UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
INSTITUTO DE GEOCIÊNCIAS E CIÊNCIAS EXATAS

Programa de Pós-Graduação em Matemática em Rede Nacional

A INVERSA GENERALIZADA DE UMA MATRIZ E APLICAÇÕES

José Norberto Reinprecht

Rio Claro - SP
2022

Universidade Estadual Paulista "Júlio de Mesquita Filho"
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

A INVERSA GENERALIZADA DE UMA MATRIZ E APLICAÇÕES

José Norberto Reinprecht

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação – Mestrado Profissional em Matemática em Rede Nacional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Rio Claro.

Orientadora
Profa. Dra. Marta Cilene Gadotti

Rio Claro - SP
2022

R373i

Reinprecht, José Norberto
A inversa generalizada de uma matriz e aplicações / José Norberto Reinprecht. -- Rio Claro, 2022
106 p.

Dissertação (mestrado profissional) - Universidade Estadual Paulista (Unesp), Instituto de Geociências e Ciências Exatas, Rio Claro
Orientadora: Marta Cilene Gadotti

1. Matemática Aplicada. 2. Decomposição em Valores Singulares. 3. Pseudo-inversa de uma matriz. 4. Sistemas Lineares. 5. Métodos dos Mínimos Quadrados. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Geociências e Ciências Exatas, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

TERMO DE APROVAÇÃO

José Norberto Reinprecht

A INVERSA GENERALIZADA DE UMA MATRIZ E APLICAÇÕES

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação – Mestrado Profissional em Matemática em Rede Nacional, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", pela seguinte banca examinadora:

Profa. Dra. Marta Cilene Gadotti
Orientadora

Profa. Dra. Carina Alves Severo
IGCE/UNESP/Rio Claro (SP)

Prof. Dr. Luciano Aparecido Magrini
IFSP/Votuporanga (SP)

Rio Claro, 14 de outubro de 2022

Aos meus pais (em memória).

Agradecimentos

Em primeiro lugar agradecer a minha orientadora, Profa. Dra. Marta Cilene Gadotti, pela sua disponibilidade, apoio e incentivo. Pela paciência e inestimável ajuda durante todo o desenvolvimento deste trabalho.

Ao professor Thiago e ao colega do curso André, pelo auxílio na parte computacional durante a realização deste trabalho.

Aos professores que aceitaram participar da banca examinadora deste trabalho.

Aos professores do Profmat pela dedicação em suas aulas e pelos incentivos na busca de qualificação.

Aos colegas do curso e a todos que, diretamente ou indiretamente, contribuíram para a realização desta dissertação de Mestrado.

E finalmente, a minha esposa pela compreensão da minha ausência em vários momentos durante estes últimos anos.

A todos a minha gratidão.

Se você quer ser bem sucedido, precisa ter dedicação total, buscar seu último limite e dar o melhor de si.

Ayrton Senna.

Resumo

A resolução de sistemas lineares tem aplicação nos mais diversos campos da ciência. Em geral, as resoluções de sistemas lineares envolvendo a inversa de uma matriz estão restritas às matrizes quadradas e não singulares. O objetivo deste trabalho é abordar a busca de soluções de sistemas lineares envolvendo matrizes não quadradas ou mesmo matrizes quadradas mas singulares, através do uso das inversas de Moore-Penrose, também denominadas de pseudo-inversas.

Palavras-chave: Matemática Aplicada. Decomposição em Valores Singulares. Pseudo-inversa . Método dos Mínimos Quadrados. Sistemas Lineares .

Abstract

The resolution of linear systems has application in the most diverse fields of science. In general, the resolutions of linear systems involving the inverse of a matrix are restricted to square and non-singular matrices. The objective of this work is to approach the search for solutions of linear systems involving non-square matrices or even square but singular matrices, through the use of inverse Moore-Penrose matrices, also known as pseudo inverses.

Keywords: Applicated Math. Singular Value Decomposition. Pseudo-inverse. Minimum Squares Method. Linear Sistems.

Sumário

1 Introdução

O objetivo deste trabalho consiste no estudo da inversa de Moore-Penrose ou pseudo-inversa de uma matriz, e algumas aplicações na resolução de sistemas lineares. A escolha deste tema teve por motivação a ausência de tal assunto nos conteúdos dos cursos de graduação do Ensino Superior. Além disso, o tema pode ser explorado no Ensino Médio como está descrito nas considerações finais.

O texto apresentado está desenvolvido de modo objetivo e didático. Dentre as aplicações, o leitor tomará conhecimento da existência: de inversas de matrizes retangulares, de inversas de matrizes quadradas e singulares e de sistemas lineares inconsistentes apresentando solução (uma melhor solução) assuntos não estudados em cursos regulares de graduação.

O conteúdo deste trabalho está organizado na seguinte sequência:

Capítulo 1. Preliminares - apresenta uma revisão de conceitos básicos da Álgebra Linear, indispensáveis ao desenvolvimento deste trabalho. E por se tratar de uma revisão deixaremos de apresentar as provas dos teoremas e proposições. Em algumas situações os conceitos são ilustrados através de exemplos.

Capítulo 2. Teoria Espectral - apresenta o estudo de autovalores e autovetores de um operador linear, suas propriedades e resultados importantes (dando uma particular importância à matriz associada ao operador linear); a sua aplicação na diagonalização de matrizes; o conceito de operador autoadjunto, e finaliza com a demonstração do chamado Teorema Espectral, um importante resultado envolvendo operador autoadjunto.

Capítulo 3. Decomposição em Valores Singulares (DVS) - apresenta de maneira breve e prática a decomposição de qualquer matriz real (trabalho restrito somente a estas matrizes), no produto de uma matriz ortogonal, uma matriz diagonal e uma outra matriz ortogonal. A DVS é de fundamental importância no estudo da inversa de Moore-Penrose de uma matriz.

Capítulo 4. A Inversa Generalizada de uma Matriz - apresenta o estudo da inversa generalizada de uma matriz, e em particular, uma das mais importantes inversas generalizadas, a de Moore-Penrose ou pseudo-inversa. Contém um algoritmo de obtenção de uma inversa generalizada,e também outros dois algoritmos algébricos e iterativos para a determinação da pseudo-inversa de uma matriz.

Capítulo 5. Aplicações da Inversa Generalizada - apresenta o uso da inversa generalizada em situações, tais como: em sistemas de equações lineares $Ax = b$, sendo A uma matriz retangular ou quadrada e singular, na obtenção da solução de um sistema linear, e no cálculo de uma solução de "melhor ajuste" (mínimos quadrados) para um sistema de equações lineares inconsistentes.

Capítulo 6. Considerações Finais - apresenta relatos de alguns pontos relevantes que devem ser destacados.

2 Preliminares

Neste capítulo trazemos as definições e resultados clássicos da Álgebra Linear necessários ao entendimento dos capítulos seguintes. O leitor poderá encontrar as demonstrações dos teoremas e proposições que foram omitidas nas referências [1], [2], [3], [4], [5], [6] e [13] que foram utilizadas para a elaboração do capítulo.

2.1 Um breve relato sobre a Álgebra Linear.

A Álgebra Linear é uma das Àreas mais importantes, versáteis e úteis da matemática, e tem se tornado nos últimos anos uma parte essencial da base matemática de que necessitam matemáticos, engenheiros, físicos e outros cientistas. Na Matemática, sua importância dificilmente pode ser subestimada, quando se compreende que é impossível atacar qualquer problema sem perfeita compreensão dos fenômenos lineares.

A Álgebra Linear surgiu através do estudo detalhado de sistemas de equações lineares algébricas ou diferenciais, e se utiliza de alguns conceitos e estruturas fundamentais da matemática como vetores, espaços vetoriais, sistemas de equações lineares, matrizes e transformações lineares.

A ideia abstrata de Espaços Vetoriais generaliza o conceito de vetores nos espaços bidimensional e tridimensional de duas maneiras. Primeiro, espaços vetoriais pode ter dimensão maior que 3. E segundo, definimos espaços vetoriais não só apenas para vetores do plano ou do espaço mas com diferentes objetos matemáticos, por exemplo números, matrizes, polinômios, funções.

Na sua forma atual, a Álgebra Linear começou a ser estudada em meados do século XIX, evoluiu ao longo de muitos anos e teve a contribuições de várias pessoas. Apesar do crédito ser dado, geralmente, ao matemático alemão H. Grassmann, como sendo o primeiro a introduzir a idéia de Espaço Vetorial em 1862, devemos mencionar também a contribuição do matemático italiano Giuseppe Peano, que em seu livro "Cálculo Geométrico" tornou claro o trabalho de Grassmann e descreveu as propriedades para um Espaço Vetorial da maneira como hoje o conhecemos. [5]

2.2 Espaços vetoriais

Os espaços vetoriais são um dos mais importantes exemplos de estruturas algébricas (nome dado a um conjunto com algumas operações definidas sobre ele satisfazendo certas propriedades) e como a definição envolve a estrutura de um *corpo*, recordemos o significado deste conceito.

Definição 2.1. Um **corpo** consiste de um conjunto não vazio munido de duas operações, denotadas por $+$ e $\cdot$, que satisfazem as propriedades a seguir:

(a) As duas operações são comutativas;
(b) As duas operações são associativas;
(c) Vale a distributiva de $\cdot$ sobre $+$;
(d) Existem elementos neutros 0 para a adição e 1 para a multiplicação;
(e) Todo elemento diferente de 0 tem inverso multiplicativo;

$(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$ e $(\mathbb{C},+,\cdot)$ são exemplos de corpos, pois os conjuntos dos racionais, reais ou complexos com as operações usuais de adição e multiplicacão satisfazem todas as propriedades citadas acima.

A definição de um espaço vetorial envolve um conjunto não vazio, um corpo arbitrário $\mathbb{K}$ (daqui em diante, $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$) e duas operações, denominadas de adição e multiplicação por um escalar, satisfazendo algumas propriedades.

Definição 2.2. Um **espaço vetorial** V sobre o corpo $\mathbb{K}$ é um conjunto não vazio V munido de duas operações:

Adição : " + " $\quad \begin{array}{rcl} V\times V & \to & V. \\ (u,v) & \mapsto & u+v. \end{array}$

Multiplicação : " $\cdot$ " $\quad \begin{array}{rcl} \mathbb{K}\times V & \to & V. \\ (a,v) & \mapsto & av. \end{array}$

Estas operações devem satisfazer os axiomas a seguir:

Adição

A_1 : $u+v=v+u$, $\forall u,v\in V.$

A_2 : $(u+v)+w=u+(v+w)$, $\forall u,v,w\in V.$

A_3 : Existe um vetor em V, denotado por 0, chamado de vetor nulo, tal que qualquer que seja u em V, temos $u+0=u$.

A_4 : Para todo vetor $u\in V$, existe um vetor, denotado por $-u$ tal que $u+(-u)=0$.

Multiplicação por escalar

M_1 : $(\alpha.\beta)u=\alpha.(\beta.u)$, $\forall\alpha,\beta\in\mathbb{K}$ e $u\in V.$

M_2 : $\alpha.(u+v)=\alpha u+\alpha v$, $\forall\alpha\in\mathbb{K}$ e $u,v\in V.$

M_3 : $(\alpha+\beta)u=\alpha u+\beta u$, $\forall\alpha,\beta\in\mathbb{K}$ e $u\in V.$

M_4 : Dado $1\in\mathbb{K}$, então vale $1.u=u$, $\forall u\in V.$

Os elementos do espaço vetorial V são chamados de vetores; e os do corpo $\mathbb{K}$ (podendo ser $\mathbb{R}$ ou $\mathbb{C}$) de escalares.

Quando consideramos o corpo dos escalares como sendo os reais, dizemos que $(V,+,\cdot)$ é um espaço vetorial real; no caso do corpo dos escalares ser os números complexos, dizemos que $(V,+,\cdot)$ é um espaço vetorial complexo.

A seguir apresentaremos alguns de exemplos de espaços vetoriais.

Exemplo 2.3. O conjunto $\mathbb{R}^n$ de todas n-uplas ordenadas de números reais, com as operações usuais de adição e multiplicação por um escalar definidas por:

Adição : $a+b=(a_1,\dots,a_n)+(b_1,\dots,b_n)=(a_1+b_1,\cdots,a_n+b_n)$, $\forall a,b\in\mathbb{R}^n.$

Multiplicação : $\lambda a=\lambda(a_1,\dots,a_n)=(\lambda a_1,\dots,\lambda a_n)$, $\forall a\in\mathbb{R}^n$ e $\forall\lambda\in\mathbb{R}.$

Exemplo 2.4. O conjunto $M_{m\times n}(\mathbb{R})$ das matrizes reais $m \times n$ com as operações usuais de adição de matrizes e produto de um escalar por uma matriz, definidas abaixo:

Adição: $A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$, $\forall A, B \in M_{m\times n}(\mathbb{R})$.

Multiplicação: $k.A = k(a_{ij}) = (ka_{ij})$, $\forall A \in M_{m\times n}(\mathbb{R})$ e $\forall\ k \in \mathbb{R}$.

Exemplo 2.5. O conjunto $P_n(\mathbb{R})$ de todos os polinômios $p(t) = a_0 + a_1t + \cdots + a_nt^n$ com coeficientes reais de grau menor ou igual a n (só igual a n não conteria o elemento neutro da adição representado pelo polinômio nulo), com as operações usuais de adição de polinômios e a multiplicação de um escalar por um polinômio.

É possível que um espaço vetorial esteja contido num outro espaço vetorial. Recordemos como reconhecer dentro de um espaço vetorial V subconjuntos W que sejam também espaços vetoriais.

Definição 2.6. Sejam V um espaço vetorial e W um subconjunto, não vazio, de V. Se W é um espaço vetorial com as operações de adição e multiplicação por escalar definidas em V, então dizemos que W é um **subespaço vetorial** de V.

Em geral, devemos verificar todos axiomas de espaço vetorial para mostrar que um conjunto W com duas operações formam um espaço vetorial. Entretanto, se W é um subconjunto do espaço vetorial V então alguns axiomas não precisam ser verificados, pois estes "herdam" de V. Entretanto, é necessário verificarmos que W é fechado em relação às duas operações definidas, pois é possível que a soma de dois vetores em W, ou a multiplicação de um vetor de W por um algum escalar possa produzir um vetor que não esteja em W.

O teorema a seguir estabelece um critério simples de identificarmos quando um subconjunto $W \subseteq V$ é um subespaço do espaço vetorial V.

Teorema 2.7. *Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um subconjunto W é subespaço de V se, e somente se,*

(a) W é não vazio.

(b) Quaisquer que sejam $u, v \in W$, então $u + v \in W$.

(c) Para todo $u \in W$, temos que $ku \in W$, para todo $k \in \mathbb{K}$.

As condições (b) e (c) garantem que W é fechado em relação as operações de adição e multiplicação por escalar.

Corolário 2.8. *W é um subespaço de V se, e somente se,*

(a) $0 \in W$ (ou seja, $W \neq \phi$).

(b) $v, w \in W$ implica $av + bw \in W$, para todo $a, b \in \mathbb{K}$.

Utilizando o Corolário 1.8 podemos verificar facilmente que os subconjuntos definidos nos exemplos a seguir constituem subespaços vetoriais.

Exemplo 2.9. O conjunto formado somente pelo vetor nulo e o próprio espaço vetorial V são subespaços de V. Portanto, todo espaço vetorial admite pelo menos estes dois subespaços chamados subespaços triviais.

Exemplo 2.10. Os conjuntos $W_S = \{ A \in M_n(\mathbb{R}) : A^t = A \}$ das matrizes reais simétricas de ordem n e $W_{S'} = \{ A \in M_n(\mathbb{R}) : A^t = -A \}$ das matrizes reais antissimétricas de ordem n são subespaços do espaço vetorial $M_n(\mathbb{R})$ das matrizes quadradas reais de ordem n.

Exemplo 2.11. Consideremos o sistema linear homogêneo de m equações e n incógnitas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}.$$

O conjunto W de todas as soluções deste sistema é um subespaço do espaço vetorial do $\mathbb{R}^n$, chamado **espaço das soluções**.

Teorema 2.12. *Se S_1 e S_2 são subespaços vetoriais de um espaço vetorial V, então $S_1 \cap S_2$ é um subespaço de V.*

Corolário 2.13. *Seja V um espaço vetorial. Então, a intersecção de um conjunto arbitrário de subespaços de V é um subespaço vetorial de V.*

Exemplo 2.14. Sejam W_1, W_2 e W_3 subespaços do espaço vetorial $M_n(\mathbb{R})$ das matrizes reais de ordem n, definidos abaixo:

W_1 : o subespaço vetorial das matrizes triangulares superiores de ordem n.

W_2 : o subespaço vetorial das matrizes triangulares inferiores de ordem n.

W_3 : o subespaço vetorial das matrizes diagonais de ordem n.

Verificamos facilmente que: $W_1 \cap W_2 = W_3$.

Como a intersecção de subespaços é sempre um subespaço, seria previsível admitirmos que a união de subespaços também seria um subespaço, mas isto, em geral, não é verdadeiro, como mostra o enunciado do teorema a seguir.

Teorema 2.15. *Sejam S_1 e S_2 subespaços vetoriais de V. Então, $S_1 \cup S_2$ é subespaço vetorial de V se, e somente se, $S_1 \subset S_2$ ou $S_2 \subset S_1$.*

Definição 2.16. Sejam W_1 e W_2 subespaços vetoriais de um espaço vetorial V. Definimos a **soma de W_1 e W_2** como sendo

$$W_1 + W_2 = \{ w_1 + w_2 \ : \ w_1 \in W_1 \ \text{ e } \ w_2 \in W_2 \}.$$

Teorema 2.17. *Sejam W_1 e W_2 subespaços de um espaço vetorial V. Então,*

(a) $W_1 + W_2$ é um subespaço de V.

(b) $W_1 + W_2$ é o menor subespaço de V contendo W_1 e W_2.

(c) $W_1 \cup W_2 \subset W_1 + W_2$.

Definição 2.18. Sejam W_1 e W_2 subespaços vetoriais de um espaço vetorial V. Dizemos que:

(a) $w_1 + w_2$ é **soma direta** quando $W_1 \cap W_2 = \{0\}$, e denotamos por $W_1 \oplus W_2$.

(b) O espaço vetorial V é a soma direta dos subespaços W_1 e W_2 quando $V = W_1 + W_2$ e $W_1 \cap W_2 = \{0\}$.

Teorema 2.19. *Sejam W_1 e W_2 subespaços vetoriais de um espaço vetorial V. Então, $V = W_1 \oplus W_2$ se, e somente se, todo elemento $v \in V$ se escreve de maneira única na forma $v = w_1 + w_2$, com $w_1 \in W_1$ e $w_2 \in W_2$.*

Exemplo 2.20. Sejam $W_1 = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a, b \in \mathbb{R} \right\}$ e $W_2 = \left\{ \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix} : b, d \in \mathbb{R} \right\}$

subespaços do espaço vetorial $M_2(\mathbb{R})$.

Então, $W_1 + W_2 = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} ; a, b, d \in \mathbb{R} \right\}$, ou seja, consiste do conjunto das matrizes cujo elemento da segunda linha e primeira coluna é igual a 0.

Portanto, o espaço vetorial $M_2(\mathbb{R})$ não é soma direta dos subespaços W_1 e W_2.

Exemplo 2.21. O espaço vetorial $\mathbb{R}^3$ é soma direta dos subespaços
$W_1 = \{(a, b, c) \in \mathbb{R}^3 : \ b = c = 0\}$ e $W_2 = \{ \ (a, b, c) \in \mathbb{R}^3 : \ a + b + c = 0 \ \}$.

Exemplo 2.22. O espaço vetorial $M_n(\mathbb{R})$ das matrizes reais de ordem n é a soma direta dos subespaços W_S e $W_{S'}$ das matrizes simétricas e antissimétricas de ordem n, respectivamente.

Exemplo 2.23. Sejam W_1 ,$W_2 \subset M_n(\mathbb{R})$, sendo W_1 o subespaço das matrizes triangulares superiores e W_2 o subespaço das matrizes triangulares inferiores. Então, $M_n(\mathbb{R}) = W_1 + W_2$. Mas, não é uma soma direta, pois $W_1 \cap W_2$ é o subespaço formado pelas matrizes diagonais (veja Exemplo 1.14.). Portanto, $U \cap W \neq \{0\}$.

A seguir recordemos as definições de combinação, dependência e independência linear, usados para construir os conceitos de base e de coordenadas de vetores em diferentes bases.

Definição 2.24. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um vetor $v \in V$ é uma **combinação linear** dos vetores $v_1, v_2, \ \ldots, v_n$ de V, se existirem escalares $a_1, a_2, \ldots, a_n \in \mathbb{K}$, também chamados de coeficientes da combinação linear , tais que

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n \ .$$

Definição 2.25. Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um conjunto, não vazio, de vetores $\{v_1, v_2, \ldots, v_n\} \subset V$ é chamado de **linearmente dependente** sobre $\mathbb{K}$, e denotamos simplificadamente por LD, se existirem $a_1, a_2, \ldots, a_n \in \mathbb{K}$, nem todos nulos, tais que $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$.

Caso contrário, dizemos que o conjunto de vetores é **linearmente independentes**, e denotamos simplificadamente por LI.

As definições acima equivalem a afirmar que: se a equação

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0.$$

admite somente a solução trivial $a_1 = a_2 = \cdots = a_n = 0$, então o conjunto de vetores $v_1, v_2, \ldots, v_n$ é LI. Caso existam outras soluções além da trivial, o conjunto de vetores é LD.

Proposição 2.26.

(a) Todo conjunto finito de vetores de um espaço vetorial que contém o vetor nulo é linearmente dependente.

(b) O conjunto $\{v\}$ formado por um único vetor, não nulo, de um espaço vetorial é linearmente independente.

(c) Um conjunto formado por dois vetores $\{v_1, v_2\}$ de um espaço vetorial é linearmente dependente se, e somente se, um deles é múltiplo escalar do outro.

(d) Se $\{v_1, \dots, v_k\}$ é um conjunto linearmente dependente, então qualquer conjunto finito contendo $v_1, \dots, v_k$ é também LD.

Teorema 2.27. *Seja $S = \{v_1, v_2, \dots, v_n\}$ um conjunto de vetores de um espaço vetorial V, com $n \geq 2$. Então,*

(a) S é LD se, e somente se, pelo menos um dos vetores de S pode ser escrito como uma combinação linear dos demais vetores em S.

(b) Se S é LI, então qualquer subconjunto de S também é LI.

(c) Se S é LI e $S \cup \{v\}$ é LD, então v é uma combinação linear dos vetores de S.

Teorema 2.28. *Seja $S = \{v_1, v_2, \dots, v_n\}$ um subcconjunto finito do espaço vetorial V. O conjunto formado por todas as combinações lineares de S, que denotamos por $\big[S\big]$, é um subespaço de V.*

$$\big[S\big] = \{\ a_1v_1 + a_2v_2 + \dots + a_nv_n\ ;\ a_i \in \mathbb{K}\ \}.$$

Corolário 2.29. *O subespaço $\big[S\big]$ é o menor subespaço de V que contém S.*

Em outras palavras, se W é um outro subespaço de V contendo S, então $\big[S\big] \subset W$.

Definição 2.30. Sejam V um espaço vetorial e $S = \{v_1, v_2, \dots, v_n\}$ um subconjunto de V. O subespaço $\big[S\big]$ das combinações lineares dos vetores do conjunto S é chamado de **subespaço gerado por** S, ou que, $v_1, v_2, \dots, v_n$ geram o subespaço $\big[S\big]$.

Os elementos $v_1, v_2, \dots, v_n$ de S são chamados geradores de $[S]$, ou que S é o conjunto gerador de $[S]$.

Definição 2.31. Um espaço vetorial V é **finitamente gerado** se existir um subconjunto finito $S \subset V$ tal que $\big[S\big] = V$.

Por conveniência definimos que, se $S = \emptyset$, então $\big[\emptyset\big] = \{0\}$.

Exemplo 2.32. O conjunto $S = \{\ (1,0,0)\ ,\ (0,1,0)\ \}$ gera o subespaço

$$W = \{\ (a,b,0) : a, b \in \mathbb{R}\ \} \subset \mathbb{R}^3.$$

Exemplo 2.33. O subespaço $W = \{A \in M_2(\mathbb{R}) : A^t = A\}$ das matrizes simétricas é um subespaço de $M_2(\mathbb{R})$ gerado pelo subconjunto

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}.$$

Teorema 2.34. *Seja* $S = \{ v_1, v_2, \ldots, v_n \}$ *um conjunto de vetores* LI *em um espaço vetorial* V. *Então, cada vetor* $v \in \left[S \right]$ *se escreve de maneira única como combinação linear dos vetores* S.

Teorema 2.35. *Seja* V *um espaço vetorial tal que* $V = \left[v_1, v_2, \ldots, v_n \right]$. *Então, todo conjunto com mais de* n *vetores em* V *é* LD.

Assim, todo conjunto de vetores LI em V possui no máximo n vetores.

2.3 Base, dimensão e coordenadas

O conceito de base de um espaço vetorial consiste em escolhermos um conjunto, com o menor número de vetores, que possam gerar todo o espaço vetorial. Assim, caso retiremos um vetor qualquer desse conjunto, os vetores restantes não geram mais o espaço todo.

Dimensão, usualmente, associamos a algo geométrico; mas, através do conceito de base podemos dar uma definição algébrica para dimensão de um espaço vetorial.

Definição 2.36. Seja $V \neq \emptyset$ um espaço vetorial sobre um corpo $\mathbb{K}$. Uma **base** de V é um subconjunto B de vetores linearmente LI que gera V.

Um espaço vetorial $V \neq \{0_V\}$ sempre possui uma infinidade de bases diferentes, pois se $\{v_1,\ v_2,\ \ldots\ , v_n\}$ é uma base de V, então $\{\alpha v_1,\ \alpha v_2,\ \ldots\ , \alpha v_n\}$ com $\alpha \in \mathbb{K}$ também é uma base de V.

Teorema 2.37. *Seja* $B = \{v_1, v_2, \ldots, v_n\}$ *uma base de um espaço vetorial* V. *Então, todo vetor em* V *pode ser escrito de maneira única como combinação linear de* B.

Teorema 2.38. *Seja* $B = \{v_1, v_2, \ldots, v_n\}$ *uma base de um espaço vetorial* V. *Então, qualquer conjunto com mais de* n *vetores é* LD.

Este teorema garante que qualquer conjunto de $n+1$ ou mais vetores é LD, e portanto, não forma uma base de V. Por outro lado, qualquer conjunto com $n-1$ ou menos vetores não é suficiente para gerar V, e então, também não é uma base de V.

Teorema 2.39. *Se* V *um espaço vetorial gerado pelo conjunto* $S = \{v_1, v_2, \ldots, v_n\}$, *então podemos extrair do conjunto* S *uma base de* V.

Definição 2.40. Seja V um espaço finitamente gerado. Definimos **dimensão** de V, e denotamos por $dim\ V$ ou $dim_K V$ para indicar também o corpo, como sendo o número de vetores de uma base qualquer de V.

No caso de $V = \{ 0 \}$, dizemos que o espaço vetorial V tem dimensão nula. E quando um espaço vetorial V não é finitamente gerado, dizemos que possui dimensão infinita.

Teorema 2.41. *Se* V *é um espaço vetorial de dimensão* n. *Então,*

(a) qualquer base de V *possui* n *elementos.*

(b) todo conjunto de n *vetores* LI *forma uma base de* V.

Teorema 2.42. *(Teorema do complementamento da base) Seja V um espaço vetorial de dimensão n. Se os vetores $u_1, u_2, \ldots, u_k$ são linearmente independentes em V, com $k < n$. Então existem vetores $u_{k+1}, \ldots, u_n$ em V, tais que $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$ formam uma base de V.*

Exemplo 2.43. O conjunto $B = \{\ (1,0,0), (0,1,0,), (0,0,1)\ \}$ é linearmente independente e gera o espaço $\mathbb{R}^3$. Portanto, B é uma base de $\mathbb{R}^3$, denominada **base canônica**, e *dim* $\mathbb{R}^3 = 3$.

Exemplo 2.44. Consideremos o espaço vetorial real $M_2(\mathbb{R})$ das matrizes quadradas reais de ordem 2. O conjunto $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ é linearmente independente, gera $M_2(\mathbb{R})$.

Logo, B é uma base de $M_2(\mathbb{R})$, denominada **canônica**, e *dim* $M_2(\mathbb{R}) = 4$.

Teorema 2.45. *Se W é um subespaço de um espaço vetorial V e dim $V = n$, então dim$W \leq n$.*

No caso da dim $W = n$, então $W = V$.

Teorema 2.46. *Sejam U e W subespaços de um espaço vetorial V de dimensão finita, então*

$$dim\ (U + W) = dim\ U + dim\ V - dim\ (U \cap V).$$

Teorema 2.47. *Sejam V um espaço vetorial de dimensão finita e W um subespaço de V. Então, existe um subespaço U de V tal que $V = U \oplus W$.*

Definição 2.48. Uma **base ordenada** de espaço vetorial V é uma base em que é considerada a ordem dos vetores da base.

Definição 2.49. Sejam V um espaço vetorial finitamente gerado e B uma base ordenada de V formada pelos vetores $v_1, v_2, \ldots, v_n$. Todo vetor em V se escreve de maneira única ($Teorema$ 1.37) como $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$. Os escalares $c_1, c_2, \ldots, c_n$ são chamados **coordenadas** de v em relação à base B, e denotamos por

$$[v]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 \ c_2 \ \ldots \ c_n \end{bmatrix}^T .$$

Dada uma base de um espaço vetorial, uma troca na ordem em que escrevemos esses vetores, acarreta uma troca na ordem das coordenadas, produzindo um vetor de coordenadas diferentes. Então, para evitar tal situação introduzimos a convenção de ordenar a base.

Exemplo 2.50. Consideremos a base $B = \{\ (1,1,1),\ (1,1,0),\ (1,0,0)\ \}$ do $\mathbb{R}^3$. Determinemos as coordenadas do vetor $(2,1,4)$ com relação à base B.

Devemos obter a , b e c tais que

$$(2,1,4) = a(1,1,1) + b(1,1,0) + c(1,0,0) = (a + b + c\ ,\ a + b\ ,\ a).$$

Equivale ao sistema linear $\begin{cases} a+b+c=2 \\ a+b=1 \\ a=4 \end{cases}$ cuja solução é $a=4$, $b=-3$ e $c=1$.

Desse modo, as coordenadas de $(2,1,4)$ com relação à base B são dadas por $(4,-3\,,1)$.

$$\begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}_B \quad ou \quad (2,1,4)^T = (4,-3,1)^T_B.$$

Uma base conveniente para um determinado problema pode não ser conveniente para um outro, de forma que um procedimento comum no estudo de Espaços Vetoriais é a mudança de uma base para outra. Nesta seção recordaremos problemas relativos à mudança de base.

Consideremos $B=\{v_1,v_2,...,v_n\}$ e $B'=\{u_1,u_2,...,u_n\}$ duas bases ordenadas de um mesmo espaço vetorial V. Um vetor v em V pode ser escrito como combinação linear de B e $B^{'}$.

Sejam $v=x_1v_1+x_2v_2+\cdots+b_nx_n$ e $v=y_1u_1+y_2u_2+\cdots+y_nu_n$ as representações de v como combinação linear dos vetores das duas bases B e $B^{'}$, respectivamente .

Então, $$\Big[v\Big]_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad \Big[v\Big]_{B'} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

.

Como os vetores de $B^{'}$ são vetores de V, e B é uma base de V, então os vetores de $B^{'}$ se escrevem como combinação linear de B.

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n. \\ u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n. \\ &\vdots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n. \end{aligned}$$

Assim, $v=y_1u_1+y_2u_2+\cdots+y_nu_n$.

Logo, $v=y_1(a_{11}v_1+\cdots+a_{n1}v_n)+y_2(a_{12}v_1+\cdots+a_{n2}v_n)+\cdots+y_n(a_{1n}v_1+\cdots+a_{nn}v_n)$.

Mas, $v=x_1v_1+x_2v_2+\cdots+b_nx_n$.

Então, $$\begin{cases} x_1 = a_{11}y_1 + a_{12}b_2 + \cdots + a_{1n}b_n \\ x_2 = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \qquad \vdots \\ x_n = a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{cases} \tag{1.1}$$.

Portanto, podemos escrever que

$$\Big[\ v\ \Big]_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \Big[M\Big]_{B\to B'} \cdot \Big[\ v\ \Big]_{B'} \ ,$$

sendo a matriz $\Big[M\Big]_{B\to B'}$ a transposta da matriz dos coeficientes do sistema (1.1).

Definição 2.51. A matriz $\left[M\right]_{B\to B'}$ é denominada **matriz mudança de base** B para a base $B^{'}$. Se B e B' são bases ordenadas de um espaço vetorial V, então para todo v em V temos

$$\left[\ v\ \right]_B = \left[M\right]_{B\to B'} \cdot \left[\ v\ \right]_{B'}.$$

A matriz $\left[\ M\ \right]_{B\to B'}$ podemos expressa-la em termos de vetores colunas, em que as colunas desta matriz são os vetores da base atual em relação à nova base, isto é

$$\left[M\right]_{B\to B'} = \left[\ [\ u_1\ ]_B \quad [\ u_2\ ]_B \quad \cdots \quad [\ u_n\ ]_B\ \right].$$

Exemplo 2.52. Determinemos a matriz mudança da base B para a base $B^{'}$ sendo $B = \{\ u_1 = (1,0)\ ,\ u_2 = (0,1)\ \}$ e $B^{'} = \{\ v_1 = (2,1)\ ,\ v_2 = (3,2)\ \}$ bases de $\mathbb{R}^2$, e também as coordenadas do vetor $v = (1,-2)_B$ na base B em relação à base $B^{'}$.

a) Matriz mudança de base B para $B^{'}$:

$$\left[M\right]_{B\to B'} = \left[\ [u_1]_{B'} \quad [u_2]_{B'}\ \right].$$

Escrevendo os vetores da base B como combinação linear dos vetores de $B^{'}$, temos:

$$(1,0) = a_{11}(2,1) + a_{21}(3,2) = (\ 2a_{11} + 3a_{21}\ ,\ a_{11} + 2a_{21}\ ).$$
$$(0,1) = a_{12}(2,1) + a_{22}(3,2) = (\ 2a_{12} + 3a_{22}\ ,\ a_{12} + 2a_{22}\ ).$$

Resolvendo os sistemas:

$$\begin{cases} 2a_{11} + 3a_{21} = 1 \\ a_{11} + 2a_{21} = 0 \end{cases} \quad ,\ \text{obtemos } a_{11} = 2 \ \text{ e } \ a_{21} = -1.$$

$$\begin{cases} 2a_{12} + 3a_{22} = 0 \\ a_{12} + 2a_{22} = 1 \end{cases} \quad ,\ \text{obtemos } a_{12} = -3 \ \text{ e } \ a_{22} = 2\ .$$

Portanto, $\left[M\right]_{B\to B'} = \left[\ [u_1]_{B'} \quad [u_2]_{B'}\ \right] = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}.$

b) Obtenção das coordenadas do vetor $v = (1,-2)_B$ na base B, para a base $B^{'}$:

$$\left[\ v\ \right]_{B'} = \left[M\right]_{B\to B'} \cdot \left[\ v\ \right]_{B'}.$$

$$\left[\ v\ \right]_{B'} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2+6 \\ -1-4 \end{bmatrix} = \begin{bmatrix} 8 \\ -5 \end{bmatrix}.$$

Portanto, $v_{B'} = (8,-5)^T$.

Teorema 2.53. *Se M é a matriz mudança de base B para uma base $B^{'}$ de um espaço vetorial V de dimensão finita, então M é invertível e M^{-1} é a matriz mudança de base $B^{'}$ para B .*

$$\left[M^{-1}\right]_{B'\to B} = \left[M\right]_{B\to B'}.$$

Teorema 2.54. *Sejam $B^{'} = \{v_1, v_2, \ldots, v_n\}$ e $B = \{e_1, e_2, \ldots, e_n\}$ bases de $\mathbb{R}^n$, sendo B a base canônica . Se os vetores da base $B^{'}$ forem escritos em forma de colunas, então*

$$\left[M\right]_{B\to B'} = \left[\ [v_1]_B \quad [v_2]_B \quad \cdots \quad [v_n]_B\ \right].$$

2.4 Produto interno e ortogonalidade

No estudo da Geometria Analítica no $\mathbb{R}^n$ define-se o comprimento de um vetor, e o ângulo entre vetores através do produto escalar. O objetivo agora é estendermos estes conceitos para os espaços vetoriais sobre um corpo $\mathbb{K}$ (reais ou complexos). Porém devido ao objetivo deste trabalho e que se realizado nos próximos capítulos vamos considerar o corpo dos reais.

O conceito de produto interno vai possibilitar que um espaço vetorial de estrutura totalmente algébrica adquira uma estrutura geométrica, o que permitirá falarmos em ângulo entre vetores, projeção de um vetor ortogonalmente sobre outro, comparar tamanho entre vetores e outros conceitos a espaços vetoriais quaisquer.

Nesta seção, V representará um espaço vetorial de dimensão finita (a menos que seja mencionado o contrário) pois alguns teoremas e proposições que serão mencionados não serem verdadeiros para espaços vetoriais de dimensão infinita. E o corpo envolvido no espaço vetorial V será o corpo dos reais.

Definição 2.55. Seja V um espaço vetorial sobre um corpo $\mathbb{R}$. Uma aplicação

$$\begin{array}{rcl} \langle\ ,\ \rangle : V \times V & \longrightarrow & \mathbb{R}. \\ (u\ ,v) & \longmapsto & \langle u, v\rangle. \end{array}$$

que satisfaz as seguintes propriedades:

(1) *Simetria*: $\langle u,v\rangle = \langle v,u\rangle$, para todo $u, v \in V$.

(2) *Distributividade*: $\langle u+w,v\rangle = \langle u,v\rangle + \langle w,v\rangle$, para todo $u, v, w \in V$.

(3) *Homogeneidade*: $\langle au,v\rangle = a\langle u,v\rangle$, para todo $u, v \in V$ e $a \in \mathbb{R}$.

(4) *Positividade*: $\langle v,v\rangle \geqslant 0$ e $\langle v,v\rangle = 0$ se, e somente se $v = 0$, para todo $v \in V$.

define um **produto interno** sobre V.

Um espaço vetorial sobre o corpo $\mathbb{R}$ com produto interno, denotamos por $(\ V\ ,\langle\cdot\ ,\cdot\rangle\ )$, e é chamado de **espaço euclidiano**.

Observação 2.56. O produto interno em $\mathbb{R}^n$ também é chamada de **produto escalar**.

Propriedade 2.57. Atráves das propriedades de simetria, distributividade e homogeneidade podemos verificar que:

(a) $\langle u,0\rangle = \langle 0,u\rangle = 0$, para todo $u \in V$.

(b) $\langle u,av\rangle = a\langle u,v\rangle$, para todo $u, v \in V$ e a $\in \mathbb{R}$.

(c) $\langle u,v+w\rangle = \langle u,v\rangle + \langle u,w\rangle$, para todo $u, v \in V$.

Exemplo 2.58. Sejam $u = (u_1, u_2, \ldots, u_n)$ e $v = (v_1, v_2, ..., v_n)$ vetores do espaço vetorial $\mathbb{R}^n$. O produto interno dado por

$$\langle u,v\rangle = u_1.v_1 + \cdots + u_n.v_n.$$

é chamado de **produto interno euclidiano**, ou **canônico** do $\mathbb{R}^n$.

Exemplo 2.59. Consideremos o espaço vetorial $\mathbb{R}^2$, os vetores $u = (u_1, u_2)$ e $v = (v_1, v_2)$ em $\mathbb{R}^2$. A aplicação definida por

$$\langle u,v\rangle = 3u_1.v_1 + 2u_2.v_2\ ,\ \text{ para todo } u, v \in \mathbb{R}^2$$

define um produto interno em $\mathbb{R}^2$.

Este produto interno é diferente do produto canônico do $\mathbb{R}^2$. Assim, podemos observar a existência de mais de um produto interno num mesmo espaço vetorial.

Definição 2.60. Definimos **traço de uma matriz** $A = (a_{ij})$ de ordem n, e denotamos por trA, como a soma dos elementos da sua diagonal principal.

Exemplo 2.61. Consideremos $V = M_{m\times n}(\mathbb{R})$. A aplicação definida por
$$\langle A, B\rangle = tr(B^tA), \ \ \forall A, B \in M_n(\mathbb{R}).$$
é um produto interno $M_{m\times n}(\mathbb{R})$.

Definição 2.62. Seja V um espaço euclidiano. A **norma** de um vetor v em relação a esse produto interno é definida por
$$\| v \| = \sqrt{\langle v, v\rangle} \ \geq \ 0 \ , \quad \forall v \in V.$$

Um espaço vetorial euclidiano de uma norma $\| \, . \, \|$ é denominado **espaço normado**.

Quando $\| v \| = 1$ dizemos que v é um **vetor unitário**.

Para todo $v \neq 0$, o vetor $u = \frac{v}{\|v\|} \in V$, e é chamado **vetor normalizado**, e $\| u \| = 1$. Pois,
$$\| u \| = \sqrt{\langle \frac{v}{\| v \|}, \frac{v}{\| v \|} \rangle} = \sqrt{\frac{1}{\| v \|^2} . \| v \|^2} \ = \ 1.$$

Observação 2.63.

(a) A norma depende do produto interno considerado.
(b) É sempre possível extrair a raiz quadrada de $\langle v, v\rangle$ pois este número é não negativo.
(c) A norma de v representa o comprimento do vetor v.

Proposição 2.64. *Em um espaço normado V são válidas as seguintes propriedades:*

(a) $\| v \| = 0$ *se, e somente se,* $v = 0$.
(b) $\| \lambda v \| = |\lambda| \, \| v \| \ , \ \forall \lambda \in \mathbb{K}$ *e* $\forall u, v \in V$.
(c) $\| u + v \| \ \leq \ \| u \| + \| v \| \ , \ \forall u, v \in V$. *(desigualdade triangular).*

Proposição 2.65. *(Desigualdade de Cauchy-Schwarz) Se V é um espaço normado, então,* $| \langle u, v\rangle | \leq \| u \| \, \| v \| \ , \ \forall u, v \in V$.

Exemplo 2.66. Consideremos o espaço vetorial $V = M_{m\times n}(\mathbb{R})$ e o produto interno definido por $\langle A, B\rangle = tr(A^tB) \ , \ \forall A, B \in M_{n\times n}(\mathbb{R})$. Calculemos a norma da matriz
$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 0 & 3 \end{bmatrix}.$$
$$\| A \|^2 = \ tr(A^tA)$$
$$A^t.A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \end{bmatrix} . \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -2 & 10 \end{bmatrix}.$$

Logo, $\| A \|^2 = \ tr(A^tA) \ = \ 5 + 10 = 15$. Portanto $\| A \| = \sqrt{15}$.

Definição 2.67. Seja V um espaço vetorial. Consideremos uma aplicação $d : V \times V \to \mathbb{R}$, que associa a cada par ordenado $(u, v) \in V \times V$ um número real $d(u, v)$ satisfazendo as seguintes condições:

(a) *Positividade* : $d(u, v) \geqslant 0 \quad$ e $\quad d(u, v) = 0 \iff u = v, \quad \forall u, v \in V$.

(b) *Simetria* : $d(u, v) = d(v, u), \ \forall u, v \in V$.

(c) *Desigualdade triangular* : $d(u, v) \leqslant d(u, w) + d(v, w), \ \forall u, v, w \in V$.

Nestas condições, dizemos que d é uma **métrica** sobre V e que $d(u, v)$ é a distância do vetor u ao vetor v.

Um espaço euclidiano munido de uma métrica é chamado de **espaço métrico**.

Observação 2.68. Em um espaço euclidiano V a **métrica** ou **distância** é dada por

$$\text{d(u,v)} = \| u - v \| = \sqrt{\langle u - v \ , u - v \rangle} \ , \quad \forall u, v \in V.$$

Definição 2.69. Em espaço vetorial euclidiano V definimos **ângulo entre dois vetores**, não nulos, u e v em V como sendo o valor $\theta \in [0, \pi)$, tal que

$$\cos\theta = \frac{\langle u,v \rangle}{\|u\|\|v\|}.$$

Quando $\langle u, v \rangle = 0$, ou seja, $cos\ \theta = 0$, o que acarreta $\theta = \frac{\pi}{2}$, dizemos que os vetores u e v são ortogonais.

Definição 2.70. Em um espaço vetorial euclidiano V definimos a **projeção ortogonal** de um vetor u sobre um vetor não nulo v, denotado por $proj_v u$, como sendo o vetor

$$proj_v u = \frac{\langle u \ , \ v \rangle}{\|v\|} v \ , \quad u, v \in V.$$

Observe que a projeção ortogonal de um vetor u sobre um vetor v não nulo é um múltiplo escalar do vetor v.

Teorema 2.71. *Seja $v \in \mathbb{R}^n$ um vetor não nulo. Então, $u - proj_v u$ é ortogonal a v, para qualquer $u \in \mathbb{R}^n$.*

Demonstração. Calculando o produto escalar de v com $u - proj_v u$, temos:

$$\langle v \ , \ u - proj_v u \rangle = \langle v \ , \ u \rangle - \langle v \ , \ proj_v u \rangle = \langle v \ , \ u \rangle - \langle v \ , \ \frac{\langle v \ , \ u \rangle}{\|v\|} v \rangle =$$

$$= \langle v \ , \ u \rangle - \frac{\langle v \ , \ u \rangle}{\|v\|} \langle v, v \rangle = 0.$$

Portanto, $u - proj_v u$ é ortogonal a v.

□

Definição 2.72. Em um espaço euclidiano V dizemos que:

(a) um conjunto $S = \{v_1, v_2, \ldots, v_n\} \subset V$ é **ortogonal** , se $v_i \perp v_j$ para todo $i \neq j$.

(b) um vetor $u \in V$ é ortogonal a um conjunto não vazio $S \subset V$, se u for ortogonal a todos os vetores de S, e denotamos por $u \perp S$.

(c) um conjunto S é chamado de **ortonormal** quando, além de ortogonal, satisfaz $\| v_i \| = 1$, para todo $1 \leqslant i \leqslant n$.

Definição 2.73. Sejam V um espaço euclidiano e S um conjunto não vazio de vetores de V. O conjunto

$$S^{\perp} = \{ \ u \in V : \langle u, v \rangle = 0 \ , \ \forall v \in V \ \}$$

é denominado **ortogonal** de S.

Teorema 2.74. *Seja V um espaço euclidiano. Se $S = \{ v_1, v_2, \ldots, v_n \}$ é um conjunto ortogonal em V, tais que $v_i \neq 0$, para $i = 1, 2, \ldots, n$. Então, S é linearmente independente em V.*

Propriedade 2.75. Num espaço vetorial euclidiano V valem as seguintes propriedades:

(a) $0 \perp v$, para todo $v \in V$.

(b) Se $u \perp v$, então $v \perp u$.

(c) Se $v \perp u$, para todo $u \in V$, então $v = 0$.

(d) Se $u \perp w$ e $v \perp w$, então $(u + v) \perp w$.

(e) Se $u \perp v$ e λ é um escalar, então $u \perp \lambda v$.

Teorema 2.76. *(Teorema de Pitágoras). Sejam V um espaço vetorial euclidiano. Então, os vetores $u, v \in V$ são ortogonais se, e somente se,*

$$\| u + v \|^2 = \| u \|^2 + \| v \|^2.$$

Proposição 2.77. *Um conjunto ortonormal $\{u_1, u_2, \ldots, u_n\}$ é LI se, para qualquer $v \in V$, o vetor $w = v - \langle v, u_1 \rangle u_1 - \langle v, u_2 \rangle u_2 - \cdots - \langle v, u_n \rangle u_n$ é ortogonal a cada um dos u_i.*

As bases ortogonais são de fundamental importância em espaços com produto interno, e o teorema a seguir mostra que sempre é possível obtermos tais bases.

Teorema 2.78. *(Processo de Ortogonalização de Gram-Schmidt). Todo espaço vetorial euclidiano de dimensão finita admite uma base ortonormal.*

Demonstração. A prova é feita por indução sobre a dimensão do espaço.

Seja V um espaço vetorial euclidiano de dimensão finita.

Se $dim\ V = 1$, então existe $v_1 \in V$ tal que $V = \left[\ v_1 \ \right]$.

Como $v_1 \neq 0$, consideremos $u_1 = \frac{v_1}{\|v_1\|}$, e então $\{u_1\}$ é uma base ortonormal de V.

Se $dim\ V = 2$, então existem $v_1, v_2 \in V$, tais que $V = \left[\ v_1, v_2 \ \right]$.

Consideremos $u_1 = \frac{v_1}{\|v_1\|}$ e determinemos um vetor u_2 ortogonal a u_1 e de norma 1.

Pela Proposição 1.77. basta tomarmos o vetor não nulo $u_2' = v_2 - \langle v_2, u_1 \rangle u_1$, (pois v_1 e v_2 são LI) , e normalizá-lo obtendo u_2.

Logo, $\{u_1, u_2\}$ forma uma base ortonormal de V.

Por hipótese de indução, suponhamos que o teorema seja válido para todo espaço euclidiano de dimensão $n - 1$, e vamos provar que o teorema é válido para todo espaço euclidiano de dimensão n.

Se $dim\ V = n \geq 2$, então existem $v_1, v_2, \ldots, v_n \in V$ que forma uma base de V.

Mas, $U = \left[\ v_1, v_2, \ldots, v_n \ \right]$ é um subespaço de dimensão $n - 1$. Então, usando a hipótese de indução, é possível tomarmos uma base $\{u_1, u_2, \ldots, u_{n-1}\}$ ortonormal de U.

Como $v_n \notin V$, então pela Proposição 1.77. o vetor

$$u_n' = v_n - \langle v_n, u_1 \rangle u_1 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}$$

é não nulo e ortogonal a todos os vetores de U. Normalizando u_n' obtemos u_n.

Portanto, $\{u_1, u_2, \ldots, u_{n-1}, u_n\}$ forma uma base ortonormal de V.

□

Teorema 2.79. *O conjunto $S^\perp$ é um subespaço de V, mesmo que S não o seja.*

No caso de S ser um subespaço, então $S \cap S^\perp = \{0\}$.

Teorema 2.80. *Sejam V um espaço vetorial euclidiano de dimensão finita e S um subespaço de V. Então, V é soma direta de S e $S^{\perp}$, isto é,*

$$V = S \oplus S^{\perp}.$$

Assim, todo elemento em V pode ser escrito de modo único como sendo, $v = u + u'$, com $u \in S$ e $u' \in S^{\perp}$. E neste caso, em que S é um subespaço vetorial de V, o conjunto $S^{\perp}$ é denominado **complemento ortogonal** de S em V.

Corolário 2.81. *Sejam V um espaço vetorial de dimensão finita munido de um produto interno. Se S um subespaço de V, então $dimV = dimS + dimS^{\perp}$.*

Teorema 2.82. *Sejam V um espaço vetorial com um produto interno, S e W subespaços vetoriais de V. Então,*

(a) $(S^{\perp})^{\perp} = S$.

(b) $(S \cap W)^{\perp} = S^{\perp} + W^{\perp}$.

(c) $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$.

2.5 Transformações lineares

Uma transformação linear é um tipo particular de aplicação em que o seu domínio e contradomínio são espaços vetoriais, e que leva combinações lineares de vetores de um espaço em combinações lineares de vetores em outro espaço. Na Física, Engenharia e várias áreas de exatas são inúmeras as aplicações envolvendo este conceito.

Nesta seção recordaremos os principais conceitos e resultados envolvendo as transformações lineares em espaços vetoriais arbitrários, e a relação fundamental entre estes espaços de dimensão n e o $\mathbb{R}^n$ $(isomorfismo)$, permitindo efetuarmos cálculos vetoriais em espaços arbitrários aparentemente diferentes, porém dotados da mesma estrutura algébrica do $\mathbb{R}^n$, possibilitando restringirmos o estudo somente a este último.

Definição 2.83. Sejam U e V espaços vetoriais sobre um mesmo corpo $\mathbb{K}$. A aplicação $T : U \to V$ é denominada **transformação linear** se, satisfaz:

(a) $T(u_1 + u_2) = T(u_1) + T(u_2)$, $\forall u_1, u_2 \in U$.

(b) $T(k.u_1) = k.T(u_1)$, $\forall u_1 \in U$ e $\forall k \in \mathbb{K}$.

Estas duas condições podem ser condensadas em:

$$T(k_1.u_1 + k_2.u_2) = k_1.T(u_1) + k_2.T(u_2), \quad \forall u_1, u_2 \in U \text{ e } \forall k_1, k_2 \in \mathbb{K}.$$

A transformação linear $T : U \to U$ é denominada de **operador linear** sobre o espaço vetorial U.

Denotamos por $L(U, V)$ o conjunto de todas as transformações lineares $T : U \to V$; e por $L(U)$ o conjunto de todos os operadores lineares $T : U \to U$.

Exemplo 2.84. Consideremos o espaço vetorial $P_n(\mathbb{R})$. A aplicação

$$T : P_n(\mathbb{R}) \to P_n(\mathbb{R}) \text{ definida por } T(p) = p\,' ,$$

sendo $p\,'$ a derivada do polinômio p, para todo $p \in P_n(\mathbb{R})$, é uma transformação linear.

Pois, para $\forall\ p, q\ \in P_n(\mathbb{R})$ e $k \in \mathbb{R}$, temos:

(a) $T(p+q)(x) = (p+q)'(x) = p'(x) + q'(x) = T[(p)(x)] + T[(q)(x)]$.

(b) $T(kp)(x) = (kp)'(x) = kp'(x) = kT[p(x)]$.

Exemplo 2.85. Dada uma matriz $A \in M_{m\times n}(\mathbb{R})$ podemos sempre associá-la a uma transformação linear $T_A : \mathbb{R}^n \to \mathbb{R}^m$ cuja imagem $T_A(v)$ do vetor $v \in \mathbb{R}^n$ é dada pelo produto da matriz $[A]_{m\times n}$ com a matriz coluna $[v]_{n\times 1}$ formada pelas coordenadas do vetor v, ou seja,

$$T_A : \mathbb{R}^n \to \mathbb{R}^m \ \text{ tal que } \ T_A(v) = [A].[v]\ .$$

T_A é uma transformação linear, pois $\forall u, v \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, temos:

(a) $T(u+v) = A(u+v) = Au + Av = T(u) + T(v)$.

(b) $T(\lambda v) = A(\lambda v) = \lambda(Av) = \lambda T(v)$.

Propriedade 2.86. Sejam U , V espaços vetoriais sobre $\mathbb{K}$, $T : U \to V$ uma transformação linear, 0_u e 0_v os vetores nulos dos espaços vetoriais U e V, respectivamente. Então,

(a) $T(0_u) = 0_v$.

(b) $T(-u) = -T(u),\ \forall u \in U$.

(c) $T(\sum\limits_{i=1}^{\text{n}} k_i u_i) = \sum\limits_{i=1}^{\text{n}} k_i T(u_i),\ \forall u_i \in U$ e $\forall k_i \in \mathbb{K}$, para $i \in \{1, \ldots, n\}$.

Teorema 2.87. *Sejam U, V espaços vetoriais, $\{v_1, v_2, \ldots, v_n\}$ uma base de V e $\{u_1, u_2, \ldots, u_n\}$ vetores de U. Então, existe uma única transformação linear $T : V \to U$ tal que $T(v_i) = u_i$, para todo $i \in \{1, 2, \ldots n\}$.*

Exemplo 2.88. Consideremos $B = \{(1,0), (1,1)\}$ uma base de $\mathbb{R}^2$, e os vetores $(3,1,0)$ e $(1,2,-1)$ de $\mathbb{R}^3$. Então, existe uma transformação linear

$$T : \mathbb{R}^2 \to \mathbb{R}^3 \ \text{ tal que } \ T(1,0) = (3,1,0) \ \text{ e } \ T(1,1) = (1,2,-1).$$

Escrevendo (x,y) como uma combinação linear da base dada, e aplicando T, encontramos a transformação linear T.

Como $(x,y) = (x-y)(1,0) + y(1,1)$ e T uma transformação linear, então

$$T(x,y) = (x-y).T(1,0) + y.T(1,1) = (x-y).(3,1,0) + y.(1,2,-1) =$$
$$= (3x - 3y + y,\ x - y + 2y,\ -y).$$

Portanto, $\quad T(x,y) = (3x - 2y,\ x + y,\ -y)$.

Definição 2.89. Sejam U e V espaços vetoriais sobre um corpo $\mathbb{K}$ e $T : U \to V$ uma transformação linear.

(a) O conjunto de todos os vetores em V que são imagens por T de pelo menos um vetor em U é denominado de **imagem da transformação linear** T, e denotamos por $Im(T)$. Ou seja,

$$Im(T) = \{\ v \in V\ :\ v = T(u)\ ,\ \text{u} \in U\ \}.$$

(b) O conjunto de todos os vetores $u \in U$ tais que $T(u) = 0$ é denominado **núcleo da transformação linear de** T, e denotamos por $Ker(T)$. Ou seja,

$$Ker(T) = \{ \ u \in U : \ T(u) = 0 \ \}.$$

No caso de $Ker(T) = \{0\}$, dizemos que T é não-singular. Caso contrário, T é singular.

Exemplo 2.90. Consideremos a transformação linear

$$T : P_1(\mathbb{R}) \to \mathbb{R} \ \text{ tal que } \ p(x) = \int_0^1 p_1(x)\, dx \ .$$

Então, dado $p(x) \in P_1(\mathbb{R})$ temos que $p(x) = ax + b$, com a , b $\in \mathbb{R}$.
Mas , $T(\ p(x)\ ) = \int_0^1 p(x)\ dx \ = \ \int_0^1 (ax+b)\ dx \ = \ [\ a\frac{x^2}{2} + bx\ ]_0^1 \ = \ \frac{a}{2} + b - 0 = \frac{a}{2} + b.$
Como, $Ker(T) = \{ \ p(x) \in P_1(x) \ : \ T(p(x)) = 0 \ \}$, então $\frac{a}{2} + b = 0$.
Logo, $b = -\frac{a}{2}$, e temos então que $p(x) = ax - \frac{a}{2}$.
Portanto, $Ker(T) = \{ \ p(x) \in P_1(x) \ : \ p(x) = ax - \frac{a}{2} \ \}$.

Agora, como para todo $a \in \mathbb{R}$, existe $p(x) = a \in P_1(\mathbb{R})$, tal que

$$T(\ p(x)\ ) = \int_0^1 p(x)\ dx \ = \ \int_0^1 a\ dx \ = \ ax|_0^1 \ = \ a - 0 \ = \ a.$$

Logo , $\mathbb{R} \subset Im(T)$.
Mas, pela definição de $Im(T)$ temos que $Im(T) \subset \mathbb{R}$.
Portanto, $Im(T) = \mathbb{R}$.

Teorema 2.91. *Seja $T : U \to V$ uma transformação linear. Então,*
(a) $Ker(T)$ é um subespaço de U.
(b) $Im(T)$ é um subespaço de V.

Definição 2.92. Seja T uma transformação linear entre espaços vetoriais de dimensão finita. Definimos:
(a) O **posto de** T como sendo a dimensão de $Im(T)$, e denotamos por $post(T)$.

$$post(T) = dim\ Im(T).$$

(b) A **nulidade de** T como sendo a dimensão de $Ker(T)$, e denotamos por $nul(T)$.

$$nul(T) = dim\ Ker(T).$$

Teorema 2.93. *Se $T : U \to V$ é uma transformação linear, sendo U um espaço vetorial de dimensão finita, então*

$$\dim\ Ker\ (T) + dim\ Im(T) = dim\ U, \ \textit{ou seja}, \ nul(T)\ +\ post(T)\ =\ dim\ U.$$

Definição 2.94. Sejam U e V espaços vetoriais sobre um corpo $\mathbb{K}$ e T uma transformação linear de U em V. Dizemos que:
(a) A transformação T é **injetora** se $u \neq v$ então $T(u) \neq T(v)$, para todo $u, v \in U$.
(b) A transformação T é **sobrejetora** se, $Im\ (T) = V$.
(c) A transformação T é **bijetora** se, T é injetora e também sobrejetora.

Teorema 2.95. *Seja $T : U \to V$ uma transformação linear. Então, $Ker(T) = \{0\}$ (T é não singular) se, e somente se, T é injetora.*

Teorema 2.96. *Seja $T : U \to V$ uma transformação linear, e $dimU = dimV$. Então, T é injetora (T é não singular) se, e somente se, T é sobrejetora.*

Definição 2.97. Sejam U e V espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$. Uma transformação linear $T : U \to V$ bijetora é denominada um **isomorfismo** de U em V.

Quando existe um isomorfismo de U em V, dizemos que o espaço vetorial U é **isomorfo** ao espaço vetorial V, e denotamos por $U \simeq V$.

Um isomorfismo $T : U \to U$ é denominado um **automorfismo** de U.

Teorema 2.98. *Seja V um espaço vetorial real sobre o corpo $\mathbb{K}$, com $dimV = n$, então V é isomorfo ao espaço vetorial $\mathbb{K}^n$.*

Exemplo 2.99. O espaço vetorial dos polinômios $P_n(\mathbb{R})$ é isomorfo ao $\mathbb{R}^{n+1}$.

Exemplo 2.100. O espaço vetorial $M_2(\mathbb{R})$ das matrizes quadradas reais de ordem 2 é isomorfo ao $\mathbb{R}^4$.

Teorema 2.101. *Dois espaços vetoriais de dimensões finitas são isomorfos se, e somente se, possuem a mesma dimensão.*

No Exemplo 1.85 vimos que dada uma matriz $A \in M_{m\times n}(\mathbb{R})$ sempre pode ser associada a uma transformação linear $T_A : \mathbb{R}^n \to \mathbb{R}^m$ cuja imagem $T_A(v)$ do vetor v do $\mathbb{R}^n$ é dada pelo produto da matriz $[A]_{m\times n}$ com a matriz coluna $[v]_{n\times 1}$ das coordenadas do vetor v com respeito à base canônica do espaço $\mathbb{R}^n$.

$$T_A(v) = [A].[v] \ .$$

O objetivo agora é estabecermos a recíproca, isto é, fixadas as bases, toda transformação linear $T : U \to V$, sendo U e V espaços vetoriais finitamente gerado, sobre o mesmo corpo $\mathbb{K}$ associa uma única matriz, que permite através dela calcularmos $T(u)$ para todo u em U, usando a forma matricial.

Estudamos matrizes de transformações lineares por duas razões básicas, uma teórica e outra bastante prática. A primeira, muitas vezes, as respostas para as questões teóricas sobre a estrutura de transformações lineares arbitrárias em espaços de dimensão finita podem ser obtidas simplesmente estudando as transformações matriciais. E a segunda, que estas matrizes tornam possíveis calcularmos as imagens de vetores usando multiplicação matricial, e estes cálculos podem ser efetudados rapidamente através de computadores.

Definição 2.102. Sejam U e V espaços vetoriais de dimensão finita, sobre um corpo $\mathbb{K}$, com bases ordenadas $B = \{u_1, u_2, \ldots, u_n\}$ e $B' = \{v_1, v_2, \ldots, v_m\}$, respectivamente.

Dada $T : U \to V$ uma transformação linear, então $T(u_1), T(u_2), \ldots, T(u_n)$ são vetores de V. Logo, cada um destes vetores podem ser escritos de maneira única como combinação linear dos elementos da base B'. Ou seja,

$$\begin{cases} T(u_1) = a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m \\ T(u_2) = a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_m \\ \qquad\qquad \vdots \\ T(u_n) = a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m \end{cases} ,$$

com $a_{ij} \in \mathbb{K}$, para $i \in \{1, 2, \ldots, n\}$ e $j \in \{1, 2, \ldots, m\}$.

A transposta da matriz formada pelos coeficientes das combinações lineares acima é chamada de **representação matricial de T em relação às bases B e B'**, e denotamos por $[T]_{B \to B'}$, isto é,

$$[T]_{B \to B'} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

No caso de $U = V$ e $B = B'$ denotamos a matriz acima por $[T]_B$. Além disso, se B e B' são as bases canônicas dos espaços vetoriais U e V, respectivamente, denotamos simplesmente por [T] .

Como a matriz $[T]_{B \to B'}$ depende das bases B e B', uma mesma transformação linear pode ser representada por infinitas matrizes, bastando mudarmos as bases. Porém, uma vez fixada as bases B e B', respectivamente, do domínio e do contra-domínio, uma transformação linear fica bem definida através de sua representação matricial, ou seja, dada uma transformação linear só existe uma única matriz associada a essa transformação linear, e vice-versa, dada uma matriz existe uma única transformação linear associada a ela. Assim, o estudo de uma transformação linear se restringe ao um estudo matricial; e a transformação linear sobre um vetor u de U se comporta simplesmente como uma multiplicação, pela direita, da representação matricial de T pela matriz das coordenadas de u na base B, como garante o teorema a seguir.

Teorema 2.103. *Sejam $T : U \to V$ uma transformação linear entre espaços finitamente gerados, B e B' as bases de U e V, respectivamente. Então, para qualquer vetor $u \in U$, temos*

$$\Big[T\Big]_{B \to B'} \Big[u\Big]_B = \Big[T(u)\Big]_{B'}.$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix},$$

onde $u = (\ \alpha_1, \alpha_2, \cdots, \alpha_n\ )_B$ *e* $T(u) = (\ \beta_1, \beta_2, \cdots, \beta_n\ )_{B'}$.

Exemplo 2.104. Sejam B e B' as bases canônicas de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente e T a transformação linear definida por

$$T : \mathbb{R}^3 \to \mathbb{R}^2, \quad \text{dada por} \quad T(x, y, z) = (2x + y,\ x - z)\ .$$

Como T está aplicada nas bases canônicas, a matriz da transformação linear se torna mais fácil de ser obtida, pois podemos escrever a expressão $T(x, y, z)$ na forma

$$T(x, y, z) = (2x\ ,\ x) + (y\ ,\ 0) + (0\ ,\ -z).$$
$$T(x, y, z) = x(2\ ,\ 1) + y(1\ ,\ 0) + z(0\ ,\ -1).$$

Assim, T pode ser escrita na forma matricial como sendo:

$$T(x, y, z) = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y & x - z \end{bmatrix}.$$

Desse modo, a transformação T pode ser representada pela matriz

$$[\ T\ ] \ = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} ,$$

e através dela podemos calcular $T(x, y, z)$ para qualquer $(x, y, z) \in \mathbb{R}^3$.

Tomando-se, como exemplo, o vetor $u = (3, -1,\ 2)$ então $T(3\ , -1\ , 2) = (5, 1)$, que também pode ser obtido aplicando o Teorema 1.103. .

$$[\ T\ ][u]_{B'} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}^t.$$

Exemplo 2.105. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ a transformação linear definida por

$$T(x, y) = (x - y,\ x + 2y,\ y),$$

sendo $B = \{(1, 0)\ ,\ (1, 1)\}$ e $B' = \{(0, 0, 1)\ ,\ (1, 1, 0)\ ,\ (1, 0, 1)\}$ as bases de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente. Determinemos a matriz $[T]_{B\to B'}$.

$$T(1, 0) = (1, 1, 0) = a_{11}.(0, 0, 1) + a_{12}.(1, 1, 0) + a_{13}.(1, 0, 1).$$

$$\begin{cases} a_{12} + a_{13} = 1 \\ a_{12} = 1 \\ a_{11} + a_{13} = 0 \end{cases} . \quad \text{Logo, } a_{12} = 1\ ,\ a_{13} = 0\ \ e\ \ a_{11} = 0.$$

Então, $T(1, 0) = 0.(0, 0, 1) + 1.(1, 1, 0) + 0.(1, 0, 1)$.

$$T(1, 1) = (0, 3, 1) = a_{21}.(0, 0, 1) + a_{22}.(1, 1, 0) + a_{23}.(1, 0, 1).$$

$$\begin{cases} a_{22} + a_{23} = 0 \\ a_{22} = 3 \\ a_{21} + a_{23} = 1 \end{cases} . \quad \text{Logo, } a_{22} = 3\ ,\ a_{23} = -3\ \ e\ \ a_{21} = 4.$$

Então, $T(1, 1) = (0, 3, 1) = 4.(0, 0, 1) + 3.(1, 1, 0) - 3.(1, 0, 1)$.

$$\text{Portanto,} \quad [T]_{B\to B'} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 3 \\ 0 & -3 \end{bmatrix} .$$

Exemplo 2.106. Seja T a transformação linear do exemplo anterior e consideremos as bases canônicas $B = \{(1, 0), (0, 1)\}$ e $B' = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

$$T(1, 0) = (1, 1, 0) = a_{11}.(0, 0, 1) + a_{12}.(1, 1, 0) + a_{13}.(1, 0, 1).$$

$$\begin{cases} a_{12} + a_{13} = 1 \\ a_{12} = 1 \\ a_{11} + a_{13} = 0 \end{cases} . \quad \text{Logo, } a_{12} = 1\ ,\ a_{13} = 0\ \ e\ \ a_{11} = 0.$$

Então, $T(1, 0) = 0.(0, 0, 1) + 1.(1, 1, 0) + 0.(1, 0, 1)$.

$$T(0, 1) = (-1, 2, 1) = a_{21}.(0, 0, 1) + a_{22}.(1, 1, 0) + a_{23}.(1, 0, 1)$$

$$\begin{cases} a_{22} + a_{23} = -1 \\ a_{22} = 2 \\ a_{21} + a_{23} = 1 \end{cases} . \quad \text{Logo, } a_{22} = 2\ ,\ a_{23} = -3\ \ e\ \ a_{21} = 4.$$

Então, $T(0,1) = 4.(0,0,1) + 2.(1,1,0) - 3.(1,0,1)$.

Portanto, $[T]_{B\to B'} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 2 \\ 0 & -3 \end{bmatrix}$.

Podemos obter a matriz de qualquer transformação linear, desde que os espaços vetoriais envolvidos na transformação seja finitamente gerado. E utilizando o conceito de coordenadas, podemos trabalhar com quaisquer bases dos espaços vetoriais envolvidos na transformação linear.

Exemplo 2.107. Sejam $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ e $B' = \{1 \,,\ t,\ t^2\}$, as bases canônicas de $M_2(\mathbb{R})$ e $P_2(\mathbb{R})$, respectivamente. Consideremos a transformação linear $T : M_2(\mathbb{R}) \to P_2(\mathbb{R})$ definida por

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = (a+b) + (c-d)t + (b+2c+d)t^2.$$

Assim, temos:

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = T(\, a,b,c,d\,)_B \;=\; (\, a+b,\ c-d,\ b+2c+d\,)_{B'}.$$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = a(1,0,0) + b(1,0,1) + c(0,1,2) + d(0,-1,1).$$

Logo, podemos escrever

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}_B = \begin{bmatrix} a+b \\ b-c \\ b+c+d \end{bmatrix}_{B'}.$$

E a matriz $[\ T\ ] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 2 & 1 \end{bmatrix} \in M_{3\times 4}(\mathbb{R})$, representa a transformação linear dada.

Assim, por exemplo:

$$\left[T\left(\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \right) \right]_{B'} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 2 & 1 \end{bmatrix} . \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}_B .$$

Portanto, $\left[T\left(\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \right) \right]_{B'} = 2 - 4t + t^2.$

Definição 2.108. Dizemos que uma matriz A é **equivalente por linha** a uma matriz B, e denotamos por $A \sim B$, quando B pode ser obtida de A por uma sequência finita de operações denominadas **operações elementares com linhas**, que são:

(a) Trocar duas linhas quaísquer.

$L_i \longleftrightarrow L_j$ (troca da linhas i com a linha j).

(b) Substituir uma linha por ela mesma multiplicada por um escalar não nulo.

$L_j \longrightarrow kL_j$ (troca da linha j por k vezes a linha j).

(c) Substituir uma linha por ela mesma mais o produto de um escalar, não nulo, por outra linha.
$L_j \longrightarrow L_j + kL_i$ (troca da linha j pela soma da linha j com k vezes a linha i).

Proposição 2.109. *Dada uma matriz $A_{m\times n}(\mathbb{K})$ podemos obter uma matriz escalonada B equivalente a A.*

Teorema 2.110. *As linhas não-nulas (quando considerada como vetores de $\mathbb{R}^n$) de uma matriz na forma escalonada são linearmente independentes.*

Definição 2.111. Podemos escrever os vetores com parênteses e vírgulas ou em forma matricial como vetores linhas ou vetores colunas.

Assim, na matriz $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, os vetores

$v_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $v_2 = (a_{21}, a_{22}, \ldots, a_{2n})$, $\cdots$, $v_n = (a_{m1}, a_{m2}, \ldots, a_{mn})$ em $\mathbb{K}^n$ formados pelas linhas de A são chamados **vetores linhas** de A. Enquanto que, os vetores $u_1 = (a_{11}, a_{21}, \ldots, a_{m1})$, $u_2 = (a_{12}, a_{22}, \ldots, a_{m2})$, $\cdots$, $u_n = (a_{m1}, a_{m2}, \ldots, a_{mn})$ em $\mathbb{K}^n$ são os **vetores colunas** de A.

O subespaço de $\mathbb{K}^n$ gerado pelos vetores linhas de A é denominado **espaço linha** de A; o subespaço de $\mathbb{K}^n$ gerado pelos vetores colunas de A é o **espaço coluna** de A, que também é um subespaço de $\mathbb{K}^n$.

O conjunto solução de um sistema homogêneo de equações $AX = 0$, é um subespaço de $\mathbb{K}^n$, denominado **espaço nulo** de A.

Teorema 2.112.

(a) As operações elementares com linhas não alteram o espaço linha, e nem o espaço nulo de uma matriz, ou seja, geram o mesmo espaço vetorial.

(b) Se uma matriz está na forma escalonada por linhas, então os vetores linhas formam uma base do espaço linha de A, e os vetores colunas formam uma base do espaço coluna de A.

As operações elementares com linhas afetam o espaço coluna de uma matriz, mas não alteram as relações de dependência linear entre os vetores colunas.

Exemplo 2.113. Consideremos a matriz $A = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ na forma escalonada por linha .

Pelo teorema anterior, os vetores $v_1 = (1, 0, -1, 0)$, $v_2 = (0, 3, 1, 5)$, $v_3 = (0, 0, 2, 4)$ formam uma base do espaço linha de A; os vetores $u_1 = (1, 0, 0, 0)$, $u_2 = (0, 3, 0, 0)$, $u_3 = (-1, 1, 2, 0)$, $u_4 = (0, 5, 4, 0)$ não formam base do espaço coluna de A, pois o vetor $u_4 = (0, 5, 4, 0)$ é combinação linear dos demais.

Como operações elementares com linhas não alteram o espaço linha de uma matriz, podemos encontrar uma base do espaço linha (ou uma base do espaço coluna) encontrando uma base do espaço linha (ou uma base do espaço coluna) de qualquer forma escalonada por linhas da matriz.

Teorema 2.114.

(a) O posto de uma matriz é o posto da transformação linear representada por ela.

(b) O posto de uma matriz é igual à quantidade de linhas (ou colunas) LI da matriz.

(c) Se uma matriz A de ordem n possui posto estritamente menor que n, então A é singular.

Teorema 2.115. *Seja $T : U \to V$ uma transformação linear e B e B' bases de U e V, respectivamente. Então,*

(a) $dim\ Im(T) = post\ [T]_{B\to B'}$.

(b) $dim\ Ker(T) = nul\ [T]_{B\to B'} =$ nº de colunas de $[T]_{B\to B'} - post\ [T]_{B\to B'}$.

Exemplo 2.116. Consideremos B e B' as bases canônicas de $\mathbb{R}^4$ e $\mathbb{R}^3$, e seja $T : \mathbb{R}^4 \to \mathbb{R}^3$ a transformação linear tal que

$$T(x, y, z, t) = (x + y - z + 2t\ ,\ y + z - t\ ,\ 3x + 4y - 2z + 5t)\ .$$

Determinemos as bases para $Im(T)$ e $Ker(T)$.

Podemos escrever T como sendo:

$$T(x, y, z, t) = x(1, 0, 3) + y(1, 1, 4) + z(-1, 1, -2) + t(2, -1, 5)\ .$$

Logo, $Im(T) = [\ (1, 0, 3), (1, 1, 4), (-1, 1, -2), (2, -1, 5)\ ]$.

Determinemos uma base de $Im(T)$ escalonando a matriz $[\ T\ ]^T_{B\to B'}$.

$$[\ T\ ]_{B\to B'} = \underset{L_3 \to L_3 - 3L_1.}{\begin{bmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & -1 \\ 3 & 4 & -2 & 5 \end{bmatrix}} \sim \underset{L_3 \to L_3 - L_2.}{\begin{bmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \end{bmatrix}} \sim \begin{bmatrix} 1 & 1 & -2 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Portanto, uma base da imagem de T é formada pelos dois primeiros vetores da matriz inicial, isto é, $B = \{(1, 0, 3), (1, 1, 4)\}$ é uma base para $Im(T)$, e $post(T) = 2$.

Para calcularmos uma base do núcleo de T temos que resolver o sistema homogêneo

$$[\ T\ ]_{B\to B'}.[v] = 0, \quad \text{ou seja,} \quad \begin{cases} x + y - z + 2t = 0 \\ y + z - t = 0 \\ 3x + 4y - 2z + 5t = 0 \end{cases}.$$

Escalonando o sistema temos:

$$\underset{L_3 \to L_3 - 3L_1.}{\begin{bmatrix} 1 & 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 3 & 4 & -2 & 5 & 0 \end{bmatrix}} \sim \underset{\substack{L_1 \to L_1 - L_2. \\ L_3 \to L_3 - L_2.}}{\begin{bmatrix} 1 & 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \end{bmatrix}} \sim \begin{bmatrix} 1 & 0 & -2 & 3 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Então, $\begin{cases} x - 2z + 3t = 0 \\ \quad y + z - t = 0 \end{cases}$, (sistema indeterminado com duas variáveis livres).

Fazendo $z = a$ e $t = b$, temos: $x = 2a - 3b$ e $y = b - a$. E o conjunto solução do sistema é dado por

$$(2a - 3b \ , b - a \ , a \ , b) = a(2, -1, 1, 0) + b(-3, 1, 0, 1) \ , \ \ a, b \in \mathbb{R}.$$

Portanto, $\{(2, -1, 1, 0), (-3, 1, 0, 1)\}$ formam uma base de $Ker(T)$.

Exemplo 2.117. Sejam $B = \{ \ (0, 1, 1), \ (1, 0, 0), \ (0, 0, 1) \ \}$ e $B' = \{ \ (1, 0), \ (1, 1) \ \}$ bases de $\mathbb{R}^3$ e $\mathbb{R}^2$, respectivamente. Determinemos uma base do núcleo e uma da imagem de T da transformação linear

$$T : \mathbb{R}^3 \to \mathbb{R}^2 \ \text{ definida por } \ T(x, y, z) = (x + y - z \ , \ 2x - y - z) \ .$$

$T(0, 1, 1) = (0, -2) = 2.(1, 0) - 2.(1, 1) = (2, -2)_{B'}$.
$T(1, 0, 0) = (\ 1, \ 2) = -1.(1, 0) + 2.(1, 1) = (-1, 2)_{B'}$.

$T(0, 0, 1) = (-1, -1) = 0(1, 0) - 1.(1, 1) = (0, -1)_{B'}$.

Portanto , $\quad [\ T \ ]_{B \to B'} = \begin{bmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \end{bmatrix}$.

Determinemos uma base de $Im(T)$ escalonado a matriz $[\ T \ ]_{B \to B'}$.

$$\begin{bmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -1 \end{bmatrix}.$$

$L_2 \to L_2 + L_1. \qquad L_1 \to \frac{1}{2}.L_1. \qquad L_1 \to L_1 + \frac{1}{2}.L_2.$

Logo, os dois primeiros vetores da matriz inicial formam uma base da imagem de T. Portanto $B = \{(2, -2), (-1, 2)\}$ forma uma base de $Im(T)$.

Escalonando o sistema $\begin{cases} x + y - z = 0 \\ 2x - y - z = 0 \end{cases}$, temos:

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ 2 & -1 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & -3 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -\frac{1}{3} & 0 \end{bmatrix}.$$

$L_2 \to L_2 - 2.L_1. \qquad L_2 \to -\frac{1}{3}.L_2.$

Então, $\begin{cases} x + y - z = 0 \\ \quad y - \frac{1}{3}z = 0 \end{cases}$ (sistema indeterminado com uma variável livre).

Fazendo $z = 3a$, temos $y = a$. Logo, $x + a - 3a = 0$ e $x = 2a$, $a \in \mathbb{R}$.

Portanto, $B = \{(2, 1, 3)\}$ é uma base de $Ker(T)$.

3 Teoria Espectral

As classes de escalares e vetores conhecidas como "autovalores" e "autovetores" são conceitos relacionados com transformações lineares especiais por suas características peculiares. A ideia surgiu no estudo do movimento rotacional e, mais tarde, foi usada para classificar tipos de superfícies e nas resoluções de certas equações diferenciais. No século XX, foi aplicada à matrizes e às transformações matriciais e hoje é aplicada em diversas áreas.

Neste capítulo apresentaremos os conceitos, propriedades e resultados importantes envolvendo autovalores e autovetores de uma matriz, e sua aplicação no estudo da diagonalização de matrizes, baseados nas referências [1], [2], [3], [5], [6], e [22].

3.1 Autovalores e autovetores de uma transformação linear

Definição 3.1. Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$ e $T : V \to V$ um operador linear. Um vetor v, não nulo, em V, é um **autovetor** de T, se existir um escalar $\lambda \in \mathbb{K}$, tal que

$$T(v) = \lambda v.$$

Neste caso, dizemos que λ é um **autovalor de** T associado ao vetor v, ou v é um autovetor de T associado a λ.

Vejamos agora algumas consequências da definição anterior e exemplos.

Propriedade 3.2. Se v é um autovetor de T associado ao autovalor λ, então qualquer vetor $u = av$, não nulo, tambem é autovetor de T associado a λ.

Demonstração.

$$T(v) = \lambda v, \quad \text{então} \quad T(av) \;=\; aT(v) \;=\; a(\lambda v) = \lambda(av).$$

Portanto, av é autovetor associado ao autovalor λ. □

Proposição 3.3. *O conjunto de todos os autovetores associados ao autovalor λ, acrescido do vetor nulo, denotado por $V(\lambda)$, forma um subespaço do espaço vetorial V, denominado* ***autoespaço de*** λ.

$$V(\lambda) = \{\ v \in V \ : \ T(v) = \lambda v\ ,\ \lambda \in \mathbb{K}\ \}.$$

Demonstração.

Sejam $v_1, v_2 \in V(\lambda)$, então $T(v_1) = \lambda v_1$ e $T(v_2) = \lambda v_2$.

Assim, para quaisquer $a, b \in K$, temos:

$T(av_1 + bv_2) = T(av_1) + T(bv_2) = aT(v_1) + bT(v_2) = a(\lambda v_1) + b(\lambda v_2) = \lambda(av_1 + bv_2)$.
Logo, $av_1 + bv_2$ é um autovetor associado à λ, ou seja, $av_1 + bv_2 \in V(\lambda)$.

Portanto, $V(\lambda)$ é um subespaço de V. $\square$

Exemplo 3.4. O operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$T(x, y, z) = (2x + y,\ y - z,\ 2y + 4z), \text{ para todo } (x, y, z) \in \mathbb{R}^3,$$

admite os autovalores $\lambda_1 = 2$ e $\lambda_2 = 3$.

De fato, $\lambda_1 = 2$ é um autovalor de T associado ao autovetor $v_1 = (1, 0, 0)$, pois

$$T(v_1) = T(1, 0, 0) = (2, 0, 0) = 2(1, 0,\ 0) = 2v_1.$$

Os vetores da forma v = (a, 0, 0), a em $\mathbb{R}^*$, são autovetores de T associados ao autovalor 2.

E o autoespaço $V(2) = \{\ v \in \mathbb{R}^3\ :\ v = (a, 0, 0),\ a \in \mathbb{R}\} = \Big[(1, 0, 0)\Big]$.

$\lambda_2 = 3$ é um autovalor de T associado ao autovetor $v_2 = (1, 1, -2)$, pois

$$T(v_2) = T(1, 1, -2) = (3, 3, -6) = 3(1, 1, -2) = 3v_2\ .$$

Os vetores da forma v = (a, a, -2a), a em $\mathbb{R}^*$, são autovetores de T associados ao autovalor 3.

E o autoespaço $V(3) = \{\ v \in \mathbb{R}^3\ :\ v = (a, a, -2a),\ a \in \mathbb{R}\} = \Big[(1, 1, -2)\Big]$.

Exemplo 3.5. Seja $T : \mathbb{C}^2 \to \mathbb{C}^2$ um operador linear definido por $T(x, y) = (-y\ , x)$, para todo $(x, y) \in \mathbb{C}^2$.

$\lambda_1 = -i$ é um autovalor de T associado ao autovetor $v_1 = (1\ , i)$, pois

$$T(v_1) = T(1\ , i) = (-i\ , 1) = (-i\ , -i^2) = -i(1\ , i) = -iv_1\ .$$

Todo vetor da forma (a, ai), a em $\mathbb{C}^*$ é um autovetor do operador T associado ao autovalor $-i$.

$$V(-i) = \{v \in \mathbb{C}^2 :\ v = (a\ ,\ ai)\ , \forall a \in \mathbb{C}\} = \Big[(1,\ i)\Big].$$

$\lambda_2 = i$ é um autovalor de T associado ao autovetor $v_2 = (1\ , -i)$, pois

$$T(v_2) = T(1\ , -i) = (i\ , 1) = (i\ , -i^2) = i(1\ , -i) = iv_2\ .$$

Todo vetor da forma $(a, -ai)$, a em $\mathbb{C}^*$ é um autovetor do operador T associado ao autovalor i.

$$V(i) = \{v \in \mathbb{C}^2 :\ v = (a\ ,\ -ai)\ , \forall a \in \mathbb{C}\} = \Big[(1,\ -i)\Big]\ .$$

3.2 Autovalores e autovetores de uma matriz

Também podemos definir autovalor e autovetor para uma matriz.

Definição 3.6. Dada uma matriz quadrada A de ordem n, dizemos que λ é um autovalor associado ao autovetor v de A, se λ e v forem, respectivamente, o autovalor e autovetor do operador linear $T_A : \mathbb{C}^n \to \mathbb{C}^n$ cuja matriz correspondente é a matriz A, em relação à base canônica, isto é, $T(v) = Av$.

Assim, um autovalor λ em $\mathbb{C}$ e um autovetor v em $\mathbb{C}^n$ de A são soluções da equação

$$Av = \lambda v\ ,\ \ v \neq 0.$$

Usando o fato de que a matriz identidade I_n é tal que $I_n v = v$, então podemos escrever a equação anterior na forma matricial como sendo

$$Av = \lambda I_n v,$$

ou seja ,

$$(A - \lambda I_n)v = 0 \ .$$

Esta última igualdade resulta em um sistema homogêneo de n equações com n incógnitas, sendo n a ordem da matriz A, isto é, num sistema compatível.

As n incógnitas representam as coordenadas do vetor v, e a matriz $A-\lambda I$ representa os coeficientes das equações.

Pelo fato do sistema ser homogêneo, caso o determinante da matriz $A-\lambda I$ seja diferente de 0, obteremos uma única solução para o sistema, a trivial, que não nos interessa, pois teríamos $v = 0$.

O sistema $(A - \lambda I_n)v = 0$ admite solução não trivial, isto é, vetores não nulos se, e somente se,

$$det \ (A - \lambda I_n) = 0 \ .$$

E esta igualdade é denominada **equação característica da matriz A**, e expandindo o primeiro membro desta equação obtemos um polinômio de grau n em λ denominado **polinômio característico**.

$$P(\lambda) = \lambda^n \ + \ a_1\lambda^{n-1} \ + \ a_2\lambda^{n-2} + \ \cdots \ + \ a_{n-1}\lambda \ + \ a_n.$$

Proposição 3.7. *Seja A uma matriz quadrada de ordem n. Então, os autovalores de A são as raízes do polinômio característico $P(\lambda) = det\ (A - \lambda I_n)$.*

Definição 3.8.

Sejam A uma matriz de ordem n e λ um autovalor de A, definimos:

(a) **Multiplicidade algébrica do autovalor** λ como sendo a multiplicidade da raiz λ no polinômio característico.

(b) **Multiplicidade geométrica do autovalor** λ como sendo a dimensão do autoespaço associado ao autovalor λ.

Observação 3.9. A multiplicidade geométrica de um autovalor λ não excede sua multiplicidade aritmética.

Nos exemplos a seguir veremos que os autovalores de uma matriz podem ser obtidos utilizando-se da Proposição 2.7. ; e os autovetores associados a estes autovalores são obtidos resolvendo-se os sistemas lineares homogêneos indeterminados extraídos da equação matricial $Av = \lambda.v$, sendo v não nulo, para cada um dos autovalores encontrados.

Exemplo 3.10. Determinemos todos os autovalores e autovetores, e uma base de cada autoespaço da matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.

Obtemos os autovalores de A encontrando as raízes do polinômio característico

$$P(\lambda) = det(A - \lambda I_2).$$

$$A - \lambda I_2 = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix}.$$

Assim, $P(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)-6 = \lambda^2-3\lambda-4 = (\lambda+1)(\lambda-4).$

Logo, $P(\lambda) = (\lambda+1)(\lambda-4)$ admite os valores $\lambda_1 = -1$ e $\lambda_2 = 4$ como raízes.
Portanto, -1 e 4 são autovalores da matriz A, ambos de multiplicidade algébrica 1.

Os autovetores (x, y) em $\mathbb{R}^2$ associados a estes autovalores são encontrados resolvendo-se os sistemas obtidos em $Av = \lambda v$, $v \neq 0$, para cada um destes autovalores, isto é,

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} . \begin{bmatrix} x \\ y \end{bmatrix} = \lambda . \begin{bmatrix} x \\ y \end{bmatrix}.$$

Para $\lambda = -1$, temos: $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -1 \begin{bmatrix} x \\ y \end{bmatrix}.$

$$\begin{bmatrix} x+2y \\ 3x+2y \end{bmatrix} = \begin{bmatrix} -x \\ 3x+2y+y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 2x+2y \\ 3x+3y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Logo, $\begin{cases} 2x+2y = 0 \\ 3x+3y = 0 \end{cases} \iff \begin{cases} x+y = 0 \end{cases}$.

Resolvendo o sistema homogêneo indeterminado, temos $y = -x$.
Logo, todo vetor da forma $v = (x, -x)$ em $\mathbb{R}^2$, com $x \neq 0$, é um autovetor associado ao autovalor -1.
Mas, $v = (x, -x) = x(1, -1)$. Assim, $B_{V(-1)} = \{(1, -1)\}$ constitui uma base do autoespaço $V(-1)$, ou seja $V(-1) = \big[(1, -1)\big]$.
E o autovalor -1 possui multiplicidade geométrica igual a 1.

Para $\lambda = 4$, temos: $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 4 \begin{bmatrix} x \\ y \end{bmatrix}.$

$$\begin{bmatrix} x+2y \\ 3x+2y \end{bmatrix} = \begin{bmatrix} 4x \\ 4y \end{bmatrix} \iff \begin{bmatrix} x+2y-4x \\ 3x+2y-4x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} -3x+2y \\ 3x-2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Logo, $\begin{cases} -3x+2y = 0 \\ 3x-2y = 0 \end{cases} \iff \begin{cases} 3x-2y = 0 \end{cases}$.

Resolvendo o sistema linear homogêneo indeterminado, temos $2y = 3x$. Assim, $y = \frac{3}{2}x$.
Logo, todo vetor da forma $u = (x, \frac{3}{2}x)$ em $\mathbb{R}^2$, com $x \neq 0$, é um autovetor associado ao autovalor 4.
Mas, $u = (x, \frac{3}{2}x) = \frac{x}{2}(2, 3)$. Assim, $B_{V(4)} = \{(2, 3)\}$ constitui uma base do autoespaço $V(4)$, ou seja $V(4) = \big[(2, 3)\big]$.
E o autovalor 4 possui multiplicidade geométrica igual a 1.
Neste exemplo podemos observamos que:
Os vetores $v = (1, -1)$ e $u = (2, 3)$ das bases dos autoespaços $B_{V(-1)}$ e $B_{V(4)}$, respectivamente, são LI.
Logo, $B' = \{(1, -1), (2, 3)\}$ forma uma base do espaço vetorial $V = \mathbb{R}^2$. Além disso, $V(-1) \oplus V(4) = \mathbb{R}^2$.
Isto nem sempre ocorre, e é interessante sabermos quais as condições para que isto aconteça.

Aproveitando o exemplo, determinemos a matriz de T com relação à base B':

$T(v) = T(1,-1) = -(1,-1) = (-1,1) = \mathbf{-1}.(1,-1) + \mathbf{0}.(2,3) = (-1,0)_{B'}$.
$T(u) = T(2,3) = 4(2,3) = (8,12) = \mathbf{0}.(1,-1) + \mathbf{4}.(2,3) = (0,4)_{B'}$.

Observamos que, $\big[T\big]_{B'} = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ é uma matriz diagonal formada pelos autovalores de A.

Mais adiante, veremos que as matrizes $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ e $\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ são semelhantes. Porém, trabalharmos com a matriz diagonal se torna mais interessante, pois ela é mais simples de ser manipulada.

Exemplo 3.11. Determinemos todos os autovalores e autovetores, e uma base de cada autoespaço do operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$T(x,y,z) = (2x+y,\ y-z,\ 2y+4z).$$

Inicialmente, encontremos a representação matricial de T, em relação à base canônica do $\mathbb{R}^3$:

$$T(x,y,z) = (2x+y,\ y-z,\ 2y+4z) = x(2,0,0) + y(1,1,2) + z(0,-1,4).$$

$$A = \big[T\big]_B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}.$$

Obtenção dos autovalores da matriz A:

$$P(\lambda) = det(A - \lambda I_3).$$

$$A - \lambda I_3 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{bmatrix}.$$

Então, $P(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)(4-\lambda) + 2(2-\lambda) =$

$$= (2-\lambda)(\lambda^2 - 5\lambda + 6) = (2-\lambda)(\lambda-2)(\lambda-3)\ .$$

Logo, $P(\lambda)$ admite duas raízes: $\lambda_1 = 3$ e $\lambda_2 = 2$ (raiz dupla).

Portanto, os autovalores de T são: 3 de multiplicidade algébrica 1 e 2 de multiplicidade algébrica 2 .

Obtenção dos autovetores $(x,y,z) \in \mathbb{R}^3$ associados aos autovalores encontrados:

$$Av = \lambda v$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Para $\lambda = 3$, temos:

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff \begin{bmatrix} 2x+y \\ y-z \\ 2y+4z \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \\ 3z \end{bmatrix}.$$

$$\begin{bmatrix} 2x+y-3x \\ y-z-3y \\ 2y+4z-3z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} -x+y \\ -2y-z \\ 2y+z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Então, $\begin{cases} -x+y=0 \\ -2y-z=0 \\ 2y+z=0 \end{cases} \iff \begin{cases} x-y=0 \ \ (1) \\ 2y+z=0 \ \ (2) \end{cases}.$

Resolvendo o sistema linear acima temos: da equação (1) que $y = x$, que substítuido na equação (2) resulta $z = -2x$.

Assim, o sistema admite como solução: $y = x$, $z = -2x$ e x é qualquer número real diferente de 0.

Os vetores da forma $v = (x, x, \ -2x) \in \mathbb{R}^3$, $x \neq 0$, são autovetores de T associado ao autovalor 3.

Mas, $v = (x, x, \ -2x) = x(1, 1, -2)$. Logo, $B_{V(3)} = \{ \ (1, 1, -2) \ \}$ é uma base do autoespaço $V(3)$, ou seja, $V(3) = \big[(1, 1, -2)\big]$.

E o autovalor 3 possui multiplicidade geométrica igual a 1.

Para $\lambda = 2$, temos:

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff \begin{bmatrix} 2x+y \\ y-z \\ 2y+4z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}.$$

$$\begin{bmatrix} 2x+y-2x \\ y-z-2y \\ 2y+4z-2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} y \\ -y-z \\ 2y+2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Então, $\begin{cases} y=0 \\ -y-z=0 \\ 2y+2z=0 \end{cases} \iff \begin{cases} y=0 \\ y+z=0 \end{cases}.$

Resolvendo o sistema linear acima temos: $y = 0$, $z = 0$ e x é qualquer real, $x \neq 0$.

Os vetores da forma $u = (x, 0, 0) \in \mathbb{R}^3$, $x \neq 0$, são autovetores de T associado ao autovalor 2 .

Mas, $u = (x, 0, 0) = x(1, 0, 0)$. Logo, $B_{V(2)} = \{ \ (1, 0, 0) \ \}$ é uma base do autoespaço $V(2)$, ou seja, $V(2) = \big[(1, 0, 0)\big]$.

E o autovalor 2 possui multiplicidade geométrica igual a 1.

Analisando este exemplo como fizemos com o exemplo anterior, observamos que: os autoespaços encontrados forneceram somente dois vetores linearmente independentes $(1, 2, -4)$ e $(1, 0, 0)$, e portanto o conjunto formado por eles não formam uma base de $V = \mathbb{R}^3$.

Exemplo 3.12. Determinemos os autovalores e autovetores, e a base de cada auto espaço da matriz triangular superior $\text{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$.

Obtenção dos autovalores da matriz A:

$$P(\lambda) = det(A - \lambda I_3).$$

$$A - \lambda I_3 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 3-\lambda & 0 & 2 \\ 0 & 3-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{bmatrix}.$$

Então,

$$P(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (1-\lambda)(1-\lambda)(3-\lambda) = (1-\lambda)^2.(3-\lambda)\ .$$

Logo, $P(\lambda)$ admite raízes: $\lambda_1 = 1$ (raiz dupla) e $\lambda_2 = 3$.

Portanto, os autovalores de T são: 1 de multiplicidade algébrica 2 e 3 de multiplicidade algébrica 1 .

Obtenção dos autovetores $(x, y, z) \in \mathbb{R}^3$ associados aos autovalores encontrados: $Av = \lambda v$.

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Para $\lambda = 1$, temos:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff \begin{bmatrix} x+2z \\ y+z \\ 3z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

$$\begin{bmatrix} x+2z-x \\ y+z-y \\ 3z-z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 2z \\ z \\ 2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad \text{Então, } \left\{ z = 0 \right. \ .$$

O sistema linear apresenta duas variáveis livres x e y. E admite a solução $z = 0$, x e y números reais quaisquer, não ambos nulos.

Assim, todo vetor da forma $v = (\ x,\ y,\ 0\ )$ em $\mathbb{R}^3$, com $x \neq 0$ e $y \neq 0$ é autovetor da matriz A associado ao autovalor 1 .

Mas, $v = (\ x,\ y,\ 0\ ) = x(1,0,0) + y(0,1,0)$. Logo, $B_{V(1)} = \{\ (1,0,0), (0,1,0)\ \}$ é uma base do autoespaço $V(1)$, ou seja, $V(1) = \big[(1,1,0), (0,1,0)\big]$.

E o autovalor 1 possui multiplicidade geométrica igual a 2.

Para $\lambda = 3$, temos:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff \begin{bmatrix} x+2z \\ y+z \\ 3z \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \\ 3z \end{bmatrix}.$$

$$\begin{bmatrix} x+2z-3x \\ y+z-3y \\ 3z-3z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} -2x+2z \\ -2y+z \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

$$\text{Então ,} \quad \begin{cases} -2x + 2z = 0 \\ -2y + \ z = 0 \end{cases} \iff \begin{cases} -x + z = 0 \ \ (1) \\ -2y + z = 0 \ \ (2) \end{cases}.$$

Resolvendo o sistema linear obtemos da equação (1) que $z = x$, e da equação (2) que $z = 2y$. Portanto, $x = z = 2y$.

Assim, todo vetor da forma $u = (2y,\ y,\ 2y)$, com $y \neq 0$, é uma autovetor da matriz A associado ao autovalor 3.

Mas, $u = (2y,\ y,\ 2y) = y(2,\ 1,\ 2)$. Logo, $B_{V(3)} = \{\ (2,1,2)\ \}$ é uma base do autoespaço $V(3)$, ou seja, $V(3) = \Big[(2,\ 1,\ 2)\Big]$.

E o autovalor 3 possui multiplicidade geométrica igual a 1.

Neste exemplo, observamos que o conjunto de vetores $(1,1,0)$, $(0,1,0)$ e $(2,1,2)$ obtidos das bases dos autoespaços $V(1)$ e $V(3)$ são LI, e formam uma base do espaço vetorial $V = \mathbb{R}^3$. Além disso, $V(1) \oplus V(3) = \mathbb{R}^3$.

Observação 3.13. Se A é uma matriz triangular (superior ou inferior), então os seus autovalores são os elementos de sua diagonal principal.

De fato, pois o determinante de uma matriz triangular A é dado pelo produto dos elementos de sua diagonal principal, e como a matriz $A - \lambda I$ é triangular sendo sua diagonal formada pelos elementos da forma $a_{ii} - \lambda$, para $i = 1, 2, \ldots n$. Então,

$$p(\lambda) = det(A - \lambda I_n) = (a_{11} - \lambda)(a_{22} - \lambda) \cdot\ \ldots\ \cdot (a_{nn} - \lambda).$$

Logo, $a_{11}, a_{22}, \cdots, a_{nn}$ são raízes de $p(\lambda)$.

Portanto, os autovalores de uma matriz triangular são os elementos de sua diagonal principal.

Proposição 3.14. *Sejam A uma matriz de ordem n com autovalor λ associado ao autovetor v. Então,*

(a) A não é invertível se, e somente se, 0 é autovalor de A.

(b) Se A é invertível, então λ^{-1} é um autovalor de A^{-1} associado ao vetor v.

(c) λ^n é um autovalor de A^n e v é o seu autovetor associado.

Demonstração.

(a) Se A é uma matriz invertível, então $detA \neq 0$.

Suponhamos por contraposição que A admita um autovalor nulo.

Assim, da equação $det(A - \lambda I_n) = 0$, temos $det\ A = 0$.

Logo, A não invertível, que é absurdo, pois admitimos que A fosse invertível.

Portanto, 0 não é autovalor de A.

(b) Com A é invertível. Logo, pela (i) , $\lambda \neq 0$.

Seja λ um autovalor de A associado ao autovetor v, então $Av = \lambda v$.

Multiplicando por A^{-1} a ambos os lados desta última igualdade resulta:

$$v = A^{-1}\lambda v = \lambda A^{-1} v.$$

Portanto, $A^{-1}v = \lambda^{-1}v$, ou seja, λ^{-1} é autovalor de A^{-1}.

(c) Demonstremos por indução sobre n.

Para $n = 1$ a proposição é válida, pois $Av = \lambda v$.

Suponhamos por indução que a proposição seja válida para todo $n - 1$, com $n > 2$, ou seja, $A^{n-1}v = \lambda^{n-1}v$.

Multiplicando à esquerda ambos os membros desta última igualdade por A, temos:

$$A^n v = A(A^{n-1}v) = A(\lambda^{n-1}v) = \lambda^{n-1}(Av) = \lambda^{n-1}(\lambda v) = (\lambda^{n-1}\lambda)v = \lambda^n v$$

Portanto, $A^n v = \lambda^n v$.

□

Teorema 3.15. *Se $v_1, v_2, \ldots, v_n$ são autovetores de uma matriz A associados a autovalores distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente, então $v_1, v_2, \ldots, v_n$ são LI.*

Demonstração. Demonstraremos por indução sobre n.

Se $n = 1$, temos $a_1 v_1 = 0$, então que v_1 é LI, pois $v_1 \neq 0$. Logo, o teorema é válido.

Suponhamos por hipótese de indução que o teorema seja válido para $n-1$ vetores, ou seja, os vetores $v_1, v_2, \cdots v_{n-1}$ são LI, e vamos mostrar que também é válido para n vetores.

Seja $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0,\ onde\ a_i \in K,\ i = 1, 2, \ldots, n.$ (2.5)

Aplicando T na relação (2.5), obtemos pela linearidade:

$$a_1 T(v_1) + a_2 T(v_2) + \cdots + a_n T(v_n) = 0.$$

Mas, por hipótese, $T(v_i) = \lambda_i v_i$, para $i = 1, 2, \ldots, n$. Logo,

$$a_1 \lambda_1 v_1 + a_1 \lambda_2 v_2 + \cdots + a_n \lambda_n v_n. \tag{2.6}$$

Por outro lado, multiplicando a relação (2.5) por λ_n, temos:

$$a_1 \lambda_n v_1 + a_2 \lambda_n v_2 + \cdots + a_n \lambda_n v_n = 0. \tag{2.7}$$

Agora, subtraindo (2.7) de (2.6):

$$a_1(\lambda_1 - \lambda_n) v_1 + a_2(\lambda_2 - \lambda_n) v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n) v_n + a_n(\lambda_n - \lambda_n) v_n = 0.$$
$$a_1(\lambda_1 - \lambda_n) v_1 + a_2(\lambda_2 - \lambda_n) v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n) v_n = 0.$$

Por indução, cada um dos coeficientes acima é 0. Como os λ_i são distintos, então $\lambda_i - \lambda_n \neq 0$, $i \neq n$, para $i = 1, 2, \ldots, n-1$.

Portanto, $a_1, a_2, \ldots, a_{n-1} = 0$. Substituindo em (2.5), temos, $a_n v_n = 0$.

Logo, $a_n = 0$, pois $v_n \neq 0$ e, portanto, $v_1, v_2, \ldots, v_n$ são LI.

□

Corolário 3.16. *Seja $T : V \to V$ um operador linear. Se dim $V = n$ e T possui n autovalores distintos, então V possui uma base formada por autovetores de T.*

Demonstração. Pelo teorema anterior, n autovalores distintos implicam na existência de um conjunto de n autovetores $\{\ v_1, v_2, \ldots, v_n\}$ LI.

Como $\Big[\ v_1, v_2, \ldots, v_n\ \Big] \subset V$ e $dim \Big[\ v_1, v_2, \ldots, v_n \Big] = n = dim\ V$, temos que $\Big[\ v_1, v_2, \ldots, v_n\ \Big] = V$. Logo, $\{\ v_1, v_2, \ldots, v_n\}$ é uma base de V.

□

Teorema 3.17. *Se a matriz A, de ordem n, possui autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$, distintos. Então, a soma dos autoespaços de A é direta, isto é, para cada $j = 1, 2, \ldots, n$, temos*

$$V(\lambda_i) \cap [\ V(\lambda_1) + \cdots + V(\lambda_{i-1}) + V(\lambda_{i+1}) + \cdots + V(\lambda_n)\ ] = \{0\}.$$

Demonstração. Provaremos por indução sobre o número de autovalores.

Inicialmente mostremos que $V(\lambda_1) \cap V(\lambda_2) = \{0\}$.

Fixemos $u_1^{(1)}, \ldots, u_{m_1}^{(1)}$ uma base de $V(\lambda_1)$ e $u_1^{(2)}, \ldots, u_{m_2}^{(2)}$ uma base de $V(\lambda_2)$.

Se $v \in V(\lambda_1) \cap V(\lambda_2$, então

$$v = a_1^{(1)} u_1^{(1)} + \cdots + a_{m_1}^{(1)} u_{m_1}^{(1)} \ = \ a_1^{(2)} u_1^{(2)} + \cdots + a_{m_2}^{(2)} u_{m_2}^{(2)}. \tag{3.1}$$

Logo, $T(u)$ é dado por

$$a_1^{(1)} T(u_1^{(1)}) + \cdots + a_{m_1}^{(1)} T(u_{m_1}^{(1)}) \ = \ a_1^{(2)} T(u_1^{(2)}) + \cdots + a_{m_2}^{(2)} T(u_{m_2}^{(2)})$$

ou seja,

$$a_1^{(1)} \lambda_1 u_1^{(1)} + \cdots + a_{m_1}^{(1)} \lambda_1 u_{m_1}^{(1)} \ = \ a_1^{(2)} \lambda_2 u_1^{(2)} + \cdots + a_{m_2}^{(2)} \lambda_2 u_{m_2}^{(2)}. \tag{3.2}$$

Multiplicando a equação (2.1) por λ_1 e subtraindo-a da equação (2.2), obtemos:

$$a_1^{(2)} (\lambda_2 - \lambda_1) u_1^{(2)} + \cdots + a_{m_2}^{(2)} (\lambda_2 - \lambda_1) u_{m_2}^{(2)} = 0.$$

Mas, $u_1^{(2)}, \dots, u_{m_2}^{(2)}$ é uma base de $V(\lambda_2)$, então:

$$a_1^{(2)}(\lambda_2 - \lambda_1) = \cdots = a_{m_2}^{(2)}(\lambda_2 - \lambda_1) = 0.$$

E como $\lambda_1 \neq \lambda_2$, resulta que $a_1^{(2)} = \cdots = a_{m_2}^2 = 0$. Segue-se de (2.1) que $v = 0$.

Suponhamos agora, por indução, que a soma de $n-1$ autoespaços de T referentes a $n-1$ autovetores distintos seja direta. E vamos mostrar que este resultado é válido quando T apresentar n autovalores distintos.

Para cada $j = 1, 2, \dots, n$ escolhemos uma base B_j de $V(\lambda_j)$ formada por vetores que denotaremos por, $v_1^{(j)}, v_2^{(j)}, \dots, v_{m_j}^{(j)}$. Observe que, cada $v_i^{(j)}$ é uma autovetor associado ao autovalor λ_j e que m_j é a multiplicidade geométrica deste autovalor.

Se $u \in V(\lambda_j) \cap [\ V(\lambda_1) + \cdots + V(\lambda_{j-1}) + V(\lambda_{j+1}) + \cdots + V(\lambda_n)\ ]$.

Então,

$$u = a_1^{(j)}v_1^{(j)} + \cdots + a_{m_j}^{(j)}v_{m_j}^{(j)} = a_1^{(1)}v_1^{(1)} + \cdots + a_{m_{j-1}}^{(j-1)}v_{m_{j-1}}^{(j-1)} + a_1^{(j+1)}v_1^{(j+1)} + \cdots + a_{m_n}^{(n)}v_{m_n}^{(n)}. \quad (3.3)$$

Assim, $T(u)$ é dado por

$$a_1^{(j)}T(v_1^{(j)}) + \cdots + a_{m_j}^{(j)}T(v_{m_j}^{(j)}) = a_1^{(1)}T(v_1^{(1)}) + \cdots + a_{m_{j-1}}^{(j-1)}T(v_{m_{j-1}}^{(j-1)}) + \\ + a_1^{(j+1)}T(v_1^{(j+1)}) + \cdots + a_{m_n}^{(n)}T(v_{m_n}^{(n)}).$$

Ou seja,

$$\begin{aligned} a_1^{(j)}\lambda_j v_1^{(j)} + \cdots + a_{m_j}^{(j)}\lambda_j v_{m_j}^{(j)} &= a_1^{(1)}\lambda_1 v_1^{(1)} + \cdots + a_{m_{j-1}}^{(j-1)}\lambda_{m_{j-1}} v_{m_{j-1}}^{(j-1)} + \\ &\quad + a_1^{(j+1)}\lambda_{j+1} v_1^{(j+1)} + \cdots + a_{m_n}^{(n)}\lambda_n v_{m_n}^{(n)}. \qquad (3.4) \end{aligned}$$

Multiplicando a equação (2.3) por λ_j e subtraindo-a da equação (2.4), temos:

$$a_1^{(j)}(\lambda_1 - \lambda_j)v_1^{(1)} + \cdots + a_{m_{j-1}}^{(j-1)}(\lambda_{j-1} - \lambda_j)v_{m_{j-1}}^{(j-1)} + \\ + \ a_1^{(j+1)}(\lambda_{j+1} - \lambda_j)v_1^{(j+1)} + \cdots + a_{m_n}^{(n)}(\lambda_n - \lambda j)v_{m_n}^{(n)} = 0.$$

Usando a hipótese de indução e o fato que $\lambda_j \neq \lambda_i$, quando $i \neq j$, obtemos:

$$a_1^{(i)} = \cdots = a_{m_i}^{(i)} 0 \text{ , para todo } i = 1, 2, \dots, n.$$

Deste resultado e da equação (2.3) resulta que $u = 0$.

□

Exemplo 3.18. A matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ (Exemplo 2.10.) admite dois autovalores distintos: $\lambda_1 = -1$ e $\lambda_2 = 4$ e os respectivos autoespaços $V(-1) = \big[(1,-1)\big]$ e $V(4) = \big[(2,3)\big]$. E vimos que, $V(-1) \oplus V(4) = \mathbb{R}^2$.

Convém observar que o teorema não garante a recíproca, ou seja, que se a soma dos autoespaços de uma matriz é direta então todos os seus autovalores são distintos.

Esse fato, percebemos no Exemplo 2.12., com a matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ em que a soma dos autoespaços $V(1) = \big[(1,1,0),(0,1,0)\big]$ e $V(3) = \big[(2,1,2)\big]$ é direta, isto é, $V(1) \oplus V(3) = \mathbb{R}^3$. No entanto, nem todos os autovalores de A apresentam multiplicidade algébrica igual a 1, ou seja, não temos três autovalores distintos, apenas 2.

Definição 3.19. Sejam A e B matrizes de ordem n . Dizemos que, B é **semelhante** ou **similar** a A ; ou que B é obtida de A por uma transformação de semelhança, que denotamos por $A \sim B$, se existir uma matriz invertível P tal que $B = P^{-1}AP$.

Teorema 3.20. *Sejam $A, B \in M_{n\times n}(\mathbb{K})$. Se B é semelhante a A, então o polinômio característico de A é idêntico ao de B.*

Demonstração. Sejam $P_A(\lambda)$ e $P_B(\lambda)$ os polinômios característicos da matrizes A e B, respectivamente.

Como $A \sim B$, existe uma matriz invertível P tal que $A = P^{-1}BP$.

Mas, $P_A(\lambda) = det\ (A - \lambda I) = det\ (PBP^{-1} - \lambda I) = det\ (PBP^{-1} - \lambda PP^{-1}) =$
$= det\ (PB - P\lambda)P^{-1} = det\ [P(B - \lambda I)P^{-1}] =$
$= det\ P.\ det\ (B - \lambda I).\ det\ P^{-1} = det\ P.\ det\ (B - \lambda I).\ \frac{1}{detP} =$
$= det\ (B - \lambda I) = P_B(\lambda).$

□

Corolário 3.21. *Sejam $A, B \in M_{n\times n}(\mathbb{K})$ matrizes semelhantes. Então,*

(a) A e B representam o mesmo operador linear.
(b) $det\ A = det\ B$.
(c) possuem os mesmos autovalores, com a mesma multiplicidade; e também os mesmos autovetores associados a estes autovalores.
(d) A é invertível se, e somente se, B é invertível.
(e) $trA = trB$.
(f) $post(A) = post(B)$.

Observação 3.22. Dados um operador $T : V \to V$, B e C duas bases distintas de V, com $dimV < \infty$, então as matrizes $\left[T\right]_B$ e $\left[T\right]_C$ são semelhantes e vale:

$$\left[T\right]_B = M_{B\to C} \left[T\right]_C M_{C\to B}.$$

onde $M_{B\to C}$ é a matriz de mudança de base de B para C e $M_{C\to B}$ é a matriz de mudança de base de C para B.

3.3 Diagonalização de operadores

Muitos sistemas podem ser representados por matrizes, e em determinadas situações, não é fácil de serem manipuladas, então devemos encontrar matrizes mais simples possível para representar tais sistemas de modo a minimizar tempo e custo operacional. Uma das formas de obter esta simplificação consiste em diagonalizar a matriz dada, quando possível.

A diagonalização de matrizes quadradas consiste na obtenção de uma matriz diagonal que seja semelhante à matriz original dada. O interessante da obtenção de uma matriz diagonal é que todos os seus elementos fora da sua diagonal principal são iguais a zero, o que implica, matematicamente, em uma redução significativa no tempo de processamento utilizando estas matrizes. Além disso, uma matriz diagonal semelhante preserva todas as condições mencionadas no Corolário 2.21 em relação à matriz original; dentre essas condições, uma importante é preservar os autovalores, e veremos que estes irão compor a diagonal dessa matriz, o que significa que podemos determinar facilmente essas matrizes.

Dado um operador linear T em V, vimos que, a cada base B de V corresponde uma matriz $[T]_B$ que representa T nesta base. O objetivo agora será obter, quando possível, uma outra base B' de V de modo que a matriz do operador linear T nesta nova base seja uma matriz diagonal.

Definição 3.23. Dizemos que um operador linear T definido sobre um espaço vetorial V de dimensão finita sobre um corpo $\mathbb{K}$ é **diagonalizável**, se for possível representá-lo por uma matriz diagonal em alguma base de V.

Dado um operador T, os teoremas a seguir caracterizam base associada ao operador que se pretende diagonalizar.

Teorema 3.24. *Um operador $T : V \to V$ admite uma base B em relação à qual a matriz $\begin{bmatrix} T \end{bmatrix}_B$ é diagonal se, e somente se, essa base B for formada por autovetores de T.*

Demonstração. Suponhamos que T é diagonalizável, ou seja, exista uma base $B = \{v_1, v_2, \dots v_n\}$ de V tal que $\begin{bmatrix} T \end{bmatrix}_B$ é diagonal.

$$\begin{bmatrix} T \end{bmatrix}_B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Tomando-se um elemento v_i da base B, observe que, da maneira como a matriz que representa T com relação à base B, a coluna i da matriz $\begin{bmatrix} T \end{bmatrix}_B$ é formada pelos coeficientes $T(v_i)$ escrito como combinação linear dos elementos da base B, isto é,

$$T(v_i) = 0v_1 + \cdots + \lambda_i v_i + \cdots + 0v_n = \lambda_i v_i.$$

Assim, $T(v_i) = \lambda_i v_i$, para $i = 1, 2, \dots, n$. Logo, v_i é um autovetor de T associado ao autovalor λ_i.

Portanto, a base B é formada por autovetores de T.

Agora, suponhamos que $B = \{v_1, v_2, \dots v_n\}$ seja uma base para V formada por autovetores de T, ou seja, $T(v_i) = \lambda_i v_i$, para $i = 1, 2, \dots, n$, sendo $\lambda_1, \lambda_2, \dots, \lambda_n$ os respectivos autovalores do operador T.

Como as imagens dos elementos v_i da base B pelo operador linear T, escritos como combinação linear dos elementos de B, são da forma $T(v_i) = \lambda_i v_i$, resulta que

$$\begin{bmatrix} T \end{bmatrix}_B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

é a matriz que representa T com relação à base B, que é uma matriz diagonal. Portanto, T é um operador diagonalizável.

$\square$

Exemplo 3.25. A matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ (Exemplo 2.10.) admite a base $B' = \{(1, -1), (2, 3)\}$ de $\mathbb{R}^2$ formada de autovetores associados pelos autovalores -1 e 4, respectivamente. Logo, pelo teorema anterior, A é diagonalizável, e a matriz diagonal $\begin{bmatrix} T \end{bmatrix}_{B'} = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ formada pelos seus autovalores diagonaliza A.

As matrizes $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ e $\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ são semelhantes. Porém, a matriz diagonal se torna mais interessante, pois é mais simples de ser manipulada.

Exemplo 3.26. A matriz $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}$ do Exemplo 2.11. admite somente dois autovetores LI. Logo, não existe uma base de $\mathbb{R}^3$ constituída só de autovetores.

Portanto, A não é diagonalizável.

Exemplo 3.27. Consideremos a matriz $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & -5 \\ 0 & 1 & 1 \end{bmatrix}$.

Seu polinômio característico:

$$P(\lambda) = \begin{vmatrix} 4-\lambda & 0 & 0 \\ 0 & -1-\lambda & -5 \\ 0 & 1 & 1-\lambda \end{vmatrix} = (4-\lambda)(-1-\lambda)(1-\lambda) + 5(4-\lambda) =$$

$$= (4-\lambda)(-1-\lambda)(1-\lambda) + 5(4-\lambda) = (4-\lambda)(\lambda^2+4).$$

Portanto, $\lambda_1 = 4$, $\lambda_2 = 2i$ e $\lambda_3 = -2i$ são raízes de $P(\lambda)$.

Se A é uma matriz sobre o corpo $\mathbb{R}$, então A tem somente o autovalor 4.

Logo, um único autovetor e, portanto, A não é diagonalizável.

No caso de A ser uma matriz sobre o corpo $\mathbb{C}$, então A têm três autovalores distintos: 4 , 2i e -2i.

Então, A admite três auvetores LI e, portanto, A é diagonalizável.

Definição 3.28. Uma matriz A de ordem n é chamada de **diagonalizável** se existir alguma matriz invertível P tal que $P^{-1}AP$ é uma matriz diagonal. Neste caso, dizemos que a matriz P **diagonaliza** A.

Teorema 3.29. *Seja V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{K}$ e T um operador linear sobre V. Então, T é diagonalizável se, e somente se:*

(a) o polinômio característico de T possui todas as suas raízes em $\mathbb{K}$.

(b) a multiplicidade algébrica de cada autovalor de T é igual a sua multiplicidade geométrica.

Demonstração. (a) Se T é um operador diagonalizável, então V possui uma base formada por autovetores de T (Teorema 2.24).

Seja $B = \{v_{11}, v_{12}, \dots, v_{1r_1}, v_{21}, v_{22}, \dots, v_{2r_2}, \dots, v_{k1}, v_{k2}, \dots, v_{kr_k}\}$ essa base, de modo que em cada $\{v_{i1}, v_{i2}, \dots, v_{ir_i}\}$ estão todos os autovetores associados ao autovalor λ_i , para $i = 1, 2, \dots, k$

A matriz que representa T em relação à base B é a matriz diagonal:

$$\big[T\big]_B = \begin{bmatrix} \lambda_1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & \lambda_1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & \lambda_k & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & \lambda_k \end{bmatrix}.$$

$P(\lambda) = det\left(\left[T\right]_B - \lambda I\right) = (\lambda_1 - \lambda)^{r_1}(\lambda_2 - \lambda)^{r_2}\cdots(\lambda_k - \lambda)^{r_k}$ é o polinômio característico de T cujas raízes estão todas em $\mathbb{K}$.

Para cada índice i, seja U_i , $i = 1, 2, \ldots, k$, o subespaço gerado por $\{v_{i1}, v_{i2}, \ldots, v_{ir_i}\}$.

Um elemento $u \in U_i$ é combinação linear de $v_{i1}, v_{i2}, \ldots, v_{ir_i}$ e, portanto, é um autovetor associado ao autovalor λ_i, ou seja, $u \in V(\lambda_i)$. Logo, $U_i \subset V(\lambda_i)$.

Agora, tomando $u \in V(\lambda_i)$ como combinação linear dos elementos de B:

$$u = a_{11}v_{11} + \cdots + a_{kr_k}v_{kr_k}.$$

Multiplicando por λ_i e utilizando que $T(u) = \lambda_i u$, uma vez que $u \in V(\lambda_i)$, temos:

$$\begin{aligned}
&\lambda_i a_{11}v_{11} + \cdots + \lambda_i a_{1r_1}v_{1r_1} + \cdots + \lambda_i a_{k1}v_{k1} + \cdots + \lambda_i a_{kr_k}v_{kr_k} = \\
&= \lambda_i u = T(u) = a_{11}T(u_{11}) + \cdots + a_{kr_k}T(u_{kr_k}) = \\
&= \lambda_1 a_{11}v_{11} + \cdots + \lambda_1 a_{1r_1}v_{1r_1} + \cdots + \lambda_k a_{k1}v_{k1} + \cdots + \lambda_k a_{kr_k}v_{kr_k}.
\end{aligned}$$

Comparando a primeira e a última combinação linear acima, obtemos:

$$\begin{gathered}
\lambda_i a_{11} = \lambda_1 a_{11}, \ \ldots \ , \lambda_i a_{1r_1} = \lambda_1 a_{1r_1}. \\
\vdots \\
\lambda_i a_{k1} = \lambda_k a_{k1}, \ \ldots \ , \lambda_i a_{kr_k} = \lambda_k a_{kr_k}.
\end{gathered}$$

E assim, $a_{11} = \cdots = a_{1r_1} = \cdots = a_{(i-1)1} = \cdots = a_{(i-1)r_{(i-1)}} =$

$$= a_{(i+1)1} = \cdots = a_{(i+1)r_{(i+1)}} = \cdots = a_{k1} = \cdots = a_{kr_k} = 0.$$

Logo, $u = a_{i1}v_{i1} + \cdots + a_{ir_i}v_{ir_i}$ o que implica que $u \in U_i$.

Assim, $V(\lambda_i) \subset U_i$.

Dessa forma, mostramos que $U_i = V(\lambda_i)$, para $i = 1, 2, \ldots, k$.

Portanto, os autovalores de T têm a mesma multiplicidade algébrica e geométrica.

Reciprocamente, sabemos que os autovalores de T são as raízes do polinômio característico de T.

Como V tem dimensão n, o polinômio característico de T terá grau n, e se ele possuir todas as suas raízes em $\mathbb{K}$, então pelo Teorema Fundamental da Álgebra temos que T possui n autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$, não necessariamente, distintos.

Mas a multiplicidade algébrica de cada autovalor de T é igual a sua multiplicidade geométrica, então é possível associarmos a um mesmo autovalor λ de T um conjunto LI formado por autovetores, que será uma base para o autoespaço $V(\lambda)$.

Além disso, os autovetores associados a autovalores distintos também serão LI, e como a dimensão de V é n, temos que $\{v_1, v_2, \ldots v_n\}$ é uma base para V.

Portanto, T é um operador diagonalizável.

□

Exemplo 3.30. A matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ (Exemplo 2.12) possui autovalor 1 com multiplicidade 2, auto espaço $V(1) = \left[(1,0,0),(0,1,0)\right]$, um autovalor 3 e autoespaço $V(3) = \left[(2,1,2)\right]$.

Observamos que, os autovalores possuem multiplicidades aritmética iguais às geométricas. Portanto, pelo teorema anterior, A é diagonalizável.

Exemplo 3.31. A matriz $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix}$ (Exemplo 2.11.) admite autovalor 2, de multiplicidade algébrica 2, mas o autoespaço associado $V(2) = \Big[(1,1,0)\Big]$ possui multiplicidade geométrica 1. Portanto, pelo teorema anterior, a matriz A não é diagonalizável.

Exemplo 3.32. Consideremos a matriz $A = \begin{bmatrix} 2 & 5 & 4 & 7 \\ 0 & 1 & 3 & 9 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

Como a matriz A é triangular, pela Observação 2.13., então os seus autovalores são os elementos da diagonal principal, $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 8$ e $\lambda_4 = -1$.

Como os autovalores são todos distintos, então A é diagonalizável. (Corolário 2.16. e Teorema 2.15.).

Exemplo 3.33. Determinemos a matriz P que diagonaliza a matriz $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 3 & 0 \\ 1 & 2 & 4 \end{bmatrix}$.

Calculemos os autovalores de A:

$$A - \lambda I = \begin{bmatrix} 4-\lambda & 2 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 2 & 4-\lambda \end{bmatrix}.$$

$$P(\lambda) = det(A - \lambda I) = (4-\lambda)(3-\lambda)(4-\lambda) - (3-\lambda) = (3-\lambda)(\lambda^2 - 8\lambda + 15) =$$
$$= (3-\lambda)(\lambda-3)(\lambda-5) = (\lambda-3)^2(\lambda-5).$$

Logo, os autovalores da matriz A são $\lambda_1 = \lambda_2 = 3$ e $\lambda_3 = 5$ que são as raízes do polinômio característico .

$$(A - \lambda I_3)v = 0 \ , \quad \text{ou seja ,} \quad \begin{bmatrix} 4-\lambda & 2 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 2 & 4-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Para $\lambda = 3$, temos $\begin{bmatrix} 4-3 & 2 & 1 \\ 0 & 3-3 & 0 \\ 1 & 2 & 4-3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} x + 2y + z = 0 \\ x + 2y + z = 0 \end{cases} \iff \Big\{ x + 2y + z = 0.$$

Resolvendo o sistema linear homogêneo, temos como solução: $z = -x-2y$, e $x, y \in \mathbb{R}$, não ambos nulos.

Assim, todo vetor $v = (x,\ y,\ -x-2y) \in \mathbb{R}^3$ com $x \neq 0$ e $y \neq 0$ é um autovetor da matriz A associado ao autovalor 3.

Mas, $v = (x,\ y,\ -x-2y) = x(1,0,-1) + y(0,1,-2)$.

Logo, $B_{V(3)} = \{(1,0,-1),(0,1,-2)\}$ é uma base do autoespaço $V(3)$.

Para $\lambda = 5$, temos $\begin{bmatrix} 4-5 & 2 & 1 \\ 0 & 3-5 & 0 \\ 1 & 2 & 4-5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

$$\begin{bmatrix} -1 & 2 & 1 \\ 0 & -2 & 0 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} -x+2y+z=0 \\ -2y=0 \\ x+2y-z=0 \end{cases} \iff \begin{cases} x-z=0 \\ y=0 \end{cases}.$$

Resolvendo o sistema linear homogêneo, obtemos $x = z$ e $y = 0$.

Assim, todo vetor $u = (x,\ 0,\ x) \in \mathbb{R}^3$ e $x \neq 0$ é um autovetor da matriz A associado ao autovalor 5.

Mas, $u = (x, 0, x) = x(1, 0, 1)$.

Logo, $B_{V(5)} = \{(1, 0, 1)\}$ é uma base do autoespaço $V(5)$.

Dessa forma, temos três autovetores $v_1 = (1, 0, -1)$, $v_2 = (0, 1, -2)$ e $v_3 = (1, 0, 1)$ linearmente independentes pois pertencem às bases dos autoespaços de A.

Construção da matriz P:

$$P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \text{ cuja inversa é } P^{-1} = \begin{bmatrix} \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}.$$

Obtenção da matriz diagonal D a partir dos autovalores de A.

$$D = P^{-1}AP = \begin{bmatrix} \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix} . \begin{bmatrix} 4 & 2 & 1 \\ 0 & 3 & 0 \\ 1 & 2 & 4 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}.$$

$$D = \begin{bmatrix} \frac{3}{2} & -3 & -\frac{3}{2} \\ 0 & 3 & 0 \\ \frac{5}{2} & 5 & \frac{5}{2} \end{bmatrix} . \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Em geral, não existe uma ordem preferencial para as colunas de P. Se trocarmos a ordem das colunas na construção da matriz P, alteramos a ordem dos autovalores na diagonal da matriz $P^{-1}AP$. Assim, por exemplo, se tivéssemos construído

$$P = \begin{bmatrix} v_1 & v_3 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & -2 \end{bmatrix} \quad \text{então} \quad D = P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Observação 3.34. A seguir provaremos importantes resultados envolvendo matrizes simétricas. Tais matrizes desempenham papel relevante no estudo da diagonalização de operadores, sendo uma de suas aplicações no reconhecimento de cônicas e quadráticas. Resumiremos alguns resultados marcantes entre os casos simétrico e não simétrico.

(a) se A é uma matriz simétrica todos os seus autovalores são reais, cada autovalor possui multiplicidade aritmética igual a geométrica, e A é diagonalizável.

(b) se A é uma matriz não-simétrica seus autovalores nem sempre são todos reais, e mesmo que todos sejam reais é possível que a matriz A não seja diagonalizável. É o caso em que um autovalor tenha multiplicidade geométrica menor que à aritmética.

(c) se a matriz A é simétrica então autovetores associados a autovalores distintos são ortogonais, o que não acontece sendo A não-simétrica.

Definição 3.35. Um operador $T : V \to V$ sobre um corpo $\mathbb{K}$ admite um operador adjunto $T^* : V \to V$ quando

$$\langle\ T(u)\ ,\ v\ \rangle = \langle\ u\ , T^*(v)\ \rangle\ ,\ \ \forall u, v \in V.$$

Dizemos que T é autoadjunto quando $T^* = T$.

Observação 3.36. Seja B uma base ortonormal de V e $T : V \to V$ um operador autoadjunto, então

(a) Quando $\mathbb{K} = \mathbb{R}$, $\big[T\big]_B$ é simétrica e denominamos o operador T autoadjunto de simétria.

(b) Quando $\mathbb{K} = \mathbb{C}$, $\big[T\big]_B$ é hermitiana, isto é, $\big[T\big]_B^T = \big[T\big]_B^T$ e denominamos o operador T autoadjunto hermitiano.

Portanto, temos um critério prático para determinarmos se um dado operador T é autoadjunto. Basta considerarmos qualquer base ortonormal B de V e verificarmos se a matriz $\big[T\big]_B$ é simétrica (caso real) ou é hermitiana (caso complexo).

Exemplo 3.37. Consideremos B a base canônica do $\mathbb{R}^3$, que é uma base ortornormal, e o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$T(x, y, z) = (3x + y - 2z,\ x + 2y + 4z,\ -2x + 4y + 5z).$$

Então, $\big[T\big] = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 2 & 4 \\ -2 & 4 & 5 \end{bmatrix}$ é uma matriz simétrica.

Portanto, T é um operador autoadjunto.

Exemplo 3.38. Determinemos um operador em $\mathbb{R}^2$ que seja autoadjunto.

De acordo com a definição, basta tomarmos qualquer matriz simétrica de ordem 2. Por exemplo, considerando a matriz simétrica $A = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$, podemos definir um operador autoadjunto como sendo $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T_A(x, y) = (5x + 3y, 3x + 2y)$, sendo B a base canônica de $\mathbb{R}^2$.

Teorema 3.39. *Se $T : V \to V$ é um operador autoadjunto, então os seus autovalores são reais.*

Demonstração. Seja λ um autovalor associado ao autovetor v, então $Av = \lambda v$.

Suponhamos $\lambda \in \mathbb{C}$ seja um autovalor de T associado ao autovetor v, ou seja,

$$T(v) = \lambda v.$$

Como T é um operador autoadjunto, temos:

$$\langle T(v), v\rangle = \langle v, T(v)\rangle.$$

Mas, $\langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle$ e $\langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$.

Assim, $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$, então $(\lambda - \overline{\lambda})\langle v, v\rangle = 0$.

Como $\langle v, v\rangle \neq 0$, temos que, $(\lambda - \overline{\lambda}) = 0$. Logo, $\lambda = \overline{\lambda}$.

Portanto, λ é real. □

Teorema 3.40.

Seja T é um operador autoadjunto. Se $\lambda_1, \lambda_2, \ldots, \lambda_n$ são autovalores de T dois a dois distintos, então os autovetores associados $v_1, v_2, \ldots, v_n$ são dois a dois ortogonais.

Demonstração. Seja $i \neq j$, então

$$\lambda_i\langle v_i, v_j\rangle = \langle \lambda_i v_i, v_j\rangle = \langle T(v_i), v_j\rangle = \langle v_i, T(v_j)\rangle = \langle v_i, \lambda_j v_j\rangle = \lambda_j\langle v_i, v_j\rangle.$$

Logo, $\lambda_i\langle v_i, v_j\rangle = \lambda_j\langle v_i, v_j\rangle$, então $\lambda_i\langle v_i, v_j\rangle - \lambda_j\langle v_i, v_j\rangle = 0$.

Assim, $(\lambda_i - \lambda_j)\langle v_i, v_j\rangle = 0$.

Como, $\lambda_i \neq \lambda_j$, então $\langle v_i, v_j\rangle = 0$.

Portanto, v_i é ortogonal a v_j, para todo $i \neq j$. □

Teorema 3.41. *Seja T um operador autoadjunto. Então, a dimensão do autoespaço associado a um autovalor é igual à multiplicidade deste autovalor.*

Proposição 3.42. *Se T é um operador autoadjunto e $\lambda_1, \lambda_2, \ldots, \lambda_k$ são os autovalores de T. Então, $V = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_k)$. Além disso, para todo $i \neq j$, $V(\lambda_i)$ e $V(\lambda_j)$ são ortogonais.*

Demonstração. Para provarmos que V é soma direta dos autoespaços

$$V(\lambda_1),\ V(\lambda_2),\ \ldots\ ,V(\lambda_k),$$

vamos mostrar que se $v_1 + v_2 + \cdots + v_k = 0$, com $v_j \in V(\lambda_j)$, para $j = 1, 2, \ldots, k$, então todos os v_j são nulos, para todo $j = 1, 2, \ldots, k$.

Suponhamos que $0 = v_1 + v_2 + \cdots + v_k$, $v_j \in V(\lambda_j)$, para $i = 1, 2, \ldots, n$. Mas, $\langle 0, v_j\rangle = \langle v_1 + v_2 + \cdots + v_k\ , v_j\rangle = \langle v_1, v_j\rangle + \langle v_2, v_j\rangle + \cdots + \langle v_j, v_j\rangle + \cdots + \langle v_k, v_j\rangle$.

Como $\langle v_i, v_j\rangle = 0$, para todo $i \neq j$.

Então, $\langle v_j, v_j\rangle = 0$, o que implica que $v_j = 0$, $\forall j$.

Portanto, V é soma direta dos autoespaços $V(\lambda_1), V(\lambda_2), \ldots, V(\lambda_k)$.

E todo vetor v em V, pode ser escrito de maneira única como sendo $v = v_1 + v_2 + \cdots + v_k$ com $v_j \in V(\lambda_j)$.

□

O próximo teorema, um dos mais importantes da Álgebra Linear, garante que todo operador autoadjunto é diagonalizável.

Teorema 3.43. *(Teorema Espectral) Sejam V um espaço vetorial de dimensão n e $T : V \to V$ um operador autoadjunto. Então, existe uma base B ortonormal de V formada por autovetores de T. Em particular, T é ortogonalmente diagonalizável.*

Demonstração. Demonstraremos por indução sobre a dimensão do espaço vetorial V.

Se $dim\ V = 1$, então, existe $v \neq 0$ em V, e todo vetor de V é um múltiplo escalar de v, ou seja, $V = \left[v\right]$.

Então $T : V \to V$ é tal que $T(v) = kv$. Logo, v é um autovetor de T. Portanto, $B = \{\frac{v}{\|v\|}\}$ é uma base ortonormal de V.

Suponhamos por hipótese de indução que o teorema seja verdadeiro para todo espaço vetorial com dimensão menor que a dimensão de V.

Seja $v_1 \in V$, $v_1 \neq 0$, um autovetor de T, e consideremos o subespaço $S = \left[v_1\right]$, cuja dimensão é 1.

Como $V = S \oplus S^{\perp}$, então $dim\ S^{\perp} = n - 1 < dim\ V = n$.

Consideremos $T|_{S^{\perp}} : S^{\perp} \to S^{\perp}$, a restrição do operador auto-adjunto T no subespaço $S^{\perp}$, e verifiquemos que para todo $v \in S^{\perp}$ temos $T(v) \in S^{\perp}$.

Seja $u \in S$, então $u = cv_1$.

Como T é operador autoadjunto segue que

$$\begin{aligned}\langle T(v), u\rangle = \langle T(v), cv_1\rangle = c\langle v, T(v_1)\rangle = \\ = c\ \langle v, \lambda_1 v_1\rangle = c\lambda_1\langle v, v_1\rangle = 0\ \text{, pois } v \in S^{\perp} \text{ e } v_1.\end{aligned}$$

Assim, $T|_{S^{\perp}}$ é um operador linear em $S^{\perp}$, também autoadjunto possuindo os mesmos autovalores e autovetores de T.

Pela hipótese de indução o espaço vetorial $S^{\perp}$ tem uma base $B = \{v_2, v_3, \ldots, v_n\}$ ortonormal de autovetores de T.

Como $V = S \oplus S^{\perp}$, então $B' = \{\frac{v_1}{\|v_1\|}, v_2, \ldots, v_n\}$ é uma base ortonormal de autovetores de T. □

Em outras palavras, toda matriz simétrica é diagonalizável. Caso seus autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$ sejam todos distintos, os autovetores associados $\{v_i\}$ serão ortogonais, isto é, $\langle v_i \, , v_j\rangle = 0$, se $i \neq j$, e a matriz P será neste caso formada pelos n autovetores $\{v_i\}$ normalizados (para garantir que P seja ortogonal), e na mesma ordem dos autovalores.

Observação 3.44. O Teorema Espectral nos permite escrever uma matriz simétrica real A de ordem n na forma $A = PDP^T$, sendo P uma matriz ortogonal e D uma matriz diagonal. Os elementos na diagonal de D são os autovalores de A, e as colunas de P são os vetores ortonormais $u_1, u_2, \ldots, u_n$. Usando a representação linha-coluna do produto, temos:

$$A = PDP^T = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}.$$

$$A = \begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n v_n \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}.$$

$A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T$, que é chamada **decomposição espectral da matriz A**.

Podemos trocar a matriz A por um operador $T : V \to V$ autoadjunto, que teremos a mesma conclusão.

Exemplo 3.45. Determinemos a decomposição espectral da matriz simétrica

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}.$$

Polinômio característico de A é dado por

$$P(\lambda) = \begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - 4\lambda + 3) = (\lambda-2)(\lambda-1)(\lambda-3).$$

Logo, os autovalores de A são: $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 3$.

Determinando os autovetores de A obtemos:

$v_1 = (0, 1, 0)$ associado ao autovalor 2; $v_2 = (1, 0, -1)$ associado ao autovalor 1; e $v_3 = (1, 0, 1)$ associado ao autovalor 3.

Então, $u_1 = \frac{v_1}{\|v_1\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $u_2 = \frac{v_2}{\|v_2\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ e $u_3 = \frac{v_3}{\|v_3\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.

Logo, $u_1^T = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$; $u_2^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}$ e $u_3^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$.

$$A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \lambda_3 u_3 u_3^T.$$

Portanto,

$$A = 2. \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} . \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + 1. \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix} + 3. \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

é a decomposição espectral da matriz A.

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{3}{2} & 0 & \frac{3}{2} \\ 0 & 0 & 0 \\ \frac{3}{2} & 0 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}.$$

A decomposição espectral expressa explicitamente uma matriz simétrica em termos de seus autovalores e autovetores. Isto nos fornece uma maneira de se construir uma matriz com autovalores e autovetores ortonormais dados.

Exemplo 3.46. Determinemos uma matriz simétrica A de ordem 2 com autovalores $\lambda_1 = 2$ e $\lambda_2 = -3$, e autovetores ortogonais associados $v_1 = (1, -2)$ e $v_2 = (2, 1)$, respectivamente.

$$u_1 = \frac{v_1}{\|v_1\|} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix} \Rightarrow u_1^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix} \quad \text{e} \quad u_2 = \frac{v_2}{\|v_2\|} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \Rightarrow u_2^T \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

$$A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T \; = \; 2. \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix} - 3. \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{2}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{8}{5} \end{bmatrix} - \begin{bmatrix} \frac{12}{5} & \frac{6}{5} \\ \frac{6}{5} & \frac{3}{5} \end{bmatrix}. \qquad \text{Portanto,} \qquad A = \begin{bmatrix} -2 & -2 \\ -2 & 1 \end{bmatrix}.$$

Para verificarmos quando um operador linear pode ser diagonalizável ou não, vimos que isto pode ser feito através da obtenção de uma base de autovetores, ou mostrando a inexistência desta base. Para espaços vetoriais de baixa dimensão este processo é conveniente. Entretanto, para espaços de dimensão elevada este procedimento pode ser trabalhoso, e nestes casos vamos utilizar um outro método que veremos a seguir.

Definição 3.47. Se $P(x) = a_n x^n + \cdots + a_1 x + a_0$ um polinômio e A uma matriz quadrada de ordem n. Definimos $P(A)$ como sendo a matriz $\;P(A) = a_n A^n + \cdots + a_1 A + a_0 I_n$.

Definição 3.48. Seja A uma matriz quadrada de ordem n. O polinômio

$$m(x) = x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0$$

é chamado de polinômio minimal de A, quando:

(a) $m(A) = 0$, ou seja, a matriz A anula $m(x)$.

(b) $m(x)$ é o polinômio de menor grau entre aqueles que são anulados pela matriz A.

Os próximos teoremas não serão provados, mas as provas podem ser encontradas na referência [2].

Teorema 3.49. *Sejam V um espaço vetorial de dimensão n, $T : V \to V$ um operador linear e B uma base qualquer de V. Então, T é diagonalizável se, e somente se, o polinômio minimal de $\left[T\right]_B$ é da forma $\;m(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_r)$, com $\lambda_1, \lambda_2, \cdots, \lambda_r$ autovalores distintos.*

Desta forma, o problema da determinação se um operador linear T é diagonalizável se reduz a encontrar o polinômio minimal de T.

Teorema 3.50. *(Cayley-Hamilton) Sejam $T : V \to V$ um operador linear, B uma base de V e $P(\lambda)$ o polinômio característico de T. Então, $P(\,[T]_B) = 0$.*

O Teorema Cayley-Hamilton garante que o polinômio característico é um candidato a ser o polinômio minimal, pois ele satisfaz a Definição 2.48.(a) .

Exemplo 3.51. Consideremos a matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ (Exemplo 2.10) que admite autovalores: $\lambda_1 = -1$ e $\lambda_2 = 4$, associados aos autovetores $v_1 = (1,-1)$ e $v_2 = (2,3)$, respectivamente.

O polinômio característico de A é dado por $P(\lambda) = \lambda^2 - 3\lambda - 4 = (\lambda+1)(\lambda-4)$, e o Teorema de Cayley-Hamilton garante que a matriz A é um zero de $P(\lambda)$.

De fato, $P(A) = A^2 - 3A - 4I_2 = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}^2 - 3\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} - 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} =$

$$= \begin{bmatrix} 7 & 6 \\ 9 & 10 \end{bmatrix} - \begin{bmatrix} 3 & 6 \\ 9 & 6 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Neste caso, $P(\lambda) = m(\lambda) = (\lambda+1)(\lambda-4)$, e a matriz A é diagonalizável (Teorema 2.49.).

Considerando a base $B = \{(1,-1),(2,3)\}$ formada pelos autovetores A temos que nesta base $\begin{bmatrix} T \end{bmatrix}_B = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$.

Teorema 3.52. *Os polinômios característicos e minimal têm os mesmos fatores irredutíveis. Em particular, $m(\lambda)$ divide o polinômio característico $P(\lambda)$ de A.*

Assim, este teorema garante que: qualquer fator irredutível de um dos polinômios deve dividir o outro polinômio. Em particular, como um fator linear é irredutível, $m(x)$ e $P(x)$ têm os mesmos fatores lineares. Logo, têm as mesmas raízes.

Teorema 3.53. *Um escalar λ é um autovalor de uma matriz A se, e somente se, λ é uma raiz do polinômio minimal de A.*

Exemplo 3.54. Consideremos $T : V \to V$ um operador linear e B uma base de V. Suponhamos que o polinômio característico de T seja $P(\lambda) = (\lambda-2)^3(\lambda-1)^2(\lambda-5)$. Então, seu polinômio minimal deverá ser um dos polinômios da forma:

$$m(\lambda) = (\lambda-2)^r(\lambda-1)^s(\lambda-5)\ ,\ \ r \in \{1,2,3\} \text{ e } s \in \{1,2\}.$$

Como o polinômio minimal é o de menor grau, verifiquemos inicialmente para $r = s = 1$, ou seja, $P_1(\lambda) = (\lambda-2)(\lambda-1)(\lambda-5)$.

Se $P_1([T]_B) = 0$, então ele é o minimal. Caso contrário, testamos o próximo de menor grau dos polinômios que restaram.

Teorema 3.55. *Sejam $\lambda_1, \lambda_2, \ldots, \lambda_r$ autovalores distintos de um operador linear T. Então, T é diagonalizável se, e somente se, o polinômio $(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_r)$ anular a matriz de T.*

Exemplo 3.56. Verifiquemos se o operador linear $T : \mathbb{R}^4 \to \mathbb{R}^4$ definido por
$T(x,y,z,t) = (x,\ y+z,\ 3z,\ 3t)$ é diagonalizável.

$$T(x,y,z,t) = (x,\ y+z,\ 3z,\ 3t) = x(1,0,0,0) + y(0,1,0,0) + z(0,1,3,0) + t(0,0,0,3)$$

$$\big[T\big]_B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Como a matriz é triangular, então os autovalores de T são: $\lambda_1 = 1$, $\lambda_2 = 3$, ambos de multiplicidades algébricas 2.

Logo, o polinômio característico de T é dado por $P(\lambda) = (\lambda - 1)^2(\lambda - 3)^2$.

Desta forma, os possíveis candidatos a polinômio minimal são:

$P_1(\lambda) = (\lambda - 1)(\lambda - 3)$, $\quad P_2(\lambda) = (\lambda - 1)^2(\lambda - 3)$, $\quad P_3(\lambda) = (\lambda - 1)(\lambda - 3)^2$

ou $P_4(\lambda) = (\lambda - 1)^2(\lambda - 3)^2$.

Testando a matriz $\big[T\big]_B$ em $P_1(\lambda)$, temos:

$$P_1(\ [T]_B\ ) = (\ [T]_B - 1.I\ )(\ [T]_B - 3.I\ ).$$

$$P_1(\ [T]_B\ ) = \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) . \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \right).$$

$$P_1(\ [T]_B\ ) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} . \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo, $P_1(\lambda) = (\lambda - 1)(\lambda - 3)$ é o polinômio minimal de T. Portanto, a matriz é diagonalizável.

4 Decomposição em Valores Singulares

Neste capítulo apresentaremos de uma forma breve e prática a decomposição em valores singulares de uma matriz não quadrada, conhecida como DVS.

A abordagem será com respeito a matrizes reais, porém também pode ser feita para o caso complexo. Este capítulo é baseado nas referências [5], [6], [11], [16], [18], [19], e [24].

4.1 Um breve relato histórico

A Decomposição em Valores Singulares é um assunto relativamente recente. Os matemáticos Eugenie Beltrami e Camille Jordan podem ser considerados os pais da DVS. Em 1873, Beltrami publica um trabalho relativo a esta decomposição considerando apenas matrizes quadradas, não singulares e com valores singulares distintos. Em 1874, Jordan, independentemente de Beltrami, publica um trabalho em que considera os valores singulares para forma bilineares, em que reduz uma forma bilinear para uma forma diagonal por substituições ortogonais. A generalização da DVS para matrizes retangulares e complexas somente aparecem nos trabalhos de Carl Eckart e Gale Young, em 1936. A utilização deste método na computação ocorre a partir de 1960. Mais recentemente, o matemático norte-americano Gene Golub também estudou a aplicabilidade deste método em diversas áreas, produzindo um algoritmo estável e eficaz para calculá-la.

Sabemos que toda matriz simétrica A pode ser fatorada como $A = PDP^T$, em que P é uma matriz ortogonal de ordem $n \times n$ de autovetores de A e D é uma matriz diagonal cujas entradas são os autovalores associados aos autovetores colunas de P. Esta decomposição podemos chamá-la de DAV "Decomposição em Autovalores da matriz A".

A Decomposição em Valores Singulares (DVS) conhecida também, abreviamente por SVD - Singular Value Decomposition - consiste numa extensão da teoria de diagonalização de matrizes simétricas $n \times n$ (DAV) para matrizes arbitrárias real ou complexa de ordem $m \times n$. Formalmente, a decomposição em valores singulares de uma matriz A de ordem $m \times n$ é uma fatoração na forma $A = U\Sigma V^T$, onde U e V são matrizes quadradas ortogonais e ortonormais de ordem $m \times m$ e $n \times n$, respectivamente. A matriz $\Sigma = \left[\delta_{ii}\right]$ é retangular de ordem $m \times n$ (mesma ordem da matriz A), em que as entradas diagonais σ_{ii} são números reais não negativos, arranjados em ordem decrescente de magnitude chamados de valores singulares de A. As m colunas de U são os autovetores de AA^T, denominadas de vetores singulares à esquerda; as n colunas V são os autovetores de A^TA, chamados de vetores singulares à direita.

A importância da decomposição em valores singulares é a sua variedade de aplicações, tais como: na compressão, no armazenamento e na transmissão de informações digitalizadas que formam a base de muitos dos melhores algoritmos computacionais disponíveis para resolução de sistemas lineares; no cálculo da pseudo-inversa de uma matriz, no ajuste de dados utilizando mínimos quadrados. De um modo geral, em todas as aplicações a utilização da decomposição em valores singulares (DVS), consiste no fato que através desta fatoração matricial podemos reduzir significativamente a dimensão do espaço inicial, podendo reter apenas os dados que são importantes, desprezando os demais.

4.2 A DVS de uma matriz

Definição 4.1. Sejam $A \in M_{m\times n}(\mathbb{R})$ e $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda p > 0$, para algum $p \leq min\{m, n\}$, os autovalores de A^TA ou AA^T. Definimos **valores singulares não nulos de A** como sendo $\sigma_k = \sqrt{\lambda_k}$, para $k \in \{1, 2, \ldots, p\}$.

Proposição 4.2. *Os valores singulares de A são os comprimentos dos vetores*

$$Av_1,\ Av_2,\ \ldots,\ Av_n.$$

Demonstração. De fato,

$$\| Av_j \|^2 = \langle Av_j, Av_j\rangle = (Av_j)^TAv_j = v_j^T(A^TA)v_j = v_j^T\lambda_jv_j = v_j^T\sigma_j^2v_j = \sigma_j^2v_j^Tv_j = \sigma_j^2$$

Assim, $\| Av_j \|^2 = \sigma_j^2$.

Portanto, $\| Av_j \| = \sigma_j$, para todo $j = 1, 2, \ldots, p$.

□

Definição 4.3. A **diagonal principal de uma matriz retangular** A de ordem $m \times n$ é definida como sendo a fileira de entradas que começa no canto superior esquerdo a_{11} e se estende diagonalmente até onde fôr possível. Dizemos que as entradas nessa diagonal principal são as entradas diagonais da matriz.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2m} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$\{ a_{11}, a_{22}, \ldots, a_{mm}\}$ forma a diagonal principal da matriz A e $\{ b_{11}, b_{22}, \ldots, b_{nn}\}$ a diagonal principal da matriz B.

O próximo teorema garante a existência da decomposição em valores singulares de qualquer matriz na chamada forma reduzida de decomposição.

Teorema 4.4. *(Existência da DVS, na forma reduzida, de uma matriz) Toda matriz real retangular A de ordem $n \times m$ e posto r pode ser decomposta em*

$$A_{n\times m} = U_{n\times r}\Sigma_{r\times r}V_{r\times m}^T,$$

em que $U \in M_{n\times r}(\mathbb{R})$ e $V \in M_{m\times r}(\mathbb{R})$ são ortonormais nas colunas, e $\Sigma_{r\times r} = diag(\sigma_i)$ (matriz diagonal de valores singulares de A), com $\lambda_i > 0$. Sendo $\lambda_1, \lambda_2, \ldots, \lambda_r$ os autovalores não nulos da matriz A^TA ou AA^T, U e V matrizes de r autovetores ortonormais por coluna, respectivamente das matrizes A^TA ou AA^T.

Demonstração. Suponhamos a matriz A seja decomposta em $A = U\Sigma V^T$, sendo U e V matrizes quadradas ortonormais nas colunas de ordem n e m, respectivamente, e $\Sigma = [\sigma_{ij}]$ uma matriz diagonal de ordem $n \times m$, tal que $\sigma_{ii} = \sqrt{\lambda_i}$, para todo $i = 1, 2, 3. \ldots, m$.

Por conveniência e sem perda de generalidade admitamos que $m \leq n$.

Construindo matrizes ortogonais U e V com acréscimos de $n-r$ e $m-r$ vetores, respectivamente, às colunas U e V, ortonormais aos primeiros r deles. Assim, uma nova representação para Σ é dada por

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

A matriz A^TA é simétrica, pois $(A^TA)^T = A^TA$, e possui os autovalores $\lambda_1, \lambda_2, \ldots, \lambda_m$.

Por suposição $A = U\Sigma_m V^T$, então

$$A^TA = (U\Sigma_m V^T)^T(U\Sigma_m V^T) = V\Sigma_m U^T U\Sigma_m V^T = V\Sigma_m\Sigma_m V^T = V\Sigma_m^2 V^T.$$

Como V é ortogonal, segue que, $V^TA^TAV = V^TV\Sigma_m^2V^TV = \Sigma_m^2$.

Assim, $V^TA^TAV = \Sigma_m^2$.

Portanto, $V^TA^TAV = \Sigma_m^2 = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}$, em que V é a matriz dos autovetores de A^TA em suas colunas.

Assumindo uma situação restrita, em apenas r dos λ_i sejam não nulos, então a matriz $\Sigma_{n\times m}$ será dada por

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_r} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Essencialmente, Σ e Σ_r são idênticas nas primeiras r linhas e colunas. Ocorre que a matriz Σ possui $n-r$ linhas e $m-r$ colunas de zeros a mais que a matriz Σ_r.

Podemos ver que $\Sigma^T\Sigma = \Sigma_m^2$, sendo $\Sigma_m^2 = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$.

E podemos verificar que $(V^TA^T)AV = \Sigma^T\Sigma$, ou seja, $(AV)^TAV = \Sigma^T\Sigma$. Fazendo $W = AV$, temos $W^TW = \Sigma^T\Sigma$, sendo $W = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}$.

Como $W^TW = \Sigma^T\Sigma$, conclui-se que

$$w_i^Tw_i = \begin{cases} \lambda_i\ , & se\ i \leq r \\ 0\ , & se\ i > r \end{cases} \quad \text{e} \quad w_i^Tw_i = 0,\ \text{se } i \neq j,\ \text{para } j = 1, 2, \ldots, m.$$

Se $w_i = 0$, para $i > r$, as primeiras colunas de W são LI. Logo, concluímos que $r \leq n$.

Agora definindo $u_i = \frac{1}{\sqrt{\lambda_i}} \cdot w_i = \frac{1}{\sqrt{\lambda_i}} \cdot Av_i$, para $i = 1, 2, \ldots, r$, e tomando $u_{r+1}, u_{r+2}, \ldots, u_m$, se $r < m$, ortogonais entre si, e os demais u_i tais que a matriz quadrada $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$ de ordem n seja ortogonal.

Assim, pela definição de u_i , vemos que a expressão $u_i \cdot \sqrt{\lambda_i} = Av_i$ pode ser expressa matricialmente por $U\Sigma = AV$.

Como V é ortogonal, temos $V^{-1} = V^T$, o que resulta em

$$AV = U\Sigma, \text{ o que implica, } A = UAV^T.$$

Considerando apenas as r primeiras linhas de U, as r primeiras linhas e colunas de Σ e as r primeiras de V, o resultado ainda continua válido, o que finaliza a demonstração. □

Observação 4.5. A Decomposição em Valores Singulares de uma matriz $A \in M_{m\times n}(\mathbb{R})$ sendo $m \geq n$ e posto r, na forma reduzida, pode ser obtida efetuando-se a sequência descrita a seguir:

1. Construímos a matriz A^TA e determinamos os seus autovalores $\lambda_1, \lambda_2, \ldots, \lambda_r$.
2. Calculamos os valores singulares σ_i da matriz A, sendo $\sigma_i = \sqrt{\lambda_i}$, para $i = 1, 2, \ldots, r$, com $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$.
3. Construímos a matriz diagonal $\Sigma_r = \begin{bmatrix}\sigma_{ij}\end{bmatrix}$ quadrada de ordem r cujas entradas da diagonal principal são os valores singulares, sendo $\sigma_{ii} = \sqrt{\lambda_i}$, para $i = 1, 2, \ldots, r$, com $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. No caso de não existir r valores singulares não nulos, completamos com linhas e colunas nulas até formamos uma matriz de dimensão $r \times r$.
4. Obtemos os autovetores $v_1, v_2, \ldots, v_r$ de A^TA associados aos autovalores λ_1 , $\lambda_2, \ldots, \lambda_r$, respectivamente.
5. Construímos a matriz $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}_{n\times r}$ e caso não seja ortonormal, normalizamos esta matriz.
6. Construímos a matriz ortonormal $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}_{m\times r}$ determinado pelos vetores $u_i = \frac{1}{\sigma_i}.Av_i$, para $i = 1, 2, \ldots, r$ (ou encontramos os autovetores u_i da matriz AA^T). Caso o número de vetores u_i seja menor que r, determinemos através do processo de ortogonalização de Gram-Schmidt vetores u_i necessários para obtermos uma matriz $m \times r$, e normalizamos u_i, caso não estejam.
7. Escrevemos a matriz $A_{m\times n} = U_{m\times r}\Sigma_{r\times r}V^T_{r\times n}$.

No caso da decomposição em valores singulares de uma matriz A de ordem $m \times n$ sendo $n \geq m$ o processo sofre alterações na construção das matrizes U e V. Inicialmente, obtemos os autovalores e autovetores da matriz AA^T. Estes autovetores formaram a matriz $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$, e a matriz $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ será construída sendo $v_i = \frac{1}{\sigma_i}A^Tu_i$ (ou obtendo-se os autovetores da matriz A^TA).

Exemplo 4.6. Determinemos a DVS, na forma reduzida, da matriz $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

O $post(A) = 2$, então a DVS da matriz A tem a forma $A_{3\times 2} = U_{3\times 2}\Sigma_{2\times 2}V^T_{2\times 2}$.

Calculemos de A^TA.

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Determinemos os autovalores de A^TA.

$$P(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)(2-\lambda) - 1 = \lambda^2 - 4\lambda + 3 = 0.$$

Logo, os autovalores de A^TA (em ordem decrescente) são: $\lambda_1 = 3$ e $\lambda_2 = 1$.

Construção da matriz $\Sigma_{2\times 2}$:

Determinemos os valores singulares $\sigma_i = \sqrt{\lambda_i}$ da matriz A.

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3} \quad \text{e} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{1} = 1.$$

Portanto, $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}.$

Construção da matriz $V_{2\times 2}$:

Determinemos os autovetores associados aos autovalores obtidos.

Consideremos o sistema $(A^TA - \lambda I_2)X = 0$.

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Fazendo $\lambda = 3$ no sistema temos:

$$\underset{L_2 \to L_2 + L_1}{\begin{bmatrix} -1 & 1 & | & 0 \\ 1 & -1 & | & 0 \end{bmatrix}} \sim \begin{bmatrix} -1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \qquad \text{Então, } \left\{ -x_1 + x_2 = 0. \right.$$

Escolhendo $x_2 = a$ (variável livre), temos $x_1 = a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ a \in \mathbb{R}^*.$

Assim, $w_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é um autovetor de A^TA associado ao autovalor $\lambda_1 = 3$.

Fazendo $\lambda = 1$ no sistema temos:

$$\underset{L_2 \to L_2 - L_1}{\begin{bmatrix} 1 & 1 & | & 0 \\ 1 & 1 & | & 0 \end{bmatrix}} \sim \begin{bmatrix} 1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \qquad \text{Então, } \left\{ x_1 + x_2 = 0. \right.$$

Escolhendo $x_2 = a$ (variável livre), temos $x_1 = -a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -a \\ a \end{bmatrix} = a \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \ a \in \mathbb{R}^*.$

Assim, $w_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ é um autovetor de A^TA associado ao autovalor $\lambda_2 = 1$.

Os vetores w_1 e w_2 são ortogonais, mas não ortonormais, normalizando-os temos:

$$v_1 = \tfrac{1}{\|w_1\|}.w_1 = \tfrac{1}{\sqrt{2}}.\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{e} \quad v_2 = \tfrac{1}{\|w_2\|}.w_2 = \tfrac{1}{\sqrt{2}}.\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Portanto, $V = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$

Construção da matriz $U_{3\times 2}$:

Determinemos $u_i = \frac{1}{\sigma_i} A v_i$.

$$u_1 = \tfrac{1}{\sigma_1}.Av_1 = \tfrac{1}{\sqrt{3}}.\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.\tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \tfrac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}.$$

$$u_2 = \tfrac{1}{\sigma_2}.Av_2 = \tfrac{1}{1}.\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.\tfrac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}.$$

Portanto, $U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{bmatrix}.$

Podemos facilmente verificar que, $A = U\Sigma V^T$.

$$U\Sigma V^T = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{bmatrix}.\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}.\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \frac{\sqrt{3}}{\sqrt{2}} & \frac{\sqrt{3}}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = A.$$

Exemplo 4.7. Fatoremos, na forma reduzida, a matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ usando a decomposição em valores singulares.

O $post(A) = 2$, então a DVS da matriz A tem a forma $A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 2}V^T_{2\times 3}$.

Calculemos de AA^T.

$$AA^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Determinemos os autovalores de AA^T.

$$P(\lambda) = \left|AA^T - \lambda I\right| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0.$$

Logo, os autovalores de AA^T (em ordem decrescente) são: $\lambda_1 = 3$, $\lambda_2 = 1$.

Construção da matriz $\Sigma_{2\times 2}$:

Determinemos os valores singulares $\sigma_i = \sqrt{\lambda_i}$ de A.

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3} \quad \text{e} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{1} = 1.$$

Logo, $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}.$

Construção da matriz $U_{2\times 2}$:

Determinemos dos autovetores associados aos autovalores obtidos.

Consideremos o sistema $\quad (AA^T - \lambda I_3)X = 0.$

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Fazendo $\lambda_1 = 3$ no sistema, obtemos:

$$\underset{L_2 \to L_2 - L_1}{\begin{bmatrix} -1 & 1 & | & 0 \\ 1 & -1 & | & 0 \end{bmatrix}} \sim \begin{bmatrix} -1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \qquad \text{Então,} \quad \left\{ -x_1 + x_2 = 0. \right.$$

Escolhendo $x_2 = a$ (variável livre), temos $x_1 = a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ a \in \mathbb{R}^*.$

Assim, $w_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é um autovetor de AA^T associado ao autovalor $\lambda_1 = 3$.

Fazendo $\lambda_2 = 1$ no sistema temos:

$$\begin{bmatrix} 1 & 1 & | & 0 \\ 1 & 1 & | & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \qquad \text{Então,} \quad \left\{ x_1 + x_2 = 0. \right.$$

Escolhendo $x_2 = a$ (variável livre), temos $x_2 = a$ e $x_1 = -a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -a \\ a \end{bmatrix} = a \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$

Assim, $w_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ é um autovetor de AA^T associado ao autovalor $\lambda_2 = 1$.

Os vetores w_1 e w_2 e são ortogonais, mas não ortornormais. Normalizando-os temos:

$$u_1 = \tfrac{1}{\|w_1\|}.w_1 = \tfrac{1}{\sqrt{2}}. \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{e} \quad u_2 = \tfrac{1}{\|w_2\|}.w_2 = \tfrac{1}{\sqrt{2}}. \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Portanto, $\quad U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$

Construção da matriz $V_{3\times 2}$:

Determinemos $\quad v_i = \frac{1}{\sigma_i}.A^T u_i.$

$$v_1 = \tfrac{1}{\sigma_1}.Av_1 \ = \tfrac{1}{\sqrt{3}}. \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} . \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ = \ \tfrac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}.$$

$$v_2 = \tfrac{1}{\sigma_1}.Av_2 \ = \tfrac{1}{1}. \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} . \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \ = \ \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Assim, $V = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Portanto, $A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$.

A seguir apresentamos o teorema da DVS de uma matriz A de ordem $m \times n$ na chamada forma completa, em que U e V são matrizes quadradas de ordem m e n, respectivamente. A matriz Σ é retangular da mesma ordem da matriz A.

Teorema 4.8. *(DVS na forma completa) Toda matriz A de ordem $m \times n$ e posto r pode ser fatorada em*

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}, \quad \text{em que}$$

U_m é uma matriz quadrada de ordem m ortonormal nas colunas formada pelos autovetores de AA^T, V_n éuma matriz quadrada de ordem n ortonormal nas colunas formada pelos autovetores de A^TA. A matriz $\Sigma_{m\times n}$ é retangular diagonal de ordem $m \times n$ cujas entradas diagonais são os valores singulares da matriz A, ou seja, $\sigma_{ii} = \sqrt{\lambda_i}$, para $i = 1, 2, \ldots, r$, com $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$.

A prova deste teorema pode ser encontrada nas referências: $[24], (p.367 - 368)$ e [6],(p.101).

Observação 4.9. A Decomposição em Valores Singulares de uma matriz $A \in M_{m\times n}(\mathbb{R})$ sendo $m \geq n$ e posto r, na forma completa, pode ser obtida de modo análogo ao descrito na Observação 3.5. As matrizes $U_{m\times m}$, $V_{n\times n}$ e $\Sigma_{m\times n}$ poderão ser completadas, se necessário, com linhas e colunas nulas de modo a serem representadas nas formas $m \times m$, $n \times n$ e $m \times n$, respectivamente, como veremos no exemplo a seguir.

Exemplo 4.10. Determinemos a DVS, na forma completa, da matriz $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

O $post(A) = 1$ e a DVS da matriz A tem a forma $A_{3\times 2} = U_{3\times 3}\Sigma_{3\times 2}V^T_{2\times 2}$.

Calculemos de A^TA.

$$A^TA = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}.$$

Determinemos os autovalores de A^TA.

$$P(\lambda) = \left|A^TA - \lambda I\right| = \begin{vmatrix} 9-\lambda & -9 \\ -9 & 9-\lambda \end{vmatrix} = (9-\lambda)^2 - 81 = \lambda(\lambda - 18) = 0.$$

Logo, os autovalores de A^TA (em ordem decrescente) são: $\lambda_1 = 18$ e $\lambda_2 = 0$.

Construção da matriz $\Sigma_{3\times 2}$:

Determinemos o único valor singular σ_1 da matriz A.

$$\sigma_i = \sqrt{\lambda_i} = \sigma_1 = \sqrt{\lambda_1} = \sqrt{18} = 3\sqrt{2} \ .$$

Neste caso, devemos completar a matriz $\Sigma_{3\times 2}$ com 2 linhas e 1 coluna de zeros.

Portanto, $\Sigma = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Construção da matriz $V_{2\times 2}$:

Determinemos os autovetores associados aos autovalores obtidos.

Consideremos o sistema $(A^TA - \lambda I_2)X = 0$.

$$\begin{bmatrix} 9-\lambda & -9 \\ -9 & 9-\lambda \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Fazendo $\lambda_1 = 18$ no sistema obtemos:

$$\underset{L_1 \to \frac{1}{9}L_1}{\left[\begin{array}{cc|c} -9 & -9 & 0 \\ -9 & -9 & 0 \end{array}\right]} \sim \underset{L_2 \to 9L_1 + L_2}{\left[\begin{array}{cc|c} 1 & 1 & 0 \\ -9 & 9 & 0 \end{array}\right]} \sim \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Então, $\left\{ x_1 + x_2 = 0. \right.$

Escolhendo $x_2 = a$ (variável livre), temos $x_1 = -a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -a \\ a \end{bmatrix} = a \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $a \in \mathbb{R}^*$.

Assim, $w_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ é um autovetor de A^TA associado ao autovalor $\lambda_1 = 18$.

Fazendo $\lambda_2 = 0$ no sistema temos:

$$\underset{L_1 \to \frac{1}{9}L_1}{\left[\begin{array}{cc|c} 9 & -9 & 0 \\ -9 & 9 & 0 \end{array}\right]} \sim \underset{L_2 \to 9L_1 + L_2}{\left[\begin{array}{cc|c} 1 & -1 & 0 \\ -9 & 9 & 0 \end{array}\right]} \sim \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Então, $\left\{ x_1 - x_2 = 0. \right.$

Escolhendo $x_2 = a$ (variável livre), temos $x_1 = a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Assim, $w_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é um autovetor associado ao autovalor $\lambda_2 = 0$.

Os vetores w_1 e w_2 são ortogonais, mas não são ortonormais, Normalizando-os temos:

$$v_1 = \frac{1}{\|w_1\|}.w_1 = \frac{1}{\sqrt{2}}.\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad ; \quad v_2 = \frac{1}{\|w_2\|}.w_2 = \frac{1}{\sqrt{2}}.\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Portanto, $V = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Construção da matriz $U_{3\times 3}$:

$$u_i = \frac{1}{\sigma_i} \cdot Av_i.$$

$$u_1 = \frac{1}{\sigma_1} \cdot Av_1 = \frac{1}{3\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -2 \\ 4 \\ -4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}.$$

Neste caso devemos determinar u_2 e u_3 de modo que $\{u_1, u_2, u_3\}$ forme uma base ortonormal de $\mathbb{R}^3$, ou seja, $\langle u_1, u_2 \rangle = 0$ e $\langle u_1, u_3 \rangle = 0$. Determinemos os vetores $x = (x_1, x_2, x_3)$ tais que $\langle u_1^T, x \rangle = 0$.

$\langle (-\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}), (x_1, x_2, x_3) \rangle = 0 \ \Rightarrow \ -\frac{1}{3}x_1 + \frac{2}{3}x_2 - \frac{2}{3}x_3 = 0$. Logo, $-x_1 + 2x_2 - 2x_3 = 0$.

Portanto, $x_1 = 2x_2 - 2x_3$.

Assim, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_3 \\ 0 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$.

Escolhendo os vetores $u_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ e $u_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ observamos que u_1 , u_2 e u_3 formam uma base de $\mathbb{R}^3$, porém não são ortogonais, pois $\langle u_2, u_3 \rangle = -4$.

Aplicando-se o processo de ortogonalização de Gram-Schmidt, determinemos vetores u_1', u_2' *e* u_3' de modo a formar uma base ortonormal de $\mathbb{R}^3$.

Seja $u_1' = \frac{1}{\|u_1\|} \cdot u_1$.

Logo, $u_1' = (-\frac{1}{3}, \frac{2}{3}, -\frac{2}{3})$.

Como $r_2 = u_2 - \langle\ u_2, u_1'\ \rangle \cdot u_1' = (2, 1, 0) - 0 \cdot u_1 = (2, 1, 0)$.

Então, $u_2' = \frac{1}{\|r_2\|} \cdot r_2 = \frac{1}{\sqrt{5}} \cdot (2, 1, 0)$.

Logo, $u_2' = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0)$.

Como $r_3 = u_3 - \langle\ u_3, u_1'\ \rangle \cdot u_1' - \langle u_3, u_2' \rangle \cdot u_2'$.

$$r_3 = (-2, 0, 1) - 0 \cdot u_1' - \langle (-2, 0, 1), (\tfrac{2}{\sqrt{5}}, \tfrac{1}{\sqrt{5}}, 0) \rangle \cdot (\tfrac{2}{\sqrt{5}}, \tfrac{1}{\sqrt{5}}, 0)$$

$$r_3 = (-2, 0, 1) - (-\tfrac{4}{\sqrt{5}}) \cdot (\tfrac{2}{\sqrt{5}}, \tfrac{1}{\sqrt{5}}, 0) = (-2, 0, 1) + (\tfrac{8}{5}, \tfrac{4}{5}, 0) = (-\tfrac{2}{5}, \tfrac{4}{5}, 1).$$

Então, $u_3' = \frac{1}{\|r_3\|} \cdot r_3 = \frac{5}{3\sqrt{5}} \cdot (-\frac{2}{5}, \frac{4}{5}, 1)$.

Logo, $u_3' = (-\frac{2}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}, \frac{4}{3\sqrt{5}})$.

Assim, $U = \begin{bmatrix} u_1' & u_2' & u_3' \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix}$.

Portanto, $A = U\Sigma V^T = \begin{bmatrix} -\frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix} \cdot \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

5 A Inversa Generalizada

Neste capítulo trataremos do conceito de inversa generalizada de uma matriz, e estudaremos um dos tipos mais importantes deste conceito por suas fortes propriedades algébricas, com a denominação de inversa de Moore-Penrose, ou simplesmente, pseudo-inversa, em que é possível garantir sua unicidade, e consequentemente manter todas as propriedades de matriz inversa no sentido clássico.

O termo inversa generalizada apareceu nos trabalhos do matemático sueco Erik Ivar Fredholm em 1903, em que chamou de "pseudo-inversa" a inversa generalizada de um operador integral. O conceito de inversa generalizada de matrizes surgiu em 1920 nos trabalhos do matemático norte-americano Eliakim Hasting Moore que define a inversa generalizada por meio de projeção de matriz e provou sua unicidade. Pouco se fez nos próximos anos, até aproximadamente 1950, quando surgiram alguns trabalhos envolvendo relações entre inversa generalizada e solução de problemas de mínimos quadrados. Em 1955, o britânico Sir Roger Penrose, físico, matemático e filósofo da ciência (ganhador do Prêmio Nobel de Física de 2020) prova que toda matriz admite uma única inversa generalizada desde que satisfaça algumas equações, denominadas de "condições de Penrose". Mostra também que sua definição de inversa generalizada coincide com a de Moore, daí hoje ser conhecida como "inversa de Moore-Penrose", ou simplesmente pseudo-inversa. ([7] e [8]).

Este capítulo é baseado nas referências: [9], [10], [14], [15], [17], [18], [20], [21] e [22].

5.1 Caracterização da inversa generalizada de uma matriz

A ideia de se obter a inversa de uma matriz A de ordem $m \times n$ também é utilizada em importantes aplicações de Álgebra Linear, como por exemplo, na determinação da solução de Mínimos Quadrados de sistemas de equações lineares inconsistentes.

O estudo de matrizes nos cursos básicos estabelece que uma matriz quadrada A de ordem n admite inversa, quando existir uma outra matriz A^{-1} também quadrada de ordem n, tal que $A.A^{-1} = A^{-1}.A = I_n$, sendo I_n a matriz identidade de ordem n. Assim, a existência da inversa está restrita somente às matrizes quadradas e não-singulares, ou equivalente a dizer, matrizes que possuem posto igual a n (posto completo). A partir destas condições verificam-se a validade das seguintes propriedades:

P_1. Se A^{-1} existe, então ela é única. P_2. $detA^{-1} = \frac{1}{detA}$.

P_3. $(A^{-1})^{-1} = A$. P_4. $(A^T)^{-1} = (A^{-1})^T$.

P_5. $(AB)^{-1} = B^{-1}A^{-1}$.

Em situações problemas, como por exemplo, que envolvam a solução de sistemas de equações lineares do tipo $Ax = b$, sendo A uma matriz quadrada de ordem n e não-singular, sabemos que o sistema é consistente e determinado (admite uma única solução) que é dada por

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

Entretanto, existem situações que esses requisitos não se cumprem, isto é, quando a matriz A não é quadrada, ou mesmo sendo quadrada apresenta posto incompleto. Uma abordagem para tratar destas situações irá envolver a utilização do conceito de inversa generalizada de uma matriz.

Uma matriz A, não quadrada, do tipo $m \times n$ não possui inversa nas condições que vimos anteriormente. Mas, pode admitir uma matriz A_d^- do tipo $n \times m$, tal que $AA_d^- = I_m$, e também uma outra matriz A_e^- do tipo $n \times m$ tal que $A_e^- A = I_n$. Estas matrizes A_d^- e A_e^- recebem os nomes de inversa generalizada de A à direira, e inversa generalizada de A à esquerda, respectivamente.

Exemplo 5.1. Consideremos a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}_{2\times 3}$.

Uma inversa generalizada à direita de A é a matriz $A_d^- = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}_{3\times 2}$, pois

$$A.A_d^- = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}_{2\times 3} . \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}_{3\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2} = I_2.$$

No exemplo a seguir baseado no que é feito para o caso tradicional faremos o cálculo de obtenção de uma inversa generalizada.

Exemplo 5.2. Calculemos uma inversa generalizada à esquerda para a matriz

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}.$$

$$\underset{L_3 \to L_3 - 2L_1}{\begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 2 & | & 0 & 1 & 0 \\ 2 & 5 & | & 0 & 0 & 1 \end{bmatrix}} \sim \underset{L_2 \to \frac{1}{2}L_2}{\begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 2 & | & 0 & 1 & 0 \\ 0 & 5 & | & 0 & 5 & 1 \end{bmatrix}} \sim \underset{L_3 \to L_3 - 5L_2}{\begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & | & 0 & \frac{1}{2} & 0 \\ 0 & 5 & | & 0 & 5 & 1 \end{bmatrix}} \sim$$

$$\sim \begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & | & 0 & \frac{1}{2} & 0 \\ 0 & 0 & | & 0 & \frac{5}{2} & 1 \end{bmatrix}. \quad \text{Logo,} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{5}{2} & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}_{2\times 3} . \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}_{3\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2}.$$ Consequentemente,

Assim, $A_e^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ é a inversa generalizada de A à esquerda.

Convém destacar que uma matriz pode admitir mais de uma inversa generalizada à direita (ou à esquerda).

A matriz $A_e'^{-} = \frac{1}{45}\begin{bmatrix} 29 & -20 & 8 \\ -10 & 10 & 5 \end{bmatrix}$ também é uma inversa generalizada de A à esquerda, pois

$$A_e'^{-}.A = \frac{1}{45}\begin{bmatrix} 29 & -20 & 8 \\ -10 & 10 & 5 \end{bmatrix} . \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = \frac{1}{45}\begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

A seguir a definição que caracteriza uma inversa generalizada de uma matriz.

Definição 5.3. Seja A uma matriz de ordem $m \times n$ e posto não nulo. Uma **inversa generalizada** de A ou **g-inversa** de A, denotada por A^-, é uma matriz de ordem $n \times m$ tal que $x = A^-b$ é uma solução do sistema de equações lineares $Ax = b$, para todo b que torne o sistema consistente.

Proposição 5.4. *A inversa generalizada A^- de uma matriz A existe se, e somente se, $AA^-A = A$.*

Demonstração. Escolhemos y como sendo a $i-ésima$ coluna a_i da matriz A. A equação $Ax = a_i$ é claramente consistente e, portanto, $x = A^-a_i$ é uma solução, ou seja, $AA^-a_i = a_i$, para todo i, o que equivale a dizer que $AA^-A = A$.

Por outro lado, se A^- é tal que $AA^-A = A$ e $Ax = y$ é consistente, então $AA^-Ax = x$ ou $A(A^-y) = y$. Portanto, $x = A^-y$ é uma solução, o que prova a proposição. □

A seguir, do resultado desta proposição, uma definição equivalente de inversa generalizada de uma matriz.

Definição 5.5. A matriz A^- é uma **inversa generalizada** de A se $AA^-A = A$.

Definição 5.6. A matriz A^- é chamada de **inversa generalizada reflexiva** de A quando $A^-AA^- = A^-$.

5.2 Algoritmo de obtenção de uma inversa generalizada

5.2.1 Algoritmo de Searle

A inversa generalizada de uma matriz A pode ser obtida através do algoritmo descrito abaixo, denominado de Algoritmo de Searle (1971).

1. Encontre uma submatriz quadrada B de posto igual ao da matriz A.
2. Obtenha a matriz transposta da inversa de B, isto é, $(B^{-1})^T$.
3. Construa a matriz C substituindo a submatriz B escolhida em A pela matriz $(B^{-1})^T$ completando com "zeros" o restante da matriz até chegar na dimensão da matriz A.
4. A transposta da matriz C é a inversa generalizada de A, isto é, $A^- = C^T$.

Exemplo 5.7. Determinemos uma inversa generalizada da matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix}$.

Como $det\ A = 0$, e o determinante da submatriz $B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2 \neq 0$, então temos $post(A) = 2$.

Calculemos $(B^{-1})^T$:

$$B^{-1} = \tfrac{1}{detB}.adjB = -\tfrac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Assim, $(B^{-1})^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

Substítuindo $(B^{-1})^T$ na matriz A no lugar da submatriz B e completando com zeros os demais elementos, temos: $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Determinando a transposta da matriz anterior, temos: $A^- = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

De fato, A^- é a inversa generalizada de A, pois

$$AA^-A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix} . \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix} . \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix} = A.$$

Exemplo 5.8. Determinemos uma inversa generalizada da matriz $A = \begin{bmatrix} 3 & 2 & 2 \\ 5 & 1 & 1 \end{bmatrix}$.

O posto de A é igual a 2, pois o determinante da submatriz $B = \begin{bmatrix} 3 & 2 \\ 5 & 1 \end{bmatrix} = -7 \neq 0$.

Calculemos $(B^{-1})^T$:

$$B^{-1} = \tfrac{1}{detB}.adjB = -\tfrac{1}{7}\begin{bmatrix} 1 & -2 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{7} & \frac{2}{7} \\ \frac{5}{7} & -\frac{3}{7} \end{bmatrix}.$$

Assim, $(B^{-1})^T = \begin{bmatrix} -\frac{1}{7} & \frac{5}{7} \\ \frac{2}{7} & -\frac{3}{7} \end{bmatrix}$.

Substítuindo $(B^{-1})^T$ na matriz A no lugar da submatriz B e completando com zeros os demais elementos, temos: $\begin{bmatrix} -\frac{1}{7} & \frac{5}{7} & 0 \\ \frac{2}{7} & -\frac{3}{7} & 0 \end{bmatrix}$.

Achando a transposta da matriz anterior, temos: $A^- = \begin{bmatrix} -\frac{1}{7} & \frac{2}{7} \\ \frac{5}{7} & -\frac{3}{7} \\ 0 & 0 \end{bmatrix}$.

De fato, A^- é a inversa generalizada de A, pois

$$AA^-A = \begin{bmatrix} 3 & 2 & 2 \\ 5 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{7} & \frac{2}{7} \\ \frac{5}{7} & -\frac{3}{7} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 & 2 \\ 5 & 1 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 & 2 \\ 5 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 2 \\ 5 & 1 & 1 \end{bmatrix} = A.$$

Proposição 5.9. *Seja A uma matriz $m \times n$. Então, A^- existe, e a classe de inversas generalizadas a partir de qualquer inversa A^- é dada por*

$$A^- + U - A^-A\,UA^- \quad ,$$

sendo U uma matriz arbitrária $n \times m$.

ou

$$A^- + V(I - AA^-) + (I - A^-A)W \quad ,$$

sendo V e W matrizes arbitrárias $n \times m$.

5.3 A inversa de Moore-Penrose

Definição 5.10. Seja A uma matriz real de ordem $m \times n$. Dizemos que a matriz real $A^\dagger$ de ordem $n \times m$ é a **inversa de Moore-Penrose** de A, ou simplesmente **pseudo-inversa** de A, se as condições a seguir, chamadas de equações de Penrose, forem satisfeitas:

(1) $AA^\dagger A = A$.
(2) $A^\dagger AA^\dagger = A^\dagger$.
(3) $AA^\dagger$ é simétrica, ou seja, $(AA^\dagger)^T = AA^\dagger$.
(4) $A^\dagger A$ é simétrica, ou seja, $(A^\dagger A)^T = A^\dagger A$.

Embora, $AA^\dagger$ desempenha o papel de matriz identidade na pré multiplicação por A e na pós-multiplicação por $A^\dagger$, respectivamente, nas propriedades (1) e (2), estas matrizes só serão iguais à identidade em situações especiais, no caso da matriz A for quadrada e não-singular. E nestes casos, $A^{-1} = A^\dagger$.

Exemplo 5.11. Mostremos que a matriz $A^\dagger = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix}$ é a pseudo-inversa da matriz $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$.

Devemos mostrar que $A^\dagger$ e A satisfazem as quatro condições da definição de inversa de Moore-Penrose.

(1) $$AA^\dagger A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{3}\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

Logo, $AA^\dagger A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = A$

(2) $$A^\dagger AA^\dagger = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} =$$

$$= \tfrac{1}{9}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} = \tfrac{1}{9}\begin{bmatrix} 6 & 3 & -3 \\ 3 & 6 & 3 \end{bmatrix} = \tfrac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix}.$$

Logo, $AA^\dagger A = A^\dagger$

(3) $AA^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \cdot \tfrac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} = \tfrac{1}{3}. \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$ é simétrica.

(4) $A^\dagger A = \tfrac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \tfrac{1}{3}\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ é simétrica.

Portanto, $A^\dagger$ é a matriz inversa de Moore-Penrose de A.

Teorema 5.12. *Se $A \in M_{m\times n}(\mathbb{R})$, então $A^\dagger = V\Sigma^+U^T$ existe é única, sendo U e V matrizes quadrada ortonormais de ordem m, e n, respectivamente, e Σ^+ uma matriz diagonal $m \times n$ cujas entradas diagonais δ_{ii} são os inversos dos valores singulares de A.*

Demonstração. Existência: Utilizando-se da decomposição em valores singulares da matriz A podemos construir uma possível pseudo-inversa de A.

Seja $A = U\Sigma V^T$, em que U e V são matrizes quadradas ortonormais de ordem m e n, respectivamente, sendo V^T a matriz transposta de V.

Sejam $\lambda_1, \lambda_2, \ldots, \lambda_p$ os autovalores não nulos de A^TA e $\sigma_i = \sqrt{\lambda_i}$ $(i = 1, 2, \ldots, p)$, os valores singulares de A. Lembremos que

$$\Sigma = \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}_{m\times n} \quad \text{em que} \quad C = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}.$$

$$\text{Consideremos} \quad \Sigma^+ = \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix}_{n\times m} \quad \text{então} \quad C^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix}_{p\times p}.$$

Como os autovalores escolhidos λ_i de A^TA são tais que $\lambda_i > 0$, então $\sigma_i > 0$. Logo, as matrizes Σ e Σ^+ estão bem definidas.

É evidente que a matriz C é invertível, pois $\sigma_1, \sigma_2, \ldots, \sigma_p$ são números reais positivos e $p \leq min\{m, n\}$, e C^{-1} é a sua inversa.

Mostremos que Σ^+ e Σ satisfazem as condições de Penrose, ou seja, uma é pseudo-inversa da outra.

$$\Sigma\,\Sigma^+\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix}\Sigma = I_p\,\Sigma = \Sigma.$$

$$\Sigma^+\Sigma\ \Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix} \Sigma^+ \ = \ I_p\ \Sigma^+ \ = \Sigma^+.$$

$\Sigma\ \Sigma^+ = I_p$ e $\Sigma^*\ \Sigma = I_p$. Logo, $\Sigma\ \Sigma^+$ e $\Sigma^+\ \Sigma$ são simétricas.

Portanto, satisfazem as condições de Penrose.

Mostremos que a pseudo-inversa de A é dada por $X = V\Sigma^+U^T$, pois A e X também satisfazem as condições de Penrose.

$$AXA = (U\Sigma V^T)(\ V\Sigma^+U^T)\ U\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T = A\ .$$

$$XAX = \ V\Sigma^+U^T\ U\Sigma V^T\ V\Sigma^+U^T = \ V\Sigma^+\Sigma\Sigma^+U^T = V\Sigma^+U^T = X\ .$$

$$\begin{aligned}(AX)^T &= (U\Sigma V^T\ V\Sigma^+U^T)^T = (\ U\Sigma\Sigma^+U^T\ )^T = U(\Sigma\Sigma^+)^TU^T = U\Sigma\Sigma^+U^T \\ &= U\Sigma V^T\ V\Sigma^+U^T = AX.\end{aligned}$$

$$\begin{aligned}(XA)^T &= (V\Sigma^+U^T\ U\Sigma V^T)^T = (V\Sigma^+\Sigma V^T)^T = V(\Sigma^+\Sigma)^TV^T = V\Sigma^+\Sigma V^T = \\ &= V\Sigma^+U^T\ U\Sigma V^T = XA.\end{aligned}$$

Assim, X é a pseudo-inversa de A . Portanto, podemos dizer que $A^\dagger = V\Sigma^+U^T$.

Unicidade: Suponhamos que X e Y sejam pseudo-inversas de A. Logo, satisfazem as condições de Penrose. Em particular que $AXA = A$, $XAX = X$ e $AYA = A$.

Assim, $AX = (AX)^T = X^TA^T = X^T(AYA)^T = X^TA^TY^TA^T = (AX)^T(AY)^T =$
$= AXAY = (AXA)Y = AY$. Logo, $AX = AY$.

Por outro lado, temos:

$$XA = (XA)^T = A^TX^T = (AYA)^TX^T = A^TY^TA^TX^T = (YA)^T(XA)^T =$$
$$= YAXA = Y(AXA) = YA\ . \quad \text{Logo,}\ \ XA = YA.$$

Então, $X = XAX = XAY = YAY = Y$.

Portanto, a pseudo-inversa de A é única.

□

Seja $A \in M_{m\times n}(\mathbb{R})$ e consideremos a matriz quadrada $A.A^T$ simétrica de ordem m. Suponhamos que $A.A^T$ seja invertível, então *posto* $(AA^T) = m$ (número de linhas de A).

Logo, a matriz $A.A^T$ admite inversa à direita $(A.A^T)^{-1}$, e

$$(A.A^T)(A.A^T)^{-1} = I_m.$$
$$A.[\ A^T(AA^T)^{-1}\ ] = I_m.$$

Portanto, $A^T(AA^T)^{-1}$ é a pseudo-inversa de A à direita, ou seja,

$$A^\dagger = A^T(AA^T)^{-1}.$$

De modo análogo, considerando a matriz quadrada A^TA simétrica de ordem n, e supondo-a invertível, então *posto* $(A^TA) = n$ (número de colunas de A).

Logo, a matriz A^TA admite inversa à esquerda $(A^TA)^{-1}$, e

$$(A^TA)^{-1}(A^TA) = I_n.$$
$$[\ (A^TA)^{-1}A^T\ ].A = I_n.$$

Portanto, $(A^TA)^{-1}A^T$ é a pseudo-inversa de A à esquerda, ou seja,

$$A^\dagger = (A^TA)^{-1}A^T.$$

É fácil de verificarmos que $A^\dagger = (A^TA)^{-1}A^T$ satisfaz as condições de Penrose. De fato,

(1) $AA^\dagger A = A[(A^TA)^{-1}A^T]A = AA^{-1}(A^T)^{-1}A^TA = I_m.I_mA = A$.
Logo, $AA^\dagger A = A$.

(2) $A^\dagger AA^\dagger = [(A^TA)^{-1}A^T]AA^\dagger = A^{-1}(A^T)^{-1}A^TAA^\dagger = I_n.I_n.A^L = A^\dagger$.
Logo, $A^\dagger AA^L = A^L$.

(3) $(AA^\dagger)^T = [A(A^TA)^{-1}A^T]^T = [AA^{-1}(A^T)^{-1}A^T]^T = (I_m.I_m)^T = I_m$.

(4) $(A^\dagger A)^T = [(A^TA)^{-1}A^TA]^T = [A^{-1}(A^T)^{-1}A^TA]^T = [A^{-1}I_mA]^T = [A^{-1}A]^T = I_n$.
Logo, $A^\dagger A$ e $AA^\dagger$ são simétricas.

Portanto, $A^\dagger = (A^TA)^{-1}A^T$ é a inversa de Moore-Penrose de A à esquerda.

Analogamente, podemos mostrar que $A^\dagger = A^T(AA^T)^{-1}$ também satisfaz as condições de Penrose e, portanto é a inversa de Moore-Penrose à direita.

5.4 Obtenção de uma pseudo-inversa.

Dada uma matriz $A \in M_{m\times n}(\mathbb{R})$, a **inversa de Moore-Penrose** ou **pseudo-inversa** $A^\dagger \in M_{n\times m}(\mathbb{R})$ pode ser obtida da seguinte forma:

(a) Se $post(A) = n$, então $A^\dagger = (A^TA)^{-1}A^T$.

(b) Se $post(A) = m$, então $A^\dagger = A^T(AA^T)^{-1}$.

(c) Se $post(A) < min\{m, n\}$, então $A^\dagger = (U\Sigma V^T)^{-1} = (V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^+U^T$.

(d) Se $m = n$ e $post(A) = n$, então $A^\dagger = A^{-1}$.

É fácil observamos que a matriz $A^\dagger = 0$, de ordem $n \times m$, é a pseudo-inversa da matriz $A = 0$ de ordem $m \times n$.

No caso de uma matriz A ter posto 1, sua inversa $A^\dagger$ também pode ser obtida como sendo $A^\dagger = \frac{1}{tr(A^TA)}A^T$.

Exemplo 5.13. Determinemos a inversa de Moore-Penrose da matriz

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{bmatrix}.$$

A é uma matriz real 3×2 e $post(A) = 2 = n$, então $A^\dagger = (A^TA)^{-1}A^T$.

Como $A^TA = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 3 \\ 3 & 5 \end{bmatrix}$.

Logo, $A^\dagger = (A^TA)^{-1}A^T = \begin{bmatrix} 10 & 3 \\ 3 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{bmatrix}^T = \frac{1}{41}\begin{bmatrix} 5 & -3 \\ -3 & 10 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{bmatrix}$.

Portanto, $A^\dagger = \frac{1}{41}\begin{bmatrix} 5 & -6 & 12 \\ -3 & 20 & 1 \end{bmatrix}$.

Exemplo 5.14. Determinemos, a inversa de Moore-Penrose da matriz
$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}.$$

A é uma matriz real 2×3 e $post(A) = 2 = m$, então $A^{\dagger} = A^T(AA^T)^{-1}$.

Como $$AA^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} . \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}.$$

Logo, $$A^{\dagger} = A^T(AA^T)^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} . \tfrac{1}{9} \begin{bmatrix} 5 & -1 \\ -1 & 2 \end{bmatrix}.$$

Portanto, $$A^{\dagger} = \tfrac{1}{9} \begin{bmatrix} 5 & -1 \\ 4 & 1 \\ -2 & 4 \end{bmatrix}.$$

Exemplo 5.15. Determinemos a pseudo-inversa da matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$.

A é uma matriz real 2×3 e $post(A) = 1 < min\{m, n\}$, então , $A^{\dagger} = V\Sigma^{+}U^T$.
A DVS da matriz A é da forma $A_{2\times 3} = U_2 \Sigma_{2\times 3} V_3^T$.
Calculemos de $A^T A$.

$$A^T A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} . \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 10 & 15 \\ 10 & 20 & 30 \\ 15 & 30 & 45 \end{bmatrix}.$$

Determinemos os autovalores de $A^T A$:

$$P(\lambda) = \left| A^T A - \lambda I_3 \right| = \begin{vmatrix} 5-\lambda & 10 & 15 \\ 10 & 20-\lambda & 30 \\ 15 & 30 & 45-\lambda \end{vmatrix} =$$

$$= -\lambda^3 + 70\lambda^2 - 1225\lambda + 13500 + 225\lambda + 900\lambda + 100\lambda - 13500 =$$

$$= -\lambda^3 + 70\lambda^2 = \lambda^2(70 - \lambda) = 0.$$

Logo, os autovalores de $A^T A$ (em ordem decrescente) são: $\lambda_1 = 70$, $\lambda_2 = 0$ e $\lambda_3 = 0$.

Os valores singulares de A: $\sigma_i = \sqrt{\lambda_i}$.

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{70} \quad ; \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{0} = 0 \quad \text{e} \quad \sigma_3 = \sqrt{\lambda_3} = \sqrt{0} = 0.$$

Construção da matriz Σ:

$$\Sigma_{2\times 3} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{70} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad \text{Então,} \quad A^{+}_{3\times 2} = \begin{bmatrix} \frac{1}{\sqrt{70}} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Construção da matriz V:

Determinemos os autovetores associados aos autovalores obtidos.

Consideremos o sistema $(A^T A - \lambda I_3)X = 0$.

$$\begin{bmatrix} 5-\lambda & 10 & 15 \\ 10 & 20-\lambda & 30 \\ 15 & 30 & 45-\lambda \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Fazendo $\lambda_1 = 70$ no sistema, obtemos:

$$\begin{bmatrix} -65 & 10 & 15 & | & 0 \\ 10 & -50 & 30 & | & 0 \\ 15 & 30 & -25 & | & 0 \end{bmatrix} \underset{\substack{L_1 \to 0,2.L_1 \\ L_2 \to 0,1.L_2 \\ L_3 \to 0,2.L_3}}{\sim} \begin{bmatrix} -13 & 2 & 3 & | & 0 \\ 1 & -5 & 3 & | & 0 \\ 3 & 6 & -5 & | & 0 \end{bmatrix} \underset{L_2 \leftrightarrow L_1}{\sim} \begin{bmatrix} 1 & -5 & 3 & | & 0 \\ -13 & 2 & 3 & | & 0 \\ 3 & 6 & -5 & | & 0 \end{bmatrix} \underset{\substack{L_2 \to 13.L_1 + L_2 \\ L_3 \to L_3 - 3.L_1}}{\sim}$$

$$\sim \begin{bmatrix} 1 & -5 & 3 & | & 0 \\ 0 & -63 & 42 & | & 0 \\ 0 & 21 & -14 & | & 0 \end{bmatrix} \underset{\substack{L_2 \to -\frac{1}{21}L_2 \\ L_2 \to \frac{1}{7}L_2}}{\sim} \begin{bmatrix} 1 & -5 & 3 & | & 0 \\ 0 & 3 & -2 & | & 0 \\ 0 & 3 & -2 & | & 0 \end{bmatrix} \underset{L_3 \to L_3 - L_2}{\sim} \begin{bmatrix} 1 & -5 & 3 & | & 0 \\ 0 & 3 & -2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Então, $\begin{cases} x_1 - 5x_2 + 3x_3 = 0 \\ 3x_2 - 2x_3 = 0 \end{cases}$.

Escolhendo $x_3 = 3a$ (variável livre), temos $x_2 = 2a$ e $x_1 = a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \\ 2a \\ 3a \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $a \in \mathbb{R}^*$.

Assim, $w_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ é um autovetor de $A^T A$ associado ao autovalor $\lambda_1 = 70$.

Fazendo $\lambda_2 = 0$ no sistema temos:

$$\begin{bmatrix} 5 & 10 & 15 & | & 0 \\ 10 & 20 & 30 & | & 0 \\ 15 & 30 & 45 & | & 0 \end{bmatrix} \underset{\substack{L_2 \to L_2 - 2.L_1 \\ L_3 \to L_3 - 3.L_1}}{\sim} \begin{bmatrix} 5 & 10 & 15 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \underset{L_1 \to \frac{1}{5}.L_1}{\sim} \begin{bmatrix} 1 & 2 & 3 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Então, $\{\, x_1 + 2x_2 + 3x_3 = 0$.

Escolhendo $x_3 = 0$ e $x_3 = a$ (variáveis livres), temos $x_1 = -2a$.

Logo, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2a \\ a \\ 0 \end{bmatrix} = a \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$.

Assim, $w_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ é um autovetor de $A^T A$ associado ao autovalor $\lambda_2 = 0$.

Determinemos o vetor w_3 de modo que os vetores w_1, w_2 e w_3 formem uma base ortogonal de $\mathbb{R}^3$. Neste caso, podemos obter w_3 efetuando o produto vetorial de w_2 e w_3.

$$w_1 \wedge w_2 = \begin{vmatrix} i & j & k \\ 1 & 2 & 3 \\ -2 & 1 & 0 \end{vmatrix} = -6j + k + 4k - 3i = -3i - 6j + 5k.$$

Portanto, $\quad w_3 = \begin{bmatrix} -3 \\ -6 \\ 5 \end{bmatrix}.$

Os vetores w_1 , w_2 e w_3 são ortogonais, mas não ortornormais. Normalizando-os temos:

$$v_1 = \frac{1}{\|w_1\|}.w_1 = \frac{1}{\sqrt{14}}.\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix} \quad ; \quad v_2 = \frac{1}{\|w_2\|}.w_2 = \frac{1}{\sqrt{5}}.\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}.$$

$$v_3 = \frac{1}{\|w_3\|}.w_3 = \frac{1}{\sqrt{70}}.\begin{bmatrix} -3 \\ -6 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{-3}{\sqrt{70}} \\ \frac{-6}{\sqrt{70}} \\ \frac{5}{\sqrt{70}} \end{bmatrix}.$$

Portanto, $\quad V = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{-2}{\sqrt{5}} & \frac{-3}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & \frac{1}{\sqrt{5}} & \frac{-6}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & 0 & \frac{5}{\sqrt{70}} \end{bmatrix}.$

Construção da matriz U: $\quad u_i = \frac{1}{\sigma_i} A v_i.$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{70}}.\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}.\frac{1}{\sqrt{14}}.\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{980}}.\begin{bmatrix} 14 \\ 28 \end{bmatrix} = \frac{1}{14\sqrt{5}}.\begin{bmatrix} 14 \\ 28 \end{bmatrix}.$$

Logo, $\quad u_1 = \frac{1}{\sqrt{5}}.\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$

Determinemos o vetor u_2 tal que $\{ u_1, u_2\}$ forme uma base ortonormal de $\mathbb{R}^2$. Neste caso, basta tomarmos $\; u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$

Assim, $\quad U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \quad$ e $\quad U^T = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$

Então, $\quad A^\dagger = V\Sigma^+U^T = \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{-2}{\sqrt{5}} & \frac{-3}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & \frac{1}{\sqrt{5}} & \frac{-6}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & 0 & \frac{5}{\sqrt{70}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{70}} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} . \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$

$$A^\dagger = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{-2}{\sqrt{5}} & \frac{-3}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & \frac{1}{\sqrt{5}} & \frac{-6}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & 0 & \frac{5}{\sqrt{70}} \end{bmatrix} . \begin{bmatrix} \frac{1}{\sqrt{70}} & \frac{2}{\sqrt{70}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1}{\sqrt{980}} & \frac{2}{\sqrt{980}} \\ \frac{2}{\sqrt{980}} & \frac{4}{\sqrt{980}} \\ \frac{3}{\sqrt{980}} & \frac{6}{\sqrt{980}} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1}{14\sqrt{5}} & \frac{2}{14\sqrt{5}} \\ \frac{2}{14\sqrt{5}} & \frac{4}{14\sqrt{5}} \\ \frac{3}{14\sqrt{5}} & \frac{6}{14\sqrt{5}} \end{bmatrix}.$$

Portanto, $\quad A^\dagger = \frac{1}{70} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}.$

Neste caso, convém observarmos, como $post(A) = 1$, então $A^\dagger$ pode ser facilmente obtida através de $\; A^\dagger = \frac{1}{tr(A^TA)} A^T.$

$$A^TA = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 10 & 15 \\ 10 & 20 & 30 \\ 15 & 30 & 45 \end{bmatrix}, \quad \text{então} \quad tr(A^TA) = 70 \; .$$

Portanto, $A^\dagger = \frac{1}{70}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$.

Propriedade 5.16. Seja $A \in M_{m\times n}(\mathbb{R})$, com $A^\dagger = V\Sigma^+U^T$ sua pseudo-inversa, então são válidas as seguintes propriedades:

(a) $(A^\dagger)^\dagger = A$.

(b) $(A^\dagger)^T = (A^T)^\dagger$.

(c) Se A é invertível, então $A^\dagger = A^{-1}$.

Demonstração.

(a) $(A^\dagger)^\dagger = (V\Sigma^\dagger U^T)^\dagger = U(\Sigma^\dagger)^\dagger V^T = U\Sigma V^T = A$.

(b) $(A^\dagger)^T = (V\Sigma^\dagger U^T)^T = U(\Sigma^\dagger)^T V^T = U(\Sigma^T)^\dagger V^T = (V\Sigma^T U^T)^\dagger =$
$= [(U\Sigma V^T)^T]^\dagger = (A^T)^\dagger$.

(c) Como A é invertível, então $m = n$ e $posto(A) = m$. Então, a matriz $\Sigma_{m\times m}$ é uma matriz diagonal sem zeros na diagonal principal. Logo, A é invertível.
Mas, $AA^\dagger = U\Sigma V^T \; V\Sigma^\dagger U^T = U\Sigma\Sigma^\dagger U^T$.
Como $\Sigma\Sigma^\dagger = I_m$, então $AA^\dagger = I_m$.
Portanto, $A^\dagger = A^{-1}$.

□

Observação 5.17. A inversa de Moore-Penrose é sem dúvida a principal inversa generalizada de matrizes, porém as condições de Penrose pode ser usadas para definir outras inversas generalizadas, as quais não possuem a garantia de unicidade, mas podem ser úteis em diversas aplicações. Existem casos de matrizes cuja inversa generalizada não satisfaz todas as condições de Penrose. Nestes casos, denotamos por $inversa - \{\,\cdot,\cdot,\cdot\,\}$ colocando entre as chaves separados por vírgula as condições de Penrose que ela satisfaz. Assim, se a matriz X é inversa de Moore-Penrose da matriz A que satisfaz somente as condições 1 e 2, denotamos por $inversa - \{1, 2\}$.

Assim a matriz $A^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ (Exemplo 4.2.) é uma inversa-{1,2,4} da matriz $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}$. De fato,

(1) $AA^-A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & \frac{5}{2} & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = A.$

(2) $A^-AA^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} =$

$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} = A^-.$

(3) $AA^- = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & \frac{5}{2} & 0 \end{bmatrix}$ não é simétrica.

(4) $A^{-}A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ é simétrica.

Portanto, A^{-} satisfaz apenas as condições (1) , (2) e (4) de Penrose.

5.5 Algoritmos de obtenção de uma pseudo-inversa.

Existem vários algoritmos algébricos para a determinação da matriz inversa de Moore-Penrose. Abordaremos dois desses algoritmos: o primeiro bastante simples proposto pelo próprio Penrose em 1956, e o segundo, denominado de algoritmo de Greville. Esta seção foi baseada nas referências [9] e [25].

5.5.1 Algoritmo de Penrose.

Consideremos uma matriz A do tipo $m \times n$ e posto igual a r , $r \leq min\{\ m, n\ \}$. A inversa de Moore-Penrose $A^{\dagger}$ de ordem $n \times m$ pode ser obtida através de um algoritmo cujos passos estão descritos a seguir:

1. Calculamos a matriz $B = A^{T}A$.
2. Definimos de $C_1 = I_n$.
3. A partir de agora, inicia-se um processo iterativo de $r-1$ passos, calculando-se $C_2, C_3, \ldots, C_r$ utilizando-se da expressão:

$$C_{i+1} = \tfrac{tr(C_iB)}{i}.I_n - C_iB \ , \quad \text{para } i = 1, 2, \ldots, r-1.$$

4. A inversa de Moore-Penrose da matriz A é dada por $A^{\dagger} = \frac{r}{tr(C_rB)}.C_rA^{T}$.

A matriz $C_{r+1}B = 0$ e $trC_rB \neq 0$.

Exemplo 5.18. Utilizando o algoritmo de Penrose, determinemos a inversa de Moore-Penrose da matriz $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

Observamos que $det\ A = 0$ e o determinante da submatriz $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \neq 0$. Então, $posto(A) = r = 2$. Logo, a obtenção da inversa Moore-Penrose de A através do algoritmo de Penrose é dada por

$$A^{\dagger} = \tfrac{r}{tr(C_rB)}.C_rA^{T} = \tfrac{2}{tr(C_2B)}.C_2A^{T}.$$

Calculemos $\quad B = A^{T}A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix}$.

Iniciando o processo de iteração temos:

$$C_1 = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ ; \ C_1.B = I_3.\begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix} \quad \text{e} \quad tr(C_1B) = 10 \ .$$

Logo, $C_2 = \frac{tr(C_1B)}{1}.I_2 - C_1B = \frac{10}{1}.\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix}$. Então,

$$C_2 = \begin{bmatrix} 8 & -3 & -1 \\ -3 & 4 & -3 \\ -1 & -3 & 8 \end{bmatrix} \quad \text{e} \quad C_2B = \begin{bmatrix} 8 & -3 & -1 \\ -3 & 4 & -3 \\ -1 & -3 & 8 \end{bmatrix} . \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 3 & -3 \\ 3 & 6 & -3 \\ -3 & 3 & 6 \end{bmatrix}$$

e $tr(C_2B) = 18$.

Assim, $A^{\dagger} = \frac{2}{tr(C_2B)}.C_2A^T$.

$$A^{\dagger} = \frac{2}{18}. \begin{bmatrix} 8 & -3 & -1 \\ -3 & 4 & -3 \\ -1 & -3 & 8 \end{bmatrix} . \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Portanto, $A^{\dagger} = \frac{1}{9} \begin{bmatrix} 1 & -5 & -4 \\ 2 & 1 & 1 \\ 1 & -4 & 5 \end{bmatrix}$.

Exemplo 5.19. Utilizando o algoritmo de Penrose, determinemos a inversa de Moore-Penrose da matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & -1 & 3 \end{bmatrix}$.

Observamos que $det\ A = 0$ e o determinante da submatriz $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \neq 0$. Então, $post(A) = r = 2$. Assim, a obtenção da inversa de Moore-Penrose de A através do algoritmo de Penrose é dada por

$$A^{\dagger} = \frac{r}{tr(C_rB)}.C_rA^T = \frac{2}{tr(C_2B)}.C_2A^T.$$

Calculemos $B = A^TA = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 3 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & 14 \end{bmatrix}$.

Iniciando o processo de iteração temos: $C_1 = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

$$C_1.B = I_3. \begin{bmatrix} 5 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & 14 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & 14 \end{bmatrix} \quad \text{e} \quad tr(C_1B) = 21 .$$

Logo, $C_2 = \frac{tr(C_1B)}{1}.I_3 - C_1B = \frac{21}{1}. \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & 14 \end{bmatrix}$.

Então, $C_2 = \begin{bmatrix} 16 & -2 & -4 \\ -2 & 19 & 2 \\ -4 & 2 & 7 \end{bmatrix}$.

$$C_2B = \begin{bmatrix} 16 & -2 & -4 \\ -2 & 19 & 2 \\ -4 & 2 & 7 \end{bmatrix} . \begin{bmatrix} 5 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & 14 \end{bmatrix} = \begin{bmatrix} 60 & 36 & 12 \\ 36 & 30 & -18 \\ 12 & -18 & 78 \end{bmatrix} \quad \text{e} \quad tr(C_2B) = 168 \ .$$

Assim, $A^\dagger = \frac{2}{tr(C_2B)}.C_2A^T \ = \ \frac{2}{168}. \begin{bmatrix} 16 & -2 & -4 \\ -2 & 19 & 2 \\ -4 & 2 & 7 \end{bmatrix} . \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 3 \end{bmatrix}.$

Portanto, $A^\dagger = \frac{1}{84}. \begin{bmatrix} 8 & 26 & -10 \\ 2 & 17 & -13 \\ 10 & 1 & 19 \end{bmatrix}.$

Exemplo 5.20. Utilizando o algoritmo de Penrose, determinemos a inversa de Moore-Penrose da matriz $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}.$

Observamos que $det\ A \neq 0$ então $posto(A) = r = 3$. Logo, a obtenção da inversa de Moore-Penrose de A através do algoritmo de Penrose é dada por

$$A^\dagger = \frac{r}{tr(C_rB)}.C_rA^T = \frac{2}{tr(C_2B)}.C_2A^T.$$

Calculemos $B = A^TA = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$

Iniciando o processo de iteração temos: $C_1 = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

$$C_1.B = I_3. \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{e} \quad tr(C_1B) = 9 \ .$$

Logo, $C_2 = \frac{tr(C_1B)}{1}.I_3 - C_1B = \frac{9}{1}. \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$

Então, $C_2 = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 6 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ e

$$C_2B = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 6 & -1 \\ 0 & -1 & 8 \end{bmatrix} . \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 3 & -3 \\ 3 & 8 & 5 \\ -3 & 5 & 7 \end{bmatrix} \quad \text{e} \quad tr(C_2B) = 26 \ .$$

Assim,

$$C_3 = \frac{tr(C_2B)}{2}.I_3 - C_2B = \frac{26}{2}. \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 11 & 3 & -3 \\ 3 & 8 & 5 \\ -3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 3 \\ -3 & 5 & -5 \\ 3 & -5 & 6 \end{bmatrix}.$$

$$C_3B = \begin{bmatrix} 2 & -3 & 3 \\ -3 & 5 & -5 \\ 3 & -5 & 6 \end{bmatrix} \begin{bmatrix} 5 & 3 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad tr(C_3B) = 3.$$

Logo,

$$A^\dagger = \tfrac{3}{tr(C_3B)}.C_3A^T = \tfrac{3}{3}. \begin{bmatrix} 2 & -3 & 3 \\ -3 & 5 & -5 \\ 3 & -5 & 6 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix}.$$

Portanto, $A^\dagger = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix}.$

5.5.2 Algoritmo de Greville

O algoritmo Greville é um outro processo iterativo desenvolvido em 1965 que permite calcular a pseudo-inversa de uma matriz A a partir das matrizes pseudo-inversas correspondentes a algumas de suas submatrizes.

Seja $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & a_{1k} & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k & \cdots & \alpha_n \end{bmatrix}$ em que

$\alpha_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}$ é a $k-ésima$ coluna de A.

Consideremos A_k a submatriz formada pelas primeiras k colunas da matriz A. Então,

$$A_k = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}.$$

Então, $A_k = \begin{bmatrix} A_{k-1} & \alpha_k \end{bmatrix}$.

A condição inicial da iteração é dada como sendo:

Se $\alpha_1 = 0$ (coluna nula), então $A_1^\dagger = 0$.

Se $\alpha_1 \neq 0$, então $A_1^\dagger = (\alpha_1^T.\alpha_1)^{-1}.\alpha_1^T$.

Definimos os vetores colunas δ_k e γ_k como sendo:

$$\delta_k = A_{k-1}^\dagger.\alpha_k \qquad \text{e} \qquad \gamma_k = \alpha_k - A_{k-1}\delta_k.$$

E um outro vetor B_k dado por:

Se $\gamma_k \neq 0$, então $B_k = \gamma_k^\dagger = (\gamma_k^T.\gamma_k)^{-1}.\gamma_k^T$.

Se $\gamma_k = 0$, então $B_k = (1 + \gamma_k^T.\delta_k)^{-1}.\delta_k^T A_{k-1}$.

Então, a matriz $A_k^\dagger$ da matriz A_k é calculada como sendo $A_k^\dagger = \begin{bmatrix} A_{k-1}^\dagger - \delta_k.B_k \\ B_k \end{bmatrix}$.

E o processo continua para $k = 2, 3 \ldots, n$.

Exemplo 5.21. Determinemos a pseudo-inversa da matriz $A = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & 2 & 1 \\ -1 & 2 & -1 & 0 \end{bmatrix}$ utilizando o algoritmo de Greville .

As matrizes colunas α_i são: $\alpha_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, $\alpha_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\alpha_3 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$ e $\alpha_4 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

1ª iteração:

$A_1 = \begin{bmatrix} \alpha_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \neq 0.$

Então, $A_1^\dagger = (\ \alpha_1^T \alpha_1\ )^{-1} \alpha_1^T$.

$$A_1^\dagger = \left(\begin{bmatrix} 0 & 1 & -1 \end{bmatrix} . \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right)^{-1} - \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}.$$

Logo, $A_1^\dagger = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

$$\delta_2 = A_1^\dagger . \alpha_2 \ = \ \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} . \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \end{bmatrix}.$$

$$\gamma_2 = \alpha_2 - A_1 \delta_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} . \begin{bmatrix} -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Daí, $B_2 = \gamma_2^\dagger = \left(\ \gamma_2^T . \gamma_2\ \right)^{-1} .\ \gamma_2^T = \left(\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} . \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)^{-1} . \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} =$

$$= \begin{bmatrix} 3 \end{bmatrix}^{-1} . \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \ = \ \begin{bmatrix} \frac{1}{3} \end{bmatrix} . \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Então, $B_2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$.

Assim, $A_2^\dagger = \begin{bmatrix} A_1^\dagger - \delta_2 B_2 \\ B_2 \end{bmatrix}$.

$\delta_2 B_2 = \begin{bmatrix} -1 \end{bmatrix} . \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$.

$A_1^\dagger - \delta_2 B_2 = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} - \begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{5}{6} & -\frac{1}{6} \end{bmatrix}$.

Portanto, $A_2^\dagger = \begin{bmatrix} \frac{1}{3} & \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$.

2ª iteração:

$$\delta_3 = A_2^\dagger \alpha_3 = \begin{bmatrix} \frac{1}{3} & \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} . \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{11}{6} \\ \frac{1}{3} \end{bmatrix}.$$

$$\gamma_3 = \alpha_3 - A_2\gamma_3 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 2 \end{bmatrix} . \begin{bmatrix} \frac{11}{6} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ \frac{11}{6} \\ -\frac{7}{6} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{6} \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Daí, $B_3 = \gamma_3^\dagger = \left(\gamma_3^T . \gamma_3 \right)^{-1} . \gamma_3^T = \left(\begin{bmatrix} -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} . \begin{bmatrix} -\frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{6} \end{bmatrix} \right)^{-1} . \begin{bmatrix} -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} =$

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

Então, $B_3 = \begin{bmatrix} -2 & 1 & 1 \end{bmatrix}$.

Assim, $A_3^\dagger = \begin{bmatrix} A_2^\dagger - \delta_3 B_3 \\ B_3 \end{bmatrix}$.

$$\delta_3 B_3 = \begin{bmatrix} \frac{11}{6} \\ \frac{1}{3} \end{bmatrix} . \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{11}{3} & \frac{11}{6} & \frac{11}{6} \\ -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

$$A_2^\dagger - \delta_3 B_3 = \begin{bmatrix} \frac{1}{3} & \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} - \begin{bmatrix} -\frac{11}{3} & \frac{11}{6} & \frac{11}{6} \\ -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 4 & -1 & -2 \\ 1 & 0 & 0 \end{bmatrix}.$$

Portanto, $A_3^\dagger = \begin{bmatrix} A_2^\dagger - \delta_3 B_3 \\ B_3 \end{bmatrix} = \begin{bmatrix} 4 & -1 & -2 \\ 1 & 0 & 0 \\ -2 & 1 & 1 \end{bmatrix}$.

3ª iteração:

$$\delta_4 = A_3^\dagger . \alpha_4 = \begin{bmatrix} 4 & -1 & -2 \\ 1 & 0 & 0 \\ -2 & 1 & 1 \end{bmatrix} . \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix}.$$

$$\gamma_4 = \alpha_4 - A_3 . \delta_4 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ -1 & 2 & -1 \end{bmatrix} . \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Daí, $B_4 = \left(1 + \delta_4^T . \delta_4 \right)^{-1} \delta_4^T . A_3$.

$$B_4 = \left(1 + \begin{bmatrix} 7 & 2 & -3 \end{bmatrix} . \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix} \right)^{-1} . \begin{bmatrix} 7 & 2 & -3 \end{bmatrix} . \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ -1 & 2 & -1 \end{bmatrix} =$$

$$= \left(1 + \begin{bmatrix} 62 \end{bmatrix} \right)^{-1} . \begin{bmatrix} 5 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 63 \end{bmatrix}^{-1} . \begin{bmatrix} 5 & 1 & 7 \end{bmatrix} = \begin{bmatrix} \frac{1}{63} \end{bmatrix} . \begin{bmatrix} 5 & 1 & 7 \end{bmatrix}.$$

Então, $B_4 = \begin{bmatrix} \frac{5}{63} & \frac{1}{63} & \frac{1}{9} \end{bmatrix}$.

Mas, $A_4^\dagger = \begin{bmatrix} A_3^\dagger - \delta_4 B_4 \\ B_4 \end{bmatrix}$.

$$\delta_4 B_4 = \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix} . \begin{bmatrix} \frac{5}{63} & \frac{1}{63} & \frac{1}{9} \end{bmatrix} = \begin{bmatrix} \frac{5}{9} & \frac{1}{9} & \frac{7}{9} \\ \frac{10}{63} & \frac{2}{63} & \frac{2}{9} \\ -\frac{5}{21} & -\frac{1}{21} & -\frac{1}{3} \end{bmatrix}.$$

$$A_3^\dagger - \delta_4 B_4 = \begin{bmatrix} 4 & -1 & -2 \\ 1 & 0 & 0 \\ -2 & 1 & 1 \end{bmatrix} - \begin{bmatrix} \frac{5}{9} & \frac{1}{9} & \frac{7}{9} \\ \frac{10}{63} & \frac{2}{63} & \frac{2}{9} \\ -\frac{5}{21} & -\frac{1}{21} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{31}{9} & -\frac{10}{9} & -\frac{25}{9} \\ \frac{53}{63} & -\frac{2}{63} & -\frac{2}{9} \\ -\frac{37}{21} & \frac{22}{21} & \frac{4}{3} \end{bmatrix}.$$

Portanto $A^\dagger = \begin{bmatrix} \frac{31}{9} & -\frac{10}{9} & -\frac{25}{9} \\ \frac{53}{63} & -\frac{2}{63} & -\frac{2}{9} \\ -\frac{37}{21} & \frac{22}{21} & \frac{4}{3} \\ \frac{5}{63} & \frac{1}{63} & \frac{1}{9} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} \frac{31}{3} & -\frac{10}{3} & -\frac{25}{3} \\ \frac{53}{21} & -\frac{2}{21} & -\frac{2}{3} \\ -\frac{37}{7} & \frac{22}{7} & 4 \\ \frac{5}{21} & \frac{1}{21} & \frac{1}{3} \end{bmatrix}$ é a inversa de Moore-Penrose da matriz A.

Exemplo 5.22. Determinemos a pseudo-inversa da matriz $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ utilizando o algoritmo de Greville.

As matrizes colunas α_i são: $\alpha_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $\alpha_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\alpha_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ *e* $\alpha_4 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

1ª iteração:

$A_1 = \begin{bmatrix} \alpha_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0.$ Então, $A_1^\dagger = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$.

Assim, $\delta_2 = A_1^\dagger.\alpha_2 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$.

$$\gamma_2 = \alpha_2 - A_1\delta_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.\begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Daí, $B_2 = \gamma_2^\dagger = \left(\gamma_2^T.\gamma_2 \right)^{-1}.\ \gamma_2^T = \left(\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}.\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right)^{-1}.\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} =$

$$= \begin{bmatrix} 2 \end{bmatrix}^{-1}.\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \end{bmatrix}.\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}.$$

Então, $B_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$.

Assim, $A_2^\dagger = \begin{bmatrix} A_1^\dagger - \delta_2 B_2 \\ B_2 \end{bmatrix}$.

$\delta_2 B_2 = \begin{bmatrix} 0 \end{bmatrix}.\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$.

$A_1^\dagger - \delta_2 B_2 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$.

Portanto, $A_2^\dagger = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$.

2ª iteração:

$$\delta_3 = A_2^\dagger \alpha_3 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2} \end{bmatrix}.$$

$$\gamma_3 = \alpha_3 - A_2\gamma_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} . \begin{bmatrix} 0 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Daí, $B_3 = \gamma_3^\dagger = \left(\gamma_3^T.\gamma_3 \right)^{-1} . \gamma_3^T = \left(\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} . \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right)^{-1} . \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} =$

$$= \begin{bmatrix} \frac{3}{2} \end{bmatrix}^{-1} . \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}.$$

Então, $B_3 = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$.

Mas, $A_3^\dagger = \begin{bmatrix} A_2^\dagger - \delta_3 B_3 \\ B_3 \end{bmatrix}$. $\qquad \delta_3 B_3 = \begin{bmatrix} 0 \\ -\frac{1}{2} \end{bmatrix} . \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{bmatrix}.$

Logo, $A_2^\dagger - \delta_3 B_3 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$

Portanto, $A_3^\dagger = \begin{bmatrix} A_2^\dagger - \delta_3 B_3 \\ B_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$

3ª iteração:

$$\delta_4 = A_3^\dagger.\alpha_4 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} . \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{5}{3} \\ \frac{4}{3} \end{bmatrix}.$$

$$\gamma_4 = \alpha_4 - A_3.\delta_4 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 0 \\ \frac{5}{3} \\ \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ \frac{4}{3} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Então, $B_4 = \left(\delta_4^T.\delta_4 \right)^{-1} \delta_4^T$.

$$B_4 = \left(\begin{bmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} . \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \right)^{-1} . \begin{bmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \end{bmatrix}^{-1} . \begin{bmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} =$$

$= \begin{bmatrix} \frac{3}{4} \end{bmatrix} . \begin{bmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$. $\qquad$ Então, $B_4 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Mas, $A_4^\dagger = \begin{bmatrix} A_3^\dagger - \delta_4 B_4 \\ B_4 \end{bmatrix}$. $\qquad \delta_4 B_4 = \begin{bmatrix} 0 \\ \frac{5}{3} \\ \frac{4}{3} \end{bmatrix} . \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -\frac{5}{6} & \frac{5}{6} & \frac{5}{6} \\ -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}.$

$$A_3^\dagger - \delta_4 B_4 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ -\frac{5}{6} & \frac{5}{6} & \frac{5}{6} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -1 & 0 \end{bmatrix}.$$

Portanto, $A^\dagger = \begin{bmatrix} 0 & 0 & 0 \\ \frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 \\ 3 & -1 & -1 \\ 2 & -2 & 0 \\ -1 & 1 & 1 \end{bmatrix}$ é a inversa de Moore-Penrose da matriz A.

6 Aplicações da Inversa Generalizada

Neste capítulo apresentaremos duas aplicações envolvendo o conceito de inversa generalizada; uma aplicação a ser abordada envolve o cálculo de solução de sistemas não quadrados e a outra envolve problemas de mínimos quadrados. As referências utilizadas foram [9], [10], [11], [15], [20] , [22] e [24].

6.1 A inversa generalizada na resolução de sistemas lineares

Nesta seção veremos uma aplicações da inversa generalizada aos sistemas lineares com coeficientes constantes. Num sistema de equações lineares da forma $Ax = b$, quando a matriz A é não singular ela admite inversa A^{-1}, e o sistema admite uma única solução. Neste caso, a solução é obtida bastando multiplicar à esquerda por A^{-1} ambos os membros da equação por $Ax = b$.

$$A^{-1}Ax = A^{-1}b \quad \Rightarrow \quad x = A^{-1}b$$

No caso da matriz A ser singular, ou não ser uma matriz quadrada podemos utilizar a inversa generalizada de A para a resolução do sistema.

Consideremos o sistema de equações lineares $Ax = b$, em que A é uma matriz $m \times n$ de números reais, X um vetor coluna $n \times 1$ de variáveis reais e b um vetor coluna $m \times 1$ de números reais. Lembremos que: dado o sistema de equações lineares $Ax = b$, se existir ao menos um vetor x^* que satisfaça o sistema, dizemos que o sistema é **consistente**. Caso contrário, o sistema é **inconsistente**.

Os teoremas a seguir estabelecem a aplicação da inversa generalizada A^- na resolução de sistemas de equações lineares.

Teorema 6.1. *Seja $Ax = b$ um sistema linear consistente. Então,*
$x^* = A^-b$ é uma solução do sistema se, e somente se, $AA^-b = b$.

Demonstração. Consideremos $x^* = A^-b$ uma solução da equação $Ax = b$.

Pré multiplicando por AA^- ambos os membros da equação $Ax = b$, temos:

$$AA^-Ax = AA^-b.$$
$$Ax = AA^-b.$$

Logo, $b = AA^-b$.

Agora, consideremos $x^* = A^-b$, então $Ax^* = AA^-b$.

Mas, por hipótese $AA^-b = b$. Assim, $Ax^* = b$.

Portanto, $x^* = A^- b$ é solução do sistema $Ax = b$.

□

Teorema 6.2. *A solução geral de um sistema consistente $Ax = b$ é dada por*

$$x = A^- b + (I_n - A^- A)h \ ,$$

sendo A^- qualquer inversa generalizada de A, I_n a matriz identidade de ordem n e h um vetor coluna n-dimensional chamado vetor parâmetro.

Demonstração.

$$Ax = A[\ A^- b + (I_n - A^- A)h\ ] \ = \ AA^- b + A(I_n - A^- A)h = \ AA^- b + (AI_n - AA^- A)h.$$

Pelo fato do sistema $Ax = b$ ser consistente, então $AA^- b = b$ (Teorema 5.1.), e como $AA^- A = A$ (Definição 4.5.), então temos

$$Ax = b + (A - A)h = b.$$

Logo, x é solução do sistema $Ax = b$, para todo h.

Caso $AA^- = I_n$, o sistema tem solução única, caso contrário o sistema possui infinitas soluções, parametrizadas pelo vetor h.

□

Exemplo 6.3. Vamos resolver o sistema linear $\begin{cases} x_1 + x_2 = 1 \\ x_2 + x_3 \ = 3 \\ 2x_1 + x_2 \ = 0 \end{cases}$.

Então temos:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \ ; \ A^- = A^\dagger = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix} \ \text{(Exemplo 4.20)} \ \text{ e } \ b = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}.$$

Condição de consistência: $AA^- b = b$.

$$AA^- b = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = b.$$

Logo, o sistema é consistente, e como $AA^- = I_3$, então sua solução é dada por

$$x^* = A^- b$$

$$x^* = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Portanto, $x_1 = -1$, $x_2 = 2$ e $x_3 = 1$.

Exemplo 6.4. Vamos resolver o sistema linear $\begin{cases} x_1 + 2x_3 = 1 \\ 2x_1 + x_2 + x_3 = 2 \\ -x_2 + 3x_3 = 0 \end{cases}$.

Então temos:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & -1 & 3 \end{bmatrix} \;;\; A^- = A^\dagger = \tfrac{1}{84}\begin{bmatrix} 8 & 26 & -10 \\ 2 & 17 & -13 \\ 10 & 1 & 19 \end{bmatrix} \text{(Exemplo 4.19.)} \;\text{ e }\; b = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Condição de consistência: $AA^-b = b$.

$$AA^-b = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & -1 & 3 \end{bmatrix} \cdot \tfrac{1}{84}\begin{bmatrix} 8 & 26 & -10 \\ 2 & 17 & -13 \\ 10 & 1 & 19 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = b.$$

Logo, o sistema é consistente e sua solução é dada por

$$x^* = A^-b + (I_3 - A^-A)h.$$

$$A^-A = \tfrac{1}{84}\begin{bmatrix} 8 & 26 & -10 \\ 2 & 17 & -13 \\ 10 & 1 & 19 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & -1 & 3 \end{bmatrix} = \begin{bmatrix} \frac{5}{7} & \frac{3}{7} & \frac{1}{7} \\ \frac{3}{7} & \frac{5}{14} & -\frac{3}{14} \\ \frac{1}{7} & -\frac{3}{14} & \frac{13}{14} \end{bmatrix} \neq I_3.$$

Assim, o sistema admite infinitas soluções.

$$I_3 - A^-A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{5}{7} & \frac{3}{7} & \frac{1}{7} \\ \frac{3}{7} & \frac{5}{14} & -\frac{3}{14} \\ \frac{1}{7} & -\frac{3}{14} & \frac{13}{14} \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & -\frac{3}{7} & -\frac{1}{7} \\ -\frac{3}{7} & \frac{9}{14} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{1}{14} \end{bmatrix}.$$

Então,

$$x^* = \tfrac{1}{84}\begin{bmatrix} 8 & 26 & -10 \\ 2 & 17 & -13 \\ 10 & 1 & 19 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{2}{7} & -\frac{3}{7} & -\frac{1}{7} \\ -\frac{3}{7} & \frac{9}{14} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{1}{14} \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}.$$

$$x^* = \tfrac{1}{84}\begin{bmatrix} 60 \\ 36 \\ 12 \end{bmatrix} + \begin{bmatrix} \frac{2}{7}h_1 - \frac{3}{7}h_2 - \frac{1}{7}h_3 \\ -\frac{3}{7}h_1 + \frac{9}{14}h_2 + \frac{3}{14}h_3 \\ -\frac{1}{7}h_1 + \frac{3}{14}h_2 + \frac{1}{14}h_3 \end{bmatrix} = \begin{bmatrix} \frac{5}{7} \\ \frac{3}{7} \\ \frac{1}{7} \end{bmatrix} + \begin{bmatrix} \frac{2}{7}h_1 - \frac{3}{7}h_2 - \frac{1}{7}h_3 \\ -\frac{3}{7}h_1 + \frac{9}{14}h_2 + \frac{3}{14}h_3 \\ -\frac{1}{7}h_1 + \frac{3}{14}h_2 + \frac{1}{14}h_3 \end{bmatrix}.$$

Portanto, $$x^* = \begin{bmatrix} \frac{5}{7} + \frac{2}{7}h_1 - \frac{3}{7}h_2 - \frac{1}{7}h_3 \\ \frac{3}{7} - \frac{3}{7}h_1 + \frac{9}{14}h_2 + \frac{3}{14}h_3 \\ \frac{1}{7} - \frac{1}{7}h_1 + \frac{3}{14}h_2 + \frac{1}{14}h_3 \end{bmatrix}.$$

Assim, $x_1 = \frac{5}{7} + \frac{2}{7}h_1 - \frac{3}{7}h_2 - \frac{1}{7}h_3$, $x_2 = \frac{3}{7} - \frac{3}{7}h_1 + \frac{9}{14}h_2 + \frac{3}{14}h_3$ e $x_3 = \frac{1}{7} - \frac{1}{7}h_1 + \frac{3}{14}h_2 + \frac{1}{14}h_3$, com h_1 e $h_3 \in \mathbb{R}$.

Exemplo 6.5. Vamos resolver o sistema linear $\begin{cases} x_1 + 2x_2 + x_3 = 3 \\ x_1 + x_2 \; = 2 \\ x_2 + x_3 \; = 0 \end{cases}$.

Então temos:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \;;\; A^- = A^\dagger = \tfrac{1}{9}.\begin{bmatrix} 1 & 5 & -4 \\ 2 & 1 & 1 \\ 1 & -4 & 5 \end{bmatrix} \text{(Exemplo 4.18)} \;\text{ e }\; b = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}.$$

Condição de consistência: $AA^-b = b$.

$$AA^-b = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tfrac{1}{9}\begin{bmatrix} 1 & 5 & -4 \\ 2 & 1 & 1 \\ 1 & -4 & 5 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} = \tfrac{1}{9}\begin{bmatrix} 6 & 3 & 3 \\ 3 & 6 & -3 \\ 3 & -3 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} =$$

$$= \frac{1}{9}\begin{bmatrix} 24 \\ 21 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} = b\ . \quad \text{Portanto, o sistema é inconsistente.}$$

6.2 Solução de mínimos quadrados.

Em várias situações, os sistemas lineares aparecem como modelos matemáticos de vários fenômenos. Ocorre que alguns sistemas lineares não possuem soluções, isso acontece em aplicações nas quais erros de medição afetam os coeficientes das equações de um sistema consistente de tal modo que o sistema se torna inconsistente.

Suponhamos que o sistema $Ax = b$ seja inconsistente de m equações e n incógnitas em que essa inconsistência ocorreu por erros de medição nos coeficientes de A. Como não é possível obtermos uma solução exata, vamos procurar um solução que seja a mais "próxima possível" de ser uma solução, no sentido de que $\| b - Ax \|$ (em relação ao produto euclidiano de $\mathbb{R}^m$) seja mínimo. Sendo Ax uma aproximação de b e $\| b - Ax \|$ o erro dessa aproximação, então quanto menor o erro, melhor a aproximação. Esse processo é denominado **Método dos Mínimos Quadrados**, e é uma técnica que permite obter informações de sistemas inconsistentes, de forma aproximada.

A terminologia mínimos quadrados se deve ao fato de que, esse método minimiza a soma dos quadrados dos erros obtidos na aproximação.

Suponhamos que na forma matricial $b - Ax = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$, entao podemos observar que: minimizar $\| b - Ax \|$ implica também minimizar $\| b - Ax \|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$.

Definição 6.6. Seja $Ax = b$ um sistema linear de m equações e n incógnitas. Um vetor x que minimiza $\| b - Ax \|$ em relação ao produto interno canônico de $\mathbb{R}^m$ é chamado de uma **solução de mínimos quadrados** do sistema linear, e que $b - Ax$ é o **vetor erro de mínimos quadrados** e $\| b - Ax \|$ é o **erro de mínimos quadrados**.

Teorema 6.7. *Dado um sistema de equações lineares $Ax = b$, com $A_{m\times n}$, então $x = A^\dagger b$ é um minimizador de $\| Ax - b \|$.*

Além disso é o minimizador da norma euclidiana $\| \ . \ \|_2$, ou seja, $A^\dagger b$ é a solução do sistema $Ax = b$.

Portanto, a melhor solução aproximada do sistema linear inconsistente $Ax = b$ é dada por $x^* = A^\dagger b$. A prova deste teorema a seguir pode ser encontrado nas referências [15] e [20].

Exemplo 6.8. Encontremos a melhor solução aproximada do sistema inconsistente

$$\begin{cases} 2x_1 + 3x_2 = 1 \\ 3x_1 + 5x_2 = -2 \\ -2x_1 - 4x_2 = -1 \end{cases} .$$

Solução:

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ -2 & -4 \end{bmatrix}, \quad post(A) = 2 \qquad \text{e} \qquad b = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

Então, $\quad A^{\dagger} = (A^T A)^{-1} A^T.$

$$A^T A = \begin{bmatrix} 2 & 3 & -2 \\ 3 & 5 & -4 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ -2 & -4 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 29 & 50 \end{bmatrix} \quad \text{e} \quad (A^T A)^{-1} = \tfrac{1}{9} \begin{bmatrix} 50 & -29 \\ -29 & 17 \end{bmatrix}.$$

Logo, $\quad A^{\dagger} = \tfrac{1}{9} \begin{bmatrix} 50 & -29 \\ -29 & 17 \end{bmatrix} \begin{bmatrix} 2 & 3 & -2 \\ 3 & 5 & -4 \end{bmatrix} = \tfrac{1}{9} \begin{bmatrix} 13 & 5 & 16 \\ -7 & -2 & -10 \end{bmatrix}.$

Assim, a melhor solução aproximada x^* do sistema é dado por

$$x^* = A^{\dagger} b = \tfrac{1}{9} \begin{bmatrix} 13 & 5 & 16 \\ -7 & -2 & -10 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} = \tfrac{1}{9} \begin{bmatrix} -13 \\ 7 \end{bmatrix}.$$

Portanto, $\quad x^* = \begin{bmatrix} -\frac{13}{9} \\ \frac{7}{9} \end{bmatrix}.$

O teorema a seguir permite calcularmos uma melhor aproximação de um sistema de uma outra maneira.

Teorema 6.9. *Se W é um subespaço de dimensão finita de um espaço vetorial V com produto interno e b um vetor de V, então $proj_W\, b$ é a melhor aproximação de b à W, no sentido de que $\| b - proj_W\, b \| < \| b - w \|$ qualquer que seja o vetor w em W distinto de $proj_W\, b$.*

Demonstração. Dado qualquer vetor $w \in W$ podemos escrever

$$b - w = (b - proj_W\, b) + (proj_W\, b - w)\ .$$

Sendo uma diferença de vetores de W, o vetor $(proj_W\, b - w) \in W$ e, como $b - proj_W\, b$ é ortogonal a W, então $b - proj_W\, b$ e $proj_W\, b - w$ são ortogonais.

Assim, pelo Teorema de Pitágoras ($Teorema$ 1.76) , temos:

$$\| b - w \|^2 = \| b - proj_W\, b \|^2 + \| proj_W\, b - w \|^2.$$

Como $b \neq proj_W\, b$ segue que $\| proj_W\, b - w \| > 0$ e, portanto,

$$\| b - proj_W\, b \|^2 < \| b - w \|^2.$$

Como as normas são não negativas, segue que $\quad \| b - proj_W\, b \| < \| b - w \|.$

□

Uma maneira de encontrarmos alguma solução de um sistema linear $Ax = b$ pelo método dos mínimos quadrados consiste em calcularmos os vetores p , sendo $p = proj_W\, b$, em que W é o espaço coluna de A , e resolvermos o sistema linear $Ax = p$, denominado sistema normal associado ao sistema $Ax = b$. Portanto, a ideia é considerarmos os vetores pertencentes ao espaço coluna de A que sejam "mais próximos" possíveis de b e cujo sistema linear associado ao $Ax = b$ possua solução.

O teorema a seguir permite evitarmos o cálculo da projeção reescrevendo $Ax = p$ na forma $A^T Ax = A^T b$.

Teorema 6.10. *Dado qualquer sistema linear* $Ax = b$*, o sistema normal associado* $A^TAx = A^Tb$ *é consistente, e todas as soluções de* $A^TAx = A^Tb$ *são soluções de mínimos quadrados de* $Ax = b$*. Além disso, se* W *for o espaço coluna de* A *e* x *uma solução de mínimos quadrados qualquer de* $Ax = b$*, então a projeção ortogonal de* B *em* W *é* $proj_W b = Ax$*.*

Demonstração. Pelo teorema anterior $proj_W\ b$ é a melhor aproximação de b em W, então $Ax = proj_W\ b$.

Assim, $b - Ax = b - proj_W\ b$.

Multiplicando-se ambos os membros da equação por A^T, temos:

$$A^T(b - Ax) = A^T(b - proj_W\ b).$$

Como $b - proj_W\ b$ é o componente de b que é ortogonal ao espaço coluna de A, segue que este vetor está no espaço nulo de A^T. Então,

$$A^T(b - proj_W\ b) = 0.$$

Assim, $$A^T(\ b - Ax\ ) = 0.$$

Portanto, $$A^TAx = A^Tb.$$

□

Assim, dado qualquer sistema linear $Ax = b$, o sistema normal associado dado por $A^TAx = A^Tb$ é consistente, e todas as suas soluções são soluções de mínimos quadrados de $Ax = b$.

Exemplo 6.11.

(a) Encontremos todas as soluções de mínimos quadrados do sistema inconsistente

$$\begin{cases} 2x_1 + 3x_2 = \ \ 1 \\ 3x_1 + 5x_2 = -2 \\ -2x_1 - 4x_2 = -1 \end{cases}.$$

(b) Encontremos o vetor erro e o erro.

Solução:

(a) É conveniente expressarmos o sistema na forma matricial.

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ -2 & -4 \end{bmatrix} \qquad \text{e} \qquad b = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

Segue que, $$A^TA = \begin{bmatrix} 2 & 3 & -2 \\ 3 & 5 & -4 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ -2 & -4 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 29 & 50 \end{bmatrix}.$$

$$A^Tb = \begin{bmatrix} 2 & 3 & -2 \\ 3 & 5 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}.$$

Utilizando o sistema normal associado $A^TAx = A^Tb$ temos:

$$\begin{bmatrix} 17 & 29 \\ 29 & 50 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}.$$

Resolvendo este matricial consistente $A^TAx = b$ temos que $x = (A^TA)^{-1}b$.

Assim, $$x = \begin{bmatrix} 19 & 29 \\ 29 & 50 \end{bmatrix}^{-1} \begin{bmatrix} -2 \\ -3 \end{bmatrix} = \frac{1}{850-841} \begin{bmatrix} 50 & -29 \\ -29 & 19 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \end{bmatrix}.$$

$$x = \frac{1}{9} \begin{bmatrix} -13 \\ 7 \end{bmatrix} = \begin{bmatrix} -\frac{13}{9} \\ \frac{7}{9} \end{bmatrix}.$$

Portanto, $x_1 = -\frac{13}{9}$ e $x_2 = \frac{7}{9}$.

(b) O vetor erro é dado por

$$b - Ax = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} -\frac{13}{9} \\ \frac{7}{9} \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} - \begin{bmatrix} -\frac{5}{9} \\ -\frac{4}{9} \\ -\frac{2}{9} \end{bmatrix}.$$

Portanto, $$b - Ax = \begin{bmatrix} \frac{14}{9} \\ -\frac{14}{9} \\ -\frac{7}{9} \end{bmatrix}.$$

O erro é dado por

$$\| b - Ax \| = \sqrt{(\tfrac{14}{9})^2 + (\tfrac{-14}{9})^2 + (\tfrac{-7}{9})^2} = \sqrt{\frac{441}{81}} .$$

Portanto, $\| b - Ax \| \approx 2,33$.

Consideremos uma tabela de n pontos (x_i, y_i), em que y_i foram obtidos experimentalmente, e queremos expressar estes pontos por uma reta $y = a_1x + a_2$ que melhor se ajuste a eles. A obtenção deste ajuste, pelo método dos mínimos quadrados, consiste em resolvermos um sistema linear inconsistente de n equações à duas incógnitas a_1 e a_2 a serem determinadas, como veremos no exemplo a seguir.

Exemplo 6.12. Seja o conjunto de pontos

$$\{ (1\ ,\ 3.8),\ (2\ ,\ 7.3),\ (3\ ,\ 10.2),\ (4\ 12.7),\ (\ 5,\ 15.5)\ \}.$$

Determinemos a função $g(x) = a_1.x + a_2$ que melhor se ajusta no sentido de mínimos quadrados aos pontos desse conjunto.

$$g(x) = a_1.x + a_2.$$

Assim, temos:

$g(1) = 1.a_1 + a_2 = 3.8$; $g(2) = 2.a_1 + a_2 = 7.3$; $g(3) = 3.a_1 + a_2 = 10.2$;
$g(4) = 4.a_1 + a_2 = 12.7$; $g(5) = 5.a_1 + a_2 = 15.5$.

Resolvendo o sistema linear $$\begin{cases} a_1 + a_2 = 3.8 \\ 2a_1 + a_2 = 7.3 \\ 3a_1 + a_2 = 10.2 \\ 4a_1 + a_2 = 12.7 \\ ax_1 + a_2 = 15.5 \end{cases}$$, temos:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} \quad \text{e} \quad A^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Utilizando o sistema normal associado $A^T Ax = A^T b$, temos:

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 55 & 15 \\ 15 & 5 \end{bmatrix}.$$

$$(A^T A)^{-1} = \tfrac{1}{275-225} \begin{bmatrix} 5 & -15 \\ -15 & 55 \end{bmatrix} = \tfrac{1}{50} \begin{bmatrix} 5 & -15 \\ -15 & 55 \end{bmatrix} = \begin{bmatrix} 0.1 & -0.3 \\ -0.3 & 1.1 \end{bmatrix}.$$

$$A^T b = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3.8 \\ 7.3 \\ 10.2 \\ 12.7 \\ 15.5 \end{bmatrix} = \begin{bmatrix} 177.3 \\ 49.5 \end{bmatrix}.$$

Assim, $$\begin{bmatrix} 0.1 & -0.3 \\ -0.3 & 1.1 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 177.3 \\ 49.5 \end{bmatrix}.$$

Como a solução do sistema normal é dado por $x = (A^T A)^{-1} A^T b$. Então,

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0.1 & -0.3 \\ -0.3 & 1.1 \end{bmatrix} \cdot \begin{bmatrix} 177.3 \\ 49.5 \end{bmatrix} = \begin{bmatrix} 2.88 \\ 1.26 \end{bmatrix}.$$

Assim, $a_1 = 2.88$ e $a_2 = 1.26$. Portanto, $g(x) = 2.88x + 1.26$.

O vetor erro é dado por

$$b - Ax = \begin{bmatrix} 3.8 \\ 7.3 \\ 10.2 \\ 12.7 \\ 15.5 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 2.88 \\ 1.26 \end{bmatrix} = \begin{bmatrix} 3.8 \\ 7.3 \\ 10.2 \\ 12.7 \\ 15.5 \end{bmatrix} - \begin{bmatrix} 4.14 \\ 7.02 \\ 9.90 \\ 12.78 \\ 15.66 \end{bmatrix} = \begin{bmatrix} -0.34 \\ 0.28 \\ 0.30 \\ -0.08 \\ -0.16 \end{bmatrix}.$$

O erro é dado por

$$\| b - Ax \| = \sqrt{(-0.34)^2 + (0.28)^2 + (0,30)^2 + (-0.08)^2 + (-0.16)^2} = \sqrt{0.316} \, .$$

Portanto, $\| b - Ax \| \approx 0.56$.

7 Considerações Finais

O conceito da pseudo-inversa é um assunto relativamente recente e nas últimas décadas este conceito encontrou uma série de aplicações em muitas áreas da ciência e tornou-se uma ferramenta útil para matemáticos aplicados, engenheiros e físicos que trabalham, por exemplo, com análise de dados, problemas de otimização, solução de equações integrais lineares e outras.

O tema pseudo-inversa é pouco explorado em trabalhos de conclusão de curso de graduação e dissertação de mestrado, talvez pelo assunto não ser ministrado no curso regular de gradução.

Este trabalho teve por objetivo apresentar de modo didático e ilustrado através de exemplos aplicações da pseudo-inversa na resolução de sistemas lineares. E este contéudo também pode ser aplicado aos alunos da 2a. série do Ensino Médio como uma atividade complementar ao estudo das matrizes e na resolução de sistemas lineares consistentes através da inversa de uma matriz. O cálculo da pseudo-inversa para resolução do sistema linear por estes alunos, seria feito através do que vimos na Seção 4.4. (a),(b),(d), e no caso do item (c) trocaria-se a DVS pelo algoritmo de Penrose.

Referências

[1] PULINO, P. *Álgebra Linear e suas Aplicações (Notas de aulas)*. Campinas: Edição Eletrônica, Unicamp, 2012.

[2] LIPSCHUTZ, S. *Álgebra Linear*. São Paulo: Editora McGraw-Hill, 1972.

[3] BOLDRINI J.L. ; COSTA S.I.R. ; FIGUEIREDO V.L. ; WETZLER H.G. . *Álgebra Linear*. São Paulo: Harper Row do Brasil, 1980.

[4] STEINBRUCH, A. ; WINTERLE P. *Introdução à Álgebra Linear*. São Paulo: Editora McGraw-Hill, 1987.

[5] HOWARD, A. ; RORRES C. *Álgebra Linear com Aplicações*. Porto Alegre: Bookman, 2012.

[6] MALAJOVICH, G. *Álgebra Linear*. Rio de Janeiro: Edição Eletrônica Preliminar, 2010.

[7] MOORE, E. H. *On the reciprocal of the general algebraic matrix*, Bull, Amer. Math. Soc. , v.26, p. 394 - 395, 1920.

[8] PENROSE, R. *A generalized inverse for matrices*. Proc. Cambridge Philos Soc., v.51, p. 406-413, 1955.

[9] BEN-ISRAEL, A. ; GREVILLE, T.N.E. . *Generalized inverses: Theory and Aplications*. New Jersey: Rutcor-Rutgers Rd, 2001.

[10] THOME, N. *La inversa generalizada de Moore-Penrose y aplicaciones, Vol.21*. València: Publicaciones Electrónicas Sociedade Matemática Mexicana, 2019.

[11] BHAYA, A. *Decomposição em valores singulares, pseudoinversa, norma*. Rio de Janeiro: Edição Eletrônica UFRJ.

[12] SALAS, J. ; TORRENTE, A. ; VILLASEÑOR E.J.S. . *Álgebra Lineal - Tema 14*. Madrid: Edição Eletrônica, Universidad Carlos III, 2015.

[13] FALEIROS, A. C. *Álgebra Linear e Aplicação*. Santo André: Edição Eletrônica Univ.Federal do ABC, 2009.

[14] SOTO, J. *Pseudoinversa de una Matriz*. Múrcia: Universidad Católica de Múrcia, URL: youtube.com/watch?v=g6ED94Wjrlg, 2016.

[15] GRAYBILL, F.A. *Introduction to matrices with applications*. Belmont, Califórnia, Wadsworth Pub. Cia, 1969.

[16] OSSANI, P.C. *Decomposição em Valores Singulares.*
URL: youtube.com/watch?v=13CX9YqQ2FQ, 2020.

[17] DUARTE, F. ; MACHADO. J.A.T. . *Matrizes Pseudoinversas: Aspectos Matemáticos e Aplicação ao Controlo de Manipuladores Redundantes.* Salamanca: 5ª Jornadas Hispano-Lusas de Ingeniería Eléctrica, 1997.

[18] PELLEGRINI, J. C. . *Álgebra Linear.* Edição Eletrônica Preliminar, 2015.

[19] MONGE, Juan J.Fallas ; MOLINA, Jeffry Chavarría ; QUIROS, Pablo Soto . *Descomposición en valores singulares de una matriz: un repaso por los fundamentos teóricos y sus aplicaciones en el procesamiento de imágenes.* Costa Rica: Investigación Operacional, vol.42 nº2 , pág.142-173, Inst.Tec.de Costa Rica, 2021.

[20] RAO, C.Radhakrishna ; MITRA, Sujit Kumar. *Generalized Inverse of Matrix and its Applicátion..* New York: Edição Eletrônica, Wiley & Sons Inc, 1971.

[21] CAMPBELL, S.L. ; MEYER JR, C.D.. *Generalized Inverses of Linear Transformations.* New York: Dover Pub.Inc, 1991.

[22] GRUBER, M.H.J. . *Matrix Algebra for Linear Models.* New York: John Wiley & Sons Inc , 2014.

[23] RAO, C.R. *Calculus of Generalized Inverse of Matrices.* Part I: General Theory, Sankhya Ser. A, vol 29, 1967.

[24] STRANG, G. *Álgebra Linear e suas Aplicações*: Editora CENGAGE Learning, São Paulo, tradução da 4ª edição norte-americana, 2012.

[25] PENROSE, R. *On best approximate solutions of linear matrix equations.* Proc. Cambridge Philos Soc., v.52, p. 17-19, 1956.