb{R}^m \to \mathbb{R}^n$. Ficando claras as dimensões do domínio e contradomínio, a notação A será usada para significar $A_{m \times n}$. As notações e os conceitos usuais de álgebra linear seguem os utilizados por Lima (2006).

2.1 Posto, imagem e núcleo de uma transformação linear

Seja $A: \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação linear. A imagem de A é o conjunto dos vetores $w \in \mathbb{R}^m$ tais que existe um vetor $v \in \mathbb{R}^n$, que satisfaz $A(v) = w$, representada por:

$$Im(A) = \{w \in \mathbb{R}^m; A(v) = w \text{ para algum } v \in \mathbb{R}^n\}.$$

O conjunto de todos os vetores $v \in \mathbb{R}^n$ tais que $A(v) = 0$ é definido como núcleo de A, denotado como $Ker(A)$. Isto é,

$$Ker(A) = \{v \in \mathbb{R}^n; A(v) = 0\}.$$

O posto de uma transformação linear A é definido como a dimensão do subespaço $Im(A)$. Como $A(e_i)$ é o vetor definido pela i-ésima coluna de A, então o posto de A será igual ao número de colunas linearmente independentes. Será usada a notação $r(A)$ para designar o posto de A.

Proposição 1 (Lima, 2006,p. 95). O número de colunas linearmente independentes de uma matriz A é igual ao número de linhas linearmente independentes desta matriz.

Corolário 1 O posto de A é igual ao posto de sua transposta, isto é, $r(A) = r(A^t)$.

Teorema 1 (Do núcleo e da imagem – Lima,2006, p.68): Se $A: \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear, então a soma das dimensões da imagem de A e de seu núcleo é igual à dimensão do domínio $\mathbb{R}^n$, isto é, $dim(Im(A)) + dim(Ker(A)) = n.$

Proposição 2 (Lima, 2006, p. 63) Seja $A: \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Para todo $y \in Im(A)$, o conjunto $V = \{x \in \mathbb{R}^n; Ax = y\}$, formado pelas soluções do sistema linear $Ax = y$ é uma variedade afim em $\mathbb{R}^n$, paralela ao $Ker(A)$.

A noção de subespaço vetorial abrange retas, planos e seus análogos multidimensionais apenas nos casos em que os conjuntos contêm a origem. No caso de retas, planos, dentre outros que não passam pela origem, tem-se a noção de variedade afim. Um subconjunto $V \subset \mathbb{R}^n$ chama-se uma *variedade afim* quando a reta que une dois pontos quaisquer de V está contida em V.

Algebricamente, a Proposição 2 significa que, para cada $y \in Im(A)$, obtém-se todas as soluções $x \in \mathbb{R}^n$ do sistema linear $Ax = y$. Logo se encontra uma "solução particular" x_0 desse sistema e a solução geral denotada por $x = x_0 + z$ é a soma dessa solução particular com a "solução geral z do sistema homogêneo associado" $Ax = 0$. Naturalmente, esta última é um elemento qualquer do núcleo de A. Se $y \notin Im(A)$ então o sistema $Ax = y$, evidentemente, não possui solução.

Corolário 2 $r(A) \leq menor\{n,m\}$.

Segue das definições se $A:\mathbb{R}^n \to \mathbb{R}^m$ tem posto linha completo, então $r(A) = m$ e a aplicação é sobrejetiva; e se A tem posto coluna completo, $r(A) = n$ e a aplicação é injetiva.

A transposta de A deve ser uma transformação linear $A^t:\mathbb{R}^n \to \mathbb{R}^m$ tal que, para cada $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ quaisquer, tem-se $\langle Ax, y\rangle = \langle x, A^t y\rangle$. Isso só se verifica porque está sendo considerada a base canônica e (que é ortonormal).

Seque que se $x \in Ker(A)$ então, para todo $y \in \mathbb{R}^m$, $\langle Ax, y\rangle = 0 = \langle x, A^t y\rangle$ e, consequentemente, $Ker(A^t)$ é perpendicular à $Im(A)$ e o $Ker(A)$ é perpendicular à $Im(A^t)$. Isso pode ser visualizado na Figura 1.

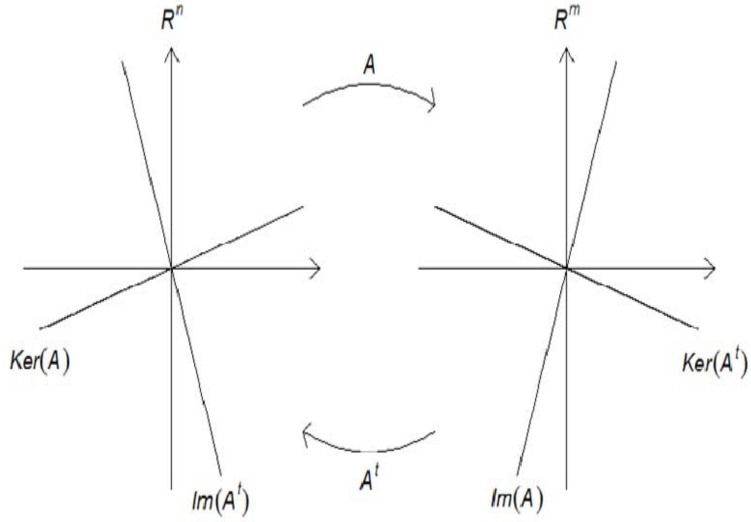

FIGURA 1 Representação geométrica do núcleo e da imagem de A e A^t

Outra observação importante, que tem uma interpretação geométrica interessante, é que se x_1 e x_2 são tais que $Ax_1 = y$ e $Ax_2 = y$, então, $Ax_2 - Ax_1 = 0$, que implica em $A(x_2 - x_1) = 0$, e, portanto, $x_2 - x_1 \in Ker(A)$. Isso pode ser visualizado na Figura 2.

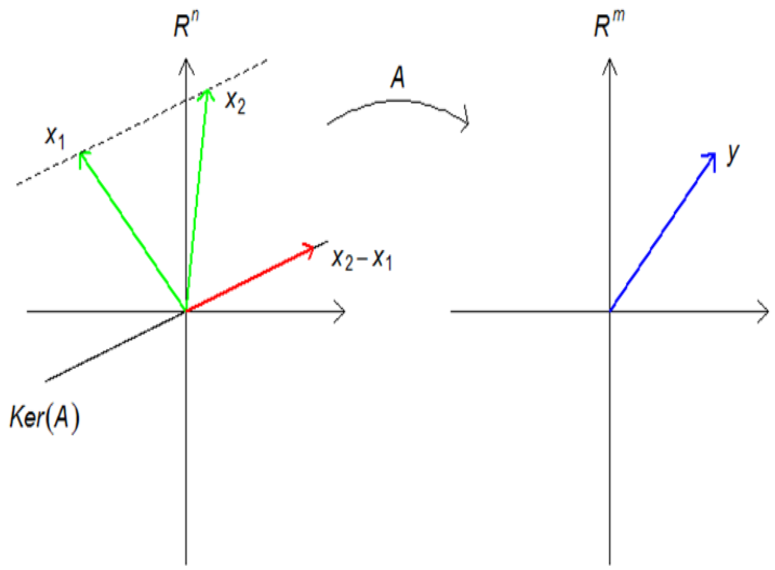

FIGURA 2 Geometria de $Ax_1 = y$ e $Ax_2 = y$

Em outras palavras: se $A(x) = y$, então, para todo $w \in Ker(A)$, $A(x+w) = y$.

Uma matriz quadrada tal que $A^2 = A$ é chamada matriz de projeção, projetor ou idempotente. Lembre-se que uma matriz será idempotente se $AA = A^2 = A$. Logo se tem que um projetor A restrito à $Im(A)$ é a identidade, ou seja, $A(Az) = A^2 z = Az$, sendo $z \in \mathbb{R}^n$ um vetor qualquer.

$I - A$ também é um projetor, pois,

$$(I - A)^2 = I - 2A + A^2 = I - A.$$

Como

$$A((I - A)z) = A(z - Az) = Az - A^2 z = Az - Az = 0,$$

segue que $Im(I-A) \subset Ker(A)$, em que I é matriz identidade.

$$z \in Ker(A) \Rightarrow (I-A)z = z - Az = z - 0 \Rightarrow z \in Im(I-A),$$

de onde segue que $Im(I-A) = Ker(A)$.

Considerando que $z \in Ker(I-A)$,

$$0 = (I-\ A)z = z - Az = 0 \Rightarrow Az = z$$

$$z \in Im(A) \Rightarrow Ker(I-A) \subset Im(A).$$

Se $w \in Im(A)$,

$$\begin{aligned} w = Az \Rightarrow (I-A)w = (I-A)Az = Az - A^2 z = \\ = Az - Az = 0 \ \Rightarrow Im(A) \subset Ker(I-A), \end{aligned}$$

logo, $Im(A) = Ker(I-A)$.

Uma matriz de projeção é dita um *projetor ortogonal* se $Av - v$ é perpendicular ao subespaço $Im(A)$.

Proposição 3 Uma matriz de projeção é simétrica se, e somente se, é um projetor ortogonal.

Prova: Sejam v e w vetores quaisquer

Se $\left\langle v - Av, Aw \right\rangle$ e utilizando as propriedades de produto interno tem-se que

$$\left\langle v - Av, Aw \right\rangle = \left\langle v, Aw \right\rangle - \left\langle Av, Aw \right\rangle = \left\langle v, Aw \right\rangle - \left\langle v, A^2 w \right\rangle =$$

$$= \left\langle v, Aw \right\rangle - \left\langle v, Aw \right\rangle = 0 \ \forall v, w \text{, pois } A^2 = A.$$

Considerando que $Av - v$ é perpendicular a Aw, isto é,

$$\langle Av{-}v, Aw \rangle = 0 \Rightarrow \langle Av, Aw \rangle = \langle v, Aw \rangle,$$

da mesma forma $Aw - w$ é perpendicular a Av, isto é,

$$\langle Aw{-}w, Av \rangle = 0 \Rightarrow \langle Aw, Av \rangle = \langle w, Av \rangle,$$

portanto, $\langle Av, w \rangle = \langle v, Aw \rangle$. □

As matrizes de projeção encontram-se representadas na Figura 3. Note que estas matrizes são simétricas e também projetores ortogonais (perpendiculares).

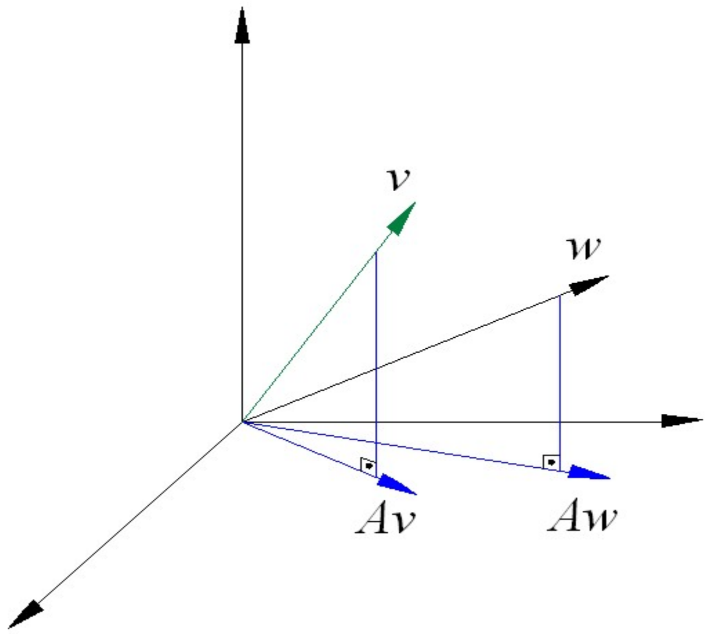

FIGURA 3 Visualização da geometria de matrizes de projeção.

2.2 Fatoração de posto completo

Fatorar uma matriz $A_{m\times n}$ consiste em obter matrizes $B_{m\times r}$ e $C_{r\times n}$ tais que $A = BC$. Uma fatoração é dita de posto completo se r é igual ao posto da matriz A.

Dwivedi (1975) propôs um algoritmo para obtenção de uma fatoração de posto completo caracterizado pela necessidade de se conhecer o posto da matriz e o número de passos necessário sendo o posto da matriz A.

Seja $A_{m\times n} = (a_{ij})$, de posto r, com $i = 1,..., p,..., m; j = 1,..., q,..., n.$ Faz-se:

(a) escolher algum elemento $a_{pq} \neq 0$;

(b) obter o produto U_1V_1,

em que: $U_1 = \dfrac{1}{a_{pq}}\begin{bmatrix} a_{1q} \\ a_{2q} \\ \cdots \\ a_{mq} \end{bmatrix}$; $V_1 = \left[a_{p1}, a_{p2}, \cdots, a_{pn}\right]$

(c) Fazer $A_1 = A - U_1V_1$

(d) Se $A_1 = 0$, o processo está encerrado, então $B = U_1$ e $C = V_1$

(e) Se $A_1 \neq 0$, repetir o mesmo processo para A_1, e assim, sucessivamente até obter $A_r = 0.$

(f) Ao final do processo, tem-se a matriz B é constituída pelos vetores colunas $U's$ e a matriz C pelos vetores linhas $V's$, de onde segue que a fatoração pode ser representada por:

$$A_{m\times n} = U_1V_1 + U_2V_2 + ... + U_rV_r = [U_1, ..., U_r]\begin{bmatrix} V_1 \\ \vdots \\ V_r \end{bmatrix} = B_{m\times r}C_{r\times n}.$$

2.3 Inversas à direita e à esquerda de uma matriz retangular

Seja A uma matriz $m \times n$ de posto m. De acordo com Rao & Mitra (1971) como AA^t é uma matriz $m \times m$ de posto m, a inversa $(AA^t)^{-1}$ existe e

$$(AA^t)(AA^t)^{-1} = I_m = A\left[A^t(AA^t)^{-1}\right] = AA_D^{-1}, \tag{2.1}$$

em que a matriz A_D^{-1} é chamada de *inversa à direita* de A.

Se $r(A) = n$ e A^tA é uma matriz $n \times n$, então a inversa $(A^tA)^{-1}$ existe e

$$(A^tA)^{-1}(A^tA) = I_n = \left[(A^tA)^{-1}A^t\right]A = A_E^{-1}A. \tag{2.2}$$

em que A_E^{-1} é chamada de *inversa à esquerda* de A.

2.4 Matriz particionada

Em certas situações é útil se trabalhar com a matriz quadrada A particionada em blocos como, por exemplo:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad (2.3)$$

em que A_{11} é uma submatriz de dimensão $m_1 \times m_1$, A_{12} é $m_1 \times m_2$, A_{21} é $m_2 \times m_1$ e A_{22} é $m_2 \times m_2$.

2.4.1 Determinante de matriz particionada

Teorema 2 Seja a matriz $A_{m\times m}$ particionada como (2.3). De acordo com Schott (2005) se $A_{12} = 0$ ou $A_{21} = 0$, então o determinante de A é dado por $|A| = |A_{11}||A_{22}|$.

Prova: Observe que

$$|A| = \begin{vmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{vmatrix} = \begin{vmatrix} A_{11} & 0 \\ A_{21} & I_{m_2} \end{vmatrix} \begin{vmatrix} I_{m_1} & 0 \\ 0 & A_{22} \end{vmatrix} = |A_{11}||A_{22}|.$$

Uma prova semelhante pode ser obtida para $|A| = |A_{11}||A_{22}|$ quando $A_{21} = 0$. □

Teorema 3 Seja A uma matriz de dimensão $m \times m$ não singular que possa ser particionada como em (2.3). Então valem as seguintes igualdades:

a) $|A| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|$, caso A_{22} seja uma matriz não singular.

b) $|A| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}|$, caso A_{11} seja uma matriz não singular.

Prova: Caso A_{22} seja uma matriz não singular, vale a identidade

$$\begin{bmatrix} I_{m_1} & -A_{12}A_{22}^{-1} \\ 0 & I_{m_2} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_{m_1} & 0 \\ -A_{22}^{-1}A_{21} & I_{m_2} \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

Tomando-se o determinante de ambos os lados da igualdade e usando o Teorema 2 obtém-se o resultado em (a). A prova de (b) é obtida de maneira similar utilizando-se a identidade:

$$\begin{bmatrix} I_{m_1} & 0 \\ -A_{21}A_{11}^{-1} & I_{m_2} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_{m_1} & -A_{11}^{-1}A_{12} \\ 0 & I_{m_2} \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \qquad \square$$

2.4.2 Complemento de Schur

Segundo Schott (2005) as matrizes $A_{11} - A_{12}A_{22}^{-1}A_{21}$ e $A_{22} - A_{21}A_{11}^{-1}A_{12}$ são denominadas complemento de Schur de A_{11} em A e complemento de Schur de A_{22} em A respectivamente, e designadas por S.

O complemento de Schur desempenha um papel importante na análise numérica e em muitas outras áreas da matemática e tem bastantes aplicações na estatística.

2.4.3 Inversa de matriz particionada

Teorema 4 (Schott, 2005,p. 256) Considere a matriz $A_{m\times m}$ não singular como em (2.3). Nesse caso a inversa de A será dada pela expressão

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1} \\ -S^{-1}A_{21}A_{11}^{-1} & S^{-1} \end{bmatrix} \tag{2.4}$$

se A_{11} e $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$ são matrizes não singulares. Do mesmo modo, se A_{22} e $S = A_{11} - A_{12}A_{22}^{-1}A_{21}$ são matrizes não singulares, então a inversa de A é dada pela expressão em (2.5):

$$A^{-1} = \begin{bmatrix} S^{-1} & -S^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}S^{-1} & A_{22}^{-1} - A_{22}^{-1}A_{21}S^{-1}A_{12}A_{22}^{-1} \end{bmatrix}. \tag{2.5}$$

2.4.4 Posto de matriz particionada

No caso de que a matriz A seja particionada como em (2.3) e, que A_{12} e A_{21} sejam ambas as matrizes nulas, ficando

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix},$$

então, o posto de A é dado por $r(A) = r(A_{11}) + r(A_{22})$. Quando A_{12} ou A_{21} forem matrizes singulares, mas se A_{11} ou A_{22} forem não singulares, de acordo com Schott (2005), o Teorema 5, enunciado abaixo, deverá ser usado para determinar o posto de A.

Teorema 5 Seja A uma matriz particionada definida como em (2.3), então:

a) Se A_{22} for não singular o posto de A será:

$$r(A) = r(A_{22}) + r(A_{11} - A_{12}A_{22}^{-1}A_{21}) .$$

b) Se A_{11} for não singular o posto de A será:

$$r(A) = r(A_{11}) + r(A_{22} - A_{21}A_{11}^{-1}A_{12}) .$$

Prova: Para demonstrar a parte (a) do Teorema 5, considerando que A_{22} é uma matriz não singular e a matriz A representada como:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$= \begin{bmatrix} I_{m_1} & A_{12}A_{22}^{-1} \\ 0 & I_{m_2} \end{bmatrix} \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I_{m_1} & 0 \\ A_{22}^{-1}A_{21} & I_{m_2} \end{bmatrix}.$$

Pelo Teorema 2 o determinante da matriz

$$\begin{bmatrix} I_{m_1} & A_{12}A_{22}^{-1} \\ 0 & I_{m_2} \end{bmatrix},$$

é igual a 1, logo, esta matriz é não singular. Do mesmo modo, a matriz

$$\begin{bmatrix} I_{m_1} & 0 \\ A_{22}^{-1}A_{21} & I_{m_2} \end{bmatrix},$$

também é não singular. Logo, o posto de A será igual:

$$\mathrm{r}\left\{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\right\} = \mathrm{r}\left\{\begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix}\right\}.$$

A parte (b) do Teorema 5 é obtida de maneira similar. $\square$

2.5 Inversa generalizada

As inversas generalizadas de matrizes apareceram na literatura em estatística e matemática como uma opção ao uso de inversas clássica matrizes, sendo que essas são usadas somente no caso de matrizes quadradas e positivas definida. A inversa generalizada é mais abrangente, valendo suas propriedades para quaisquer matrizes e, principalmente as matrizes retangulares, que são de ocorrência bastante comum em estatística.

Rao & Mitra (1971) definiram inversa generalizada de uma matriz qualquer com propriedades similares àquelas da inversa de uma matriz não singular.

Definição 1 Seja A uma matriz de dimensão $m \times n$ de posto qualquer, não nulo. Uma inversa generalizada de A, com notação A^{-}, é uma matriz $n \times m$ tal que $x = A^{-}y$ é uma solução do sistema $Ax = y$, para qualquer y que torne o sistema consistente.

Outra definição algébrica de inversa generalizada é dada por:

Definição 2 Uma inversa generalizada de A é uma matriz A^{-} que satisfaz $AA^{-}A = A$.

Proposição 4 As definições 1 e 2 são equivalentes.

Prova: Se y é tal que $Ax = y$ é consistente, $A^{-}y$ será solução do sistema e, portanto, $AA^{-}y = y$. Para todo $x \in \mathbb{R}^n$, $y = Ax$ define um sistema consistente. Logo, se tem $A\,A^{-}y = A\,A^{-}Ax = Ax = y$ e, por conseguinte, $A\,A^{-}A = A.$ $\square$

Teorema 6 (Lima,2006,p. 220): Se A é uma matriz de dimensão $m \times n$ de posto $r(A) > 0$, existem matrizes ortogonais $P_{m \times m}$ e $Q_{n \times n}$,tais que, $A = P\Delta Q^t$, em que a matriz

$$\Delta_{m\times n} = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix},$$

é uma matriz não singular e D_r uma matriz diagonal com entradas positivas, $d_{ii} = \sigma_i > 0$,com σ_i^2 autovalores de $A^t A$ ou de AA^t.

Para qualquer matriz $A_{m\times n}$ existe sempre uma inversa generalizada. Pelo Teorema 6 existem matrizes ortogonais P e Q de ordens m e n respectivamente, tais que $A = P\Delta Q^t$. Assim uma maneira de se encontrar uma inversa generalizada de A é:

$$A^- = Q\Delta^- P^t, \tag{2.6}$$

com $\Delta^- = \begin{pmatrix} D_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}$.

Utilizando-se a Definição 2, pode-se verificar que

$$AA^- A = P\Delta Q^t Q\Delta^- P^t P\Delta Q^t = P\Delta\Delta^-\Delta Q^t = P\Delta Q^t = A. \quad \square$$

Searle (1971) apresentou um algoritmo extremamente simples para a obtenção de inversas generalizadas. Dada uma matriz $A_{m\times n}$ de posto r faz-se:

i) Escolher uma submatriz qualquer em $A_{m\times n}$, de dimensão $r\times r$ dita complementar não singular de A, denotada por $M_{(r)}$, com mesmo posto de $A_{m\times n}$.

ii) Obter a transposta da inversa de M , isto é, $(M^{-1})^t$;

iii) Substituir em A os elementos de M por seus correspondentes em $(M^{-1})^t$;

iv) Fazer todos os outros elementos iguais a zero;

v) Transpor a matriz resultante.

O resultado assim obtido é uma matriz inversa generalizada de A.

Exemplo 1 Seja a matriz:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 0 & 2 \end{bmatrix}, \quad r(A) = 2.$$

Um menor complementar não singular de A com $r(M) = 2$ é $M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, então $(M^{-1})^t = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$, logo uma inversa generalizada de A é

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

Observe que podem ser obtidas diferentes inversas generalizadas de $A_{m \times n}$, dependendo da escolha de M.

Teorema 7 Segundo Rao (1962) O sistema de equações $Ax = y$ é consistente se e somente se, para alguma inversa generalizada A^{-} de A, $AA^{-}y = y$.

Prova: Suponha que o sistema seja consistente e que x^* seja solução. Então $y = Ax^*$. Pré-multiplicando esta identidade por AA^{-}, no qual A^{-} é qualquer inversa generalizada de A, tem-se que

$$AA^{-}y = AA^{-}Ax^* = Ax^* = y.$$

Supondo agora que há uma inversa generalizada de A satisfazendo $AA^{-}y = y$. Definindo-se $x^* = A^{-}y$, nota-se que

$$Ax^* = AA^{-}y = y.$$

Segue que, como $x^* = A^- y$ é solução, o sistema é consistente e, portanto a prova está completa. □

Teorema 8 Suponha que $Ax = y$ seja um sistema consistente de equações e seja A^- qualquer inversa generalizada da matriz $A_{m\times n}$. Então, para qualquer vetor h, $n\times 1$,

$$x_h = A^- y + (I_n - A^- A)h \tag{2.7}$$

é solução, e para qualquer solução x_h, existe um vetor h que satisfaz a equação 2.7.

Prova: Como $Ax = y$ seja um sistema consistente de equações, pelo Teorema 7 tem-se que $AA^- y = y$ e então,

$$Ax_h = AA^- y + A\,(I_n - A^- A)h$$

$$= y + (A - AA^- A)h = y,$$

pois, $AA^- A = A$. Então, x_h é solução do sistema independente da escolha h. Em outras palavras, se x^* é uma solução arbitrária, como $Ax^* = y$, então $A^- Ax^* = A^- y$. Consequentemente,

$$A^- y + (I_n - A^- A)x^* = A^- y + x^* - A^- Ax^* = x^*.$$

Dessa forma $x^* = x_h$, e assim a prova está completa. □

2.6 Inversa generalizada de Moore-Penrose

Moore (1920) e posteriormente Penrose (1955) em trabalhos independentes, mostraram que, para toda matriz real $A_{m \times n}$ há uma e somente uma matriz denotada por $A^+_{n \times m}$ que satisfaz as quatro condições seguintes:

i. $AA^+A = A$
ii. $A^+AA^+ = A^+$
iii. $(AA^+)^t = AA^+$
iv. $(A^+A)^t = A^+A$.

A matriz A^+ que satisfaz as quatro condições (i – iv) é denominada inversa generalizada de Moore-Penrose de A.

Teorema 9 Toda matriz A admite uma, e somente uma, inversa generalizada de Moore-Penrose A^+.

Prova: Primeiramente será provada a existência de A^+. Se A for uma matriz nula $m \times n$, então são facilmente verificadas as quatro condições (i – iv) com $A^+ = 0_{n \times m}$, a matriz nula. Mas se $A \neq 0_{n \times m}$, então $r(A) > 0$. Portanto,

$$A = P\Delta Q^t \quad \text{, em que } \Delta = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Seja a inversa de Moore-Penrose dada por: $A^+ = Q\Delta^- P^t$ em que

$$\Delta^- = \begin{pmatrix} D_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

i) $AA^+A = P\Delta Q^t Q\Delta^- P^t P\Delta Q^t = P\Delta\Delta^-\Delta Q^t = P\Delta Q^t = A$;

ii) $A^+AA^+ = Q\Delta^- P^t P\Delta Q^t Q\Delta^- P^t = Q\Delta^-\Delta\Delta^- P^t = Q\Delta^- P^t = A^+$;

iii) $AA^+ = P\Delta Q^t Q\Delta^- P^t = P\Delta\Delta^- P^t$ (simétrica);

iv) $A^+A = Q\Delta^- P^t P\Delta Q^t = Q\Delta^-\Delta Q^t$ (simétrica).

Portanto, $A^{+} = Q\Delta^{-}P^{t}$ é uma inversa de Moore-Penrose de A, e fica estabelecida a sua existência.

Para provar a unicidade de A^{+} suponha que existam duas matrizes B e C satisfazendo as condições (i – iv). Então utilizando as quatro condições tem-se que:

$$AB = (AB)^{t} = B^{t}A^{t} = B^{t}(ACA)^{t} = B^{t}A^{t}(AC)^{t}$$
$$= (AB)^{t}AC = ABAC = AC;$$

e

$$BA = (BA)^{t} = A^{t}B^{t} = (ACA)^{t}B^{t} = (CA)^{t}A^{t}B^{t}$$
$$= CA(BA)^{t} = CABA = CA.$$

Usando-se estas duas identidades e (ii), tem-se que

$$B = BAB = BAC = CAC = C.$$

Desde que B e C sejam matrizes idênticas, a inversa generalizada de Moore-Penrose é única. □

Outra maneira de se obter a inversa de Moore-Penrose de A é através da fatoração de posto completo.

Proposição 5 Se $A_{m\times n} = B_{m\times r}C_{r\times n}$ é a fatoração de posto completo de $A_{m\times n}$, então

$$A^{+} = C^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t}. \qquad (2.8)$$

Prova:

i) $AA^{+}A = BCC^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t}BC = B\,I\,I\,C = BC = A$

ii) $A^{+}AA^{+} = C^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t}BCC^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t} =$
$= C^{t}(CC^{t})^{-1}II(B^{t}B)^{-1}B^{t} = C^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t} = A^{+},$

iii) $A\,A^{+} = BCC^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t} = BI(B^{t}B)^{-1}B^{t} = B(B^{t}B)^{-1}B^{t}$ e, portanto, simétrica.

iv) $A^+A = C^t(CC^t)^{-1}(B^tB)^{-1}B^tBC = C^t(CC^t)^{-1}IC = C^t(CC^t)^{-1}C$ e, portanto, simétrica. □

Observe que a fórmula da inversa de Moore-Penrose da Proposição 5 pode ser escrita como o produto da inversa à direita de C pela inversa à esquerda de B, isto é, $A^+ = C_D^{-1}B_E^{-1}$.

As duas maneiras de se obter a inversa de Moore-Penrose enunciadas acima estão relacionadas da forma: Definindo

$$\Delta^{1/2} = \begin{pmatrix} \sqrt{D_r} & 0 \\ 0 & 0 \end{pmatrix},$$

a matriz A pode ser escrita como $A = BC = P\Delta^{1/2}\Delta^{1/2}Q^t$, em que $B = P\Delta^{1/2}$ e $C = \Delta^{1/2}Q^t$, sendo P e Q matrizes ortogonais e, B e C matrizes de posto completo. Tem-se então que:

$$\begin{aligned}
A^+ &= C^t(CC^t)^{-1}(B^tB)^{-1}B^t \\
&= Q\Delta^{1/2}(\Delta^{1/2}Q^tQ\Delta^{1/2})^{-1}(\Delta^{1/2}P^tP\Delta^{1/2})^{-1}\Delta^{1/2}P^t \\
&= Q\Delta^{1/2}(\Delta^{1/2}\Delta^{1/2})^{-1}(\Delta^{1/2}\Delta^{1/2})^{-1}\Delta^{1/2}P^t \\
&= Q(\Delta^{1/2}\Delta^{1/2})(\Delta^{1/2}\Delta^{1/2})^{-1}(\Delta^{1/2}\Delta^{1/2})^{-1}P^t \\
&= QI(\Delta^{1/2}\Delta^{1/2})^{-1}P^t \\
&= Q\Delta^{-1}P^t = Q\Delta^{-}P^t = A^+
\end{aligned}$$

□

Exemplo 2 Seja a matriz A definida abaixo:

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

Pela fatoração de posto completo, obtém-se

$$B=\begin{bmatrix}1 & 0\\1 & 0\\1 & 1\\1 & 1\end{bmatrix} \qquad C=\begin{bmatrix}1 & 1 & 0\\0 & -1 & 1\end{bmatrix}.$$

A inversa de Moore-Penrose de A pode ser computada por: $A^{+}=C^{t}(CC^{t})^{-1}(B^{t}B)^{-1}B^{t}$.

$$A^{+}=\begin{bmatrix}1/6 & 1/6 & 1/6 & 1/6\\1/3 & 1/3 & -1/6 & -1/6\\-1/6 & -1/6 & 1/3 & 1/3\end{bmatrix}.$$

Outra forma de se computar a inversa generalizada de Moore-Penrose de A é através de $A^{+}=Q\Delta^{-}P^{t}$. Para isso é necessário decompor a matriz A em $A=P\Delta Q^{t}$.

$$A=\begin{bmatrix}1 & 1 & 0\\1 & 1 & 0\\1 & 0 & 1\\1 & 0 & 1\end{bmatrix} = \begin{bmatrix}1/2 & 1/2\\1/2 & 1/2\\1/2 & -1/2\\1/2 & -1/2\end{bmatrix}\begin{bmatrix}\sqrt{6} & 0\\0 & \sqrt{2}\end{bmatrix}\begin{bmatrix}2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6}\\0 & 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}.$$

$$A^{+}=\begin{bmatrix}2/\sqrt{6} & 0\\1/\sqrt{6} & 1/\sqrt{2}\\1/\sqrt{6} & -1/\sqrt{2}\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 0\\0 & 1/\sqrt{2}\end{bmatrix}\frac{1}{2}\begin{bmatrix}1 & 1 & 1 & 1\\1 & 1 & -1 & -1\end{bmatrix}=$$

$$=\begin{bmatrix}1/6 & 1/6 & 1/6 & 1/6\\1/3 & 1/3 & -1/6 & -1/6\\-1/6 & -1/6 & 1/3 & 1/3\end{bmatrix}.$$

Chega-se ao mesmo resultado, evidenciando que as duas maneiras de se computar a inversa generalizada de Moore-Penrose são iguais.

2.6.1 Algumas propriedades da inversa generalizada de Moore-Penrose

Teorema 10

a) Se $A_{m\times n} = B_{m\times r}C_{r\times n}$ é a fatoração de posto completo, então $A^+ = C^+B^+$,

b) $(A^t)^+ = (A^+)^t$,

c) $(A^tA)^+ = A^+(A^+)^t$ e $(AA^t)^+ = (A^+)^tA^+$,

d) $r(A) = r(A^+) = r(AA^+) = r(A^+A)$.

Prova: **(a)** Como $A^+ = C^t(CC^t)^{-1}(B^tB)^{-1}B^t$, basta mostrar que $C^+ = C^t(CC^t)^{-1}$ e $B^+ = (B^tB)^{-1}B^t$.

Como B tem posto coluna completo, então $B^+ = (B^tB)^{-1}B^t$, logo,

i) $BB^+B = B(B^tB)^{-1}B^tB = B$;

ii) $B^+BB^+ = (B^tB)^{-1}B^tB(B^tB)^{-1}B^t = (B^tB)^{-1}B^t = B^+$;

iii) $BB^+ = B(B^tB)^{-1}B^t$ é uma matriz simétrica e;

iv) $B^+B = (B^tB)^{-1}B^tB$ é simétrica também.

A prova para $C^+ = C^t(CC^t)^{-1}$ segue de maneira idêntica. □

(b) Escrevendo $A = P\Delta Q^t$, logo $A^t = Q\Delta P^t$. Assim A^t pode ser escrito da forma:

$$A^t = A^t(A^t)^+A^t = Q\Delta P^t(Q\Delta P^t)^+Q\Delta P^t =$$

$$= Q\Delta P^t(P\Delta^- Q^t)Q\Delta P^t = Q\Delta\Delta^-\Delta P^t = Q\Delta P^t = A^t,$$

$$(A^t)^+ = (A^t)^+ A^t (A^t)^+ = (Q\Delta P^t)^+ Q\Delta P^t (Q\Delta P^t)^+$$

$$= P\Delta^- Q^t Q\Delta P^t P\Delta^- Q^t = P\Delta^- Q^t = (A^t)^+. \qquad \square$$

As outras duas condições seguem sem problema. Para **(c)**, desde que A^+ satisfaça as condições (i – iv) tem-se que:

i) $A^t A(A^t A)^+ A^t A = A^t AA^+ A^{+t} A^t A = A^t AA^+ (AA^+)^t A$

$$= A^t AA^+ AA^+ A = A^t AA^+ A = A^t A,$$

ii) $(A^t A)^+ A^t A(A^t A)^+ = A^+ A^{+t} A^t AA^+ A^{+t}$

$$= A^+ (AA^+)^t AA^+ A^{+t} = A^+ AA^+ AA^+ A^{+t}$$

$$= A^+ AA^+ A^{+t} = A^+ A^{+t} = (A^t A)^+,$$

iii) $A^t A(A^t A)^+ = A^t AA^+ A^{+t} = A^t (A^+ (AA^+)^t)^t$

$$= A^t (A^+ AA^+)^t = A^t A^{+t} = (A^+ A)^t,$$

iv) $(A^t A)^+ A^t A = A^+ A^{+t} A^t A = A^+ (A^+ (AA^+)^t)^t A$

$$= A^+ AA^+ A = A^+ A.$$

Assim a prova está completa. $\square$

Para a prova de **(d)** será utilizado a condição (i) e o fato de que o posto da matriz produto não pode exceder o posto de qualquer das matrizes no produto. Logo se tem que:

$$r(A) = r(AA^{+}A) \leq r(AA^{+}) \leq r(A^{+}).$$

De forma similar utilizando (ii), tem-se que

$$r(A^{+}) = r(A^{+}AA^{+}) \leq r(A^{+}A) \leq r(A). \qquad \square$$

2.7 Inversa reflexiva

Uma inversa generalizada de A é chamada reflexiva com notação A^r, segundo Rao & Mitra (1971) se satisfaz

$$AA^r A = A \text{ e } A^r AA^r = A^r . \tag{2.9}$$

Teorema 11 (Rao & Mitra, 1971, p. 28) Uma inversa generalizada A^r de A é reflexiva se, e somente se, $r(A) = r(A^r)$.

Para a obtenção da inversa reflexiva de uma matriz $A_{m\times n}$, considere a fatoração de posto completo de A, isto é, $A_{m\times n} = B_{m\times r}C_{r\times n}$ então, $A^r_{n\times m}$ pode ser expressa em função da inversa à esquerda de B e da inversa à direita de C, como

$$A^r = C_D^{-1} B_E^{-1} . \tag{2.10}$$

2.8 Inversa de quadrados mínimos

Uma inversa generalizada A^L de A é chamada de quadrados mínimos se AA^L é simétrica. A obtenção de inversas generalizadas de quadrados mínimos de A é obtida por: $A^L = (A^t A)^- A^t$, em que $(A^t A)^-$ é uma inversa generalizada de $A^t A$.

Prova:

i) Como $A^t A(A^t A)^- A^t A = A^t A$, multiplicando-se a esquerda por $(A^+)^t$,

$$(A^+)^t A^t A(A^t A)^- A^t A = (A^+)^t A^t A = AA^+ A(A^t A)^- A^t A = AA^+ A$$

$$= A(A^t A)^- A^t A = AA^L A = A.$$

ii) Desde que $AA^L A = A$ tem-se que $(AA^L)^t = AA^L$ tem-se que

$$AAA^L = AA^+ AA^L = (AA^+)^t (AA^L)^t = A^{+t} A^t A^{Lt} A^t$$

$$= A^{+t} (AA^L A)^t = A^{+t} A^t = (AA^+)^t = AA^+,$$

uma vez que AA^+ é simétrica. □

Exemplo 3 Seja $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 0 & 2 \end{bmatrix}$,então $A^t A = \begin{bmatrix} 7 & 2 & 9 \\ 2 & 2 & 4 \\ 9 & 4 & 13 \end{bmatrix}$.

Uma inversa generalizada de $A^t A$ é $(A^t A)^- = \frac{1}{10}\begin{bmatrix} 2 & -2 & 0 \\ -2 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, logo

uma inversa de quadrados mínimos de A é dada por:

$$\frac{1}{10}\begin{bmatrix} 2 & -2 & 0 \\ -2 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 2 & 2 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 0 & 2 & 0 & 4 \\ 5 & -2 & 5 & -4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

2.9 Inversa generalizada de matriz particionada

Teorema 12 Seja A uma matriz $m \times m$ particionada da forma (2.3). Suponha que $A_{12} = 0$ e $A_{21} = 0$ e que A_{11}^- e A_{22}^- são inversas generalizadas de A_{11} e A_{22} respectivamente. Logo a inversa generalizada de A será dada por:

$$A^- = \begin{bmatrix} A_{11}^- & 0 \\ 0 & A_{22}^- \end{bmatrix}. \tag{2.11}$$

Prova: O resultado de (2.11) é obtido por meio da identidade da inversa generalizada de Rao & Mitra (1971) $AA^-A = A$. Portanto, utilizando a identidade temos que $A_{11}A_{11}^-A_{11} = A_{11}$ e $A_{22}A_{22}^-A_{22} = A_{22}$, ficando provado (2.11). □

Uma maneira de se computar uma inversa generalizada de uma matriz $A_{m\times n}$ de posto r, escrito na forma particionada seguinte

$$A_{m\times n} = \begin{bmatrix} B_{\mathrm{r}\times\mathrm{r}} & C_{\mathrm{r}\times(n-\mathrm{r})} \\ D_{(m-\mathrm{r})\times\mathrm{r}} & E_{(m-\mathrm{r})\times(n-\mathrm{r})} \end{bmatrix} \tag{2.12}$$

em que B é inversível é dada pela Proposição 6.

Proposição 6 Uma inversa generalizada de A é :

$$\begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

Prova: Conforme a definição de Rao (1971) tem-se que $AA^{-}A = A$, logo,

$$AA^{-}A = \begin{bmatrix} B & C \\ D & E \end{bmatrix} \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ DB^{-1} & 0 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix} = \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} = A.$$

O resultado acima é válido se e somente se $E = DB^{-1}C$.

A demonstração desta propriedade não é trivial e geralmente não é encontrada nos livros de álgebra de matrizes.

Para cada coluna $(a_{1j},...,a_{nj})$ em que $r+1 \le j \le n$, existem r números $\beta_{1j},...,\beta_{lj}$ tais que:

$$\sum_{l=1}^{r} \begin{pmatrix} a_{1l} \\ \vdots \\ a_{ml} \end{pmatrix} \beta_{lj} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Matricialmente considere a matriz $\beta_{r\times(n-r)} = (\beta_{ij})$, logo

$$\begin{bmatrix} B_{\mathrm{r}\times\mathrm{r}} \\ D_{(m-\mathrm{r})\times\mathrm{r}} \end{bmatrix} \begin{bmatrix} \beta_{\mathrm{r}\times(n-\mathrm{r})} \end{bmatrix} = \begin{bmatrix} C_{\mathrm{r}\times(n-\mathrm{r})} \\ E_{(m-\mathrm{r})\times(n-\mathrm{r})} \end{bmatrix}$$

$$\begin{cases} B_{\mathrm{r}\times\mathrm{r}} \times \beta_{\mathrm{r}\times(n-\mathrm{r})} = C_{\mathrm{r}\times(n-\mathrm{r})} \Rightarrow \beta = B^{-1}C \\ D_{(m-\mathrm{r})\times\mathrm{r}} \times \beta_{\mathrm{r}\times(n-\mathrm{r})} = E_{(m-\mathrm{r})\times(n-\mathrm{r})} \end{cases}$$

$$E = DB^{-1}C.$$

A demonstração acima não é restritiva, isto é, uma matriz de posto r quando não estiver na forma (2.12) pode ser colocada nessa forma por transposição de linhas e colunas.

Exemplo 4 Considere a matriz A abaixo:

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

Escrevendo a matriz A de forma particionada, esta fica da forma:

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

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad E = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

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

Como $E = DB^{-1}C$, logo se obtém que

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

Teorema 13 (Marsaglia & Styan, 1974). Seja A uma matriz $m \times m$ particionada da forma (2.3). Seja A_{11}^{-} qualquer inversa generalizada de A_{11} e $S = A_{22} - A_{21}A_{11}^{-}A_{12}$ (Complemento de Schur em relação à A_{11}^{-}).

Então,

$$A^{-}=\begin{bmatrix} A_{11}^{-}+A_{11}^{-}A_{12}S^{-}A_{21}A_{11}^{-} & -A_{11}^{-}A_{12}S^{-} \\ -S^{-}A_{21}A_{11} & S^{-} \end{bmatrix} \qquad (2.13)$$

é a inversa generalizada de A definida em (2.3) para uma particular inversa generalizada S^{-} se e somente se o posto de A for dado por:

$$r(A)=r\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\right)=r(A_{11})+r(A_{22}-A_{21}A_{11}^{-}A_{12})$$

e, portanto, A^{-} é uma inversa generalizada de A para alguma inversa generalizada S^{-}.

A inversa generalizada de A é reflexiva se e somente se A_{11}^{r} e S^{r} são ambas inversas reflexivas de A_{11} e de S respectivamente.

2.9.1 Inversa generalizada de Moore-Penrose de matriz particionada

Suponha que a matriz A seja particionada como $A=\left[U \mid V\right]$, em que U tem dimensão $m\times n_1$ e V é $m\times n_2$. Em algumas situações, pode ser útil ter uma expressão de para A^{+} em termos das submatrizes U e V.

Teorema 14 De acordo com Schott (2005) dada as condições acima, e ainda, considerando que $n=n_1+n_2$, tem-se que

$$A^{+}=\begin{bmatrix} U^{+}-U^{+}V(C^{+}+W) \\ C^{+}+W \end{bmatrix} \qquad (2.14)$$

no qual $C=(I_m-UU^{+})V$, $M=\{I_{n2}+(I_{n2}-C^{+}C)V'U^{+'}U^{+}V(I_{n2}-C^{+}C)\}^{-1}$ e $W=(I_{n2}-C^{+}C)MV'U^{+'}U^{+}(I_m-VC^{+})$.

2.10 Inversas generalizadas na teoria dos Modelos Lineares

A teoria de modelos lineares é descrita em vários textos, podendo-se citar Searle (1971), Graybill (1976) e Rao (1973), entre outros.

De acordo com Searle (1971) o modelo linear de Gauss-Markov pode ser representado por

$$Y = \boldsymbol{X}\beta + \varepsilon \tag{2.15}$$

em que:

Y é um vetor $n \times 1$ de realizações de variáveis aleatórias observáveis;

$\boldsymbol{X}$ é uma matriz $n \times m$ de elementos conhecidos (matriz do delineamento) de posto r ,

β é um vetor $m \times 1$ de parâmetros desconhecidos;

ε é um vetor $n \times 1$ de componentes aleatórios com esperança zero e matriz de variâncias e covariâncias igual a um múltiplo da identidade. Para um dado modelo particular, a esperança matemática de um conjunto de observações é

$$\begin{array}{l}
E(Y_1) = x_{11}\beta_1 + x_{12}\beta_2 + \ldots + x_{1m}\beta_m \\
E(Y_2) = x_{21}\beta_1 + x_{22}\beta_2 + \ldots + x_{2m}\beta_m \\
\vdots \qquad\quad \vdots \qquad\qquad \vdots \qquad\quad \ddots \qquad \vdots \\
E(Y_n) = x_{n1}\beta_1 + x_{n2}\beta_2 + \ldots + x_{nm}\beta_m
\end{array}$$

e a variância de cada observação $V(Y_i) = \sigma^2$.

Representando matricialmente este sistema de equações tem-se:

$$\begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_n) \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix},$$

que é equivalente a: $E(Y) = \boldsymbol{X}\beta$ e $\varepsilon = Y - E(Y)$.

Observe que utilizando o conceito de esperança e variância tem-se

$$E(\varepsilon) = E(Y - E(Y)) = E(Y) - E(E(Y)) = E(Y) - E(Y) = 0$$

$$V(\varepsilon) = V(Y - E(Y)) = V(Y) - V(E(Y)) = V(Y) - 0_n = \sigma^2 I_n$$

em que $V(\cdot)$ é a matriz de dispersão padrão (matriz de variância covariância) e I_n é a matriz identidade de ordem n.

Um dos problemas que se encontra na utilização desses modelos é como estimar os parâmetros desconhecidos β_j com base nas observações y_i.

Dado o modelo linear $Y = \boldsymbol{X}\beta + \varepsilon$. Um dos métodos utilizados para estimar o vetor de parâmetros β é através do método de mínimos quadrados, o qual requer a minimização da soma de quadrados dos desvios. Na forma matricial consiste em minimizar

$$(y - \boldsymbol{X}\beta)^t (y - \boldsymbol{X}\beta) = \sum_{i=1}^{n} (y_i - (E(y_i))^2 \qquad (2.16)$$

Utilizando os conceitos de derivadas e de pontos críticos em relação aos parâmetros obtém-se o sistema

$$X^t X \hat{\beta} = X^t y \qquad (2.17)$$

o qual é denominado de sistema de equações normais.

Quando X tem posto coluna completo, $X^t X$ é uma matriz não singular e quadrada, logo tem inversa clássica e a solução única de (2.17) é

$$\hat{\beta} = (X^t X)^{-1} X^t y. \tag{2.18}$$

No entanto, nem sempre a matriz X tem posto coluna completo, como nos delineamentos experimentais e modelos superparametrizados. Neste caso $X^t X$ é uma matriz singular, logo não possui inversa clássica. Dado que a matriz $X^t X$ é singular o sistema $X^t X \hat{\beta} = X^t y$ não possui solução única. Uma solução de mínimos quadrados pode ser obtida com a utilização da inversa de Moore-Penrose.

Outra maneira de se escrever a solução das equações normais é por meio da matriz de projeção, definida nas Proposições 2 e 3. Desde que é possível escrever o sistema linear da forma $\hat{Y} = X\hat{\beta}$, então por (2.18) tem-se que

$$\hat{Y} = X\hat{\beta} = X(X^t X)^{-1} X^t y = Py, \tag{2.19}$$

em que a matriz $X(X^t X)^{-1} X^t$ em (2.19) é chamada de matriz de projeção. A inversa de Moore-Penrose é definida em função de um projetor, desta forma a matriz de projeção (2.19) pode ser reescrita como

$$X(X^t X)^{-1} X = XX^+. \tag{2.20}$$

A prova para o caso geral de (2.20) é dada no Corolário 5.

Teorema 15 As equações normais $X^t X \hat{\beta} = X^t y$ são sempre consistentes.

Prova: Primeiramente uma demonstração algébrica**.** O sistema $(X^t X)^- \hat{\beta} = X^t y$ é consistente se e somente se $X^t X(X^t X)^- X^t y = X^t y$.

$$\begin{aligned} X^t X(X^t X)^- X^t y &= X^t X(X^t X)^+ X^t y \\ &= X^t XX^+ (X^t)^+ X^t y \end{aligned}$$

desenvolvendo adequadamente e aplicando as propriedades da inversa de Moore-Penrose chega-se a

$$= X^t (X^t)^+ X^t y = X^t y. \qquad \square$$

Lema 1 A transformação linear X^t restrita ao subespaço $Im(X)$ é um isomorfismo de $Im(X)$ em $Im(X^t)$.

Prova: Seja $z \in Im(X)$ tal que $X^t z = 0$, $z = Xw$ para algum w, logo $X^t Xw = 0$. Como $\left\langle X^t Xw, x \right\rangle = 0 = \left\langle Xw, Xx \right\rangle \forall x$. Tem-se a $Im(X)$, de onde segue que X^t restrito à imagem de X é injetiva. Como $dim(X) = dim\,(X^t)$, pois o posto linha de X é igual ao seu posto coluna, X^t restrito a $Im(X)$ é um isomorfismo.

A consistência segue do fato dado y, $X^t y \in Im(X^t)$. Como X^t restrita a $Im(X)$ é um isomorfismo então existe $z \in Im(X)$ e existe β, tal que $X^t z = X^t y$, portanto, $X^t X \beta = X^t y$.

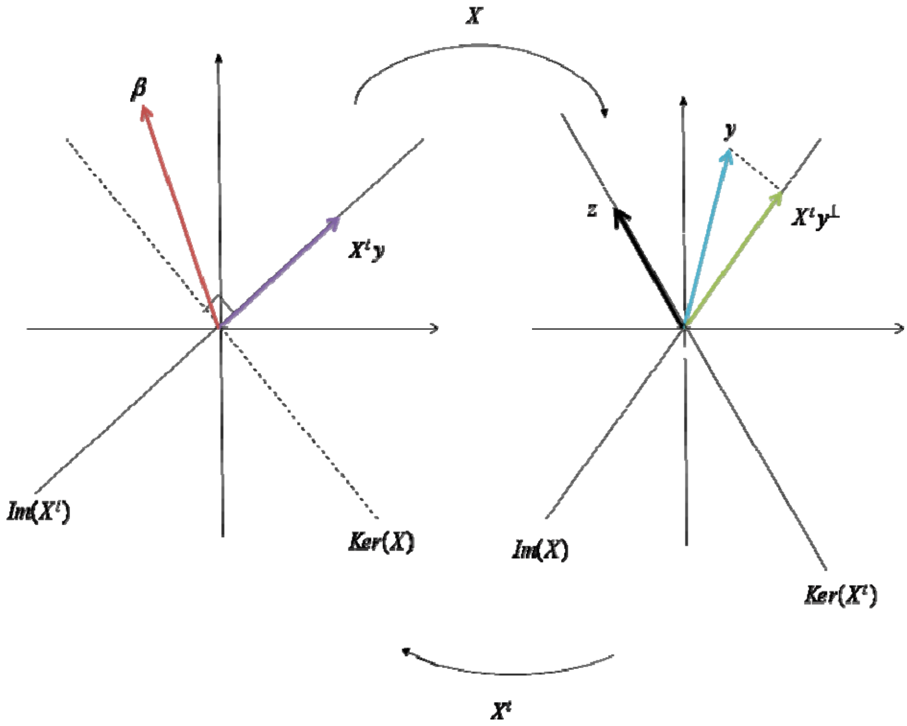

FIGURA 4 Representação geométrica das equações normais

Seja $y^{\perp}$ a projeção ortogonal do vetor de dados y na $Im(X)$. Sendo assim, existe $\beta \in \mathbb{R}^n$, tal que $X\beta = y^{\perp}$. Como $X^t y^{\perp} = X^t y$ segue que $X^t X \beta = X^t y^{\perp} = X^t y$.

2.10.1 Interpretação geométrica da soma de quadrados

Tem-se que a inversa generalizada de Moore-Penrose pode ser definida como $XX^{+} = P$. Se $\hat{y} = Py$, isso garante que o vetor de observações y pode ser decomposto na soma de quadrados de dois vetores, isto é, o vetor y do espaço coluna da matriz do delineamento X e o vetor de erros $\hat{e} = y - \hat{y}$, do complemento ortogonal do espaço coluna da matriz X, $C(X)$. Este fato está ilustrado na Figura 5, a qual mostra o espaço coluna $C(X)$ da matriz X representando o planejamento experimental utilizado, o vetor y dos valores observados, a decomposição ortogonal do vetor de observações.

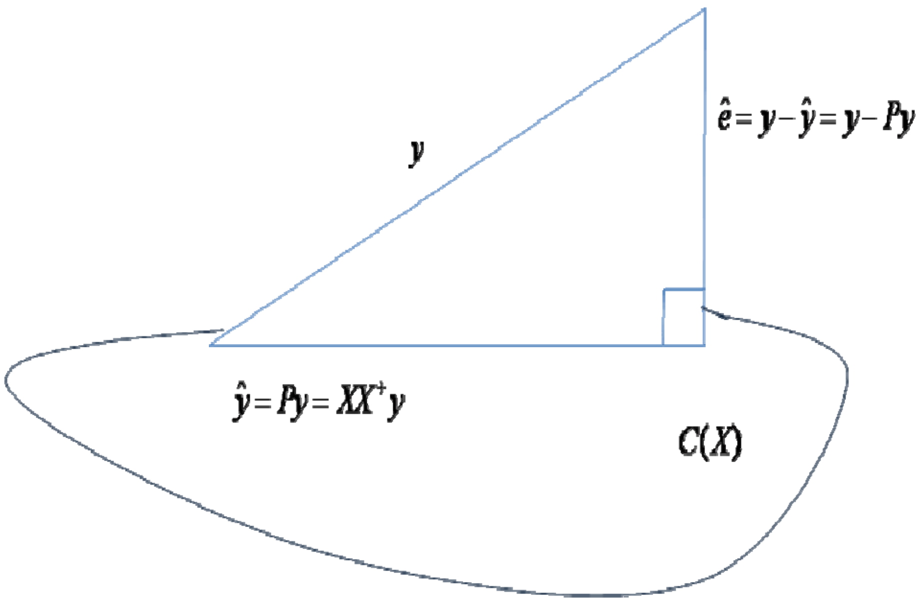

FIGURA 5 Interpretação geométrica da decomposição ortogonal do vetor de observações y nos vetores $\hat{y}$ dos parâmetros e $\hat{e}$ dos erros.

A decomposição como ilustrada na Figura 5 pode ser usada como base para aplicação do teorema de Pitágoras. Aplicando o conceito de norma (comprimento de um vetor qualquer), obtêm-se

$$\|y\|^2 = \|\hat{y}\|^2 + \|\hat{e}\|^2 = \left\|XX^+y\right\|^2 + \left\|y - XX^+y\right\|^2$$

$$y^t y = (XX^+y)^t (XX^+y) + (y - XX^+y)^t (y - XX^+y)$$

$$= y^t (X^+)^t (X^t) XX^+ y + y^t y - y^t XX^+ y - y^t (X^+)^t (X^t) + y^t (X^+)^t XX^+ y$$

$$= y^t XX^+ XX^+ y + y^t y - y^t y - y^t XX^+ y + yXX^+ XX^+ y$$

$$y^t y = y^t Py \; - \; y^t (I - P) y\,, \tag{2.21}$$

em que (2.21) representa a decomposição da soma de quadrados total ($y^t y$) do modelo linear na soma de quadrados de parâmetros $(y^t Py)$ e na soma de quadrados do erro ou resíduo $(y^t (I - P) y)$.

2.10.2 Métodos para o cálculo de $(X^t X)^-$

Como $X^t X$ é uma matriz simétrica, utilizando o teorema de diagonalização de matrizes simétricas (Teorema 6), tem-se que os autovetores de $X^t X$ formam uma base de $\mathbb{R}^n$. Sejam $\lambda_1, \ldots, \lambda_r$ os autovalores não nulos de $X^t X$ com correspondentes autovetores $v_1, \ldots, v_r$, logo, $(X^t X)v_i = \lambda_i v_i$.

Sem perda de generalidade podem-se tomar os autovetores como ortonormais, ou seja, $\langle v_i, v_j \rangle = 0,$ para todo $i \neq j$ e $\|v_i\| = 1$. Portanto, se tem que se v é um vetor qualquer de $\mathbb{R}^n$ então $v = \sum a_i v_i + w$, onde $w \in Ker(X^t X)$. Logo,

$$\begin{aligned} (X^t X)v &= X^t X(\sum a_i v_i + w) = X^t X(\sum a_i v_i) \\ &= \sum a_i X^t X v_i = \sum a_i \lambda_i v_i. \end{aligned}$$

Portanto, a matriz $X^t X$ pode ser decomposta em $X^t X = \lambda_1 v_1 v_1^t + \ldots + \lambda_r v_r v_r^t$, pois,

$$(\lambda_1 v_1 v_1^t + \ldots + \lambda_r v_r v_r^t)(\sum a_i v_i) = \sum_j \sum_i a_i \lambda_j v_j v_j^t v_i = \sum a_i \lambda_i v_j = (X^t X)v.$$

Proposição 7 Uma inversa generalizada de $X^t X$ é dada por:

$$(X^t X)^- = \frac{1}{\lambda_1} v_1 v_1^t + \ldots + \frac{1}{\lambda_r} v_r v_r^t.$$

Prova: $(X^t X) = (X^t X)(X^t X)^-(X^t X)$

$$\begin{aligned} &= (\sum \lambda_i v_i v_i^t)(\sum \frac{1}{\lambda_i} v_i v_i^t)(\sum \lambda_i v_i v_i^t) \\ &= I(\sum \lambda_i v_i v_i^t) = X^t X. \end{aligned}$$

□

Teorema 16 (Graybill, 1976, p. 32) Para qualquer matriz $X_{m\times n}$, a matriz $X(X^tX)^-X^t$ é a mesma, qualquer que seja a inversa generalizada $(X^tX)^-$.

Prova: Sejam M e N duas inversas generalizadas quaisquer de X^tX. Então $X^tXMX^tX = X^tXNX^tX$. Seja $X = B_E^{-1}C_D^{-1}$ uma fatoração de posto completo de X.

$$(C_D^{-1})^t(B_E^{-1})^tB_E^{-1}C_D^{-1}M(C_D^{-1})^t(B_E^{-1})^tB_E^{-1}C_D^{-1} =$$
$$=(C_D^{-1})^t(B_E^{-1})^tB_E^{-1}C_D^{-1}N(C_D^{-1})^t(B_E^{-1})^tB_E^{-1}C_D^{-1}.$$

Multiplicando ambos os lados da equação à esquerda por $(C_D^{-1})^-$ e à direita por C_D^{-1} e sabendo que $(C_D^{-1})^-(C_D^{-1}) = I$ e $C_D^{-1}(C_D^{-1})^- = I$ (fatoração de posto completo). O resultado será

$$(B_E^{-1})^tB_E^{-1}C_D^{-1}M(C_D^{-1})^t(B_E^{-1})^tB_E^{-1} = (B_E^{-1})^tB_E^{-1}C_D^{-1}N(C_D^{-1})^t(B_E^{-1})^tB_E^{-1}.$$

Multiplicando agora ambos os lados da equação na esquerda e direita por $((B_E^{-1})^tB_E^{-1})^{-1}$, o resultado desta operação será

$$C_D^{-1}M(C_D^{-1})^t = C_D^{-1}N(C_D^{-1})^t.$$

Agora, multiplicando ambos os lados da equação à esquerda por B_E^{-1} e à direita por $(B_E^{-1})^t$, o resultado fica

$$B_E^{-1}C_D^{-1}M(C_D^{-1})^t(B_E^{-1})^t = B_E^{-1}C_D^{-1}N(C_D^{-1})^t(B_E^{-1})^t,$$

simplificando essa expressão encontra-se

$$XMX^t = XNX^t.$$

Como M e N são inversas generalizadas arbitrárias de X^tX, o resultado é invariante. $\square$

Corolário 3 $X(X^tX)^-X^t = XX^+$, em que $(X^tX)^-$ é qualquer inversa generalizada.

Prova: Tomando $(X^tX)^- = (X^tX)^+$ tem-se que

$$X(X^tX)^-X^t = X(X^tX)^+X^t = X(X^+(X^t)^+)X^t$$
$$= XX^+XX^+ = XX^+ \qquad \square$$

Corolário 4 X^+y é solução das equações normais $X^tX\hat{\beta} = X^ty$.

Prova: As soluções das equações normais são da forma $\hat{\beta} = (X^tX)^-X^ty.$

Sabendo-se que $X(X^tX)^-X^t = XX^+$, então

$$X\hat{\beta} = X(X^tX)^-X^ty = X(X^tX)^+X^ty = XX^+y. \qquad \square$$

Proposição 8 As soluções para as equações normais $\hat{\beta} = (X^tX)^+X^ty$ e para $Y = \boldsymbol{X}\beta \Rightarrow \hat{\beta} = X^+y$ são iguais.

Prova:

$$(X^tX)^+X^ty = X^+(X^t)^+X^ty = X^+(X^+)^tX^ty$$
$$= X^+(XX^+)^ty = X^+XX^+y = X^+y. \qquad \square$$

A solução geral para o sistema de equações normais pode ser escrita como

$$\hat{\beta} = (X^tX)^-X^ty + \{I_m - (X^tX)^-X^tX\}h, \tag{2.22}$$

em que h é um vetor arbitrário $m \times 1$, ou utilizando a inversa generalizada de Moore-Penrose, como

$$\hat{\beta} = (X^tX)^+X^ty + \{I - (X^tX)^+X^tX\}h$$
$$= X^+y + \{I - X^+(X^t)^+X^tX\}h = X^+y + \{I - X^+(X^+)^tX^tX\}h$$

$$= X^{+}y + \{I - X^{+}(XX^{+})^{t}X\}h = X^{+}y + \{I - X^{+}XX^{+}X\}h =$$

$$\hat{\beta} = X^{+}y + \{I - X^{+}X\}h.$$

Exemplo 5 Suponha que sete tomateiros são tratados com três diferentes fertilizantes: três plantas receberam um tratamento, duas plantas receberam outro, e duas receberam um terceiro. A Tabela 1 mostra a altura das sete plantas, oito semanas depois de plantadas, com sementes de mesma idade.

TABELA 1 Altura das plantas de tomates (em polegadas) em três diferentes fertilizantes.

	Tratamentos	
1	**2**	**3**
74	76	87
68	80	91
77		

FONTE: Searle – 1987

Considerando y_{ij} a j – ésima observação do i – ésimo tratamento, com o modelo:

$$y_{ij} = \beta_i + \varepsilon_{ij},$$

em β_i representa o efeito do i – ésimo tratamento e ε_{ij} é o erro aleatório associado à observação y_{ij}. As observações do experimento podem ser representadas pelas equações:

$$\begin{aligned} y_{11} &= \beta_1 + \varepsilon_{11} \\ y_{12} &= \beta_1 + \varepsilon_{12} \\ y_{13} &= \beta_1 + \varepsilon_{13} \\ y_{21} &= \beta_2 + \varepsilon_{21} \\ y_{21} &= \beta_2 + \varepsilon_{21} \\ y_{22} &= \beta_2 + \varepsilon_{22} \\ y_{31} &= \beta_3 + \varepsilon_{31} \\ y_{32} &= \beta_3 + \varepsilon_{32} \end{aligned} \Rightarrow \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{bmatrix} \Rightarrow Y = \boldsymbol{X}\beta + \varepsilon$$

$$\text{Com } Y = \begin{bmatrix} 74 \\ 68 \\ 77 \\ 76 \\ 80 \\ 87 \\ 91 \end{bmatrix}; \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

A estimação dos parâmetros do modelo pode ser encontrada por $\hat{\beta} = (X^t X)^+ X^t y$.

$$\hat{\beta} = \left(\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \right)^+ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 74 \\ 68 \\ 77 \\ 76 \\ 80 \\ 87 \\ 91 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} 73 \\ 78 \\ 89 \end{bmatrix}.$$

Ou de forma equivalente por $\hat{\beta} = X^{+}y$.

$$\hat{\beta} = \left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \right)^{+} \begin{bmatrix} 74 \\ 68 \\ 77 \\ 76 \\ 80 \\ 87 \\ 91 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 74 \\ 68 \\ 77 \\ 76 \\ 80 \\ 87 \\ 91 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} 73 \\ 78 \\ 89 \end{bmatrix}.$$

Exemplo 6 (Alves) Considere o ganho de peso, em quilogramas (kg) de suínos, em função de três tipos de aditivos à ração: t_1 = adição de uréia; t_2 = adição de óleo vegetal (1%); t_3 = adição de óleo vegetal (2%).

TABELA 2 Valores do ganho de peso em diferentes aditivos aplicados à ração de suínos

Tratamentos (Aditivos à Ração)		
1	**2**	**3**
5,0	6,0	9,0
4,0	7,0	8,0
3,0	8,0	10,0
4,0	-	-

Considerando y_{ij} a j – ésima observação do i – ésimo tratamento, com o modelo:

$$y_{ij} = \beta_0 + \beta_i + \varepsilon_{ij},$$

em que β_0 é a constante geral do modelo (média). Portanto, têm-se as seguintes equações:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} = \begin{bmatrix} 5,0 \\ 4,0 \\ 3,0 \\ 4,0 \\ 6,0 \\ 7,0 \\ 8,0 \\ 9,0 \\ 8,0 \\ 10,0 \end{bmatrix} = \begin{bmatrix} 1 & \vdots & 1 & 0 & 0 \\ 1 & \vdots & 1 & 0 & 0 \\ 1 & \vdots & 1 & 0 & 0 \\ 1 & \vdots & 1 & 0 & 0 \\ 1 & \vdots & 0 & 1 & 0 \\ 1 & \vdots & 0 & 1 & 0 \\ 1 & \vdots & 0 & 1 & 0 \\ 1 & \vdots & 0 & 0 & 1 \\ 1 & \vdots & 0 & 0 & 1 \\ 1 & \vdots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \\ \varepsilon_{31} \\ \varepsilon_{32} \\ \varepsilon_{33} \end{bmatrix} \Rightarrow Y = \boldsymbol{X}\beta + \varepsilon$$

A estimação dos parâmetros do modelo pode ser encontrada através de $\hat{\beta} = (X^t X)^+ X^t y.$

$$\hat{\beta} = \begin{bmatrix} 5 \\ -1 \\ 2 \\ 4 \end{bmatrix}.$$

Utilizando-se a expressão $\hat{\beta} = X^+ y$ chega-se no mesmo resultado, isto é,

$$\hat{\beta} = \begin{bmatrix} 5 \\ -1 \\ 2 \\ 4 \end{bmatrix}.$$

2.10.3 Melhor solução aproximada

Dado um sistema inconsistente $Ax = y$, e uma solução aproximada x_a, desse sistema, define-se o erro dessa solução como sendo o vetor $e(x_a) = y - Ax_a$.

O vetor x_+ é definido como melhor solução aproximada ou solução de norma mínima se atender duas condições:

i) $\| e(x_+) \|^2 \leq \| e(x_a) \|^2$, em que x_a é qualquer outra solução aproximada.

ii) Se $\| e(x_+) \|^2 = \| e(x_a) \|^2$ então $\| x_+ \|^2 < \| x_a \|^2$

Teorema 17 A melhor solução aproximada do sistema inconsistente $Ax = y$ é dada por $x_+ = A^+ y$, em que A^+ é a inversa de Moore-Penrose da matriz A.

Prova: Deve-se mostrar que, se $x_+ = A^+ y$, então $\| e(x_+) \|^2 \leq \| e(x) \|^2$, $\forall x$.

Sejam $x_+ = A^+ y$ e x outra solução qualquer de $Ax = y$ inconsistente. Então:

i) $e(x_+) = y - Ax_+ = y - AA^+ y = (I - AA^+)\, y$, em que $I - AA^+$ é uma matriz simétrica e idempotente. Portanto,

$$\| e(x_+) \|^2 = y^t (I - AA^+) y = y^t (I - AA^+)(I - AA^+) y.$$

Sendo $e(x) = y - Ax.$ Somando e subtraindo $AA^+ y$, tem-se

$$e(x) = y - AA^+ y + AA^+ y - Ax = (I - AA^+) y + A(A^+ y - x)$$

logo, $e(x) = e(x_+) + A(A^+ y - x),$ então a norma ao quadrado de $e(x)$ é

$$\| e(x) \|^2 = \left[e(x_+) + A(A^+ y - x) \right]^t \left[e(x_+) + A(A^+ y - x) \right]$$

$$=\left[e(x_+)^t + (x_+ - x)^t A^t\right]\left[e(x_+) + A(x_+ - x)\right]$$

$$= e(x_+)^t e(x_+) + e(x_+)^t A(x_+ - x) + (x_+ - x)^t A^t e(x_+)$$

$$+ (x_+ - x) A^t A(x_+ - x) .$$

Como

$$e(x_+)^t A(x_+ - x) = (y - Ax_+)^t A(x_+ - x) =$$

$$= (y - A A^+ y)^t A(A^+ y - x) = (y^t - y^t A A^+)(AA^+ y - Ax)$$

$$= y^t A A^+ y - y^t Ax - y^t AA^+ AA^+ y + y^t AA^+ Ax$$

$$= y^t A A^+ y - y^t Ax - y^t AA^+ y + y^t Ax = 0,$$

então,

$$\| e(x) \|^2 = e(x_+)^t e(x_+) + (x_+ - x)^t A^t A(x_+ - x)$$

$$= \| e(x_+) \|^2 + \| A(x_+ - x) \|^2 \geq \| e(x_+) \|^2$$

Sabendo-se que $A^t A$ é uma matriz não negativa, então $(x_+ - x)^t A^t A(x_+ - x)$ é uma forma quadrática não negativa. Lembrando que uma forma quadrática é uma função do tipo $Q(x) = x^t Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ em que os elementos a_{ij} são constantes reais e x um vetor de variáveis aleatórias ou valores desconhecidos. Portanto, voltando ao problema anterior tem-se que $\| e(x) \|^2 \geq \| e(x_+) \|, \forall x$.

Se $x_+ = A^+ y$ então $\| x_+ \|^2 < \| x \|^2$, $\forall x \neq x_+$.

Seja $x = A^+ y + (I - AA^+)h$ uma solução aproximada qualquer de $Ax = y$ inconsistente e diferente de $x_+ = A^+ y$. Sendo $x_+ \neq x$ e que o sistema $Ax = y$ é inconsistente então $(I - AA^+)h \neq \varnothing$. Desta forma

$$\| x \|^2 = \| A^+ y + (I - AA^+)h \|^2 = \left[(A^+ y)^t + h^t (I - AA^+) \right][A^+ y + (I - AA^+)h \big] =$$

$$= y^t (A^+)^t A^+ y + y^t (A^+)^t (I - AA^+)h + h^t (I - AA^+) A^+ y +$$

$$+ h^t (I - AA^+)h] =$$

mas, $y^t (A^+)^t (I - AA^+)h = y^t (A^+)^t - y^t (A^+)^t AA^+ h = y^t (A^+)^t h -$

$$- y^t (A^+)^t A^t (A^+)^t h = y^t (A^+)^t h - y^t (A^+)^t h = 0.$$

Então,

$$\| \mathrm{x} \|^2 = y^t (A^+)^t A^+ y + h^t (I - AA^+)h = \| \mathrm{x}_+ \|^2 + h^t (I - AA^+)h.$$

Sabendo que $h^t (I - AA^+)h$ é uma forma quadrática não negativa, tem-se $\| x \|^2 \geq \| x_+ \|^2$.

No entanto, por hipótese $(I - AA^+)h \neq \varnothing$, e assim, $h^t (I - AA^+)h$ é estritamente maior do que zero consequentemente $\| x \|^2 > \| x_+ \|^2$.

Desse modo, $x_+ = A^+ y$ é a melhor solução aproximada de $Ax = y$ inconsistente, e a unicidade de A^+ garante a unicidade de $x_+ = A^+ y$. □

3 METODOLOGIA

Será apresentada uma construção geométrica da inversa de Moore-Penrose, que servirá de base para a interpretação geométrica das outras inversas generalizadas apresentadas no texto. O que o trabalho apresenta de novo é, principalmente, a geometria das inversas generalizadas de mínimos quadrados e das inversas reflexivas, além do apelo geométrico para projetores e estimadores de mínimos quadrados.

Seja a transformação linear $A:\mathbb{R}^n \to \mathbb{R}^m$. A inversa de Moore-Penrose, também conhecida como pseudo-inversa de A, será a correspondência $A^+:\mathbb{R}^m \to \mathbb{R}^n$ que associa a cada vetor de observações $y \in \mathbb{R}^m$ o vetor $A^+ y = x \in \mathbb{R}^n$ de menor norma possível dentre os que tornam mínima as distâncias $|y - Ax|$.

Se y não pertencer à imagem de A, o sistema linear não terá solução. Desta forma, a ideia é procurar em $\mathbb{R}^n$ um vetor x tal que Ax esteja o mais próximo possível de y e, dentre esses vetores x, aquele que possua a menor norma. Dado um vetor $y \in \mathbb{R}^m$, o vetor na $Im(A)$ mais próximo de y será a projeção ortogonal de y, representado por $y^\perp$, sobre $Im(A)$. Portanto $y^\perp \in Im(A)$ e $y - y^\perp \in Ker(A^t)$, pois é perpendicular a todos os outros vetores da $Im(A)$. Uma vez que $y^\perp \in Im(A)$ existe um vetor $x \in \mathbb{R}^n$, tal que $Ax = y^\perp$. Pela Proposição 2, há uma infinidade de outros vetores da forma $x + z$ com $A(x+z) = y^\perp$, em que $z \in Ker(A)$. Dentre esses vetores $x + z$, o de menor norma é $x - x^\perp$ em que $x^\perp$ é a projeção ortogonal de x sobre o

$Ker(A)$. O vetor $x - x^{\perp}$ é perpendicular ao $Ker(A)$ e, portanto, $x - x^{\perp} \in Im(A^t)$. Por conseguinte, $A^+ y = x - x^{\perp}$. O vetor $A^+ y \in \mathrm{Im}(A^t)$ é ortogonal ao $Ker(A)$ e é o único vetor da $Im(A^t)$ tal que $AA^+ y = y^{\perp}$. Note que A restrita a $Im(A^t)$ é injetiva, uma vez que $Im(A^t) \cap Ker(A) = \{0\}$. Toda essa construção está descrita geometricamente na Figura 6.

Apesar de ser intuitivamente claro que A^+ é uma transformação linear, tal fato tem que ser provado e a demonstração pode ser vista em Lima (2006), p. 204.

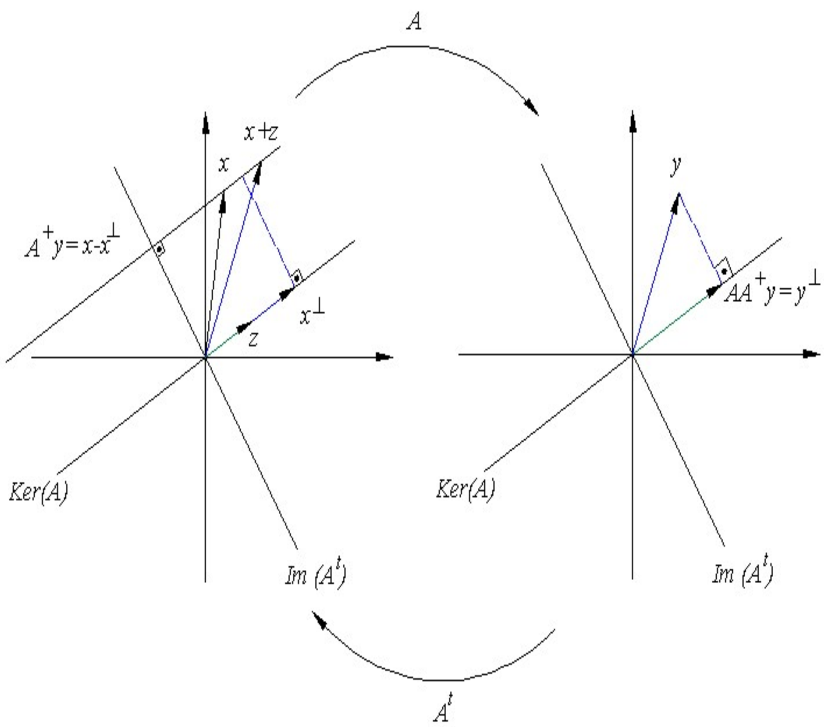

FIGURA 6 Configuração geométrica da inversa generalizada de Moore – Penrose.

4 RESULTADO E DISCUSSÃO

Nesta seção serão apresentadas algumas demonstrações anteriormente obtidas por propriedades algébricas, utilizando apenas propriedades geométricas da inversa de Moore-Penrose. Em algumas delas verifica-se que a abordagem geométrica simplifica a demonstração, além de deixar claro o significado essencial dessas propriedades. Em outras, o método geométrico é claramente outra forma que ajuda o leitor na compreensão da demonstração algébrica.

4.1 Teoria geométrica da inversa generalizada de Moore – Penrose

Proposição 9 A transformação linear $A^{+}:\mathbb{R}^{m}\to\mathbb{R}^{n}$, definida geometricamente na metodologia, é a inversa de Moore-Penrose.

Prova: Basta provar as quatro identidades que definem a inversa de Moore-Penrose.

$i)\;AA^{+}A=A$

Como $Ax=y\in Im(A)\Rightarrow y=y^{\perp}$ e, portanto,

$$AA^{+}Ax=AA^{+}y=A(x-x^{\perp})=Ax-Ax^{\perp}=Ax,$$

visto que $Ax^{\perp}\in Ker(A)$ sendo, portanto, nulo. Na Figura 7 pode-se constatar a situação configurada acima.

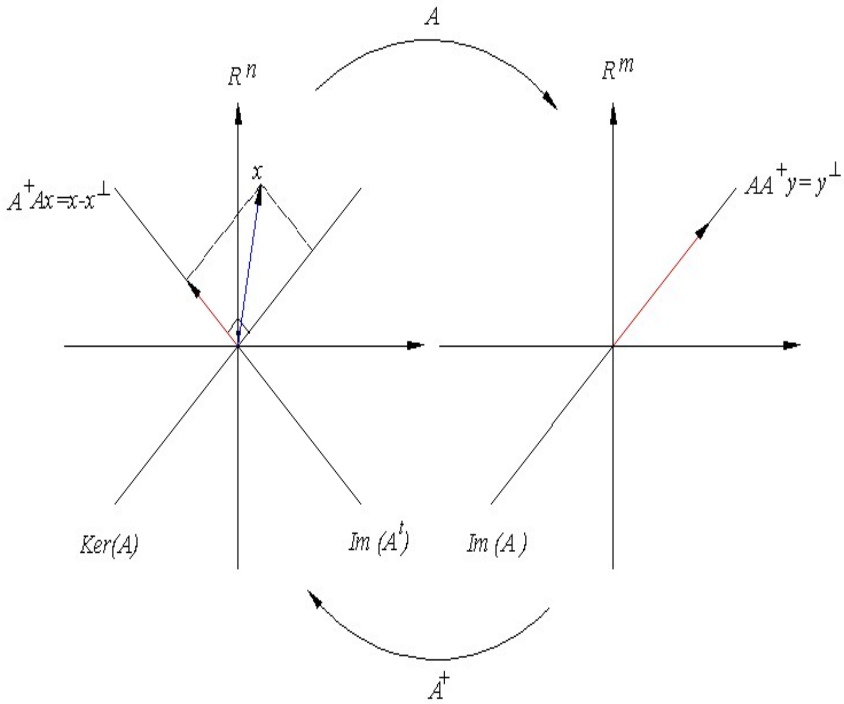

FIGURA 7 Visualização geométrica de $AA^+A = A$

$ii)$ $A^+AA^+ = A^+$

$$A^+AA^+y = A^+A(A^+y) = A^+A(x - x^\perp) =$$

$$= A^+(Ax - Ax^\perp) = A^+(Ax) \;= A^+y^\perp = A^+y\,.$$

A interpretação geométrica deste fato encontra-se na Figura 8.

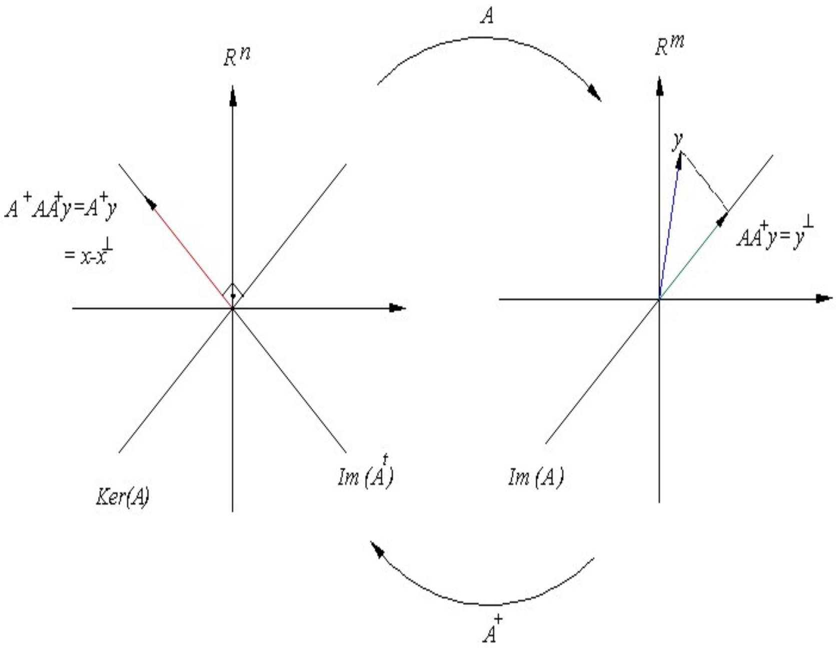

FIGURA 8 Visualização geométrica de $A^+AA^+ = A^+$

$iii)\ (A^+A)^t = A^+A$

Em termos de transformações lineares, tal fato é equivalente a mostrar que, para todo $x \in \mathbb{R}^n$ e todo $z \in \mathbb{R}^n$, $\langle A^+Ax, z\rangle = \langle x, A^+Az\rangle$.

$$\langle A^+Ax, z\rangle = \langle A^+(Ax), z\rangle = \langle x - x^\perp, z\rangle =$$

$$= \langle x - x^\perp, z_1 + z_2\rangle = \langle x - x^\perp, z_2\rangle,$$

em que $z_1 \in Ker(A)$ e $z_2 \in Im(A^t)$.

$$\langle x, A^+Az\rangle = \langle x, z - z^\perp\rangle = \langle x, z_2\rangle = \langle x - x^\perp + x^\perp, z_2\rangle = \langle x - x^\perp, z_2\rangle,$$

pois $z - z^\perp \in Im(A^t)$ e $x^\perp \perp z_2$.

$iv)$ $(AA^+)^t = AA^+$

A demonstração segue como no caso anterior. □

Corolário 5 Lima (2006) $AA^+ : \mathbb{R}^m \to \mathbb{R}^m$ é o projetor ortogonal sobre $Im(A)$.

Seja $y \in Im(A)$, logo existe $x \in \mathbb{R}^n$ tal que $Ax = y$. Portanto, pela definição de A^+ tem-se que $x = A^+y$ e $AA^+y = Ax = y$, por conseguinte, A^+A restrito à $Im(A)$ é a identidade.

Considere agora $y \in Ker(A^t)$, logo $y^\perp = 0$ e, portanto, $A^+y = A^+0 = 0 \Rightarrow AA^+y = 0.$

Corolário 6 Lima (2006) $A^+A : \mathbb{R}^n \to \mathbb{R}^n$ é a projetor ortogonal sobre $Im(A^t)$.

Seja $x \in Im(A^t)$. Se $y = Ax$ então, pela construção geométrica de Moore-Penrose, $A^+(Ax) = x$ e consequentemente A^+A restrito à $Im(A^t)$ é a identidade. Se $x \in Ker(A)$, $Ax = 0$ e, portanto $A^+Ax = 0$.

Proposição 10 Lima (2006) Um operador linear $P : \mathbb{R}^n \to \mathbb{R}^n$ é um projetor ortogonal se e somente se $P = P^+$, em que P^+ é a inversa de Moore-Penrose.

Prova: P é um projetor, isto é, $P^2 = P$ e consequentemente $PPP = P$. Tem-se então, que as quatro condições de Moore − Penrose são satisfeitas, ou seja, $P^+ = P$. □

Como $P^+ = P$ e $PP^+ = PP = P^2 = P$ é um projetor pela propriedade $PP^+PP^+ = PP^+$ e, portanto é simétrico, pois PP^+ é simétrico. Uma demonstração geométrica segue da configuração e está representado na Figura 9.

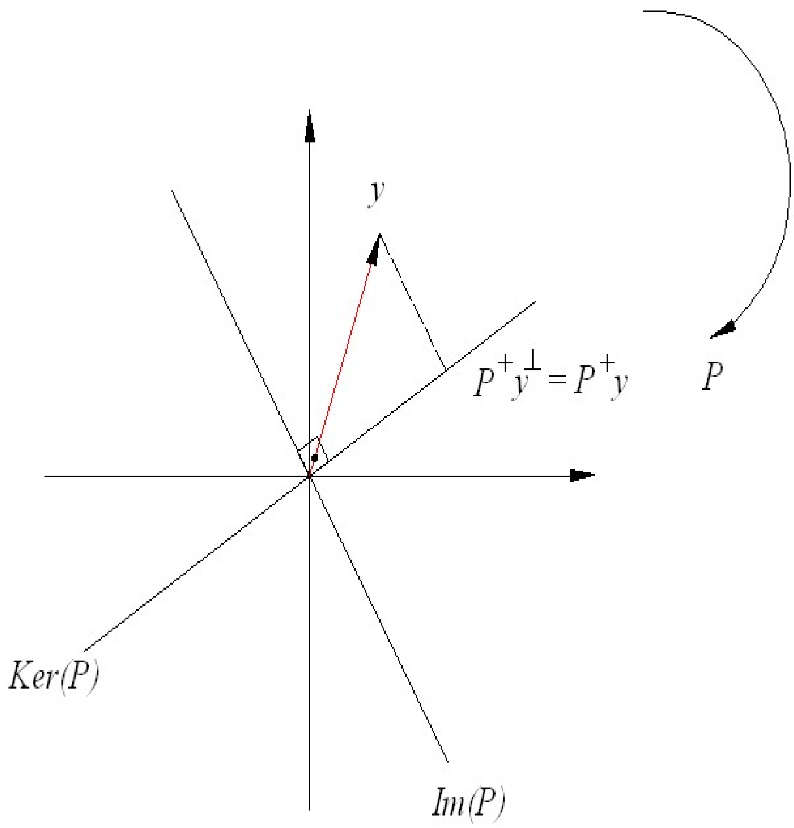

FIGURA 9 Projeção ortogonal P

4.1.1 Demonstração geométrica de algumas propriedades da inversa de Moore-Penrose

Será feito a prova geométrica de duas propriedades (a e d) da inversa generalizada de Moore-Penrose enunciadas no Teorema 10.

a) Se $A_{m\times n} = B_{m\times r}C_{r\times n}$ é a fatoração de posto completo, então $A^+ = C^+B^+$.

Prova: Tem-se $dim(Im(A)) = r$ e $dim(Im(B)) = r$. Além disso, $Im(A) \subset Im(B) \Rightarrow Im(A) = Im(B)$ e sobre a implicação que $Ker(A) \subset Ker(C)$, portanto $dim(Im(C)) = r$, tem-se pelo teorema do núcleo e da imagem que $dim(Ker(C)) + dim(Im(C)) = n$. Logo se tem que

$$dim(Ker(C)) + r = n \Rightarrow dim(Ker(C)) = n - r$$

$$dim(Ker(A)) + dim(Im(A)) = n$$

$$dim(Ker(A)) = n - r \Rightarrow Ker(A) = Ker(C).$$

Se $Ax = y$, então $A^+y = x - x^\perp$. Como B é injetiva, existe um único z_1 que aplicado por B leva em $y^\perp$, ou seja, $Bz_1 = y^\perp$. Logo, $B^+y = z_1$. A transformação linear C é sobrejetiva, logo existe z_2 tal que $z_2 = C^+z_1$.

Projetando-se ortogonalmente z_2 em $Ker(A)$ tem-se que $z_2 - z_2{}^\perp \in Im(A)$ e como $C(z_2 - z_2{}^\perp) = Cz_2 - C(z_2{}^\perp) = z_1$, portanto $C^+z_1 = z_2 - z_2{}^\perp$.

Como $A(z_2 - z_2{}^\perp) = BC(z_2 - z_2{}^\perp) = B(z_1) = y^\perp$. Então,

$$z_2 - z_2{}^\perp = A^+y = C^+(z_1) = C^+B^+(y). \qquad \square$$

A Figura 10 representa a propriedade $A^+ = C^+B^+$.

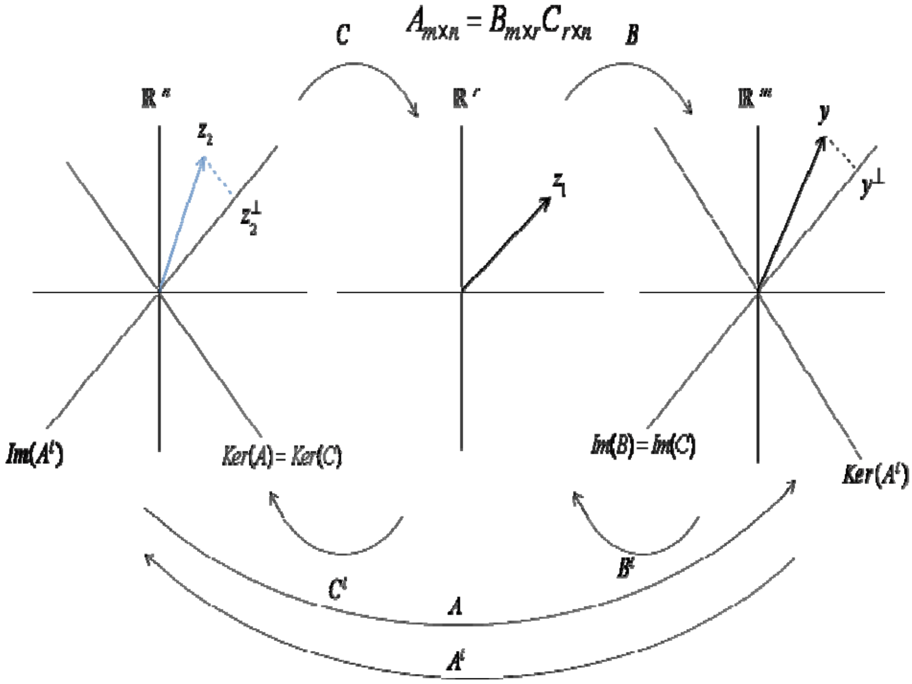

FIGURA 10 Representação geométrica de $A^+ = C^+ B^+$

b) $\quad r(A) = r(A^+) = r(AA^+) = r(A^+A).$

Prova: A dimensão da imagem da matriz produto não pode exceder a dimensão da imagem de quaisquer das matrizes no produto. Logo se tem que:

$$dim(Im(A)) = dim(Im(AA^+A)) \leq dim(Im(AA^+)) \leq dim(Im(A^+)),$$

$$dim(Im(A^+)) = dim(Im(A^+AA^+)) \leq dim(Im(A^+A)) \leq dim(Im(A)). \quad \square$$

4.2 Interpretação geométrica da inversa generalizada reflexiva

A construção geométrica da inversa generalizada de Moore-Penrose permite com poucas modificações uma interpretação geométrica para inversas generalizadas em geral que possuem o mesmo posto da matriz inicial.

A interpretação geométrica para a inversa generalizada reflexiva é baseado no fato de no lugar do subespaço $Ker(A^t)$ toma-se qualquer subespaço V, com $dim(V) = dim(Ker(A^t))$ tal que $\mathbb{R}^m = Im(A) \oplus V$ (soma direta). No lugar da $Im(A^t)$ toma-se um subespaço W qualquer tal que $\mathbb{R}^n = Ker(A) \oplus W$. A inversa generalizada de A aplicada em um vetor y é definida da forma: Projeta-se y em $Im(A)$ paralelamente ao subespaço V, obtendo-se um vetor denotado por $y^{\perp}$. Toma-se um vetor x tal que $Ax = y^{\perp}$. Projeta-se o vetor x em W paralelamente ao subespaço $Ker(A)$ (projeção ortogonal), obtendo-se o vetor $x - x^{\perp}$.

De fato A^r assim definida é uma inversa reflexiva, pois

$$AA^r A(x) = AA^r y = A(x - x^{\perp}) = Ax - Ax^{\perp} = Ax.$$

Portanto, $AA^r A = A$, de onde segue que A^r é uma inversa generalizada.

$$A^r AA^r (y) = A^r A(x - x^{\perp}) = A^r (y^{\perp}) = A^r y.$$

De onde segue que $A^r AA^r = A^r$ e, portanto, reflexiva. A interpretação geométrica da inversa generalizada reflexiva pode ser visualizada na Figura 11.

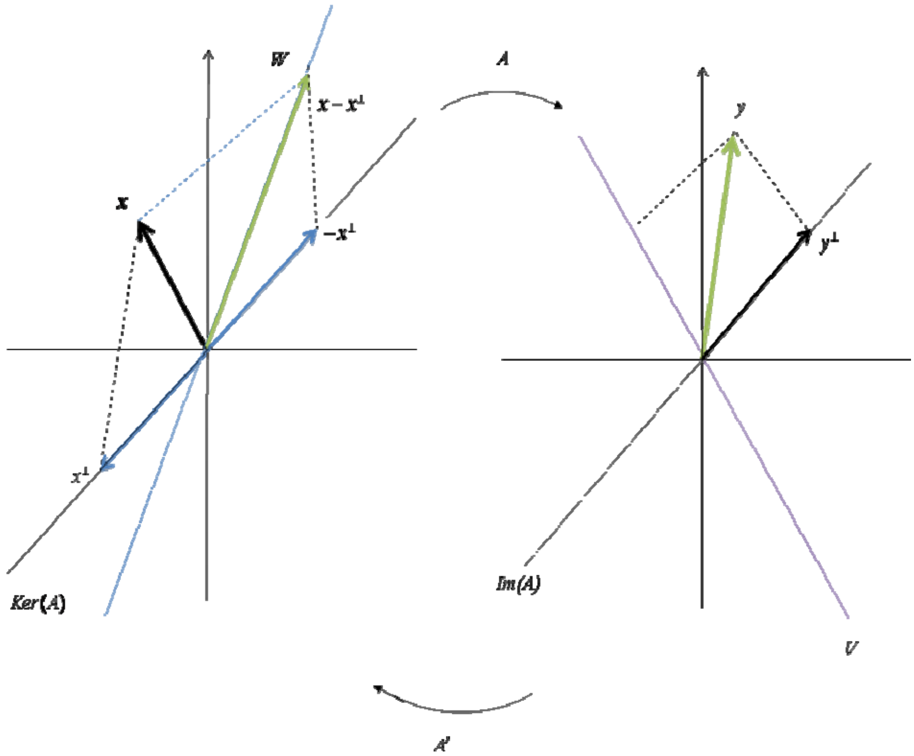

FIGURA 11 Configuração geométrica da inversa reflexiva

Com essa interpretação geométrica é possível dar outra demonstração para o Teorema 16 no qual se tem que matriz $A(A^tA)^-A^t$ é a mesma para qualquer que seja a inversa generalizada de A^tA.

Prova: Encontrar uma inversa generalizada de A^tA é equivalente a se escolher dois subespaços W e V em $\mathbb{R}^n$ onde a $dim(W) = dim(Im(A^t))$ e $dim(V) = dim(Ker(A))$ tais que $(A^tA)^-z$ é obtido da seguinte forma: Projeta-se z paralelamente a V em $Im(A^t)$, obtendo-se $z^\perp$. Toma-se $x \in \mathbb{R}^n$ tal que $A^tA(x) = z^\perp$. Projeta-se x em $Ker(A)$ paralelamente a W obtendo-se $x^\perp$. Assim $(A^tA)^-(z) = x - x^\perp$.

Como $A^t y \in Im(A^t)$ a projeção de $A^t y$ em $Im(A^t)$ é o próprio, desta forma obtém-se $x - x^{\perp}$ tal que $(A^t A)^{-} A^t y = x - x^{\perp}$, logo $A(A^t A)^{-} A^t y = A(x - x^{\perp}) = A(x)$, de onde segue que o valor de $A(A^t A)^{-} A^t$ não depende das projeções $y^{\perp}$ e $x^{\perp}$, portanto, não depende das escolhas dos subespaços V e W, consequentemente não importa a inversa generalizada de $A^t A$ escolhida. $\square$

Uma observação interessante é que se $A^r A$ é simétrica, então é uma projeção ortogonal na $Im(A)$, isto é,

$$A^r A = P_{Im(A)}. \tag{4.1}$$

Do mesmo modo, se AA^r é simétrica então é uma projeção ortogonal $Im(A^t)$, isto é,

$$AA^r = P_{Im(A^t)}. \tag{4.2}$$

Teorema 18 Seja A uma matriz qualquer $m \times n$, então $AB = AA^+$ se e somente se B é tal que $ABA = A$ e AB é simétrica.

Prova: Para se demonstrar algebricamente essa propriedade considera-se que:

Se $ABA = AA^+ A = A$ e $AB = AA^+$ é simétrica.

Considerando agora que se $AB = AA^+ AB = (AA^+)^t (AB)^t$

$$= (A^t)^+ A^t B^t A^t = (A^+)^t (ABA)^t = (A^t)^+ A^t = (AA^+)^t = AA^+. \quad \square$$

Para a demonstração geométrica, tem-se que $ABA = A \Rightarrow ABAB = AB$ e, portanto AB é uma projeção. Como AB é simétrica a projeção é ortogonal. Esta projeção ortogonal é sobre $Im(A)$. Como AA^+ é uma projeção ortogonal sobre $Im(A)$ as duas matrizes devem ser iguais.

4.3 Interpretação geométrica da inversa de quadrados mínimos

A construção é semelhante à anterior, porém agora se toma em $\mathbb{R}^n$ um subespaço W tal que $dim(W) = dim(Ker(A))$ e $\mathbb{R}^n = Im(A^t) \oplus W$ (soma direta). A inversa generalizada A^L em um vetor y é definida da forma: Projeta-se y em $Im(A)$ paralelamente ao subespaço $Ker(A^t)$, obtendo-se o vetor $y^{\perp}$ (projetor ortogonal). Toma-se um vetor x tal que $Ax = y^{\perp}$. Projeta-se o vetor x em W paralelo ao subespaço $Ker(A)$ obtendo-se o vetor $x - x^{\perp}$. A^L assim definida é uma inversa de quadrados mínimos. De fato como $AA^L A = A$ e por construção $AA^L(y) = y^{\perp}$ é a projeção ortogonal de y em $Im(A)$ relativa ao subespaço $Ker(A^t)$. Como projeções ortogonais são simétricas, tem-se que AA^L é uma matriz simétrica. A interpretação geométrica deste fato encontra-se na Figura 12.

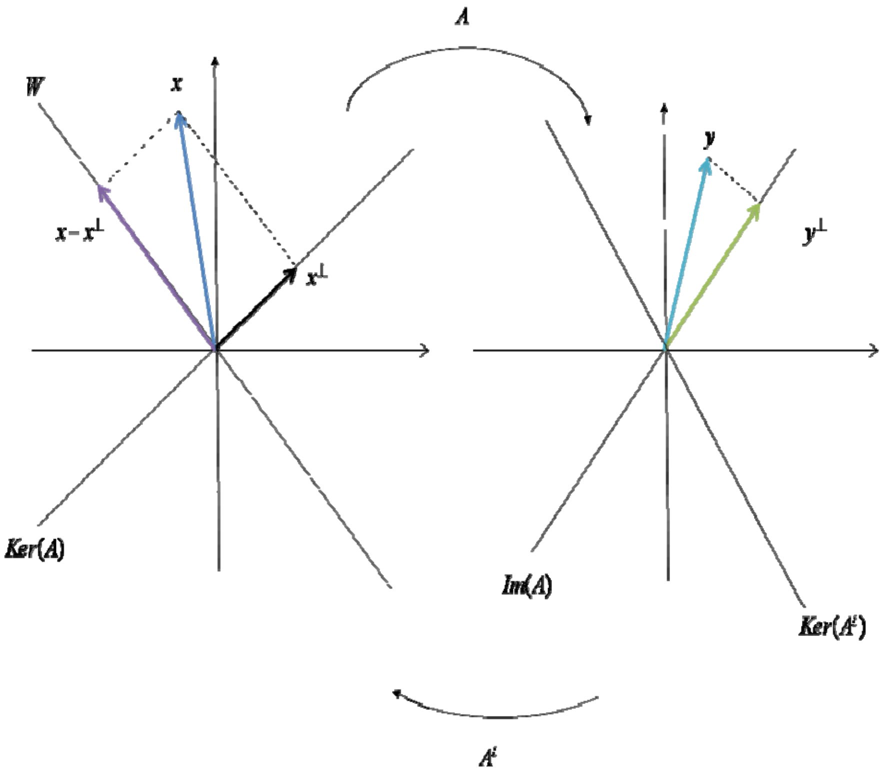

FIGURA 12 Configuração geométrica da inversa de quadrados mínimos

Proposição 11 Se A^L é uma inversa de quadrados mínimos de A, então

$$A^t A A^L = A^t . \qquad (4.3)$$

Prova: Uma demonstração algébrica é:

$$A^t A A^L = A^t (A A^L)^t = A^t (A^L)^t A^t = (A A^L A)^t = A^t$$

Demonstração geométrica:

$$A^t A A^L (y) = A^t (y^{\perp}) = A^t (y), \forall y . \qquad \square$$

5 CONCLUSÕES

1) A abordagem geométrica na teoria das inversas generalizadas é didática e conceitualmente interessante. As propriedades passam a ter um significado intuitivo. Algumas dessas propriedades podem ser facilmente demonstradas.

2) A generalização da abordagem geométrica da inversa de Moore-Penrose às demais inversas generalizadas (inversas reflexivas e de quadrados mínimos) é simples e não acrescenta nenhuma dificuldade à teoria.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, M. F. **Inversa generalizada de Moore-Penrose e soluções de norma mínima em delineamentos experimentais**. 1990. 81p. Dissertação (Mestrado em Estatística Experimentação Agronômica) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

ALVES, M. F.; MATUSO, T.; IEMMA, A. F. Formas gerais para a inversa generalizada de Moore–Penrose em delineamentos experimentais. In: SEAGRO, 2., 1987, Londrina. **Anais...** Londrina: UEL/DMA, 1987. p. 168-176.

ANTON, H. RORRES, C. **Álgebra linear com aplicações.** Porto Alegre: Bookman, 2001. 572p.

BOLDRINI, J. L.; COSTA, S. I. R.; FIGUEIREDO, V. L.; WETZLER, H. G. **Álgebra Linear**. 3. ed. São Paulo: Harper & Row Brasil, 1984. 411 p.

DWIVEDI, T.D. A method to compute the rank factors of matrix. **Sankhyã:** the Indian journal of statistics, Calcutta, v. 37, pt. 4, p. 463-464, 1975.

GRAYBILL, F. A. **An introduction to matrices with applications in the statistics.** New York: Mcgraw Hill, 1961. 463p.

GRAYBILL, F. A. **Theory and application of the linear model**. Boston: Wadsworth Publishing Company, 1976, 704p.

HOFFMAN, K.; KUNZE, R. **Álgebra Linear**. São Paulo: EDUSP/Polígono, 1971.

IEMMA, A. F. **Matrizes para estatística**: um texto para profissionais de ciências aplicadas**.** Piracicaba: ESAUQ/USP, 1988. 339p.

LIMA, E. L. **Álgebra Linear**. 7. ed. Rio de Janeiro: IMPA, 2006. 357p. (Coleção Matemática Universitária).

MARSAGLIA, G.; STYAN, G. P. H. Rank conditions for generalized inverses of partitioned matrices. **Sãnkhyã**: the indian journal of statistics, Calcutta, v. 36, pt. 4, p. 437-442, 1974.

MOORE, E. H. On the reciprocal of the general algebraic matrix. **Bulletin of the American Mathematical Society**, Lancaster, v. 26, n. 8, p.394-395, 1920.

MORAIS, A. R.; CHAVES, L. M.; COSTA, M. C. P. T. **Introdução à álgebra de matrizes**. Lavras: UFLA/FAEPE, 2001. 240p.

PENROSE, R. A generalized inverse for matrices. **Mathematical Proceedings of the Cambridge Philosophical Society**, Cambridge, v. 51, n. 3, p. 406-413, July 1955.

RAO, C. R. **Linear statistical inference and its applications**. 2. ed. New York: J. Wiley, 1973.

RAO, C. R. A note on a generalized inverse of matrix with applications to problems in mathematical statistics. **Journal of the Royal Statistics Society:** series B, methodological, Londres, v. 24, n. 1, p. 152-158, 1962.

RAO, C. R.; MITRA, S. K.**Generalized inverse of matrices and its applications**. New York: J. Wiley, 1971.

SANTOS, R. J. **Álgebra linear e aplicações**. Belo Horizonte: UFMG, 2002. 600p.

SCHOTT, J.R. **Matrix analysis for statistics.** 2. ed. New Jersey: J. Wiley, 2005. 456p.

SEARLE, S. R. **Linear models.** New York: J. Wiley, 1971. 523p.

SEARLE, S. R. **Liner models for unbalanced data.** New York: J. Wiley, 1987. 536p.

SEARLE, S. R. **Matrix algebra useful for statistics.** New York: J. Wiley, 1982. 438p.