Capítulo 8

Autopsia dos Operadores

Vamos iniciar o estudo dos operadores lineares e entender a sua estrutura. Isso será executado em duas partes: Diagonalização e Formas Canônicas.

8.1 O que faremos?

Como o observamos podemos associar infinitas matrizes a um operador linear $T : V \to V$. A principal característica que todas estas matrizes tem em comum é que são todas semelhantes. A relação de semelhança é uma relação de equivalência sobre M_n. Como sabemos todas as relações de equivalência particionam o conjunto em classes de equivaência. Para cada classe de equivalência é possível eleger um representante (canônico). Com respeito a relação de semelhança entre matrizes e o corpo $K = \mathbb{C}$, esta escolha se chama a **forma canônica de Jordan** daquela classe de equivalência. No caso em que $F = \mathbb{R}$, vamos chamar o representante de **forma canônica real de Jordan**, e para o caso, em que o corpo, é mais geral, vamos chamar o representante para cada uma das classes de equivalência de **forma racional do operador**.

Se você não captou a explicação acima não fique muito preocupado, ela esta ai para depois que você já tenha digerido o assunto deste capítulo. Vamos voltar para uma abordagem mais terrestre.

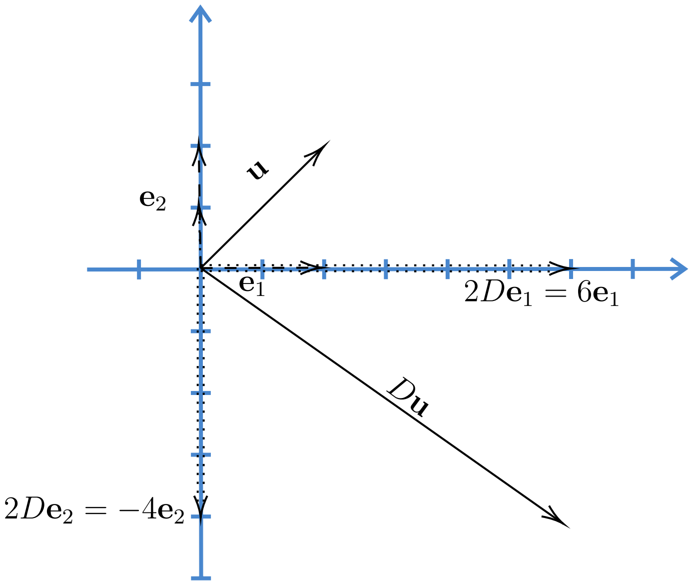
Vamos considerar duas matrizes:

$$A = \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix} \text{ e } D = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}.$$

A matriz D "de diagonal" tem o seguinte efeito sobre $\mathbf{e}_1$, ela é simplesmente uma "dilatação" pelo fator 3. O efeito sobre $\mathbf{e}_2$ é de dilatação pelo fator 2, seguida de uma reflexão com respeito à origem (ou seja, uma "reversão" no sentido do vetor, mas preservando sua direção). Por outro lado, o operador D *não* preserva a direção do vetor $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, isto é, o vetor $D\mathbf{u}$ *não* é colinear a $\mathbf{u}$.

Veja a figura abaixo que descrevem as ações de D sobre $\mathbf{e}_1$, $\mathbf{e}_2$ e $\mathbf{u}$.

Veja que não temos muito o que dizer sobre como a matriz A opera sobre um vetor. Para tentar levar a cabo o estudo de operadores lineares, procedemos



como faríamos no estudo de funções $f : A \to A$ que agem de um conjunto A nele mesmo. Nesta situação devemos procurar os pontos que são mantidos fixos, isto é, $a \in A$ tal que $f(a) = a$. Agora, lembramos que um operador linear leva uma reta que passa pela origem em uma outra reta que passa pela origem. Então, adaptando esta característica, precisamos determinar aqueles vetores $\mathbf{u}$ tal que

$$T\mathbf{u} = \lambda\mathbf{u}, \text{ para algum } \lambda \in \mathbb{R}.$$

Resolvendo este tipo de problema, encontramos aquelas direções que são mantidas fixas pelo nosso operador linear. Com isso estamos prontos para entrar na próxima seção.

8.2 Autovalores

Definição 8.1
Seja $V = \mathbb{R}^n$ um espaço vetorial sobre $\mathbb{R}$ e T um operador linear sobre V. Um autovalor de T é um escalar $\lambda \in \mathbb{R}$ se existe um vetor **não nulo** $\mathbf{u} \in V$ tal que $T\mathbf{u} = \lambda\mathbf{u}$. O vetor $\mathbf{u}$ é chamado de autovetor de T associado ao autovalor λ.

Os autovalores de um operador T também são chamados de: valor característico, raízes características, valor próprio, ou ainda de valor espectral.

Se T é um operador linear qualquer e λ é um escalar, o conjunto dos vetores $\mathbf{u}$ que satisfazem $T\mathbf{u} = \lambda\mathbf{u}$ é um subespaço vetorial de V, uma vez que é o espaço nulo da transformação $T - \lambda I$. Podemos expressar este fato por dizer que λ é um autovalor de T se este subespaço ($\mathcal{N}(T - \lambda I)$) **não** for o subespaço trivial formado apenas pelo vetor nulo. Isto implica também que $T - \lambda I$ não é injetora. Portanto, podemos resumir dizendo que λ é um autovalor T se, e somente se, $T - \lambda I$ não é invertível.

Sejam T um operador linear sobre V e α uma base de V. Considere a matriz $A = [T]^{\alpha}_{\alpha}$ de T com respeito a base α. Se queremos determinar os autovalores λ de

T, precisamos determinar também $\mathbf{u} \in V$ tal que $T\mathbf{u} = \lambda\mathbf{u}$, mas não conseguimos resolver os dois problemas simultaneamente, mas

$$T\mathbf{u} = \lambda\mathbf{u} \Leftrightarrow A\mathbf{u} = \lambda\mathbf{u} \Leftrightarrow (A - \lambda I)\,\mathbf{u} = 0.$$

Logo podemos nos concentrar em determinar os λ tais que $\det(A - \lambda I) = 0$, ao calcularmos este determinante nos deparamos com um polinômio de grau $n = \dim V$. As raízes deste polinômio, vão tornar o operador $A - \lambda I$ não invertíveis e, portanto, deve existir $\mathbf{u} \neq 0$ tal que $\mathbf{u} \in \mathcal{N}(A - \lambda I)$. Desta forma, temos um algoritmo para buscarmos os autovalores e depois os autovetores associados.

Definição 8.2
Seja $V = \mathbb{R}^n$ um espaço vetorial sobre $\mathbb{R}$ e T um operador linear sobre V. Se $\lambda \in \mathbf{K}$ é um autovalor de T o subespaço vetorial $\mathcal{N}(T - \lambda I)$ é chamado de **autoespaço** associado ao autovalor λ.

Exemplo 8.3
Considere o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 8x + 5y \\ -10x - 7y \end{bmatrix}.$$

Fixado a base canônica $\alpha = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \subset \mathbb{R}^2$ obtemos

$$A = [T]^\alpha_\alpha = \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix},$$

calculado o polinômio característico obtemos

$$\begin{aligned} \Delta_A(\lambda) = \det(A - \lambda I) &= \det \begin{bmatrix} 8-\lambda & 5 \\ -10 & -7-\lambda \end{bmatrix} \\ &= (8-\lambda)(-7-\lambda) + 50 = \lambda^2 - \lambda - 6 = 0. \end{aligned}$$

Calculando as raízes deste polinômio obtemos $\lambda = -2$ e $\lambda = 3$. Para determinarmos os autovetores associados a estes autovalores precisamos determinar $\mathcal{N}(A + I)$ e $\mathcal{N}(A - 3I)$.

Vamos determinar o geradores do subespaço $\mathcal{N}(A + 2I)$

$$\left(\begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \to \begin{bmatrix} 10 & 5 \\ -10 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}.$$

Portanto, obtemos a equação $x + \frac{1}{2}y = 0$ e daí, $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}y \\ y \end{bmatrix} = \frac{y}{2} \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Portanto, $\mathcal{N}(A + I) = \text{Span}\left\{ \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$.

Para determinar uma base para o subespaço $\mathcal{N}(A - 3I)$ devemos fazer

$$\left(\begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \to \begin{bmatrix} 5 & 5 \\ -10 & -10 \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Portanto, obtemos a equação $x + y = 0$ e daí, $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ y \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Portanto, $\mathcal{N}(A - 3I) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$.

Considerando a base $\beta = \left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\} \subset \mathbb{R}^2$ temos $P = [I]_\alpha^\beta = \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix}$. Nessa base a matriz associada ao operador T se torna

$$D = [T]_\beta^\beta = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = [I]_\beta^\alpha \, [T]_\alpha^\alpha \, [I]_\alpha^\beta = P^{-1}AP.$$

Observe que as matrizes que obtivemos neste exemplo são exatamente as matrizes A e D que iniciamos esse capítulo. Isto nos leva a uma consideração trivial, mas profunda. Muitas matrizes podem representar o mesmo operador linear, mas na primeira vista não podemos reconhecer.

Como para cada operador linear podemos associar infinitas matrizes, surge a pergunta natural, como se relacionam os polinômios característicos destas matrizes.

Proposição 8.4
Se A e B são matrizes $n \times n$ semelhantes, então

$$\Delta_A(x) = \Delta_B(x).$$

Demonstração: Veja o exercício R8.2. □

8.2.1 Fórmulas

Existem fórmulas para o polinômio característico, veja os casos em que as matrizes são de dimensão 2 e 17 3. Se

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \text{ Então}$$

$$\Delta_A(t) = t^2 - (a + d)t + (ad - bc) = t^2 - \text{tr}(A)t + \det A.$$

No caso que

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}. \text{ Então}$$

$$\Delta_A(t) = -t^3 + \text{tr}(A)t^2 - (\Delta_{11} + \Delta_{22} + \Delta_{33})\, t + \det A,$$

onde o Δ_{ii} são os cofatores associados a posição (i, i), com $i = 1, 2, 3$.

Exemplo 8.5
Considere o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $(x, y) \mapsto (3x - 5y, 2x - 3y)$. Fixado a base canônica $\alpha = \{(1, 0), (0, 1)\} \subset \mathbb{R}^2$ obtemos

$$A = [T]_\alpha^\alpha = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix},$$

calculado o polinômio característico obtemos $\Delta_A(t) = \det(A - tI) = t^2 + 1$. Procurando as raízes obtemos $i, -i \in \mathbb{C}$. Como operador atua em um espaço vetorial sobre os Reais, o procedimento falha.

Veja que se A é uma matriz $n \times n$ então, $\det(A - tI) = (-1)^n \det(tI - A)$.

Exemplo 8.6
Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por $(x, y, z) \mapsto (3x - y + z, 7x - 5y + z, 6x - 6y + 2z)$. Fixado a base canônica $\alpha = \left\{[1\,0\,0]^t, [0\,1\,0]^t, [0\,0\,1]^t\right\} \subset \mathbb{R}^3$ obtemos a matriz

$$B = [T]_\alpha^\alpha = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix},$$

calculado o polinômio característico obtemos $\Delta_B(t) = \det(tI - A) = t^3 - 12t + 16 = (t-2)^2(t+4)$. Calculando $\mathcal{N}(A + 4I)$ obtemos

$$\left(\begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix} + 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \to \begin{bmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Portanto, obtemos as equações: $x = 0$ e $y - z = 0$. Portanto, um vetor $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ z \\ z \end{bmatrix} = z \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ e, daí, $\mathcal{N}(A + 4I) = \operatorname{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$. Calculando $\mathcal{N}(A - 2I)$, isto é,

$$\left(\begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \to \begin{bmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Portanto, obtemos as equações: $x - y = 0$ e $z = 0$. Portanto, o vetor $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ y \\ 0 \end{bmatrix} = y \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ e daí $\mathcal{N}(A - 2I) = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$.

Não temos a quantidade suficiente de autovetores e portanto, este operador também não pode ser diagonalizado.

No exemplo 8.6 devemos observar que: para todo autovalor existe pelo menos um autovetor; o processo de diagonalização pode vir a falhar se acontecer de haver autovalores repetidos. O próximo exemplo mostra que uma condição para que um operador não seja diagonalizável é ter autovalores repetidos, mas esta condição não é suficiente.

Exemplo 8.7
Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por $(x, y, z) \mapsto (9x+6y+2z, -16x-11y-4z, 24x+18y+7z)$. Fixado a base canônica $\alpha = \left\{[1\,0\,0]^t, [0\,1\,0]^t, [0\,0\,1]^t\right\} \subset \mathbb{R}^3$ obtemos a matriz

$$B = [T]_\alpha^\alpha = \begin{bmatrix} 9 & 6 & 2 \\ -16 & -11 & -4 \\ 24 & 18 & 7 \end{bmatrix},$$

calculado o polinômio característico obtemos $\Delta_B(t) = \det(tI-A) = t^3-5t^2+7t-3 = (t-1)^2(t-3)$. Vamos nos preocupar com o subespaço $\mathcal{N}(A-I)$. Calculando

$$A - I = \begin{bmatrix} 8 & 6 & 2 \\ -16 & -12 & -4 \\ 24 & 18 & 6 \end{bmatrix}$$

Veja que as duas útimas linhas são multiplas da primeira! Portanto,

$$\mathcal{N}(A-I) = \text{Span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} \right\}.$$

Esse operador será diagonalizável!

Ao concluir o exemplo sem calcular o autovalor associado ao autovetor 3, estamos admitindo que autovalores associados a autovetores distintos serão sempre LI.

Lema 8.8
Seja $T : V \to V$ um operador linear do espaço vetorial V e sejam $\mathbf{u}$ e $\mathbf{v}$ autovetores de T associados aos autovalores não nulos λ e θ, respectivamente. Se $\lambda \neq \theta$ segue que $\mathbf{u}$ e $\mathbf{v}$ são linearmente independentes.

Demonstração: De fato, considere a equação vetorial $x\mathbf{u} + y\mathbf{v} = 0$. Aplicando T doqs dois lados da igualdade e lembrando que T é linear obtemos

$$x\lambda\mathbf{u} + y\theta\mathbf{v} = 0$$

Multiplicando a primeira equação vetorial por θ obtemos $x\theta\mathbf{u} + y\theta\mathbf{v} = 0$ e subtraindo, uma da outra, obtemos

$$x(\lambda - \theta)\mathbf{u} = 0 \Rightarrow x = 0, \text{ mas isso também implica } y = 0.$$

□

Portanto, podemos ver que existem duas situações onde o procedimento falha, a primeira é quando o polinômio nõ tem raízes no corpo em que estamos considerando o espaço vetorial, a segunda é que apesar de ter todas as raízes não temos a quantidade necessária de autovetores para construir uma base do espaço vetorial

V. A primeira dificuldade pode ser transposta se admitirmos que o espaço vetorial esta sobre um corpo algebricamente fechado, já na segunda dificuldade ainda não sabemos como tratar esta situação. Mas já sabemos que para ela acontecer é necessário que existam raízes repetidas. Um fato, que pode nos tranquilizar um pouco é que se fosse feito um estudo estatístico em torno de todos os polinômios de um determinado grau veriamos que é relativamente improváveis aparecerem polinômios com raízes repetidas. Mesmo assim estas duas dificuldades, serão analisadas e tentaremos apontar soluções a contento nas próximas seções.

8.3 Polinômio Mínimo

Definição 8.9
Seja T um operador linear sobre um espaço vetorial $V = \mathbb{R}^n$. Dizemos que T é **diagonalizável** se existe uma base de V tal que cada um de seus vetores é um autovetor de T.

O motivo deste nome deve ser evidente, pois se $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ é uma base ordenada de V formado por autovetores de T, isto é, $T(\mathbf{u}_i) = c_i\mathbf{u}_i$. Logo,

$$\begin{aligned} T\mathbf{u}_1 &= c_1\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 0\mathbf{u}_n \\ T\mathbf{u}_2 &= 0\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + 0\mathbf{u}_n \\ \vdots \quad & \qquad\qquad \vdots \\ T\mathbf{u}_n &= 0\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n. \end{aligned}$$

A matriz associada a T é

$$[T]_\alpha^\alpha = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}.$$

claramente não exigimos que os escalares c_i sejam distintos e nem diferentes de zero.

O outra fato que podemos ver é que se o autovalor c_i aparece d_i vezes na fatoração de $\Delta_A(x)$. Como o operador é diagonalizável deve acontecer que a dimensão de $\mathcal{N}(A - c_iI)$ é d_i, esta é a relação entre a quantidade de vezes que um autovalor se repete e a $\dim(\mathcal{N}(A - c_iI))$, como vimos no exemplo 8.6 pode acontecer que $\dim(\mathcal{N}(A - c_iI)) < d_i$.

Também podemos concluir que se todas as raízes do polinômio característico são distintas então o operador será diagonalizável.

Algumas vezes acontece, quando podemos fatorar o polinômio no produto de fatores lineares, os autovalores serem repetidos, então neste caso o polinômio característico torna-se

$$\Delta_A(x) = (x - c_1)^{d_1}(x - c_2)^{d_2} \cdots (x - c_k)^{d_k}.$$

É claro que se o corpo de escalares for os números complexos esta situação sempre acontece.

Existe uma outra forma de verificar se um operador é diagonalizável, que é calcular o polinômio mínimo associado a um operador linear.

Exemplo 8.10
Considere a nossa matriz
$$A = \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix}$$
e os polinômios $f(x) = 3x + 7$ e $g(x) = x^2 - x - 6$. Podemos avaliar estes polinômios nesta matriz

$$f(A) = 3A + 7I = 3\begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 31 & 15 \\ -30 & -14 \end{bmatrix}.$$

Já para calcularmos polinômio $g(A)$, precisamos calcular

$$A^2 = A.A = \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix}\begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix} = \begin{bmatrix} 14 & 5 \\ -10 & -1 \end{bmatrix}.$$

e

$$\begin{aligned} g(A) &= A^2 - A - 6I \\ &= \begin{bmatrix} 14 & 5 \\ -10 & -1 \end{bmatrix} - \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Veja que o polinômio $g(x)$ anula a matriz A. Segue que deve existir infinitos polinômio que anulam a matriz A.

Todo matriz admite um polinômio que a anule? Bom... é fácil de construir um argumento que garante este fato. Digamos as matrizes 2×2. Noesse caso considere o espaço vetorial $M(2,2)$ das matrizes 2×2. Este espaço tem dimensão 4, pois,
$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$
forma uma base para este espaço. Portanto, se A é uma matriz não nula, as matrizes I, A, A^2, A^3, A^4 com certeza formam um conjunto LD. Logo devem existir a_0, a_1, a_2, a_3 e $a_4 \in \mathbb{R}$ não todos nulos, tais que:

$$a_0 I + a_1 A + a_3 A^2 + a_4 A^3 + a_4 A^4 = 0$$

Então o polinômio não nulo $h(x) = a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0$ anula a matriz A.

É claro que podemos repetir este argumento para matrizes de qualquer ordem.

Dada uma matriz A podemos considerar todos os polinômios que anula a matriz, vamos chamar este conjunto de $J(A)$. E podemos considerar o polinômio de menor grau e **mônico** que anula a matriz A, chamamos este polinômio de **polinômio mínimo** de A, e o denotamos por $m_A(x)$.

Gostaria de chamar a atenção que no exemplo 8.10 o polinômio $g(x) = x^2 - x - 6$ não foi escolhido por acaso. Na verdade ele é o polinômio característico de A (veja exemplo 8.3).

Teorema 8.11
Seja A uma matriz $n \times n$. Se $\Delta_A(t)$ é o seu polinômio característico, então

$$\Delta_A(A) = 0.$$

Demonstração: Veja o exercício R8.3. □

Usando o teorema 8.11, podemos garantir criar um argumento para garantir que o polinômio mínimo sempre divide o polinômio característico.

Lema 8.12
Suponha que $A\mathbf{u} = c\mathbf{u}$, se f é um polinômio, então $f(A)\mathbf{u} = f(c)\mathbf{u}$.

A demonstração deste lema é simples. Com ele podemos concluir que se λ é um autovalor de uma matriz A, então $m_A(\lambda) = 0$.

Proposição 8.13
Seja A uma matriz, e $m_A(x)$ o seu polinômio mínimo. Suponha que o polinômio se fatores da forma

$$m_A(x) = (x - c_1)^{e_1}(x - c_2)^{e_2} \cdots (x - c_k)^{e_k}.$$

Então, A é diagonalizável se, e somente se, $e_1 = e_2 = \cdots = e_k = 1$.

Lema 8.14
Suponha que $T : V \to V$ é um operador linear sobre um espaço vetorial $V = \mathbb{R}^n$. Sejam c_1, $c_2, \ldots, c_k$ os autovalores distintos de T e sejam $W_i = \mathcal{N}(A - c_iI)$ os espaços característicos associados. Se $W = W_1 + W_2 + \cdots + W_k$, então

$$\dim W = \dim W_1 + \dim W_2 + \cdots + \dim W_k.$$

Demonstração: De fato, se β_i são bases de W_i, com $i = 1, \ldots, k$, então $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ é uma base de W. □

Exemplo 8.15
Considere a matriz

$$A = \begin{bmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{bmatrix}.$$

e verifique que esta matriz é diagonalizável.

Solução: Para calcular o polinômio característico veja que

$$\operatorname{tr}(A) = 5, \;\; A_{11} + A_{22} + A_{33} = 2 - 3 + 7 \text{ e } \det(A) = 3.$$

Usando isso temos

$$\Delta_A(t) = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3).$$

O polinômio mínimo tem as mesmas raízes que o polinômio característico. Portanto, o polinômio mínimo pode ser $g(t) = (t-1)(t-3)$ ou o próprio $\Delta_A(t)$.

Vamos verificar isso, isto é,

$$g(A) = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Portanto, $m_A(t) = (t-1)(t-3)$, isto quer dizer que o nosso operador é diagonalizável. □

8.4 Espaços Invariantes

Vamos começar discutindo as formas de decompor um espaço vetorial em somas de subespaços.

Definição 8.16
Sejam $W_1, \ldots, W_k$ subespaçs do espaço vetorial $V = \mathbb{R}^n$. Dizemos $W_1, \ldots, W_k$ são **independente** se

$$\mathbf{w}_1 + \cdots + \mathbf{w}_k = 0, \text{ com } \mathbf{w}_i \text{ em } W_i$$

implicar que cada $\mathbf{w}_i$ é o vetor nulo.

Digamos que fossem apenas dois subespaços, W_1 e W_2 e, que eles fossem independentes, neste caso é fácil de verificar que $W_1 \cap W_2 = \{\mathbf{0}\}$, pois se $\mathbf{u} \in W_1 \cap W_2$, então fazendo $w_1 = \mathbf{u} \in W_1$ e $\mathbf{w}_2 = -\mathbf{u} \in W_2$, teríamos que

$$\mathbf{w}_1 + \mathbf{w}_2 = 0.$$

Portanto, podemos concluir que W_1 e W_2 são independentes se e somente se $W_1 \cap W_2 = \{\mathbf{0}\}$.

Outro detalhe importante a observar é que

$$W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2 : \mathbf{w}_1 \in W_1 \text{ e } \mathbf{w}_2 \in W_2\}$$

é um subespaço vetorial de V.

Exemplo 8.17
Considere $W_1 = \left\{ \left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right] \in \mathbb{R}^3 : z - y = 0 \right\}$ e $W_2 = \operatorname{Span}\left\{ \left[\begin{smallmatrix} 1 \\ 0 \\ 2 \end{smallmatrix}\right] \right\}$. Mostre que estes subespaços são independentes.

Solução: Observe que para que um vetor se o vetor $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ pertence a W_1, então $x = y$ e z pode ser qualquer. Daí, podemos escrever

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Seja $\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in W_1$ e $\mathbf{v} = \begin{bmatrix} t \\ 0 \\ 2t \end{bmatrix} \in W_2$ daí

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x \\ x \\ z \end{bmatrix} + \begin{bmatrix} t \\ 0 \\ 2t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Resolvendo obtemos $x = 0$, $t = 0$ e por fim $z = 0$. E isso mostra que $\mathbf{u} = 0 = \mathbf{v}$. $\square$

No caso que $W_1, W_2, \ldots, W_k$ são subespaço vetoriais independentes, então escrevemos

$$W_1 + W_2 + \cdots + W_k \text{ assim } W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Definição 8.18
Seja $V = \mathbb{R}^n$ e T um operador linear de V. Se W é um subespaço vetorial de V, dizemos que W é invariante por T (ou é T-invariante) se para cada vetor $\mathbf{w} \in W$ o vetor $T\mathbf{w} \in W$, isto é, $T(W) \subset W$.

Exemplo 8.19
Considere $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear que para cada vetor $\mathbf{v}$ no $\mathbb{R}^3$ roda este vetor em torno do eixo z um ângulo θ. Não é difícil de ver que

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} \cos(\theta)x - \operatorname{sen}(\theta)y \\ \operatorname{sen}(\theta)x + \cos(\theta)y \\ z \end{bmatrix}.$$

Considere $W = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$, um vetor $\mathbf{w} = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \in W$. E se calcuramos $T\mathbf{w} \in W$. Por isso, W é um subespaço T-invariante.

Em geral, se λ é um autovalor de T, então $W = \mathcal{N}(T - \lambda I)$ é um subespaço T-invariante, de fato, seja $\mathbf{u} \in \mathcal{N}(T - \lambda I)$ e avaliando $T\mathbf{u}$ temos

$$(T - \lambda I)(T\mathbf{u}) = T^2\mathbf{u} - \lambda T\mathbf{u} = \lambda T\mathbf{u} - \lambda T\mathbf{u} = 0.$$

Teorema 8.20
Seja $T : V \to V$ um operador linear sobre $V = \mathbb{R}^n$ cujo polinômio mínimo é

$$m_T(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_k(t)^{n_k},$$

onde os polinômio $f_i(t)$ são da forma: $t - c_i$ ou $t^2 + a_i t + b_i$ com $a_i^2 - 4b_i < 0$ e a_i, b_i e c_i são números reais. Os polinômios $f_i(t)$ são distintos. Então,

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Onde $W_i = \mathcal{N}(f_i(T)^{n_i})$, $i = 1, 2 \ldots, k$. Além disso, $f_i(t)^{n_i}$ são os polinômios mínimos de $T|_{W_i}$ que são os opertador T restristos ao subespaço W_i.

Demonstração: A demonstração deste resultado foge do escopo deste texto. Mas uma aplica ção deste resultado no próximo exemplo. $\square$

Exemplo 8.21
Considere o operador $T\mathbb{R}^3 \to \mathbb{R}^3$ que com respeito a base canônica do $\mathbb{R}^3$ é representando pela matriz

$$C = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Obtenha a decomposição primária desse operador.

Solução: Vamos determinar o polinômio característico, para isso, veja que tr(C) $= 2$, $\Delta_{11} + \Delta_{22} + \Delta_{33} = -7 + 2 + 6 = 1$ e $\det(C) = 2$. Daí,

$$\Delta_C(x) = x^3 - 2x^2 + x - 2 = (x-2)(x^2+1)$$

Veja que $x^2 + 1$ não tem raízes reais. Calculando

$$C^2 + I = \begin{bmatrix} 5 & -5 & 0 \\ 0 & 0 & 0 \\ 10 & -10 & 0 \end{bmatrix} \Rightarrow W_1 = \mathcal{N}(C^2 + I) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Já para o autovalor 2 temos

$$C - 2I = \begin{bmatrix} 4 & -3 & -2 \\ 4 & -3 & -2 \\ 10 & -5 & -5 \end{bmatrix} \Rightarrow W_2 = \mathcal{N}(C - 2I) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}.$$

Mostramos no exemplo 8.17 que estes dois espaços são independentes. Se construirmos uma base $\beta = \left\{ \left[\begin{smallmatrix} 1 \\ 0 \\ 2 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right] \right\}$ do $\mathbb{R}^3$. Nessa base a matriz de T torna-se

$$[T]_\beta^\beta = \begin{bmatrix} \boxed{2} & 0 & 0 \\ 0 & -3 & 5 \\ 0 & -2 & 3 \end{bmatrix}$$

$\square$

8.5 Autovalores Complexos

Como já havíamos comentado, um operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$, pode não ser diagonalizável em duas situações. A primeira é quando o polinômio mínimo possuir raízes (autovalores) repetidas e a quantidade de autovetores associado a este autovalor não for em quantidade suficiente. A segunda ocorre quando o polinômio característico possui raízes complexas. Vamos analizar o que esta acontecendo e determinar uma forma padrão para este tipo de operador.

Vamos iniciar com um exemplo. Considere $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x-y \\ 2x-y \end{smallmatrix}\right]$. Seja α a base canônica do $\mathbb{R}^2$. Então

$$A = [T]^{\alpha}_{\alpha} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}.$$

O polinômio característico fica $\Delta_A(t) = t^2 - \mathrm{tr}(A)t + \det(A) = t^2+1 = (t-i)(t+i)$.

Neste momento não há nada que se pode fazer, pois o operador foi concebido para ser sobre um espaço vetorial real. Mas, vamos continuar fazendo o processo de diagonalizção como se o operador fosse $T : \mathbb{C}^2 \to \mathbb{C}^2$, nesta situação tal operador é diagonalizável. Veja que $A - iI = \begin{bmatrix} 1-i & -1 \\ 2 & -1-i \end{bmatrix}$, observe que podemos obter a segunda linha por multiplicar a primeira por $1+i$, portanto a segunda linha é múltipla da primeira, ao procurarmos o $\ker(A - iI)$, obtemos a restrição $(1-i)z = w$, isto é,

$$\begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} z \\ (1-i)z \end{bmatrix} = z \begin{bmatrix} 1 \\ (1-i) \end{bmatrix}.$$

Ao procurarmos os autovalores associados ao autovalor $-i$, temos $A + iI = \begin{bmatrix} 1+i & -1 \\ 2 & -1+i \end{bmatrix}$, e multiplicando a primeira linha por $1-i$ obtemos a segunda. Daí, que a restrição será $(1+i)z = w$ e os autovetores serão da forma:

$$\begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} z \\ (1+i)z \end{bmatrix} = z \begin{bmatrix} 1 \\ (1+i) \end{bmatrix}.$$

Portanto, o operador T com respeito a base $\beta = \{(1, 1-i), (1, 1+i)\}$ será representado pela matriz

$$B = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

A princípio tal informação nos parece totalmente irrelevante, visto que não parece ter nenhuma conexão com o nosso operador inicial. Já veremos que isso não é bem assim.

Vamos demonstrar dois pequenos resultados.

Lema 8.22
Seja A uma matriz real $n \times n$ e λ um de seus autovalores. Se $\mathbf{v}$ é um autovetor associado ao autovalor $\lambda \in \mathbb{C}$, então $\bar{\lambda}$ (o conjugado complexo) também é um autovalor de A, e $\bar{\mathbf{v}}$ é um autovalor correspondente.

Demonstração: Observe que como $\mathbf{v}$ é um autovetor associado ao autovalor λ, isto quer dizer

$$A\mathbf{v} = \lambda\mathbf{v} \text{ conjugando } \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} \Rightarrow \bar{A}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}.$$

Como A é uma matriz real obtemos que $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. Uma vez que $\mathbf{v}$ é não nulo uma vez que é um autovetor, segue que $\bar{\mathbf{v}}$ também é não nulo. $\square$

Lema 8.23
Seja A uma matriz real 2×2 com autovalores complexos $\lambda = a \pm bi$ (onde $b \neq 0$). Se $\mathbf{w} = \mathbf{u} + i\mathbf{v}$ é o autovalor de A correspondente ao $\lambda = a - bi$, então $\beta = \{\mathbf{u}, \mathbf{v}\}$ é uma base e se P é a matriz de mudança de coordenadas da base β para a base canônica, então

$$A = P \begin{bmatrix} a & -b \\ b & a \end{bmatrix} P^{-1}.$$

Demonstração: Inicialmente observe que os vetores $\mathbf{u}$ e $\mathbf{v}$ não podem ser múltiplos (reais) um do outro, pois se o fossem, digamos $\mathbf{u} = a\mathbf{v}$ com $a \in \mathbb{R}$, então $\mathbf{w} = \mathbf{u} + i\mathbf{v} = a\mathbf{v} + i\mathbf{v} = (a+i)\mathbf{v}$, como $\mathbf{w}$ é um autovetor complexo, qualquer múltiplo dele por um número complexo também o será. Logo o autovetor real $\mathbf{v}$ será um autovetor de A associado a autovalor complexo λ o que claramente não pode acorrer.

Vamos tentar escrever a matriz do operador A na base β, mas antes, veja que segue das equações $A\mathbf{w} = \lambda\mathbf{w}$ e $A\bar{\mathbf{w}} = \bar{\lambda}\bar{\mathbf{w}}$ temos

$$\begin{aligned} A(\mathbf{u} + i\mathbf{v}) &= A\mathbf{u} + iA\mathbf{v} = (a - ib)(\mathbf{u} + i\mathbf{v}) = (a\mathbf{u} + b\mathbf{v}) + i(-b\mathbf{u} + a\mathbf{v}), \\ A(\mathbf{u} - i\mathbf{v}) &= A\mathbf{u} - iA\mathbf{v} = (a + ib)(\mathbf{u} - i\mathbf{v}) = (a\mathbf{u} + b\mathbf{v}) + i(b\mathbf{u} - a\mathbf{v}). \end{aligned}$$

Somando as duas equações obtemos $2A(\mathbf{u}) = 2(a\mathbf{u} + b\mathbf{v}) \Rightarrow A(\mathbf{u}) = a\mathbf{u} + b\mathbf{v}$. Se fizermos a primeira equação menos a segunda equação obtemos $2iA(\mathbf{v}) = 2i(-b\mathbf{u} + a\mathbf{v}) \Rightarrow A(\mathbf{v}) = -b\mathbf{u} + a\mathbf{v}$. Desta forma, a matriz A na base β fica

$$B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

□

Observe que para este tipo de matriz o seu $\det(B)$ é $a^2 + b^2$ e podemos reescrever ela da seguinte forma

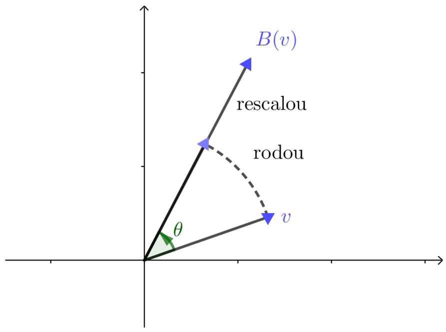
$$B = \det(B) \begin{bmatrix} \frac{a}{\det(B)} & -\frac{b}{\det(B)} \\ \frac{b}{\det(B)} & \frac{a}{\det(B)} \end{bmatrix}.$$

Nesta situação os pontos $(\frac{a}{\det(B)}, \frac{b}{\det(B)})$ estão sobre o círculo unitário, logo deve existir $\theta \in \mathbb{R}$, com $0 \leq \theta < 2\pi$, tal que $(\frac{a}{\det(B)}, \frac{b}{\det(B)}) = (\cos\theta, \operatorname{sen}\theta)$, e podemos reescrever

$$B = \det(B) \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

A matriz B nos diz que os vetores escritos na base β são primeiramente rodados θ radianos no sentido anti-horário seguindo de uma contração ou expansão dependendo apenas se $\det(B) > 1$ ou $\det(B) < 1$.

Em posse destes fatos voltamos ao exemplo inicial. Como no lema acima vamos considerar o autovetor $(1, 1+i) = (1,1) + i(0,1)$, que esta associado ao



autovalor $-i$. Considere a base $\beta' = \{(1,1),(0,1)\}$. Com respeito a esta base o operador $T(x,y) = (x-y, 2x-y)$ tem a seguinte matriz,

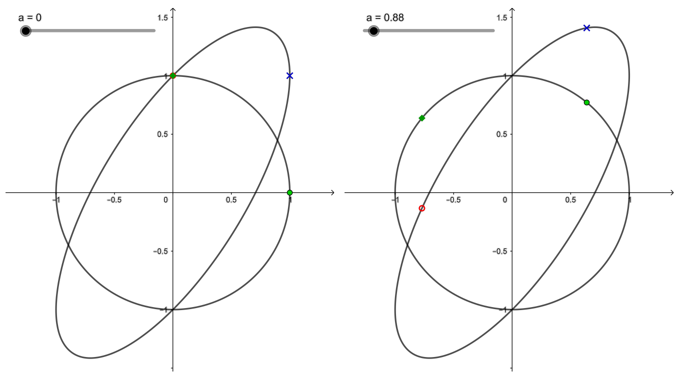
$$[T]_{\beta'}^{\beta'} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

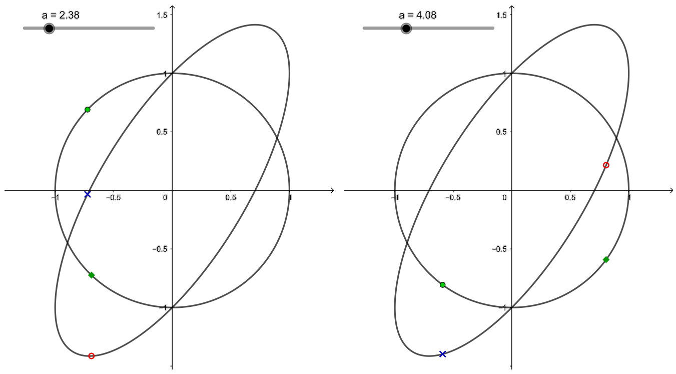
e isso quer dizer que T quando escrita na base β' faz somente uma rotação de $\pi/2$ no sentido anti-horário.

Apesar do que foi dito insinuar que os operadores com autovalores complexos fazem simplesmente uma rotação seguido de um reescalonamento do vetor, na prática o comportamento de um operador complexo é bem mais intricado. No nosso exemplo a matriz $P = [I]_{\alpha}^{\beta'} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Logo, $A = PBP^{-1}$ é

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

Considere o círculo de centro $(0,0)$ e raio 1. Vamos procurar os pontos do plano cujas coordenadas estão na base β' que são mapeados neste círculo. Não é difícil de perceber que estes pontos satisfazem $x^2 + (y-x)^2 = 1$, que é uma elipse. Isto quer dizer que os pontos da elipse são mapeados no círculo unitário por P^{-1}. Fiz uma sequência de imagem para tentar descrever o comportamento da dinâmica quando damos uma volta ao redor da círculo. O ponto $\times$ em azul é enviado (pela P^{-1}) no $\bullet$ em verde. O ponto $\bullet$ (verde) é enviado (pela B) no $\blacksquare$ (verde escuro). O ponto $\blacksquare$ em verde escuro é enviado (pela P) no $\circ$ em vermelho.





8.6 Forma Canônica de Jordan

Os operadores nilpotentes desempenham papel importante na descrição das formas canônicas de Jordan. Por isso, considere $T : V \to V$ um operador linear, dizemos que T é um operador nilpotente se $T^n = 0$ para algum inteiro positivo n, diremos que k é o seu índice de nilpotência se T^k é o operador nulo, mas existe $\mathbf{v} \in V$ tal que $T^{k-1}\mathbf{v} \neq 0$. Analogamente, esta definição se estende para as matrizes quadradas, dizemos que A é nilpotente se $A^n = 0$ para algum inteiro positivo n, dizemos, também, que k é o seu índice de nilpotência se A^k é o operador nulo, mas $A^{k-1} \neq 0$. Claramente o polinômio mínimo da matriz A será $m_A(\lambda) = \lambda^k$. Daí, segue que seu único autovalor será o zero.

Como já comentamos existem operadores que não admitem uma base na qual os operadores sejam representados como matrizes diagonais. Já sabemos que isso acontece quando o polinômio mínimo tem raízes repetidas. Além disso, os operadores diagonalizaveis precisam aparecer como casos particulares.

Exemplo 8.24
As matrizes serão usadas

$$J(\tau) = \begin{bmatrix} \tau & 1 & 0 & \cdots & 0 & 0 \\ 0 & \tau & 1 & \cdots & 0 & 0 \\ & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & \tau & 1 \\ 0 & 0 & 0 & \cdots & 0 & \tau \end{bmatrix} \text{ e } N(r) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}_{r\times r}.$$

A primeira matriz $J(\tau)$, é chamada de bloco de Jordan do autovalor τ. A segunda matriz $N(r)$ é chamada de bloco nilpotente de Jordan, consiste de 1's acima da diagonal principal e zeros no restante. seu índice de nilpotência é r (a matriz de ordem 1 consiste da matriz nula 1×1).

Observe que
$$J(\tau) = \tau I + N(r).$$

O teorema de Jordan vai nos garantir que qualquer operador T pode ser decomposto operadores, cada um dos quais é soma de um escalar com um operador nilpotente.

Um operador T pode ser posto na forma canônica de Jordan se o seu polinômio caractarístico e seu polinômio mínimo se fatoram em fatores lineares. Isso sempre acontece sobre o corpo dos números complexos. Sobre o corpo dos números reais existe uma adaptação que é chamado de forma canônica real, mas isto foge ao escopo deste texto.

Teorema 8.25
Seja $T : V \to V$ um operador linear cujos polinômios caracteristicos e mínimos são, respectivamente,

$$\Delta_T(t) = (t - c_1)^{m_1} \cdots (t - c_k)^{m_k} \text{ e } m_T(t) = (t - c_1)^{n_1} \cdots (t - c_k)^{n_k},$$

onde os c_i são escalares distintos. Então T admite uma base na qual a matriz que o representa esta na forma de blocos diagonais $J_{ij} = J(c_i)$ no qual para cada c_i, os blocos de Jordan correspondentes J_{ij} tem as seguintes propriedades:

1. *tem no mínimo um J_{ij} de ordem n_i; todos os outros J_{ij} tem ordem $\leq m_i$.*
2. *A soma das ordens de J_{ij} é m_i.*
3. *O número de J_{ij} é igual a $\dim \mathcal{N}(T - c_i I)$.*
4. *O número de J_{ij} e suas possíveis ordens é unicamente determinado por T.*

A demonstração deste teorema esta além dos objetivos deste texto, mas vamos dar alguns exemplo e explicar o procedimento para encontrar a base na qual a matriz que representa o operador fica na forma canônica de Jordan pelo menos para um operador de ordem 2.

Exemplo 8.26
Considere o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ determinado pela matriz

$$A = \begin{bmatrix} 15 & 9 \\ -16 & -9 \end{bmatrix}$$

na base canônica. Seu polinômio característico será

$$\Delta_A(t) = t^2 - (9 + 5)t - 9 = (t - 3)^2.$$

Claramente, $m_A(t) = \Delta_A(t)$, pois A não será anulado por $t - 3$. Pelo teorema 8.25 este operador admite uma base β do $\mathbb{R}^2$ na qual

$$[T]^\alpha_\alpha = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

Para determinar esta base vamos determinar o $\mathcal{N}(A - 3I)$.

$$A - 3I = \begin{bmatrix} 15 & 9 \\ -16 & -9 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 9 \\ -16 & -12 \end{bmatrix} \to \begin{bmatrix} 4 & 3 \\ 0 & 0 \end{bmatrix}.$$

Portanto, $\mathcal{N}(A-3I) = \text{Span}\left\{\begin{bmatrix}-3\\4\end{bmatrix}\right\}$. Diferente do que faziamos para diagonializar um operafor escolha qualquer vetor que não esteja em $\mathcal{N}(A-3I)$, como por exemplo $\begin{bmatrix}1\\0\end{bmatrix}$ e considere a base

$$\beta = \left\{\begin{bmatrix}1\\0\end{bmatrix}, (A-3I)\begin{bmatrix}1\\0\end{bmatrix}\right\} = \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}12\\-16\end{bmatrix}\right\}$$

Se P for a matriz mudança de coordenadas temos que

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

Veja o exercício R8.3 para entender melhor por que este procedimento funciona.

Exemplo 8.27
Suponha que o polinômio característico e o polinômio mínimo do operador T são, respectivamente,

$$\Delta_T(t) = (t+2)^4(t-5)^3 \text{ e } m_T(t) = (t+2)^2(t-5)^3.$$

Então a forma canônica de Jordan de T pode ser uma das seguintes matrizes:

$$\text{diag}\left(\begin{bmatrix}-2 & 1\\0 & -2\end{bmatrix}, \begin{bmatrix}-2 & 1\\0 & -2\end{bmatrix}, \begin{bmatrix}5 & 1 & 0\\0 & 5 & 1\\0 & 0 & 5\end{bmatrix}\right),$$
$$\text{diag}\left([-2],[-2],[-2],[-2], \begin{bmatrix}5 & 1 & 0\\0 & 5 & 1\\0 & 0 & 5\end{bmatrix}\right).$$

Obtemos a primeira matrix se T tem dois autovalores independentes, pertencentes ao autovalor -2, e obtemos a segunda matriz se T tem quatro autovalores pertencentes a -2.

Exemplo 8.28
Obtenha a forma de Jordan do operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ que na base canônica do $\mathbb{R}^3$ é representando pela seguinte matriz

$$A = \begin{bmatrix}-7 & 4 & 3\\-4 & 3 & 2\\-16 & 8 & 7\end{bmatrix}$$

Solução: Vamos iniciar calculando o polinômio característico, para isso considere: $\text{tr}(A) = 3$, $\Delta_{11} + \Delta_{22} + \Delta_{33} = 5 - 1 - 5 = -1$ e o $\det(A) = -3$, temos então

$$\Delta_A(t) = t^3 - 3t^2 - t + 3 = (t-3)(t-1)(t+1).$$

Como as raízes são distintas, esse operador é diagonalizável, logo existe uma base β do $\mathbb{R}^3$

$$[T]_\beta^\beta = \begin{bmatrix} \boxed{3} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{-1} \end{bmatrix}.$$

Veja que ele é composto por 3 bloco 1×1, portanto, este operador já esta na forma de Jordan! Isto quer dizer que o processo de diagonalização é o método para obter uma base na qual o operador se escreve na forma de Jordan.

□

8.7 Aplicação a Equações Diferenciais

Considere a equação diferencial

$$y' = \frac{dy}{dx} = ay$$

onde $y = f(x)$ é uma função desconhecida a ser determinada e a é uma constante. Como a maioria das equações diferenciais essa equação tem infinitas soluções, mas se houver condição inicial, do tipo $y(0) = 5$ por exemplo, então somos capazes de encontrar uma única solução. Sem se levar em consideração como obtemos a solução podemos admitir que a solução geral é dado por

$$y = ce^{ax}.$$

Veja que com a condição inicial podemos concluir que $c = 5$.

Sabemos do calculo que para calcular

$$e^s = 1 + s + \frac{s}{2!} + \frac{s^3}{3!} + \cdots + \frac{s^n}{n!} + \cdots$$

Um sistema de equações linear de equações diferenciais de primeira ordem é um sistema da forma

$$\begin{cases} y_1' & = & a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ y_2' & = & a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \vdots' & = & \vdots \\ y_n' & = & a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{cases}$$

onde $y_1 = f_1(x)$, $y_2 = f_2(x), \ldots, y_n = f_n(x)$ são funções a serem determinadas e os coeficientes a_{ij} são contantes.

Usando a notação matricial o sistema se torna

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$
$$Y' = AY$$

Portanto, seguindo a receita, precisamos ser capazes de calcular potências da Matriz A, pois desta forma seremos capazes de determinar e^{Ax}.

Exemplo 8.29
Considere

$$\begin{cases} y_1' &= 3y_1 \\ y_2' &= -2y_2 \end{cases}$$

em linguagem matricial se torna

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$
$$Y' = DY$$

Então precisamos calcular potências de D Vamos primeiro calcular D^2:

$$D^2 = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 3 \cdot 3 + 0 \cdot 0 & 3 \cdot 0 + 0 \cdot (-2) \\ 0 \cdot 3 - 2 \cdot 0 & 0 \cdot 0 + (-2) \cdot (-2) \end{bmatrix} = \begin{bmatrix} 3^2 & 0 \\ 0 & (-2)^2 \end{bmatrix}.$$

Em seguida, calculamos D^3:

$$D^3 = D^2 \cdot D = \begin{bmatrix} 3^2 & 0 \\ 0 & (-2)^2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 3^2 \cdot 3 & 0 \\ 0 & (-2)^2 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 3^3 & 0 \\ 0 & (-2)^3 \end{bmatrix}.$$

Não é difícil convencer-se de que, para qualquer k natural, vale a fórmula

$$D^k = \begin{bmatrix} 3^k & 0 \\ 0 & (-2)^k \end{bmatrix}. \tag{8.1}$$

E portanto, as soluções da equação são da forma

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} e^{3x} & 0 \\ 0 & e^{-2x} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

onde c_1 e c_2 são constantes que podem ser determinadas se soubermos uma condição inicial, do tipo, $y_1(0) = 3$ e $y_2(0) = -2$.

Exercícios resolvidos

R8.1. Sejam A, B e P matrizes $n \times n$, e suponha que P seja invertível. Mostre que a relação $A = PBP^{-1}$ é equivalente a $P^{-1}AP = B$.

Solução: Supondo que $A = PBP^{-1}$, temos

$$\begin{aligned} A = PBP^{-1} \quad &\Longrightarrow \quad P^{-1}AP = P^{-1}(PBP^{-1})P \\ &\Longrightarrow \quad P^{-1}AP = IBI \quad \Longrightarrow \quad P^{-1}AP = B. \end{aligned}$$

Reciprocamente, supondo que $P^{-1}AP = B$, temos

$$\begin{aligned} P^{-1}AP = B \quad &\Longrightarrow \quad P(P^{-1}AP)P^{-1} = PBP^{-1} \\ &\Longrightarrow \quad IAI = PBP^{-1} \quad \Longrightarrow \quad A = PBP^{-1}. \quad \square \end{aligned}$$

R8.2. Prove a proposição 8.4 cujo enunciado é: Se A e B são matrizes $n \times n$ semelhantes, então

$$\Delta_A(x) = \Delta_B(x).$$

Solução: Como as matrizes A e B são semelhantes, então existe uma matriz P invertível tal que $B = P^{-1}AP$. Daí

$$\begin{aligned}
\Delta_B(x) &= \det(B - xI) \\
&= \det(P^{-1}AP - xP^{-1}IP) \\
&= \det(P^{-1}(AP - xIP)) \\
&= \det(P^{-1})\det((A - xI)P) \\
&= \det(P^{-1})\det(A - xI)\det(P) = \Delta_A(x).
\end{aligned}$$

□

R8.3. Prove o teorema 8.11 para o caso 2×2.

Solução: Considere uma matriz 2×2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Pelas fórmulas obtidas na seção 8.2.1 temos que $\Delta_A(x) = x^2 - \operatorname{tr}(A)x + \det(A)$

$$\begin{aligned}
&A^2 - (a+d)A + (ad - bc)I \\
&= \begin{bmatrix} a^2 + bc & ab + db \\ ac + dc & d^2 + bc \end{bmatrix} - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad - bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

□

R8.4. Considere a matriz A 2×2 com autovalores λ repetidos e que não seja diagonalizável. Tome $\mathbf{u} \notin \mathcal{N}(A - \lambda I)$ com $\mathbf{u} \neq 0$, então $\{(A - \lambda I)\mathbf{u}, \mathbf{u}\}$ forma uma base na qual a matriz A fica na forma canônica de Jordan.

Solução: A primeira dúvida é se de fato, $\{(A - \lambda I)\mathbf{u}, \mathbf{u}\}$ é uma base. Como $\mathbf{u} \notin \mathcal{N}(A - \lambda I)$ segue que $(A - \lambda I)\mathbf{u} \neq 0$. Para garantir que este conjunto é LI, basta garantir que um vetor não é multiplo do outro. Digamos que fossem multiplos, isto é, deve existir $k \in R$, com $k \neq 0$ tal que

$$(A - \lambda I)\mathbf{u} = k\mathbf{u} \Rightarrow A\mathbf{u} = (k + \lambda)\mathbf{u}$$

Mas então, $\mathbf{u}$ seria um autovetor associado ao autovalor $k + \lambda \neq \lambda$. Mas A só tem um autovalor o que vai contrar a nossa hipótese. Portanto, esta possibilidade não pode acontecer.

Estamos dizendo que λ é o único autovalor de A, portanto,

$$\Delta_A(x) = (x-\lambda)^2 = x^2 - 2\lambda x + \lambda^2.$$

E pelo teorema de Cayley temos que

$$A^2 - 2\lambda A + \lambda^2 I = 0 \Rightarrow A^2 - \lambda A = \lambda A - \lambda^2 I.$$

Agora vamos determinar a Matriz de A com respeito a base $\{(A-\lambda I)\mathbf{u}, \mathbf{u}\}$, isto é,

$$\begin{gathered} A((A-\lambda I)\mathbf{u}) = (A^2 - \lambda A)\mathbf{u} = (\lambda A - \lambda^2 I)\mathbf{u} = \lambda(A-\lambda I)\mathbf{u} + 0\mathbf{u} \\ A\mathbf{u} = A\mathbf{u} - \lambda\mathbf{u} + \lambda\mathbf{u} = 1(A-\lambda I)\mathbf{u} + \lambda\mathbf{u} \end{gathered}$$

Isso mostra que na base $\{(A-\lambda I)\mathbf{u}, \mathbf{u}\}$ a matriz A fica na forma de Jordan

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

□

Exercícios propostos

P8.1. (a) Seja A a matriz do exemplo 8.3. Verifique que a equação $A\mathbf{v} = \lambda\mathbf{v}$ equivale ao sistema

$$\begin{cases} 8v_1 + 5v_2 - \lambda v_1 = 0 \\ -10v_1 - 7v_2 - \lambda v_2 = 0, \end{cases}$$

nas variáveis λ, v_1 e v_2 (onde v_1 e v_2 são as coordenadas de $\mathbf{v}$). Note que esse sistema é *não-linear*, em vista dos produtos λv_1 e λv_2.

(b) Observe que, em geral, se A é uma matriz $n \times n$, a equação $A\mathbf{x} = \lambda\mathbf{x}$ representa um sistema não-linear nas variáveis λ, $x_1, \ldots, x_n$, onde $x_1, \ldots, x_n$ são as coordenadas do vetor $\mathbf{x}$.

P8.2. Complete o estudo do operador do exemplo 8.28. Encontre a base β do $\mathbb{R}^3$.

P8.3. Para as matrizes abaixo, determine se A é diagonalizável. Você consegue determinar P que coloca $P^{-1}AP$ na forma de canônica de Jordan.

$$\text{a) } A = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix} \qquad \text{b) } A = \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix}$$

$$\text{c) } A = \begin{bmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{bmatrix} \qquad \text{d) } A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 5 & -5 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

P8.4. Calcule A^{10} quando
$$A = \begin{bmatrix} 8 & 5 \\ -10 & -7 \end{bmatrix}.$$

P8.5. Resolva a equação diferencial
$$\begin{cases} y_1' &= 8y_1 + 5y_2 \\ y_2' &= -10y_1 - 7y_2 \end{cases}$$
com condições iniciais: $y_1(0) = 2$ e $y_2(0) = -2$.

P8.6. (a) Mostre que se D é uma matriz diagonal com entradas não negativas na diagonal principal, então existe uma matriz S tal que $S^2 = D$. (b) Mostre que se A é uma matriz diagonalizável com autovalores não-negativos, então existe uma matriz S tal que $S^2 = A$. (c) Encontre uma matriz S tal que $S^2 = A$, se
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 4 & 0 \\ 1 & 5 & 9 \end{bmatrix}.$$

P8.7. (a) Prove que se A for uma matriz quadrada, então A e A^t tem os mesmos autovalores. (b) Moste que A e A^t não precisam ter os mesmo autovalores.

P8.8. Considere $\mathbf{u}$ autovetor de A associado ao autovalor λ e $\mathbf{v}$ autovetor de A^t associados a μ com $\lambda \neq \mu$. Prove que $\mathbf{u}$ é perpendicular com $\mathbf{v}$.

P8.9. Considere $a, b, c, e \in \mathbb{Z}$ com $a + b = c + d$. Prove que $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tem somente autovalores inteiros.

P8.10. Determine os autovalores de
$$A = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & & \vdots \\ b_1 & b_2 & \cdots & b_n \end{bmatrix}.$$

P8.11. Decida se a afirmação dada é verdadeira ou falsa. Justifique sua resposta dando um argumento lógico ou um contra-exemplo.

a) Uma matriz quadrada com vetores colunas linearmente independentes é diagonalizável.

b) Se A é diagonalizável, então existe uma única matriz P tal que $P^{-1}AP$ é uma matriz diagonal.

c) Se $\mathbf{u}_1$, $\mathbf{u}_2$ e $\mathbf{u}_3$ são vetores do auto espaço distintos de A, então é impossível expressar $\mathbf{u}_3$ como combinação linear de $\mathbf{u}_1$ e $\mathbf{u}_2$.

d) Se A é diagonalizável e invertível, então A^{-1} também é diagonalizável.

P8.12. Suponha que o polinômio característico de A seja $(t-1)(t-3)^2(t-4)^3$. Para cada afirmação responda e explique o seu raciocínio.

a) O que você pode dizer sobre a dimensão dos auto espaços de A.

b) O que você pode dizer sobre a dimensão dos auto espaço de A se você souber que A é diagonalizável.

P8.13. Calcule A^2, $A^2, \ldots, A^9$, tente identificar o padrão onde

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

P8.14. Calcule e^A, onde A é a matriz do exercícios P8.13. (DICA: $\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$ e $\operatorname{sen}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots$).

P8.15. Considere o polinômio $f(x) = x^5 + x^3 - 2x + 4$, calcule $f(A)$, onde

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

(DICA: Se $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ então o derivada destes polinômio é $f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$.)

P8.16. Se $A_1, A_2, \ldots, A_k, \ldots$ é uma sequência infinita de matrizes $n \times n$, dizemos que a sequência converge para a matriz A se as entradas na i-ésima linha e da j-ésima coluna da sequência convergem à entrada da i-ésima linha e da j-ésima coluna da matriz A, para cada i e j. Neste caso, dizemos que A é o limite da sequência e escrevemos $\lim_{k\to\infty} A_k = A$. Suponha que P é uma matriz $n \times n$ invertível e se $\lim_{k\to\infty} A_k = A$ se e somente se $\lim_{k\to\infty} P^{-1}A_k P = P^{-1}AP$.

P8.17. Se $A_1 + A_2 + \cdots + A_k + \cdots$ é uma série infinita de matrizes $n \times n$, dizemos que a série converge para a matriz A se a sequência das somas parciais converge para A no sentido do exercício refseq:matrizes. Neste caso, dizemos que A é soma da série e escrevemos $A = A_1 + A_2 + \cdots + A_k + \cdots$. Lembre-se das condições impostas para que a série geométrica

$$1 + x + x^2 + \cdots + x_k + \cdots$$

seja convergente. Qual condição sobre os autovalores de A garante a convergência da série geométrica $A_1 + A_2 + \cdots + A_k + \cdots$.