Revista Sergipana de Matemática e Educação Matemática

https://doi.org/10.34179/revisem.v10i2.22067

MATRIZES QUADRADAS DE POSTO 1: CRITÉRIO DE DIAGONALIZAÇÃO.

RANK-ONE SQUARE MATRICES: A DIAGONALIZATION CRITERION

Francisco Ferreira do Nascimento Junior
Universidade Federal do Piauí - UFPI
juniorgmath78@gmail.com

Marcos Vinicio Travaglia
Universidade Federal do Piauí - UFPI
mvtravaglia@ufpi.edu.br

Comunicado por Z. Ramos

Resumo

Neste trabalho apresentamos duas abordagens para demonstrar o resultado conhecido de que matrizes quadradas de ordem $n \geq 2$ com posto 1 são diagonalizáveis se e somente se a soma da diagonal principal (traço) é não nula. A primeira abordagem usa o critério do polinômio minimal e a segunda usa o fato de que hiperplanos de $\mathbb{R}^n$ são $n-1$ dimensionais.

Palavras-chave: Matrizes de posto 1, Diagonalização, Critério do polinômio minimal.

Abstract

In this work, we present two approaches to demonstrate the known result that square matrices of order greater than or equal to 2 with rank 1 are diagonalizable if and only if the sum of the main diagonal (trace) is non-zero. The first approach uses the criterion of the minimal polynomial, and the second approach uses the fact that hyperplanes in $\mathbb{R}^n$ are $(n-1)$-dimensional.

Keywords: Rank-1 matrices, Diagonalization, Minimal polynomial criterion.

1 Introdução

Muitos alunos de graduação, principalmente por questão de tempo, não têm a oportunidade de demonstrar um mesmo resultado sobre uma certa classe de matrizes utilizando

This open access article is licensed under
a Creative Commons Attribution-NonCommercial 4.0.

diferentes ferramentas da Álgebra Linear. Neste trabalho apresentamos duas diferentes demonstrações para o resultado já conhecido de que a classe das matrizes quadradas de posto 1 de tamanho $n \geq 2$ são diagonalizáveis se e somente se possuem traço (soma da diagonal principal) diferente de zero; veja por exemplo [7, Exercício 7.2.13, página 522].

A primeira demonstração que apresentamos deste resultado acima consiste em calcular explicitamente o polinômio característico direto pela definição – via determinante $\det(A - \lambda I)$; e depois usar o critério de diagonalização via polinômio mínimal. Até onde nós pudemos pesquisar, não encontramos esta abordagem de demonstração em outro lugar. Já a segunda abordagem usa o apelo geométrico de que hiperplanos de $\mathbb{R}^n$ são $n-1$ dimensionais. Ou, equivalentemente, no contexto da Álgebra Linear, que o complementar ortogonal de um espaço vetorial gerado por um vetor não nulo é $n-1$ dimensional (fato este que pode ser provado usando o Teorema do Núcleo e da Imagem da Álgebra Linear).

É interessante salientar que matrizes de posto 1, bem como matrizes diagonalizáveis, nos ajudam a entender como age uma transformação linear, pois podemos decompor essa ação em casos mais simples. Nesta direção temos o seguinte para matrizes de posto 1: A decomposição em valores singulares [8] nos diz que qualquer matriz $m \times n$ – quadrada ou não – pode ser decomposta como soma de matrizes de posto 1:

$$A \;=\; U\Sigma V^T \;=\; \sigma_1 u_1 v_1^T \;+\; \sigma_2\, u_2\, v_2^T \;+\; \ldots \;+\; \sigma_r\, u_r v_r^T$$

Sendo $r \leq \min\{m, n\}$ o posto da matriz A. Esta decomposição nos ajuda a diminuir a complexidade de um problema, como na compressão de dados, no processamento de sinais e na análise de dados estatísticos (Análise de Componentes Principais), à medida que formos truncando a soma acima.

Já o fato de saber que uma matriz é diagonalizável, nos ajuda a entender a ação desta matriz como transformação linear, pois neste caso existe um sistema de eixos coordenados onde a ação da transformação linear sobre um vetor consiste em multiplicar cada componente do vetor pelas entradas da diagonal; em símbolos:

$$[A]_\beta^\beta [x]_\beta \;=\; \Lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{bmatrix}$$

Desta forma uma alteração na componente x_i de x não altera as demais componentes do resultado Λx nesse sistema de eixos.

Critérios de diagonalização são importantes por exemplo no estudo de sistemas lineares de n equações diferenciais $x'(t) = Ax(t)$ cujas funções incógnitas são as n

componentes do vetor $x(t)$. Isto porque no caso de A ser diagonalizável, este sistema admitirá apenas soluções puramente exponenciais no tempo – não aparecendo fatores polinomiais no tempo. Desta forma, sendo A diagonalizável e com autovalores não positivos ($\lambda_i \leq 0$) a estabilidade das soluções $x(t)$ quando $t \to \infty$ está garantida. Entretanto, caso também tenhamos autovalores não positivos ($\lambda_i \leq 0$) mas A não é diagonalizável, o sistema admite soluções instáveis; veja por exemplo [2], [4].

O presente trabalho teve como origem as contas feitas pelo primeiro autor na obtenção da expressão do polinômio característico para matrizes quadradas de posto 1. Tal expressão é dada por

$$p(\lambda) \, = \, \lambda^{n-1} \, (\lambda \, - \, \text{soma da diagonal principal}) \; . \tag{1.1}$$

E a forma como o trabalho está organizado é a seguinte. Na próxima seção, Seção 2 chamada de 'Preliminares', é fixada a notação e apresentada a teoria mínima para enunciar e provar os resultados principais. Tais resultados estão presentes na Seção 3. Dentre esses resultados principais, está a expressão (1.1) que é o resultado base para obter os demais, e a sua prova é feita em detalhes na Seção 4. Desta forma, as seções 3 e 4 correspondem à primeira abordagem para provar o critério de diagonalização de matrizes quadradas de posto 1 para $n \geq 2$. A nossa segunda abordagem para provar este resultado é feita na Seção 5. Por questão de completude, na Seção 6 apresentamos uma prova escrita a partir das sugestões de resolução do livro [7].

2 Preliminares

Nesta seção, fixaremos a notação e recordaremos os conceitos necessários para apresentar os resultados da próxima seção. Denotaremos por $M_n(\mathbb{R})$ o conjunto das matrizes quadradas de ordem n, ou seja, que possuem n linhas e n colunas, onde as suas n^2 entradas são números reais. Para $A \in M_n(\mathbb{R})$ denotaremos as entradas desta matriz A por a_{ij} onde $1 \leq i \leq n$ e $1 \leq j \leq n$ são os índices da posição da linha e da coluna, respectivamente, do elemento a_{ij}. A matriz coluna que tem como entradas a j-ésima coluna da matriz A será denotada por A_j e a matriz linha que tem como entradas a i-ésima linha da matriz A será denotada por A_i^T, onde o símbolo ^{T}, como super-índice, indica que ocorreu a transposta da matriz em questão. Se tivermos n valores reais $x_1, x_2, \ldots, x_n$ denotaremos por $x = [x_1 \; \ldots \; x_n]^T$ a matriz coluna cujo i-ésimo elemento é x_i. O conjunto de tais matrizes colunas, chamados também de vetores, será denotado por $\mathbb{R}^n$. Denotaremos por y^T a matriz linha - vetor dual - com entradas y_j. Perceba que

o produto matricial xy^T será uma matriz $n \times n$ enquanto o produto matricial $x^T y$ será uma matriz 1×1, a qual será identificada com um número real que é exatamente o produto interno usual $\sum_{i=1}^{n} x_i y_i$ entre os vetores x e y. Note que os vetores e_j, $j = 1, \ldots, n$ formam a base canônica do espaço vetorial $\mathbb{R}^n$ onde $e_j = [0, \ldots, 0, 1, 0 \ldots, 0]^T$ com o 1 aparecendo apenas na j-ésima posição e 0 nas demais posições. Perceba que a matriz $n \times n$ com entrada 1 na linha i e coluna j e zero nas demais é dada por $e_i e_j^T$, e para a matriz identidade, a qual é denotada por I, temos $I = \sum_{i=1}^{n} e_i e_i^T$. Denota-se por 0 o vetor nulo de $\mathbb{R}^n$, ou seja $0 = [0, 0, \ldots, 0]^T$. Assim, quando x é um vetor não nulo, em símbolos $x \neq 0$, significa que pelo menos uma de suas componentes $x_1, x_2, \ldots, x_n$ é não nula.

Na seguinte definição abaixo é apresentado o objeto de estudo deste trabalho.

Definição 2.1 (matrizes da forma xy^T ou, equivalentemente, matrizes quadradas de posto zero ou um)**.** Considere dois vetores $x = [x_1\,,\, x_2\,,\, \ldots\,,\, x_n\,]^T$ e $y = [\,y_1\,,\, y_2\;\; \ldots\,,\, y_n\,]^T$ de $\mathbb{R}^n$. Denotamos por xy^T a seguinte matriz quadrada $n \times n$:

$$xy^T \overset{\text{def.}}{=} \begin{bmatrix} x_1\, y_1 & x_1\, y_2 & x_1\, y_3 & \ldots & x_1\, y_n \\ x_2\, y_1 & x_2\, y_2 & x_2\, y_3 & \ldots & x_2\, y_n \\ \vdots & \vdots & \ddots & & \vdots \\ x_n\, y_1 & x_n\, y_2 & x_n\, y_3 & \ldots & x_n\, y_n \end{bmatrix}, \tag{2.1}$$

a qual é chamada de matriz de posto zero ou um formada a partir da matriz coluna x e da matriz linha y^T.

Para ilustrar essa definição, consideremos os seguintes exemplos

Exemplo 2.2. Para $n = 2$, $x = [1\,,\, 0]^T$ e $y = [0\,,\, 1]^T$ obtém-se $xy^T = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} 0 \,,\, 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$.

Exemplo 2.3. Para $n = 2$, $x = [1\,,\, 2]^T$ e $y = [1\,,\, 1]^T$ obtém-se $xy^T = \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \,,\, 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 2 & 2 \end{smallmatrix}\right]$.

Exemplo 2.4. Para $n = 2$, $x = [5\,,\, -5]^T$ e $y = [1\,,\, 1]^T$ obtém-se $xy^T = \left[\begin{smallmatrix} 5 \\ -5 \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \,,\, 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 5 & 5 \\ -5 & -5 \end{smallmatrix}\right]$.

Exemplo 2.5. Para $n = 3$, $x = [\,4\,,\, 0\,,\, 5\,]^T$ e $y = [\,1\,,\, 2\,,\, 3\,]^T$ teremos então $xy^T = \left[\begin{smallmatrix} 4 \\ 0 \\ 5 \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \,,\, 2 \,,\, 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 4 & 8 & 12 \\ 0 & 0 & 0 \\ 5 & 10 & 15 \end{smallmatrix}\right]$.

Exemplo 2.6. Para $n = 3$, $x = [\,1\,,\, 4\,,\, -3\,]^T$ e $y = [\,1\,,\, 2\,,\, 3\,]^T$ teremos então $xy^T = \left[\begin{smallmatrix} 1 \\ 4 \\ -3 \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \,,\, 2 \,,\, 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 8 & 12 \\ -3 & -6 & -9 \end{smallmatrix}\right]$.

Antes de comentar os exemplos acima lembramo-nos da seguinte definição de Traço de uma matriz quadrada.

Definição 2.7 (traço de uma matriz quadrada)**.** Dada uma matriz $A \in M_n(\mathbb{R})$ define-se o seu traço, denotado por $\operatorname{Tr} A$, como sendo a soma de sua diagonal principal. Ou seja,

$$\operatorname{Tr} A := \sum_{i=1}^{n} a_{ii}$$

No caso particular de A ser da forma xy^T temos a seguinte observação.

Observação 2.8 (traço de matrizes quadradas da forma xy^T)**.** O traço de matrizes quadradas da forma xy^T é igual ao valor real de $y^T x$, o qual é valor do produto interno dos vetores x e y. Isto porque as entradas da diagonal principal de xy^T são $x_i y_i$; veja a diagonal principal de (2.1). Ou seja, vale

$$\operatorname{Tr}\left(xy^T\right) = y^T x = \sum_{i=1}^{n} x_i y_i$$

Conforme será mostrado na Proposição 2.10, as matrizes apresentadas nos exemplos anteriores são matrizes de posto 1, e podem ser utilizadas para ilustrar o resultado central desse trabalho que enunciamos a seguir:

Resultado Principal: *Uma matriz quadrada de tamanho $n \geq 2$ e posto 1 é diagonalizável se, e somente se, tem traço não nulo.* (2.2)

Uma prova deste resultado será apresentada no Corolário 3.6.

De acordo com o resultado acima, os Exemplos 2.2, 2.4 e 2.6 são de matrizes não diagonalizáveis, enquanto os Exemplos 2.3 e 2.5 são de matrizes diagonalizáveis.

No que se segue revisitaremos e detalharemos mais algumas definições (como a de posto e a de matriz diagonalizável) e consequências imediatas.

Para poder explicar porque matrizes da forma xy^T são chamadas de matrizes quadradas de posto zero ou um, recordemos a seguinte definição de posto de matriz.

Definição 2.9 (posto de matriz)**.** Dada uma matriz A, quadrada ou não, seu posto é definido como o número máximo de vetores-linha linearmente independentes (LI), ou equivalentemente, ao número de linhas não nulas após o escalonamento por linhas.

Lembramo-nos de que um conjunto de vetores-linha $\{l_1, \ldots, l_n\}$ é linearmente independente se nenhum de seus vetores puder ser escrito como combinação linear dos demais, e caso o conjunto for formado por uma única linha $\{l_1\}$ este será um conjunto LI por definição quando $l_1 \neq [0, 0, \ldots, 0]$. Desta forma, a matriz nula terá posto zero, pois não há subconjuntos de vetores-linha que sejam LI. Além disso, somente a matriz nula tem posto zero.

A seguinte proposição justifica porque matrizes quadradas da forma xy^T têm posto zero ou um; veja por exemplo [3, §51, Theorem 1].

Proposição 2.10. Seja M uma matriz $n \times n$. Então M é da forma (2.1) se e somente se M tem posto zero ou posto 1. Mais precisamente, vale:

(a) Se posto de M é um, então M é da forma xy^T com $x, y \in \mathbb{R}^n$ ambos não nulos;

(b) Se M é da forma xy^T, com $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ não nulos, então M tem posto um;

(c) Se posto de M é zero, então M é da forma xy^T com $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$, sendo pelo menos um deles nulo;

(d) Se M é da forma xy^T, com $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$, sendo pelo menos um deles nulo, então M tem posto zero.

Demonstração. (a) Se o posto de M é um, então é porque existe uma linha não nula, digamos que o índice desta linha seja $i_* \in \{1, 2, \ldots, n\}$, e as demais linhas serão múltiplas desta linha l_{i_*}. Isto significa que existem escalares reais c_i, com $i \in \{1, \ldots n\}$ tais que vale $l_i = c_i l_{i_*}$, para $i \in \{1, \ldots, n\}$ sendo $c_{i_*} = 1$. Desta forma basta tomar a matriz coluna x como sendo $x = [c_1, c_2, \ldots, 1, \ldots, c_n]^T$ e a matriz linha y^T como sendo $y^T = l_{i_*}$ para termos $xy^T = M$. Note que nessa construção, tivemos tanto $x \neq 0$ como $y \neq 0$.

(b) Considere a matriz M sendo da forma xy^T com x e y não nulos. Neste caso, seja i_* o índice da componente $x_{i_*} \neq 0$ da matriz coluna x. Note que a linha i_* de M, a qual é $[x_{i_*}y_1\,,\, x_{i_*}y_2\,,\, \ldots\,,\, x_{i_*}y_n]$, não é nula pois $x_{i_*} \neq 0$ e para algum $j \in \{1, \ldots, n\}$ temos $y_j \neq 0$, uma vez que $y \neq 0$. Desta forma o número máximo de linhas LI da matriz M é 1 pois, como acabamos de ver acima, a i_*-ésima linha de M é não nula; e, já as demais linhas de M são múltiplas desta. Para ver este último fato perceba que se $l_k = [x_k y_1, x_k y_2, \ldots, x_k y_n]$ é a k-ésima linha de M, então, como $x_{i_*} \neq 0$, podemos reescrever $l_k = x_k/x_{i_*}[x_{i_*}y_1\,,\, x_{i_*}y_2\,,\, \ldots\,,\, x_{i_*}y_n]$.

(c) Se o posto de M é zero vimos pela Definição 2.9 que $M = 0$. Assim o resultado segue pois a matriz quadrada nula 0 pode ser obtida fazendo $0\,y^T$ ou $x\,0^T$.

(d) Neste caso xy^T será a matriz nula e pela Definição 2.9 a matriz nula tem posto zero. □

A próxima definição trata de um conceito central na Álgebra Linear.

Definição 2.11 (matriz diagonalizável)**.** Dizemos que uma matriz $A \in M_n(\mathbb{R})$ é real diagonalizável se existir uma matriz $V \in M_n(\mathbb{R})$ inversível tal que

$$A = V \Lambda V^{-1}, \tag{2.3}$$

onde Λ é matriz diagonal com entradas reais. Ou seja, $\lambda_{ii} \in \mathbb{R}$ e $\lambda_{ij} = 0$ quando $i \neq j$.

Isto significa que no sistema de eixos coordenados dado pelas colunas da matriz V a ação da matriz A é de uma matriz diagonal; nesse caso a ação de A corresponde a apenas ampliar cada coordenada nesse sistema por um fator que é a correspondente entrada da matriz diagonal Λ. Nesse caso a matriz V é a matriz mudança de base e dizemos que V *diagonaliza* a matriz A. Matrizes diagonais são automaticamente diagonalizáveis pois pode-se escolher $V = I$ como matriz inversível. Critérios para decidir se uma matriz não diagonal é diagonalizável e, no caso afirmativo, métodos para conseguir uma matriz V inversível satisfazendo (2.3) estão relacionados com os conceitos de autovetores e autovalores, os quais desempenham um papel importante na Álgebra Linear.

Definição 2.12 (autovalor/autovetor)**.** Seja $A \in M_n(\mathbb{R})$. Dizemos que $\lambda \in \mathbb{R}$ é um autovalor de A quando existir um vetor $v \in \mathbb{R}^n \setminus \{0\}$ satisfazendo $A\,v = \lambda\,v$. Neste caso, dizemos que v é um autovetor da matriz A associado àquele autovalor λ.

Perceba que, para $v = 0$, a igualdade $A\,v = \lambda\,v$ é sempre satisfeita. A condição $v \neq 0$ é uma exigência mínima para quando formarmos a matriz V, cujas colunas são os autovetores de A, ela seja inversível. A seguinte observação decorre da definição de diagonalização Definição 2.11 e do fato de uma matriz V ser inversível se, e somente se, seus vetores colunas forem linearmente independentes.

Observação 2.13 (n autovetores LI & diagonalização)**.** Uma matriz A de tamanho $n \times n$ com autovalores reais é diagonalizável se existir um conjunto LI $\{v_1, v_2, \ldots, v_n\} \subset \mathbb{R}^n$ onde cada v_j, $j = 1, 2, \ldots, n$ é autovetor de A. Ou seja, satisfaz $Av_j = \lambda_j v_j$. Além disso, nesse caso a matriz V tendo como colunas os autovetores v_j é inversível e diagonaliza a matriz A, ou seja vale $A = V\Lambda V^{-1}$ onde

$$V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} \qquad \text{e} \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & & 0 \\ 0 & 0 & \ddots & & \\ 0 & 0 & & & \\ 0 & 0 & & 0 & \lambda_n \end{bmatrix}.$$

Reciprocamente, se A é diagonalizável, ou seja, A admite uma fatoração da forma $V\Lambda V^{-1}$, com Λ diagonal e V inversível, então as entradas da diagonal principal de Λ serão formadas por autovalores de A e os vetores-coluna da matriz V serão os correspondentes autovetores e estes serão linearmente independentes.

Como um sistema linear homogêneo $(\lambda I - A)v = 0$ admite uma solução não nula (não trivial) se, e somente se, a matriz dos coeficientes tem determinante zero, segue que λ é autovalor de A se, e somente se, λ for solução da equação polinomial $p_A(\lambda) = 0$ onde $p_A(\lambda)$ é obtido pelo desenvolvimento do determinante com variável λ dado por $\det(\lambda I - A) = 0$.

Definição 2.14 (polinômio característico)**.** Dada uma matriz quadrada A, de ordem n, então o seguinte polinômio com variável λ definido através de determinante por $p_A(\lambda) \overset{\text{def.}}{=} \det(\lambda I - A)$ é chamado de polinômio característico da matriz A.

Desta forma a obtenção dos autovalores de uma matriz A, também conhecidos como valores característicos ou pontos do espectro de A, corresponde à obtenção das raízes do seu polinômio característico. Como $\det(\lambda I - A)$ é um polinômio de grau n e cujo coeficiente de λ^n é 1 (ou seja, é um polinômio mônico) temos que ele admite fatoração da forma

$$p_A(\lambda) \;=\; (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_s)^{n_s} \qquad \text{com} \qquad n_1 + \cdots + n_s = n, \tag{2.4}$$

onde $\lambda_1, \dots, \lambda_s$ são suas raízes (sem repetição) com respectivas multiplicidades n_1, $\dots$, n_s. A multiplicidade (algébrica) de um autovalor λ_i, $i = 1, \dots, s$ de uma matriz quadrada A é definida como a multiplicidade da raiz λ_i do polinômio característico desta matriz.

Observação 2.15. [autovalores reais] Estaremos considerando neste trabalho apenas matrizes quadradas $n \times n$ com entradas reais possuindo polinômio característico com <u>raízes reais</u>, ou, equivalentemente, possuindo autovalores reais. Como veremos, esta suposição abrange as matrizes quadradas com entradas reais de posto 1 ou zero, as quais são o foco deste trabalho.

Definição 2.16 (espectro de uma matriz)**.** Dada uma matriz $A \in M_n(\mathbb{R})$ cujas raízes do seu polinômio característico serão assumidas reais. O conjunto de tais raízes, autovalores, é denotado por $\sigma(A)$. No nosso caso $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$. E se, além dos autovalores, quisermos indicar a multiplicidade (algébrica) de cada um deles, escrevemos $\sigma_{\text{multiset}}(A) = \{\underbrace{\lambda_1, \dots, \lambda_1}_{n_1 \text{ vezes}}, \underbrace{\lambda_2, \dots, \lambda_2}_{n_2 \text{ vezes}}, \cdots \underbrace{\lambda_s, \dots, \lambda_s}_{n_2 \text{ vezes}}\}$; chamado de multiconjunto por permitir repetições de um mesmo elemento. Outra notação utilizada é $\sigma_{\text{multiset}}(A) = \{\lambda_1^{(n_1)}, \lambda_2^{(n_2)}, \dots, \lambda_s^{(n_s)}\}$.

Antes de enunciar o critério de diagonalização via polinômio minimal, precisaremos das seguintes quatro definições abaixo:

Definição 2.17 (polinômio com fatores lineares distintos de uma matriz quadrada)**.** Seja $p_A(\lambda)$ o polinômio característico da matriz $A \in M_n(\mathbb{R})$, o qual é dado por $p_A(\lambda) = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_s)^{n_s}$ com $n_1 + \cdots + n_s = n$. Então define-se

$$p_A^1(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_s) \tag{2.5}$$

como o polinômio com todos os fatores lineares distintos de A.

Usamos o sobrescrito "1"em (2.5) na notação do polinômio com todos os fatores lineares distintos pois fatores lineares têm grau 1. Note também que este polinômio é obtido a partir do polinômio característico por apenas substituir cada expoente n_i, $i = 1, \dots, s$ não nulo que aparece na sua fatoração (2.4) pelo expoente 1. Perceba que para o polinômio característico temos:

$$p_A(\lambda) \;=\; \prod_{\mu \,\in\, \sigma_{\text{multiset}}(A)} (\lambda - \mu)\,,$$

enquanto que para o polinômio com todos os fatores lineares distintos temos

$$p_A^1(\lambda) \;=\; \prod_{\mu \,\in\, \sigma(A)} (\lambda - \mu)\,.$$

Nestas duas expressões acima Π é o símbolo de produtório.

Definição 2.18 (função polinomial de matriz quadrada)**.** Sejam $A \in M_n(\mathbb{R})$ e $f(\lambda)$ uma função polinomial da forma $f(\lambda) = c_m\lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0$, com coeficientes c_m, c_{m-1}, $\dots$, c_1, c_0 reais. Então define-se a matriz $f(A) \in M_n(\mathbb{R})$ como

$$f(A) \overset{\text{def.}}{=} c_m A^m + c_{m-1}A^{m-1} + \cdots + c_1 A + c_0 I. \tag{2.6}$$

Nesta expressão a potência de matriz A^k denota o produto matricial $A \cdot A \cdot \ldots \cdot A$, onde a matriz A aparece k vezes.

Observação 2.19 (função polinomial de matriz quadrada quando o polinômio está na forma fatorada)**.** Caso $f(\lambda)$ for um polinômio de grau m, ou seja, $f(\lambda) = \sum_{k=0}^{m} c_k\lambda^k$ admitindo a fatoração $\alpha\Pi_{j=1}^{m}(\lambda - \beta_j)$, com $\alpha, \beta_1, \dots \beta_m \in \mathbb{R}$, então para uma matriz $A \in M_n(\mathbb{R})$, pode-se demonstrar que a matriz $f(A)$ definida em (2.6) também é dada por $f(A) = \alpha\Pi_{j=1}^{m}(A - \beta_j I)$.

Definição 2.20. (polinômio mínimal de uma matriz quadrada) Seja A uma matriz quadrada com autovalores reais. Define-se o polinômio mínimal de A como aquele, dentre os polinômios $f(\lambda)$ mônicos ($c_m = 1$) satisfazendo a igualdade matricial $f(A) = 0$, que tenha o menor grau m. Tal polinômio $f(\lambda)$ é denotado por $p_A^{\min}(\lambda)$.

O seguinte resultado importante de Álgebra Linear estabelece um critério algébrico de diagonalização, no sentido que permite decidir se uma matriz é diagonalizável sem precisar obter seus autovetores e provar que eles são linearmente independentes. Mais precisamente, temos o seguinte critério do polinômio mínimal para diagonalização de uma matriz quadrada A; veja por exemplo [5].

Teorema 2.21 (critério de diagonalização via polinômio minimal)**.** *Seja A uma matriz quadrada com autovalores reais $\lambda_1, \ldots, \lambda_s$. Então A é diagonalizável se, e somente se,*

$$p_A^{\min}(\lambda) = p_A^1(\lambda),$$

ou, equivalentemente, se, e somente se, a seguinte equação matricial envolvendo a matriz A e seus autovalores

$$(A - \lambda_1 I) \cdot (A - \lambda_2 I) \cdot \ldots \cdot (A - \lambda_s I) = 0.$$

for verificada. Nesta expressão acima λ_1, λ_2, ..., λ_s são todos os autovalores distintos da matriz A, já o zero 0 do lado direito é a matriz zero, e, o símbolo $\cdot$ representa o produto de matrizes.

Com este critério de diagonalização do polinômio minimal para uma matriz quadrada A, finalizamos a parte dos resultados preliminares de Álgebra Linear para a primeira demonstração do Resultado Principal (2.2); veja Teorema 3.1 e Corolário 3.5 abaixo. Já para a segunda demonstração deste Resultado Principal (2.2), a qual é construtiva – veja Teorema 5.1 abaixo – usaremos a Observação 2.13 acima.

3 Resultados Principais

Esta seção dos resultados principais está dividida em duas subseções. Na primeira subseção são obtidas as expressões para o polinômio característico e para o polinômio minimal de matrizes quadradas da forma xy^T. Como consequência disto e do critério de diagonalização via polinômio minimal Teorema 2.21, é estabelecido, na segunda subseção, o critério de diagonalização das matrizes da forma xy^T.

3.1 Expressões para os polinômios característico e minimal de matrizes quadradas da forma xy^T (posto 1 ou 0)

Como primeiro resultado desta subseção, temos o seguinte:

Teorema 3.1. *Seja $n \geq 1$, então o polinômio característico de matrizes da forma xy^T com $x, y \in \mathbb{R}^n$ é dado por*

$$p_{xy^T}(\lambda) = \lambda^{n-1}\left(\lambda - \sum_{i=1}^{n} x_i y_i\right). \tag{3.1}$$

Observação 3.2. Note que as raízes do polinômio em (3.1) são reais, pois essas raízes estão no conjunto $\{0, \sum_{i=1}^{n} x_i y_i\}$ cujos elementos são números reais. Desta forma, a Observação 2.15 é cumprida, podendo-se, assim, aplicar o Teorema 2.21 para obter como resultados consequentes o Corolário 3.5 e o Corolário 3.6 abaixo.

Demonstração. A prova do Teorema 3.1 é dividida em duas partes:

• **Parte 1.** Caso $n = 1$. Neste caso a matriz xy^T é dada por x_1y_1 e seu polinômio característico é por definição $p_{xy^T}(\lambda) = \det(\lambda I - x_1y_1)$ e por estarmos no caso $n = 1$ o polinômio característico fica igual a $\lambda - x_1y_1$, mas este é exatamente o polinômio em (3.1) quando $n = 1$.

• **Parte 2.** Caso $n \geq 2$. Esta parte será dividida em dois casos:

Caso 1: $n \geq 2$ e os vetores $x = [x_1, x_2, \ldots, x_n]^T$ e/ou $y = [y_1, y_2, \ldots, y_n]^T$ possuem/possui todas as suas entradas nulas. Neste caso, xy^T será a matriz nula 0. Assim, de um lado por definição, o polinômio característico da matriz nula é $p_0(\lambda) = \det(\lambda I - 0) = \lambda^n$. Por outro lado, temos que $\sum_{i=1}^{n} x_i y_i$ é zero pois x e/ou y é/são nulos e, desta forma, a expressão em (3.1) também nos fornece $p_{xy^T}(\lambda) = \lambda^n$.

Caso 2: $n \geq 2$ com $x = [x_1, x_2, \ldots, x_n]^T$ e $y = [y_1, y_2, \ldots, y_n]^T$ não nulos. Este é o caso mais interessante e é provado em detalhes à parte na Seção 4 usando o desenvolvimento de Laplace em menores, reduzindo desta forma a ordem dos determinantes até chegar a um determinante de ordem 2×2.

□

Como consequência deste teorema temos o seguinte resultado referente ao polinômio com todos os fatores lineares distintos $p_A^1(\lambda)$ para matrizes quadradas da forma xy^T.

Corolário 3.3. Os possíveis polinômios com todos os fatores lineares distintos $p_A^1(\lambda)$ para matrizes quadradas da forma $A = xy^T$ são os seguintes conforme $n = 1$ ou $n \geq 2$ ou conforme $(\sum_{i=1}^{n} x_i y_i) \neq 0$ ou $(\sum_{i=1}^{n} x_i y_i) = 0$.

$$
\begin{array}{llllll}
(a) & p^1_{xy^T}(\lambda) & = & \lambda - x_1y_1 & \text{se} \quad n=1 & \text{e} \quad (\sum_{i=1}^n x_iy_i) \neq 0; \\
(b) & p^1_{xy^T}(\lambda) & = & \lambda & \text{se} \quad n=1 & \text{e} \quad (\sum_{i=1}^n x_iy_i) = 0; \\
(c) & p^1_{xy^T}(\lambda) & = & (\lambda - 0)(\lambda - \sum_{i=1}^n x_iy_i) & \text{se} \quad n \geq 2 & \text{e} \quad (\sum_{i=1}^n x_iy_i) \neq 0; \\
(d) & p^1_{xy^T}(\lambda) & = & \lambda & \text{se} \quad n \geq 2 & \text{e} \quad \sum_{i=1}^n x_iy_i = 0.
\end{array}
$$

3.2 Critério de diagonalização para matrizes de forma xy^T

Precisamos do seguinte resultado antes de aplicar o critério do polinômio minimal (Teorema 2.21).

Lema 3.4. Para $n \geq 1$ matrizes da forma xy^T satisfazem a seguinte equação matricial:

$$(xy^T)^2 \;=\; (\textstyle\sum_{i=1}^n x_iy_i)\, xy^T\,,$$

onde a expressão no lado esquerdo da igualdade significa que a matriz xy^T está multiplicada por ela mesma, sendo o produto desta multiplicação o matricial.

Demonstração. Trata-se de uma prova bem direta pois basta fazer a conta $(xy^T)(xy^T) = x(y^Tx)y^T = (y^Tx)xy^T$. Aqui usamos a associatividade do produto de matrizes e o fato de que y^Tx é um número (escalar) e por isso comuta com matrizes. Por último, observa-se que y^Tx é exatamente $\sum_{i=1}^n x_iy_i$.

□

Como consequência direta do Corolário 3.3, do Lema 3.4 e do critério de diagonalização via polinômio minimal (Teorema 2.21) temos o seguinte resultado a cerca da diagonalização de matrizes da forma xy^T; matrizes quadradas de posto zero ou um.

Corolário 3.5 (Critério de diagonalização para matrizes da forma xy^T)**.** Para matrizes da forma xy^T vale o seguinte:

(a) Se $n = 1$ então xy^T é diagonalizável;

(b) Se $n \geq 1$ e $xy^T = 0$ (matriz nula) então xy^T é diagonalizável.

(c) Se $n \geq 2$ e $(\sum_{i=1}^n x_iy_i) \neq 0$ então xy^T é diagonalizável.

(d) Se $n \geq 2$, $(\sum_{i=1}^{n} x_i y_i) = 0$ mas $xy^T \neq 0$ (matriz não nula) então xy^T não é diagonalizável.

Demonstração. A prova do item (a) segue do fato de que toda matriz de tamanho 1 é diagonalizável. Perceba também que se aplicarmos o critério do polinômio minimal, também obteríamos que $xy^T = x_1 y_1$ é diagonalizável pois pelo Corolário 3.3 o polinômio com todos fatores lineares distintos é $p^1_{xy^T}(\lambda) = \lambda - x_1 y_1$ e sendo, no caso $n = 1$, a matriz $xy^T = x_1 y_1$ teríamos que $p^1_{xy^T}(xy^T) = x_1 y_1 - x_1 y_1 = 0$ e assim, $p^1_{xy^T}(\lambda)$ é o polinômio minimal de xy^T. Note que isto vale também caso $x_1 y_1 = 0$.

Prova do item (b): Segue do fato das matrizes nulas serem diagonalizáveis. Isto pode ser provado pelo critério do polinômio minimal pois o polinômio característico da matriz nula 0 é $p_0(\lambda) = \det(\lambda I - 0) = \lambda^n$ assim $p^1_0(\lambda) = \lambda$ e como $p^1_0(0) = 0$ segue a sua diagonalização.

Prova do item (c): Como $(\sum_{i=1}^{n} x_i y_i) \neq 0$ temos pelo Corolário 3.3 item (c) que $p^1_{xy^T}(\lambda) = (\lambda - 0)(\lambda - \sum_{i=1}^{n} x_i y_i)$. Assim, $p^1_{xy^T}\left(xy^T\right) = \left(xy^T - 0\right)\left(xy^T - \sum_{i=1}^{n} x_i y_i\right) = (xy^T)^2 - (\sum_{i=1}^{n} x_i y_i)\, xy^T$ e tal matriz é a matriz nula devido ao Lema 3.4. Segue, portanto, pelo critério do polinômio minimal (Teorema 2.21) que nesse caso a matriz xy^T é diagonalizável.

Prova do item (d): Como $(\sum_{i=1}^{n} x_i y_i) = 0$ e $n \geq 2$, temos pelo Corolário 3.3 item (d) que $p^1_{xy^T}(\lambda) = \lambda$. Desta forma segue que $p^1_{xy^T}\left(xy^T\right) = xy^T$ e como por hipótese $xy^T \neq 0$ (não é a matriz nula) temos que $p^1_{xy^T}\left(xy^T\right) \neq 0$. Isto significa que pelo critério do polinômio minimal (Teorema 2.21) a matriz xy^T é não diagonalizável. □

Lembrando-se de que somente a matriz nula tem posto zero, podemos resumir os itens (c) e (d) do Corolário 3.5 acima como o principal resultado deste artigo:

Corolário 3.6. Uma matriz A de posto 1 com $n \geq 2$ é diagonalizável se e somente se $\mathrm{Tr}\{A\} \neq 0$.

4 Prova da expressão do polinômio característico para matrizes da forma xy^T com $n \geq 2$ e tanto $x \neq 0$ como $y \neq 0$; Caso 2 da prova do Teorema 3.1

Dados $x = [x_1\,,\, x_2\,,\, \ldots\,,\, x_n]^T$ e $y = [y_1\,,\, y_2\,,\, \ldots\,,\, y_n]^T$ vetores quaisquer de $\mathbb{R}^n \setminus \{0\}$ e $n \geq 2$. Então, para qualquer $\lambda \in \mathbb{R}$, a matriz $\lambda I - xy^T$ tem suas entradas dadas

explicitamente por

$$\lambda I \,-\, xy^T \,=\, \begin{bmatrix} \lambda - x_1y_1 & -x_1y_2 & \cdots & -x_1y_n \\ -x_2y_1 & \lambda - x_2y_2 & \cdots & -x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ -x_ny_1 & -x_ny_2 & \cdots & \lambda - x_ny_n \end{bmatrix}. \tag{4.1}$$

Estamos interessados em encontrar o polinômio característico de xy^T, portanto precisamos obter o determinante de (4.1). Para o cálculo deste determinante, iremos usar que determinantes são invariantes ao adicionar a uma linha um múltiplo de uma outra linha. Este processo de adição será feito $n-1$ vezes. Mais precisamente, faremos o seguinte procedimento:

Para cada $i \in \{1, 2, \ldots, n-1\}$ somaremos à i-ésima linha de (4.1) o resultado da n-ésima linha multiplicada por $-\frac{x_i}{x_n}$.

Aqui temos assumido sem perda de generalidade que $x_n \neq 0$, uma vez que estamos assumindo que nem x e nem y têm todas as entradas nulas. Para justificar o porquê de não perder a generalidade em assumir $x_n \neq 0$ perceba que caso tivéssemos $x_n = 0$ existiria alguma $x_{i_*} \neq 0$ e então faríamos a transposição entre x_n e x_{i_*} e também faríamos a transposição (=permutação que troca apenas dois elementos entre si e deixa os demais como estavam) entre y_n e y_{i_*}. A ação destas duas transposições não alteraria o polinômio característico. Para ver que não ocorreria alteração basta lembrar que para qualquer transposição $P : \mathbb{R}^n \to \mathbb{R}^n$ vale $P = P^{-1} = P^T$. Assim, $\det(\lambda I - Px(Py)^T) = \det(\lambda P\, I\, P^T - P\, xy^T\, P^T) = \det P \det(\lambda I - xy^T) \det P^T = \det(I - xy^T)$ pois $\det P = \det P^T = (-1)$.

Após executar o procedimento acima, estabeleceremos a seguinte igualdade para o determinante da matriz (4.1) acima:

$$\begin{vmatrix} \lambda - x_1y_1 & -x_1y_2 & \cdots & -x_1y_{n-1} & -x_1y_n \\ -x_2y_1 & \lambda - x_2y_2 & \cdots & -x_2y_{n-1} & -x_2y_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -x_{n-1}y_1 & -x_{n-1}y_2 & \cdots & \lambda - x_{n-1}y_{n-1} & -x_{n-1}y_n \\ -x_ny_1 & -x_ny_2 & \cdots & -x_ny_{n-1} & \lambda - x_ny_n \end{vmatrix} =$$

$$\begin{vmatrix} \lambda & 0 & \cdots & 0 & -\frac{x_1}{x_n}\lambda \\ 0 & \lambda & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \lambda & -\frac{x_{n-1}}{x_n}\lambda \\ -x_ny_1 & -x_ny_2 & \cdots & -x_ny_{n-1} & \lambda - x_ny_n \end{vmatrix}. \tag{4.2}$$

Aplicando o desenvolvimento de Laplace em determinantes menores, tomando como base a primeira coluna do último determinante em (4.2), obtemos que (4.2) é igual a:

$$\lambda\cdot\begin{vmatrix} \lambda & 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ 0 & \lambda & \cdots & 0 & -\frac{x_3}{x_n}\lambda \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & \lambda & -\frac{x_{n-1}}{x_n}\lambda \\ -x_ny_2 & -x_ny_3 & \cdots & -x_ny_{n-1} & \lambda - x_ny_n \end{vmatrix} + (-1)^{n+1}(-x_ny_1)\begin{vmatrix} 0 & 0 & \cdots & 0 & -\frac{x_1}{x_n}\lambda \\ \lambda & 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ 0 & \lambda & & 0 & -\frac{x_3}{x_n}\lambda \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & -\frac{x_{n-1}}{x_n}\lambda \end{vmatrix} \tag{4.3}$$

Agora, iremos aplicar novamente o desenvolvimento de Laplace no segundo determinante de (4.3) onde usaremos como base a primeira linha. Mais especificamente, teremos:

$$\begin{vmatrix} 0 & 0 & 0 & \cdots & 0 & -\frac{x_1}{x_n}\lambda \\ \lambda & 0 & 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ 0 & \lambda & 0 & & & -\frac{x_3}{x_n}\lambda \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & \ldots & 0 & \lambda & 0 & -\frac{x_{n-2}}{x_n}\lambda \\ 0 & \cdots & \cdots & 0 & \lambda & -\frac{x_{n-1}}{x_n}\lambda \end{vmatrix} = (-1)^{(1+(n-1))}\left(-\frac{x_1}{x_n}\right)\lambda\, D\,, \tag{4.4}$$

onde $D = \lambda^{n-2}$ é o determinante da submatriz quadrada diagonal (bloco vermelho acima) de ordem $n-2$. Portanto, temos o seguinte para segunda parcela de (4.3):

$$\text{Segunda Parcela de (4.3)} = -\lambda^{n-1}x_1y_1\,. \tag{4.5}$$

Analisaremos agora o primeiro determinante de (4.3). Aplicando novamente o desenvolvimento de Laplace baseado na primeira coluna:

$$\lambda\begin{vmatrix} \lambda & 0 & 0 & \cdots & -\frac{x_2}{x_n}\lambda \\ 0 & \lambda & 0 & \cdots & -\frac{x_3}{x_n}\lambda \\ 0 & 0 & \lambda & & -\frac{x_4}{x_n}\lambda \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_ny_2 & -x_ny_3 & -x_ny_4 & \cdots & \lambda - x_ny_n \end{vmatrix} = \lambda^2\begin{vmatrix} \lambda & 0 & \cdots & -\frac{x_3}{x_n}\lambda \\ 0 & \lambda & \cdots & -\frac{x_4}{x_n}\lambda \\ \vdots & \vdots & \ddots & \vdots \\ -x_ny_3 & -x_ny_4 & \cdots & \lambda - x_ny_n \end{vmatrix}$$

$$+\lambda(-1)^{n-1+1}(-x_ny_2)\begin{vmatrix} 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ \lambda & \cdots & 0 & -\frac{x_3}{x_n}\lambda \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \lambda & -\frac{x_{n-1}}{x_n}\lambda \end{vmatrix} \tag{4.6}$$

Observe que podemos proceder com o determinante em (4.6) como havíamos procedido em (4.3) – usando a primeira linha–, mas agora as matrizes consideradas possuem uma linha e uma coluna a menos. Assim, trocando n por $n-1$ e x_1 por x_2 em (4.4) temos

$$\begin{vmatrix} 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ \lambda & \cdots & 0 & -\frac{x_3}{x_n}\lambda \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \lambda & -\frac{x_{n-1}}{x_n}\lambda \end{vmatrix} = (-1)^{1+(n-2)}\left(-\frac{x_2}{x_n}\right)\lambda\,\lambda^{n-3} = (-1)^n\,\frac{x_2}{x_n}\,\lambda^{n-2} \tag{4.7}$$

Logo a segunda parcela em (4.6) fica

$$\text{Segunda parcela de (4.6)} \;=\; \lambda\, x_n y_2\,(-1)^{n+1}\,(-1)^n\,\frac{x_2}{x_n}\,\lambda^{n-2} \;=\; -\lambda^{n-1}\,x_2\,y_2 \tag{4.8}$$

Pelas contas que fizemos até aqui percebemos que cada vez que aplicamos o desenvolvimento de Laplace no segundo determinante em (4.2), reduzimos a um determinante de ordem uma unidade menor que a anterior adicionado de uma parcela da forma $-\lambda^{n-1}x_iy_i$. Portanto, se os procedimentos descritos acima forem aplicados $n-2$ vezes obteremos no final um determinante 2×2. Mais especificamente, teremos o seguinte para o determinante $\det(\lambda I - xy^T)$:

$$\det(\lambda I - xy^T) = \begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & -\frac{x_1}{x_n}\lambda \\ 0 & \lambda & 0 & \cdots & 0 & -\frac{x_2}{x_n}\lambda \\ 0 & 0 & \lambda & & 0 & -\frac{x_3}{x_n}\lambda \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & & \lambda & -\frac{x_{n-1}}{x_n}\lambda \\ -x_ny_1 & -x_ny_2 & -x_ny_3 & \cdots & -x_ny_{n-1} & \lambda - x_ny_n \end{vmatrix}$$

$$= \lambda^{n-2}\begin{vmatrix} \lambda & -\frac{x_{n-1}}{x_n}\lambda \\ -x_ny_{n-1} & \lambda - x_ny_n \end{vmatrix} - \lambda^{n-1}x_{n-2}y_{n-2} - \lambda^{n-1}x_3y_3 - \lambda^{n-1}x_2y_2 - \lambda^{n-1}x_1y_1$$

$$= \lambda^{n-2}\left(\lambda^2 - \lambda x_ny_n - \frac{x_{n-1}x_ny_{n-1}}{x_n}\lambda\right) - \lambda^{n-1}x_{n-2}y_{n-2} - \cdots - \lambda^{n-1}x_1y_1$$

$$= \lambda^n - \lambda^{n-1}x_ny_n - \lambda^{n-1}x_{n-1}y_{n-1} - \lambda^{n-1}x_{n-2}y_{n-2} - \cdots - \lambda^{n-1}x_1y_1$$

$$= \lambda^{n-1}\left(\lambda - \sum_{i=1}^{n} x_iy_i\right).$$

Acima, no primeiro determinante de tamanho $n \times n$ utilizamos um degradê de cores para enfatizar a relação entre os blocos encaixantes e as parcelas que vão surgindo à medida que vamos aplicando os desenvolvimentos de Laplace descritos acima, até parar no caso do determinante 2×2; por causa disto foi assumido $n \geq 2$. Desta forma, terminamos a prova da expressão do polinômio característico para matrizes quadradas de posto 1 com $n \geq 2$.

5 Prova do critério de diagonalização de matrizes da forma xy^T usando que hiperplanos de $\mathbb{R}^n$ são $n-1$ dimensionais

Nessa seção a diagonalização de matrizes $n \times n$ da forma xy^T – nos casos onde ela ocorre – será provada diretamente indicando como encontrar n autovetores linearmente independentes, os quais dependem de x e y. Mais precisamente, temos o seguinte resultado, o qual é coerente com o Corolário 3.5.

Teorema 5.1 (critério de diagonalização de matrizes da forma xy^T)**.** *Sejam x, y vetores (colunas) com n componentes reais. Então vale o seguinte para uma matriz quadrada $n \times n$ da forma xy^T*

(a) Se $n = 1$ então xy^T é diagonalizável.

(b) Se x e/ou y é/são o vetor nulo, $n \geq 1$, segue que xy^T é a matriz nula (posto zero) e portanto diagonalizável;

(c) Sejam $n \geq 2$ e x e y vetores não nulos. Se a soma $\sum_{i=1}^{n} x_i y_i = y^T x = \mathrm{Tr}\{xy^T\} \neq 0$, então a matriz xy^T é diagonalizável;

(d) Sejam $n \geq 2$ e x e y vetores não nulos. Se a soma $\sum_{i=1}^{n} x_i y_i = y^T x = \mathrm{Tr}\{xy^T\} = 0$, então a matriz xy^T não é diagonalizável.

Para a prova do item (c) do teorema acima iremos precisar do seguinte resultado na forma de lema; veja Lema 5.2 abaixo. Esse resultado possui uma interpretação geométrica e, no contexto da Álgebra Linear, afirma que a dimensão do espaço ortogonal ao subespaço gerado por um vetor não nulo y é $n-1$. Faremos uma demonstração construtiva, exibindo explicitamente uma base.

Lema 5.2 (hiperplanos comportam $n-1$ vetores LI)**.** Seja $y \in \mathbb{R}^n$ com $y \neq 0$ então o seguinte subespaço vetorial de $\mathbb{R}^n$

$$\mathcal{H}_y := \{v \in \mathbb{R}^n : y^T v = 0\}, \tag{5.1}$$

o qual é chamado de hiperplano passando pela origem, comporta $n-1$ vetores LI.

Exemplo 5.3. Para $n = 3$ e $y = [5\,,\,2\,,\,3]^T$ basta tomarmos $v_1 = \left[1\,,\,0\,,\,\frac{-5}{3}\right]^T$ e $v_2 = \left[0\,,\,1\,,\,\frac{-2}{3}\right]^T$. Desta forma $y^T v_1 = 5 \times 1 + 2 \times 0 + 3 \times (-5)/3 = 5 + 0 - 5 = 0$; ou seja, $v_1 \in \mathcal{H}_y$. Também temos $y^T v_2 = 15 \times 0 + 2 \times 1 + 3 \times (-2)/3 = 0 + 2 - 2 = 0$; ou seja, $v_2 \in \mathcal{H}_y$. Além disso, o conjunto $\{v_1, v_2\} \subset \mathbb{R}^3$ é LI pois a matriz de ordem 3×2 dada $V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -5/3 & -2/3 \end{bmatrix}$ possui uma submatriz de tamanho 2×2 que é a identidade; cujo determinante é não nulo. Observamos que em Geometria Analítica este é o procedimento padrão para encontrar dois vetores LI pertencentes ao plano $\pi : 5x + 2y + 3z = 0$.

Demonstração do Lema 5.2. A demonstração do Lema 5.2 acima é praticamente a formalização deste Exemplo 5.3 para o caso geral de $n \geq 2$.

Observação 5.4 (hiperplanos são $n-1$ dimensionais)**.** Para provar o Teorema 5.1 só precisaremos usar que os hiperplanos $\mathcal{H}_y$ com $y \neq 0$ comportam $n-1$ vetores LI. Não será preciso usar que estes subespaços vetoriais têm dimensão $n-1$. Mas observamos que, de fato, $\mathcal{H}_y$, com $y \neq 0$, são subespaços de $\mathbb{R}^n$ com dimensão $n-1$. Para provar isso, lembramos que a dimensão de um subespaço pode ser definida como o número máximo de vetores linearmente independentes que ele pode conter. E de fato, o hiperplano $\mathcal{H}_y$ tem dimensão $n-1$, pois não se pode aumentar de $n-1$ vetores LI, quantidade já garantida pelo Lema 5.2 acima, para n vetores LI uma vez que se fosse o caso $\mathcal{H}_y$ seria todo o $\mathbb{R}^n$ e assim $y \in \mathcal{H}_y$ e desta forma teríamos $y^T y = 0$ donde seguiria $y = 0$ o que é uma contradição já que $y \neq 0$ foi assumido.

Observação 5.5 (usando o Teorema do Núcleo e da Imagem para provar que $\mathcal{H}_y$ é $(n-1)$-dimensional, se $y \neq 0$)**.** Basta observar que a função $h_y : \mathbb{R}^n \to \mathbb{R}$ definida por $x \mapsto h_y(x) := y^T x$ é linear, e sendo $y \neq 0$, temos $\dim\{h_y(x) : x \in \mathbb{R}^n\} = 1$ (dimensão da Imagem) e como o núcleo de h_y é $\mathcal{N}(h_y) = \{x \in \mathbb{R}^n \,:\, y^T x = 0\}$ o qual é exatamente igual ao hiperplano $\mathcal{H}_y$. Do Teorema do Núcleo e da Imagem segue que $\dim \mathbb{R}^n = \dim \mathcal{N}(h_y) + \dim \mathcal{I}m(h_y)$, ou seja, $n = \dim \mathcal{H}_y + 1$. Logo $\dim \mathcal{H}_y = n - 1$.

Após essas duas observações acima, segue a demonstração do Lema 5.2. Como por hipótese $y = [y_1\,,\,\ldots\,,\,y_n] \neq 0$, podemos, sem perda de generalidade, assumir que a n-ésima componente y_n é não nula; caso contrário, faz-se uma permutação de dois elementos $P : \mathbb{R}^n \to \mathbb{R}^n$ tal que $Pe_n = e_{i_*}$, $Pe_{i_*} = e_n$ e $Pe_j = e_j$ para $j \notin \{i_*, n\}$, onde y_{i_*} é uma componente não nula. Desta forma a n-ésima componente do vetor Py será não nula. (Note que caso y_n já for não nula então $P = I$). A seguir, iremos construir $n-1$ vetores w_1, w_2, $\ldots$, w_{n-1} de $\mathbb{R}^n$ dentro do seguinte subespaço de $\mathbb{R}^n$:

$$\mathcal{H}_{Py} := \{w \in \mathbb{R}^n : (Py)^T w = 0\}$$

Para cada $j \in \{1, 2, \ldots, n-1\}$ defina w_{ij}, a qual é a i-ésima componente do j-ésimo vetor w_j, como

$$w_{ij} := \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \\ \frac{-(Py)_j}{(Py)_n} & \text{se } i = n \end{cases} \tag{5.2}$$

Desta forma, para cada $j \in \{1, 2, \ldots, n-1\}$ o vetor w_j pertence ao subspaço $\mathcal{H}_{Py}$ pois para cada $j \in \{1, 2, \ldots, n-1\}$ temos que

$$(Py)^T w_j = \sum_{i=1}^{n} (Py)_i w_{ij} = (Py)_j \times 1 + (Py)_n \times \frac{-(Py)_j}{(Py)_n} = 0\,. \tag{5.3}$$

Além do mais, o conjunto dos vetores $\mathcal{W} := \{w_1, w_2, \ldots, w_n\}$ é conjunto LI de vetores de $\mathbb{R}^n$ uma vez que a matriz W com entradas w_{ij} definidas acima é uma matriz de tamanho $n \times (n-1)$ possuindo a submatriz identidade de ordem $n-1$, cujo determinante é não nulo.

Lembrando-se de que permutações são inversíveis e vale $P^{-1} = P^T$ e definindo $v_j := P^{-1} w_j$ segue que o conjunto $\mathcal{V} := \{v_1, v_2, \ldots v_{n-1}\}$ está em $\mathcal{H}_y$ pois $y^T v_j = y^T P^{-1} w_j = y^T P^T w_j = (Py)^T w_j = 0$ por (5.3). Também vale que o conjunto $\mathcal{V}$ é LI pois sendo P inversível, P^{-1} leva um conjunto LI em outro conjunto LI. □

Demonstração do Teorema 5.1. **Prova de (a) e (b).** Como já vimos anteriormente – veja Corolário 3.5 – todo número real é diagonalizável e a matriz nula é diagonalizável, justificando esses dois primeiros itens.

Prova de (c). Neste caso estamos considerando que xy^T tem posto 1 (i.e. tanto x como y não são vetores nulos), $n \geq 2$, e a soma da sua diagonal $\sum_{i=1}^{n} x_i y_i = x^T y = y^T x$ não é zero.

Como $y \neq 0$ e $n \geq 2$ segue do Lema 5.2 que existe um conjunto $\{v_1, v_2, \ldots, v_{n-1}\}$ de vetores de $\mathbb{R}^n$ que é LI e satisfaz $y^T v_j = 0$ para cada $j \in \{1, 2, \ldots, n-1\}$. Assim, também temos que $(xy^T)v_j = x(y^T v_j) = x0 = 0 = 0v_j$ para cada $j \in \{1, 2, \ldots, n-1\}$. Desta forma, $v_1, v_2, \ldots v_{n-1}$ são autovetores LI associados ao mesmo autovalor $\lambda_1 = 0$.

E como também $x \neq 0$ podemos escolher como o n-ésimo autovetor v_n o próprio vetor x, ou seja, $v_n = x \neq 0$ pois temos que x é um autovetor da matriz xy^T com correspondente autovalor o escalar $y^T x \neq 0$. Para ver isto, basta notar que a seguinte igualdade sempre ocorre:

$$(xy^T)x \;=\; x(y^T x) = (y^T x)x$$

uma vez que o produto matricial é associativo, justificando a primeira igualdade, e comutativo quando se trata de um vetor x e um escalar $y^T x$, justificando a segunda igualdade. Desta forma, $\lambda_2 = y^T x = \mathrm{Tr}\{xy^T\} \neq 0$ é o segundo autovalor distinto.

Só falta mostrar que o conjunto $\{v_1, v_2, v_3, \ldots, v_{n-1}, v_n = x\}$, agora com n autovetores, é LI, pois de acordo com a Observação 2.13 isso garante que a matriz xy^T é diagonalizável.

A prova de que o conjunto $\{v_1, v_2, v_3, \ldots, v_{n-1}, v_n = x\}$ é LI segue a mesma linha da prova que autovetores associados a autovalores distintos são LI. De fato, se c_1, c_2, $\ldots$, c_{n-1}, c_n são coeficientes reais tais que

$$c_1\, v_1 \;+\; c_2\, v_2 \;+\; \ldots \;+\; c_{n-1}\, v_{n-1} \;+\; c_n\, x \;=\; 0 \tag{5.4}$$

então, após fazer o produto com y^T em ambos os lados da última igualdade, segue que

$$c_1\, y^T v_1 \;+\; c_2\, y^T v_2 \;+\; \ldots \;+\; c_{n-1}\, y^T v_{n-1} \;+\; c_n\, y^T x \;=\; 0\,.$$

Agora lembrando que $y^T v_j = 0$ para $j = 1, 2, \ldots, n-1$, pois $v_j \in \mathcal{H}_y$, obtemos $c_n\, y^T x \;=\; 0$ e como $y^T x = \mathrm{Tr}\{xy^T\} \neq 0$ (por hipótese) segue que $c_n = 0$. Assim a equação (5.4) fica

$$c_1\, v_1 \;+\; c_2\, v_2 \ldots \;+\; c_{n-1}\, v_{n-1} \;=\; 0 \tag{5.5}$$

e como o conjunto $\{v_1, v_2, \ldots, v_{n-1}\}$ é LI, segue que $c_1 = c_2 = \cdots = c_{n-1} = 0$. Isto juntamente com $c_n = 0$, visto acima, prova que o conjunto $\{v_1, v_2, \ldots, v_{n-1}, v_n = x\}$ é LI.

Prova de (d). Aqui estamos assumindo que $n \geq 2$, nem x e nem y são vetores nulos (ou seja a matriz xy^T não é a matriz nula), mas a soma $\sum_{i=1}^{n} x_i y_i = x^T y = y^T x = \mathrm{Tr}\{xy^T\} = 0$. Vamos mostrar que a matriz xy^T neste caso é não diagonalizável. A prova deste item é feita em três etapas. <u>Primeira etapa:</u> Embora $xy^T \neq 0$ temos que $(xy^T)(xy^T) = 0$, assim xy^T é dita ser uma matriz nilpotente de índice 2. Para ver isto, basta fazer a conta $(xy^T)(xy^T) = x(y^T x)y^T = x0y^T = 0$. <u>Segunda etapa:</u> Sendo $(xy^T)(xy^T) = 0$ temos que $\sigma(xy^T) = \{0\}$; ou seja, o conjunto de autovalores é formado apenas pelo valor zero. Para ver isto, suponha que $\lambda \in \sigma(xy^T)$ seja um autovalor de xy^T. Então existe $v \in \mathbb{R}^n \setminus \{0\}$ autovetor tal que $(xy^T)v = \lambda v$. Decorre da aplicação da matriz xy^T a ambos os lados da última igualdade o seguinte:

$$xy^T\, xy^T v = xy^T(\lambda v) \;=\; \lambda(xy^T)v = \lambda\lambda v = \lambda^2 v. \tag{5.6}$$

Como vimos na primeira etapa $(xy^T)\,(xy^T) = 0$ logo o lado esquerdo de (5.6) é o vetor nulo. Ou seja, temos $0 = \lambda^2 v$ e sendo $v \neq 0$ (autovetor) temos que $\lambda^2 = 0$, o que implica $\lambda = 0$. <u>Terceira etapa:</u> Nesta etapa é feito o argumento de contradição. Se xy^T fosse diagonalizável decorreria que $xy^T = 0$, o que não é o caso. Para ver isto, suponha que xy^T fosse diagonalizável. Neste caso, valeria $xy^T = V\Lambda V^{-1}$ sendo Λ matriz diagonal cujas entradas da diagonal necessariamente seriam os autovalores $\lambda \in \sigma(xy^T)$ e V matriz inversível cujas colunas necessariamente seriam correspondentes autovetores. Sendo a diagonal principal de Λ formada por autovalores segue da Segunda Etapa que $\Lambda = 0$ (matriz nula). Logo se xy^T fosse diagonalizável deveríamos ter $xy^T = V\Lambda V^{-1} = V0V^{-1} = 0$, ou seja $xy^T = 0$, contradizendo a hipótese deste item (d). Portanto, xy^T é não diagonalizável. $\square$

6 Prova da Diagonalização de matrizes de posto 1 segundo [7]

No livro [7] a questão da diagonalização de matrizes de posto 1 é colocada na forma de exercício no seu capítulo 7, *Autovalores e Autovetores* na sua segunda seção, 7.2 *Diagonalização por Transformações de Similaridade*, e no final do livro é feita uma sugestão de solução deste exercício. Por questão de completude, reproduziremos aqui esta solução. Tal solução considera apenas os casos não triviais de $n \geq 2$, ou seja, a matriz xy^T é não nula (posto 1). A prova é feita em duas partes. A primeira parte consiste em provar que $\lambda^{n-1}(\lambda - y^T x)$ é a expressão do polinômio característico para a matriz xy^T. Ao contrário da prova apresentada na Seção 4, a qual usa as propriedades de invariância do determinante por escalonamento e o desenvolvimento de Laplace reduzindo iterativamente até chegar ao caso de um determinante 2×2, a prova em [7] também parte da definição $p_{xy^T}(\lambda) = \det(\lambda I - xy^T)$ porém este último determinante não é escalonado e reduzido, mas sim é usada a seguinte identidade

$$\det(I + xy^T) = 1 + y^T x, \tag{6.1}$$

a qual é generalizada para

$$\det(A + zy^T) = \det(A)\,(1 + y^T A^{-1} z). \tag{6.2}$$

no caso de A ser inversível. A prova de (6.1) usa a seguinte fatoração matricial em blocos de ordem $(n+1) \times (n+1)$

$$\begin{bmatrix} I & 0 \\ y^T & 1 \end{bmatrix} \begin{bmatrix} I + xy^T & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ -y^T & 1 \end{bmatrix} = \begin{bmatrix} I & x \\ 0 & 1 + y^T x \end{bmatrix}$$

e propriedade multiplicativa de determinantes

$$\det(AB) = \det A \ \det B. \tag{6.3}$$

Já a prova de (6.2) é feita reescrevendo $\det(A + zy^T) = \det A \det(I + A^{-1} zy^T)$ e aplicando (6.3) e o caso anterior (6.1). As fórmulas (6.1) e (6.2) são apresentadas e provadas desta forma em [7] no seu capítulo 6, *Determinantes*, na sua segunda seção 6.2 *Propriedades Adicionais de Determinantes*, onde estão algumas propriedades de determinantes aplicadas a matrizes em blocos. Finalmente, a prova da expressão do polinômio característico da matriz não nula xy^T e $n \geq 2$ segue de (6.2) com a escolha $A = \lambda I$, a qual é inversível caso $\lambda \neq 0$, e $z = -x$, obtendo-se $\det(\lambda I)(1 - y^T \frac{1}{\lambda} I x) = \lambda^n \left(\frac{\lambda}{\lambda} - \frac{1}{\lambda} y^T x\right) = \lambda^{n-1}(\lambda - y^T x)$. Caso $\lambda = 0$, temos que $p_{xy^T}(0) = \det(-xy^T)$ e este determinante é zero devido a xy^T ser singular quando $n \geq 2$. Assim, em qualquer caso ($\lambda \neq 0$ ou $\lambda = 0$) a expressão $\lambda^{n-1}(\lambda - y^T x)$ para o polinômio característico fica provada desta forma.

A segunda parte da prova em [7] para o critério de diagonalização de xy^T, não usa o critério do polinômio minimal (Teorema 2.21) como se fez na Seção 3.2, mas sim, o critério

da igualdade entre as multiplicidades algébricas e geométricas para cada autovalor distinto. Este critério é bem conhecido e por isto o enunciamos na forma de lema; veja, por exemplo, [7, página 512] para uma prova.

Lema 6.1 (critério da igualdade das multiplicidades)**.** Uma matriz A é diagonalizável se, e somente se,

$$\text{alg mult}_A(\lambda_i) \;=\; \text{geo mult}_A(\lambda_i) \qquad \text{para cada } \lambda_i,\, i = 1, 2, \ldots, s \;. \tag{6.4}$$

Onde, como vimos na seção de resultados preliminares, alg $\text{mult}_A(\lambda_i) = n_i$ multiplicidade da raiz λ_i do polinômio característico e geo $\text{mult}_A(\lambda_i)$ é definida como a quantidade máxima de soluções v's LI para a equação homogênea $Av = \lambda_i v$, ou seja geo $\text{mult}_A(\lambda_i) = \dim \mathcal{N}(\lambda_i I - A)$ onde $\mathcal{N}(B)$ é o núcleo de B definido como o subespaço vetorial $\mathcal{N}(B) := \{v \in \mathbb{R}^n \,:\, B\,v \;=\; 0\,\}$. Em [7] o critério (6.4) do lema acima é apresentado na sua seção 7.2 *Diagonalização e Transformações de Similaridade*, onde logo após, o exercício 7.2.13 – de provar que xy^T (com $n \geq 2$) é diagonalizável se e somente se $y^T x \neq 0$ – é proposto.

A multiplicidade geométrica de $\lambda_1 = 0$ pode ser obtida usando o Teorema do Núcleo e da Imagem:

$$n = \dim \mathcal{N}(xy^T) + \dim \mathcal{I}m(xy^T)$$

e sendo $\dim \mathcal{I}m(xy^T) =$ posto de $A = 1$ segue que

$$\text{geo mult}_{xy^T}(\lambda_1 = 0) \;=\; \dim \mathcal{N}(xy^T) \;=\; n \,-\, 1. \tag{6.5}$$

É interessante perceber que a multiplicidade geométrica do autovalor $\lambda_1 = 0$ é sempre igual $n - 1$ independente de $y^T x$ ser não nulo ou nulo. Por outro lado, o valor da multiplicidade algébrica do autovalor $\lambda_1 = 0$ irá depender se $y^T x$ é não nulo ou nulo. Mais precisamente, temos que

$$\text{alg mult}_{xy^T}(\lambda_1 = 0) \;=\; \begin{cases} n \,-\, 1 & \text{se } y^T x \neq 0 \;, \\ \\ n & \text{se } y^T x = 0. \end{cases} \tag{6.6}$$

Estes valores para a multiplicidade algébrica do autovalor $\lambda_1 = 0$ decorrem imediatamente da expressão do polinômio característico que é $p_{xy^T}(\lambda) = \lambda^{n-1}(\lambda - y^T x)$; pois quando $y^T x \neq 0$ temos que o autovalor $\lambda_1 = 0$ tem multiplicidade algébrica $n-1$ e quando $y^T x = 0$ o autovalor $\lambda_1 = 0$ tem multiplicidade n.

Finaliza-se a prova do critério de diagonalização para matrizes de posto 1 com $n \geq 2$ combinando (6.4) com (6.5) e (6.6). Mais detalhadamente, temos os seguintes dois casos:

Caso $y^T x \neq 0$**.** Da primeira parte vimos que $p_{xy^T}(\lambda) = \lambda^{n-1}(\lambda - y^T x)$ se $y^T x \neq 0$. Assim, neste caso, $\lambda_1 = 0$ tem multiplicidade algébrica $n \,-\, 1$ e $\lambda_2 = y^T x \neq 0$ tem multiplicidade algébrica 1. Para λ_2 o critério (6.4) é facilmente verificado pois para qualquer matriz A sempre

se tem geo $\text{mult}_A(\lambda_i) \leq$ alg $\text{mult}_A(\lambda_i)$ e geo $\text{mult}_A(\lambda_i) \geq 1$. Assim, geo $\text{mult}_{xy^T}(\lambda_2) = 1 =$ alg $\text{mult}_{xy^T}(\lambda_2)$. Na verdade, este argumento serve para mostrar que para todo autovalor simples (com multiplicidade algébrica 1) tem-se que sua multiplicidade algébrica é igual a sua multiplicidade geométrica. Já para o autovalor $\lambda_1 = 0$ temos que sua multiplicidade algébrica é $n - 1$. E de acordo com (6.5) o critério (6.4) também foi verificado. Assim, a matriz xy^T é diagonalizável no caso de $y^T x = \text{Tr}\{xy^T\} \neq 0$.

Caso $y^T x = 0$. Nesse caso da primeira parte temos que $p_{xy^T}(\lambda) = \lambda^n$. Assim só há um autovalor $\lambda_1 = 0$ e sua multiplicidade algébrica é n. Note que neste caso o critério (6.4) não se verifica. De fato, por (6.6) e (6.5) temos alg $\text{mult}_A(\lambda_1 = 0) = n > n - 1 =$ geo $\text{mult}_A(\lambda_i)$. Portanto, a matriz xy^T não é diagonalizável no caso de $y^T x = \text{Tr}\{xy^T\} = 0$.

Desta forma finaliza-se a prova detalhada do critério de diagonalização sugerida em [7].

7 Conclusão

Neste trabalho apresentamos o resultado já conhecido – mas não muito difundido – que caracteriza quando uma matriz quadrada A de posto 1 é diagonalizável: para $n \geq 2$ a matriz A é diagonalizável se e somente $\text{Tr}\{A\} \neq 0$. Tal resultado é encontrado praticamente na forma de exercício (em livros ou em fóruns de discussão). Sendo A quadrada e de posto 1 tem-se que A é da forma $A = xy^T$ com $x \in \mathbb{R}^n - \{0\}$ e $y \in \mathbb{R}^n - \{0\}$. De modo geral estudamos matrizes quadradas da forma $A = xy^T$ as quais podem possuir apenas posto 0 ou 1. Já para os casos triviais (a) e (b) abaixo a matriz A é obviamente diagonalizável.

(a) A ordem é fixa em $n = 1$ mas A pode ser $A = a_{11} = 0$ (posto 0) ou $A = a_{11} \neq 0$ (posto 1); ou

(b) A matriz $A = 0$ é fixa (posto 0, aqui $x = 0$ e/ou $y = 0$) e ordem n qualquer.

Temos apresentado aqui duas demonstrações detalhadas para o critério de diagonalização destas matrizes: a primeira delas utiliza o critério do polinômio minimal onde o polinômio minimal de xy^T é obtido facilmente do polinômio característico de xy^T e este é calculado explicitamente pela definição de polinômio característico $p_{xy^T}(\lambda) := \det(\lambda I - xy^T)$, e para o cálculo deste determinante utiliza-se operações nas linhas e o desenvolvimento de Laplace em determinantes menores. Até onde nós sabemos não encontramos em outro lugar esta primeira prova feita pelo primeiro autor onde o determinante é calculado explicitamente. Já a segunda demonstração detalhada utiliza-se do fato de que para $y \neq 0$ hiperplano $\mathcal{H}_y := \{z \in \mathbb{R}^n : y^T z = 0\}$ é $n - 1$ dimensional permitindo construir, no caso de $\text{Tr}\{xy^T\} \neq 0$, uma base de n autovetores para xy^T, o que caracteriza a diagonaliziblilidade de xy^T.

8 Agradecimentos

Os autores são muito gratos pelas sugestões da(o) referee por melhorar o texto como um todo, bem como por simplificar algumas demonstrações.

Referências

[1] Bueno, Hamilton Prado. *Álgebra Linear: Um Segundo Curso.* Rio de Janeiro: SBM, 2006.

[2] Doering, Claus Ivo; Lopes, Artur Oscar. *Equações Diferenciais Ordinárias.* Coleção Matemática Universitária, 4a. Edição, IMPA, 2010

[3] Halmos, Paul Robert. *Finite-Dimensional Vector Spaces*, Springer, 2nd edition, 1958.

[4] Hirsch, M; Smale, S. *Differential Equqtions, Dynamical Systems, and Linear Algebra.* New York: Academic Press, 1974.

[5] Hoffman, Kenneth; Kunze, Ray. *Álgebra Linear.* São Paulo: Prntice-Hall, 1971.

[6] Lima, Elon Lages. *Álgebra Linear.* 8. ed. Rio de Janeiro: IMPA, 2014.

[7] Meyer, Carl D. *Matrix Analysis and Applied Linear Algebra*, SIAM, 1st edition, 2000.

[8] Strang, Gilbert. *Linear Algebra and Its Applications*, 4th Editon, Cengage Learning, 2006.

Recebido em 08 de Novembro de 2024
Revisado em 11 de Fevereiro de 2025
Aceito em 15 de Setembro 2025