Capítulo 7

Diagonalização

7.1 Diagonalização de operadores

Dado um operador linear $T : V \to V$, nosso objetivo é encontrar, se existir, uma base β de V tal que a matriz de T nessa base, $[\,T\,]_\beta^\beta$, seja diagonal. É a isso que nos referimos por *diagonalizar um operador*, ou ainda, dizer que um operador é *diagonalizável*. É evidente que matrizes diagonais são mais simples de lidar, por exemplo, quando queremos aplicar o operador a um vetor: $[\,T(v)\,]_\beta = [\,T\,]_\beta^\beta\,[\,v\,]_\beta$.

Antes de apresentar a teoria, vamos retomar um exemplo do capítulo anterior.

Considere o operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

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

Escolhendo a base

$$\beta = \{(1, 0, 1); (0, 2, 0); (0, 0, 1)\}$$

de $\mathbb{R}^3$ e fazendo as contas chegamos a

$$[\,T\,]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Note que $[\,T\,]_\beta^\beta$ é diagonal, e que cada vetor da base β é autovetor de T. De fato, você pode verificar que

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

Este exemplo indica que a procura pela base β passa por estudar os autovetores de T.

7.1.1 Bases formadas por autovetores

Dado um operador linear $T : V \to V$ suponha que

$$\beta = \{v_1, \ldots, v_n\}$$

seja uma base de V para a qual

$$[\,T\,]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Sabemos que a cada coluna i dessa matriz é exatamente $[\,T(v_i)\,]_\beta$, e assim

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

para todo i. Ou seja,

> Se $[\,T\,]_\beta^\beta$ é diagonal, então a base β é formada por autovetores de T.

Po outro lado, se

$$\beta = \{v_1, \ldots, v_n\}$$

é base V formada por autovetores de T, digamos $T(v_i) = \lambda_i v_i$ para todo i (autovetor v_i associado ao autovalor λ_i), então

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

para todo i, e logo

$$[\,T\,]_\beta^\beta = \begin{bmatrix} | & | & & | \\ [\,T(v_1)\,]_\beta & [\,T(v_2)\,]_\beta & \cdots & [\,T(v_n)\,]_\beta \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Ou seja,

> Se β é base formada por autovetores de T, então $[\,T\,]_\beta^\beta$ é diagonal. Neste caso, **a diagonal de $[\,T\,]_\beta^\beta$ será formada pelos autovalores de T na ordem em que a base de autovetores é construída**, como mostra a expressão acima.

Isso nos leva à concluir que T admite uma base β que torna sua matriz diagonal se, e somente se, essa base β é formada por autovetores T. Isso motiva a definição de *operador diagonalizável*:

Definição 7.1. *Um operador $T : V \to V$ é **diagonalizável** se existe uma base de V formada por autovetores de T.*

Exemplo 7.1. O operador T sobre $\mathbb{R}^3$ do início do capítulo,

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

é diagonalizável pois, como vimos, $\beta = \{(1, 0, 1); (0, 2, 0); (0, 0, 1)\}$ é uma base de $\mathbb{R}^3$ onde todos os elementos são autovetores de T. Neste caso, vimos que $[\,T\,]_\beta^\beta$ é diagonal. □

Exemplo 7.2. Considere o operador $S : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$S(x, y) = (\ x + y\ ,\ y\ ).$$

O polinômio característico associado é

$$p(\lambda) = \det\big([\,S\,]_{can}^{can} - \lambda \mathbf{I}_2\big) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2,$$

cuja única raíz (autovalor) é $\lambda = 1$. Resolvendo o sistema $([\,S\,]_{can}^{can} - \mathbf{I}_2)[\,v\,]_{can} = \mathbf{0}$, obtemos

$$[\,v\,]_{can} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

(verifique!). Assim, conseguimos apenas um autovetor LI, por exemplo, $v = (1, 0)$. Logo, **não é possível** construir uma base de $\mathbb{R}^2$ formada por autovetores de S. Assim, S **não** é diagonalizável. Você pode tentar conseguir uma base para a qual a matriz de S seja diagonal, e não terá sucesso! □

O exemplo anterior mostra que **nem sempre** é possível diagonalizar um operador. **Então, como saber se um operador é diagonalizável?**

Passamos agora a estudar quando conseguimos diagonalizar um operador. Ou seja, **quando conseguimos autovetores LI suficientes para formar uma base**.

Teorema 7.1. *Autovetores associados a autovalores distintos são LI.*
Mais especificamente se $\lambda_1, \ldots, \lambda_r$ são autovalores de T, distintos entre si, e $v_1, \ldots, v_r$ são autovetores correspondentes então $\{v_1, \ldots, v_r\}$ é um conjunto LI.

Imediatamente, se o operador $T : V \to V$ possui n autovalores distintos ($n = \dim V$) então existem n autovetores LI, e logo formam uma base de V.

Corolário 7.1. *Se V é um espaço vetorial de dimensão n e o operador linear $T : V \to V$ possui n autovalores distintos, então T é diagonalizável.*

Exemplo 7.3. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear definido por

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

Sua matriz na base canônica do $\mathbb{R}^3$ é

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

donde segue que o polinômio característico é

$$p(\lambda) = \det\big([\,T\,]^{can}_{can} - \lambda \mathbf{I}_3\big) = \det \begin{bmatrix} 1-\lambda & 0 & 2 \\ -1 & -\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix}.$$

Usando desenvolvimento de Laplace sobre a primeira linha, obtemos

$$\begin{aligned} p(\lambda) &= (-1)^{1+1}(1-\lambda)\det\begin{bmatrix} -\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} + (-1)^{1+3}\,2\det\begin{bmatrix} -1 & -\lambda \\ 1 & 1 \end{bmatrix} \\ &= (1-\lambda)\big[-\lambda(2-\lambda) - 1\big] + 2(-1+\lambda) \\ &= (1-\lambda)\big[\lambda^2 - 2\lambda - 3\big]. \end{aligned}$$

Assim, os autovalores de T (raízes de p) são

$$\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 3.$$

Como os três autovalores são distintos entre si, T é diagonalizável. Podemos então calcular uma base de autovetores que diagonaliza T.

Autovetor associado a $\lambda_1 = 1$. Devemos encontrar uma solução não trivial do sistema $\big([\,T\,]^{can}_{can} - \mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$ (onde $[\,v\,]_{can} = \begin{bmatrix} x & y & z \end{bmatrix}^t$), ou seja,

$$\left\{\begin{array}{rrrcl} & & 2z & = & 0 \\ -x & -y & +z & = & 0 \\ x & +y & +z & = & 0 \end{array}\right.$$

Uma solução não trivial é

$$[\,v_1\,]_{can} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \Rightarrow \qquad v_1 = (1, -1, 0).$$

Autovetor associado a $\lambda_2 = -1$. O sistema associado $\big([\,T\,]_{can}^{can} + \mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$ toma a forma

$$\left\{\begin{array}{rrrcl} 2x & & +2z & = & 0 \\ -x & +y & +z & = & 0 \\ x & +y & +3z & = & 0 \end{array}\right. .$$

Uma solução não trivial é (verifique)

$$[\,v_2\,]_{can} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \qquad \Rightarrow \qquad v_2 = (1, 2, -1).$$

Autovetor associado a $\lambda_3 = 3$. O sistema associado é $\big([\,T\,]_{can}^{can} - 3\,\mathbf{I}_3\big)[\,v\,]_{can} = \mathbf{0}$, ou seja,

$$\left\{\begin{array}{rrrcl} -2x & & +2z & = & 0 \\ -x & -3y & +z & = & 0 \\ x & +y & -z & = & 0 \end{array}\right. .$$

Uma solução não trivial é

$$[\,v_3\,]_{can} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad \Rightarrow \qquad v_3 = (1, 0, 1).$$

Portanto o conjunto

$$\beta = \{v_1, v_2, v_3\} = \{(1, -1, 0)\ ;\ (1, 2, -1)\ ;\ (1, 0, 1)\ \}$$

é uma base de $\mathbb{R}^3$ formanda por autovetores de T.

Finalmente, vimos que a diagonal de $[\,T\,]_\beta^\beta$ é formada pelos autovalores na mesma ordem em que os autovetores aparecem na base β, ou seja, **sem a necessidade de fazer contas** temos

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

□

Exemplo 7.4. Considere o operador linear sobre $\mathbb{R}^4$

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

A matriz de T na base canônica de $\mathbb{R}^4$ é

$$[\,T\,]_{can}^{can} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 5 \end{bmatrix},$$

cujos autovalores são

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = -1, \quad \lambda_4 = 5,$$

pois trata-se de uma matriz triangular inferior. Como todos os autovalores são distintos, T é diagonalizável. Isto é, é possível construir uma base β de $\mathbb{R}^4$ formada por autovetores de T, para a qual $[\,T\,]_\beta^\beta$ será diagonal. □

Atividade 7.1. Calcule uma base de $\mathbb{R}^4$ formada por autovetores de T no exercício anterior, e escreva a matriz de T nessa base.

7.1.2 Matrizes

Operadores lineares podem ser descritos por suas matrizes em determinadas bases; portanto, eles podemos falar em *matrizes diagonalizáveis* também.

Definição 7.2. *Uma matriz quadrada* $\mathbf{A}$ *de ordem* n *é dita* **diagonalizável** *se o operador associado* $T : \mathbb{R}^n \to \mathbb{R}^n$ *definido por* $[T]_{can}^{can} = \mathbf{A}$ *é diagonalizável.*

Isto é, $\mathbf{A}$ ***(***$n \times n$***) é diagonalizável se admite*** n ***autovetores LI.***

Para saber se uma matriz é diagonalizável, procedemos da mesma forma que com operadores:

- Calculamos as raízes do polinômio característico de $\mathbf{A}$,
$$p(\lambda) = \det\big(\mathbf{A} - \lambda \mathbf{I}_n\big)$$
- Se as raízes (autovalores de $\mathbf{A}$) forem todas distintas entre si, então $\mathbf{A}$ é diagonalizável e podemos calcular os autovetores resolvendo os sistemas $\mathbf{A}\mathbf{v} = \lambda_i \mathbf{v}$ (ou $(\mathbf{A} - \lambda_i \mathbf{I}_n)\mathbf{v} = \mathbf{0}$).

Exemplo 7.5. A matriz
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$
é diagonalizável pois seu polinômio característico é
$$p(\lambda) = \det\big(\mathbf{A} - \lambda \mathbf{I}_3\big) = \det \begin{bmatrix} 1-\lambda & 1 & 8 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda)(-1-\lambda),$$
cujas raízes são distintas entre si: $\lambda_1 = 1$, $\lambda_2 = 2$ e $\lambda_3 = -1$. □

Atividade 7.2. Calcule 3 autovetores LI da matriz $\mathbf{A}$ do exemplo anterior.

7.1.3 O polinômio minimal

É de interesse saber se um operador/matriz é diagonalizável antes de calcular seus autovetores, pois, caso não seja, evitamos buscar uma base de autovetores em vão. Nas seções anteriores vimos um caso onde isso é possível: quando T (ou $\mathbf{A}$) possui todos seus autovalores distintos entre si (Corolário 7.1). Mas quando T (ou $\mathbf{A}$) possui alguns autovalores iguais? Como decidir apenas pelos autovalores se T (ou $\mathbf{A}$) é diagonalizável?

A resposta passa por eliminar raízes repetidas do polinômio característico (autovalores com multiplicidade algébrica ≥ 1). Isso nos levará a um polinômio com mesmas raízes que o polinômio característico, mas com grau menor. Definiremos um polinômio de grau mínimo possível com essas características, que chamaremos de *polinômio minimal*.

Como de costume, utilizaremos matrizes do operador T em nossas contas. Iniciemos então o estudo em termos de matrizes.

Definição 7.3. *Seja*
$$q(x) = a_n x^n + \cdots + a_1 x + a_0$$
um polinômio e $\mathbf{A}$ *uma matriz quadrada. Então* $q(\mathbf{A})$ *é matriz*
$$q(\mathbf{A}) = a_n \mathbf{A}^n + \cdots + a_1 \mathbf{A} + a_0 \mathbf{I}.$$
Isto é, trocamos x *por* $\mathbf{A}$ *e* 1 *por* $\mathbf{I}$*, e operamos normalmente.*

Quando $q(\mathbf{A}) = \mathbf{0}$ *(matriz nula) dizemos que o polinômio* q ***anula*** *a matriz* $\mathbf{A}$.

Exemplo 7.6. Sejam $q(x) = x^2 - 9$, $t(x) = 2x + 3$ e $\mathbf{A} = \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix}$. Temos

$$q(\mathbf{A}) = \mathbf{A}^2 - 9\mathbf{I}_2 = \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} - 9 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

e

$$t(\mathbf{A}) = 2\mathbf{A} - 3\mathbf{I}_2 = 2 \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 8 \\ 4 & 5 \end{bmatrix}.$$

Assim, q anula $\mathbf{A}$ e t não anula $\mathbf{A}$. □

Definição 7.4. *Seja* $\mathbf{A}$ *uma matriz quadrada. O* ***polinômio minimal*** *de* $\mathbf{A}$ *é o polinômio da forma*

$$m(x) = x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0$$

(observe que $a_r = 1$*) tal que*

- $m(\mathbf{A}) = \mathbf{0}$*, isto é, m anula* $\mathbf{A}$*;*
- m *é o polinômio de menor grau que anula* $\mathbf{A}$*.*

Podemos responder se o operador linear $T : V \to V$ é diagonalizável ou não olhando para o polinômio minimal da matriz (quadrada) $[\,T\,]_\beta^\beta$ em qualquer base β de V.

Teorema 7.2. *Sejam* $T : V \to V$ *um operador linear e* β *uma base qualquer de* V*. Então* T *é diagonalizável se, e somente se, o polinômio minimal da matriz* $[\,T\,]_\beta^\beta$ *tem a forma*

$$m(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_r),$$

onde $\lambda_1, \ldots, \lambda_r$ *são os autovalores distintos de* T*.*

O teorema acima indica que precisamos apenas calcular so autovalores de T para saber se o operador é diagonalizável ou não. Isso envolve, evidentemente, calcular as raízes do polinômio característico.
O que falta é uma maneira de calcular o polinômio minimal! Veremos como fazer isso a partir do polinômio característico.

Teorema 7.3 (de Cayley-Hamilton)**.** *Sejam* $T : V \to V$ *um operador linear,* β *uma base de* V*, e*

$$p(\lambda) = \det\left([\,T\,]_\beta^\beta - \lambda\mathbf{I}\right)$$

o polinômio característico associado a T*. Então* p *anula* $[\,T\,]_\beta^\beta$*, isto é,*

$$p([\,T\,]_\beta^\beta) = \mathbf{0}.$$

NÃO se pode concluir que $p([\,T\,]_\beta^\beta) = \mathbf{0}$ pensando que "$\det\left([\,T\,]_\beta^\beta - [\,T\,]_\beta^\beta\mathbf{I}\right) = \det\mathbf{0}$"... **Esta conta está ERRADA e não faz sentido**, pois $\det\mathbf{0}$ é um número, não uma matriz nula! Deve-se primeiro calcular o polinômio $p(\lambda)$ desenvolvendo o determinante normalmente e **SÓ DEPOIS** substituir λ por $[\,T\,]_\beta^\beta$. A prova deste teorema é bem mais complicada...

O Teorema de Cayley-Hamilton diz que o polinômio característico é **candidato** a ser o polinômio minimal, pois anula a matriz de T. Agora, devemos cuidar das possíveis raízes repetidas!

Teorema 7.4. *Os polinômios minimal e característico de* $[\,T\,]_\beta^\beta$ *têm as mesmas raízes, exceto pelas multiplicidades.*

O teorema acima diz que se λ é autovalor de T, então ambos os polinômios característico e minimal possuem o fator (ou mais precisamente, são divisíveis por)

$$x - \lambda.$$

Assim, os Teoremas 7.2, 7.4 e 7.3 fornecem a seguinte maneira de calcular o polinômio minimal e decidir se T é diagonalizável ou não:

Cálculo do polinômio minimal

1. Calcule o polinômio característico $p(\lambda) = \det\big([\,T\,]_\beta^\beta - \lambda\mathbf{I}\big)$;
2. Encontre todas as raízes de p (autovalores); digamos que suas raízes distintas sejam $\lambda_1, \dots, \lambda_r$;
3. Escreva $p(x)$ na forma
$$p(x) = c\,(x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\cdots(x-\lambda_r)^{d_r}$$
(λ foi trocado por x por conveniência). Os expoentes $d_1, \dots, d_r$ são o número de vezes que cada autovalor aparece como raiz de p, ou seja, são as multiplicidades algébricas de cada autovalor;
4. O polinômio minimal será o polinômio **de menor grau**, com $c = 1$, que contém pelo menos um de cada termo $x - \lambda_i$ ($i = 1, \dots, r$), e que **anule a matriz** $[\,T\,]_\beta^\beta$.

Decidindo se T é diagonalizável ou não

5. Com o polinômio minimal $m(x)$ em mãos, T será diagonalizável se, e somente se, se **cada termo** $x - \lambda_i$ aparece uma **única** vez em $m(x)$ (Teorema 7.2).

Exemplo 7.7. Considere o operador linear sobre $\mathbb{R}^3$ dado por

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

Temos

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

e o polinômio característico é

$$p(\lambda) = \det\big([\,T\,]_{can}^{can} - \lambda\mathbf{I}_3\big) = \det\begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ -1 & 0 & 2-\lambda \end{bmatrix} = (1-\lambda)^2(2-\lambda).$$

Como o polinômio minimal tem as mesmas raízes de p, ele deve ser um dos seguintes polinômios:

- $p_1(x) = (1-x)(2-x)$
- $p_2(x) = (1-x)^2(2-x)$ (o próprio polinômio característico p).

Por definição, o polinômio minimal $m(x)$ é aquele de **menor grau** que **anula** $[T]_{can}^{can}$. Então, testamos se cada candidato $p_1(x)$ e $p_2(x)$ anula $[T]_{can}^{can}$, em ordem crescente de grau. Ao encontrarmos o primeiro polinômio que anula esta matriz, paramos a busca!

Primeiro, verificamos se $p_1(x)$ anula $[T]_{can}^{can}$: expandindo $p_1(x)$, temos

$$p_1(x) = x^2 - 3x + 2.$$

Assim,

$$\begin{aligned}
p_1([T]_{can}^{can}) &= \left([T]_{can}^{can}\right)^2 - 3[T]_{can}^{can} + 2\mathbf{I}_3 \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 4 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & 0 & -4 \end{bmatrix} = \mathbf{0}
\end{aligned}$$

(você pode também calcular $(\mathbf{I}_3 - [T]_{can}^{can})(2\mathbf{I}_3 - [T]_{can}^{can})$ e chegar ao mesmo resultado). Ou seja, $p_1([T]_{can}^{can}) = \mathbf{0}$, e portanto p_1 anula $[T]_{can}^{can}$. Assim, o polinômio minimal é

$$m(x) = (1 - x)(2 - x).$$

Como cada um dos termos do tipo $(a - x)$ aparecem uma única vez $m(x)$, concluímos que T é diagonalizável. □

Exemplo 7.8. Considere o operador linear sobre $\mathbb{R}^4$ definido por

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

Temos

$$[T]_{can}^{can} = \begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

e, como esta matriz é triangular,

$$p(\lambda) = (3 - \lambda)^2(-1 - \lambda)^2.$$

Os candidatos à polinômio minimal são

- $p_1(x) = (3 - x)(-1 - x)$ (polinômio de grau 2)
- $p_2(x) = (3 - x)^2(-1 - x)$ (polinômio de grau 3)
- $p_3(x) = (3 - x)(-1 - x)^2$ (polinômio de grau 3)
- $p_4(x) = p(x) = (3 - x)^2(-1 - x)^2$ (polinômio de grau 4).

O operador T será diagonalizável somente se $p_1(x)$ for seu polinômio minimal, dado que em todos os outros candidatos há termos de grau 1 aparecendo mais de uma vez. Temos

$$\begin{aligned}
p_1([T]_{can}^{can}) &= \left(3\mathbf{I}_4 - [T]_{can}^{can}\right)\left(-\mathbf{I}_4 - [T]_{can}^{can}\right) \\
&= \begin{bmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} -4 & 0 & 4 & 0 \\ 0 & -4 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{0}.
\end{aligned}$$

Ou seja, $p_1(x)$ anula $[T]_{can}^{can}$. Sendo o candidato de menor grau, $p_1(x)$ é o polinômio minimal. Logo, T é diagonalizável. □

Exemplo 7.9. Considere o operador linear sobre $\mathbb{R}^5$ definido por

$$[T]_{can}^{can} = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

O polinômio característico é

$$p(\lambda) = (2-\lambda)^3(1-\lambda)^2,$$

e os candidatos à polinômio minimal são

- $p_1(x) = (2-x)(1-x) \quad$ (grau 2)
- $p_2(x) = (2-x)^2(1-x) \quad$ (grau 3)
- $p_3(x) = (2-x)(1-x)^2 \quad$ (grau 3)
- $p_4(x) = (2-x)^3(1-x) \quad$ (grau 4)
- $p_5(x) = (2-x)^2(1-x)^2 \quad$ (grau 4)
- $p_6(x) = p(x) = (2-x)^3(1-x)^2 \quad$ (grau 5).

Você pode verificar que

$$p_1([T]_{can}^{can}) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \neq \mathbf{0},$$

e logo T **não** é diagonalizável. Se calcularmos $p_2([T]_{can}^{can}), \ldots, p_5([T]_{can}^{can})$ em sequência, veremos que o primeiro polinômio que anula $[T]_{can}^{can}$ é $p_3(x)$. Ou seja, $p_3(x)$ é o polinômio minimal $m(x)$. $\square$

Exemplo 7.10. Considere o operador definido pela *matriz simétrica*

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

que é uma pequena modificação do exemplo anterior. Calculando o polinômio característico, temos

$$
\begin{aligned}
p(\lambda) = \det\left([T]_{can}^{can} - \lambda \mathbf{I}_3\right) &= \det \begin{bmatrix} 2-\lambda & 0 & 1 & 0 & 0 \\ 0 & 2-\lambda & 0 & 0 & 0 \\ 1 & 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\
&= (2-\lambda)(-1)^{1+1} \det \begin{bmatrix} 2-\lambda & 0 & 0 & 0 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\
&\quad + (-1)^{1+3} \det \begin{bmatrix} 0 & 2-\lambda & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\
&= (2-\lambda)^3(1-\lambda)^2 - (2-\lambda)(1-\lambda)^2 \\
&= (2-\lambda)(1-\lambda)^2\left[(2-\lambda)^2 - 1\right] \\
&= (2-\lambda)(1-\lambda)^2(\lambda^2 - 4\lambda + 3)
\end{aligned}
$$

. As raízes da equação do segundo grau $\lambda^2 - 4\lambda + 3 = 0$ são 1 e 3. Portanto podemos escrever

$$p(\lambda) = (2-\lambda)\,(1-\lambda)^3\,(3-\lambda).$$

O candidato à polinômio minimal que não possui raízes repetidas é

$$p_1(x) = (2-x)\,(1-x)\,(3-x),$$

e, fazendo as contas, podemos verificar que

$$p_1([T]_{can}^{can}) = \mathbf{0}.$$

Ou seja, $m(x) = p_1(x)$ é o polinômio minimal. Logo T é diagonalizável. □

Os dois exemplos anteriores se parecem. Porém, o operador do primeiro exemplo não é diagonalizável, enquanto o operador do segundo exemplo sim. Isso não ocorre à toa: veremos adiante que toda matriz simétrica é diagonalizável!

Finalmente, a mesma estratégia feita para operadores pode ser aplicada à matrizes. Veja que em todos os exemplos anteriores olhamos para as matrizes dos operadores.

Exemplo 7.11. Pergunta-se se a matriz

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

é diagonalizável. Seu polinômio característico é

$$p(\lambda) = \det\left(\mathbf{A} - \lambda \mathbf{I}_2\right) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2.$$

Os candidatos à polinômio minimal são

- $p_1(x) = (1 - x)$
- $p_2(x) = p(x) = (1 - x)^2$.

Veja que

$$p_1(\mathbf{A}) = \mathbf{I}_2 - \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \mathbf{0}.$$

Daí, $p_1(x)$ não é o polinômio minimal de $\mathbf{A}$. Ou seja, o termo de grau 1 em $m(x)$ não aparece uma única vez, e logo $\mathbf{A}$ **não** é diagonalizável. □

7.1.4 COMENTÁRIO EXTRA E NÃO OBRIGATÓRIO – Diagonalização e multiplicidade geométrica – uma visão geométrica

Existe uma relação entre diagonalização e as multiplicidades geométricas dos autovalores. Retomando, a multiplicidade geométrica de um autovalor λ do operador (ou matriz) $T : V \to V$ é a dimensão do autoespaço

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}.$$

Ora, dizer que T é diagonalizável é o mesmo que dizer que existe uma base de V formada por autovetores de T. É claro que tal base deve pertencer à união dos autoespaços; isto é, se $\lambda_1, \ldots, \lambda_r$ são os autovalores **distintos** de T, então uma base $\beta = \{v_1, \ldots, v_n\}$ de V formada por autovetores é tal que

$$\{v_1, \ldots, v_n\} \subset V_{\lambda_1} \cup V_{\lambda_2} \cup \cdots \cup V_{\lambda_r}.$$

Mas como a união $V_{\lambda_1} \cup V_{\lambda_2} \cup \cdots \cup V_{\lambda_r}$ está contida em V, devemos ter

$$V = V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_r}.$$

Do Teorema 7.1, esta soma é direta, isto é,

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

Então podemos dizer que

“o operador $T : V \to V$ é diagonalizável se, e somente se, podemos decompor V como soma direta dos autoespaços associados aos autovalores de T”

ou ainda,

“o operador $T : V \to V$ é diagonalizável se, e somente se, a soma das multiplicidades geométricas dos autovalores distintos de T é igual à dimensão do espaço V”.

7.2 Diagonalização de matrizes simétricas

Seja $\mathbf{A}$ uma matriz $n \times n$ simétrica, isto é, $\mathbf{A} = \mathbf{A}^t$. Considere o operador $T : \mathbb{R}^n \to \mathbb{R}^n$ definido por $[\,T\,]_{can}^{can} = \mathbf{A}$.

Operadores lineares sobre $\mathbb{R}^n$ cuja matriz na base canônica é simétrica são especiais, pois eles **sempre** são diagonalizáveis. Tais operadores são chamados *auto-adjuntos*. Não faremos o estudo de operadores auto-adjuntos aqui, pois isso depende de conceitos ainda não vistos. Nos restringiremos à linguagem das matrizes.

Lembre-se da Definição 7.2 que a matriz $\mathbf{A}$ de ordem $n \times n$ é diagonalizável se admite n autovetores LI. O primeiro fato sobre matrizes simétricas é que seu polinômio característico possui somente raízes reais.

Teorema 7.5. *Seja* $\mathbf{A}$ *matriz simétrica. Então o polinômio característico*

$$p(\lambda) = \det\big(\mathbf{A} - \lambda\mathbf{I}\big)$$

possui somente raízes reais.

O teorema anterior não vale em geral para uma matriz **não simétrica**. Por exemplo,

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

tem polinômio característico

$$p(\lambda) = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2,$$

e a equação $p(\lambda) = 0$ possui raízes complexas $1 \pm i$.

Vimos anteriormente que há matrizes/operadores diagonalizáveis e outros não. A descoberta se uma matriz é ou não diagonalizável foi feita pelo polinômio minimal. Ocorre que para matrizes *simétricas*, temos a certeza que são diagonalizáveis!

Teorema 7.6. *Toda matriz simétrica* $\mathbf{A}$ *é diagonalizável.*

7.3 Matrizes semelhantes

Dado um operador $T : \mathbb{R}^n \to \mathbb{R}^n$ e bases α e β de $\mathbb{R}^n$, sabemos que

$$[\,T\,]^{\beta}_{\beta} = [\,I\,]^{\alpha}_{\beta}\,[\,T\,]^{\alpha}_{\alpha}\,[\,I\,]^{\beta}_{\alpha} \quad\Rightarrow\quad [\,T\,]^{\beta}_{\beta} = \big([\,I\,]^{\beta}_{\alpha}\big)^{-1}\,[\,T\,]^{\alpha}_{\alpha}\,[\,I\,]^{\beta}_{\alpha}. \tag{7.1}$$

Chamando $\mathbf{B} = [\,T\,]^{\beta}_{\beta}$, $\mathbf{A} = [\,T\,]^{\alpha}_{\alpha}$ e $\mathbf{P} = [\,I\,]^{\beta}_{\alpha}$, temos

$$\mathbf{B} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}.$$

isso motiva a seguinte definição:

Definição 7.5. *Duas matrizes quadradas* $\mathbf{A}$ *e* $\mathbf{B}$ *de mesma ordem são* ***semelhantes*** *(ou* ***similares****) se existir uma matriz* $\mathbf{P}$ *inversível tal que*

$$\mathbf{B} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}.$$

Quando $\mathbf{A}$ é diagonalizável, é semelhante à uma matriz diagonal $\mathbf{D}$, pois neste caso $\mathbf{A}$ pode ser vista como a matriz de um operador diagonalizável $T : \mathbb{R}^n \to \mathbb{R}^n$, e $\mathbf{D}$ a matriz desse operador na base β que o diagonaliza.

Ou seja, no caso de $\mathbf{A}$ ser diagonalizável, podemos escrever

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P},$$

onde $\mathbf{D}$ é matriz diagonal. Mas quem são $\mathbf{D}$ e $\mathbf{P}$? Como calculá-las?

Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ o operador definido por $[\,T\,]^{can}_{can} = \mathbf{A}$. Da expressão (7.1), vemos que

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \qquad \text{e} \qquad \mathbf{P} = [\,I\,]^{\beta}_{can} = \begin{bmatrix} | & | & \cdots & | \\ [\,v_1\,]_{can} & [\,v_2\,]_{can} & \cdots & [\,v_n\,]_{can} \\ | & | & \cdots & | \end{bmatrix},$$

onde $\beta = \{v_1, \dots, v_n\}$ é uma base de autovetores, associados aos autovalores $\lambda_1, \dots, \lambda_n$. Isto é,

Se $\mathbf{A}$ é uma matriz quadrada diagonalizável, então

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P},$$

onde $\mathbf{D}$ é a matriz diagonal formada pelos autovalores de $\mathbf{A}$ e $\mathbf{P}$ é a matriz cujas colunas são os autovetores associados (na mesma ordem).

Em particular, a escrita $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ **sempre** é possível para matrizes **A simétricas**. Em outras palavras, qualquer matriz simétrica é semelhante à uma matriz diagonal.

Vale também a recíproca: se $\mathbf{A}$ é semelhante à uma matriz diagonal, digamos $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, então $\mathbf{D}$ é formada pelos autovalores e as colunas de $\mathbf{P}$ são os autovetores associados (veja o Exercício 6, Lista 7). Desta forma, toda matriz semelhante à uma diagonal é diagonalizável (isso dá ainda mais sentido ao termo "diagonalizável").

Exemplo 7.12. Considere a matriz simétrica

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

Sabemos que $\mathbf{A}$ é diagonalizável (Teorema 7.6). Vamos calcular $\mathbf{D}$ e $\mathbf{P}$ tais que $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

Autovalores de A: Temos

$$p(\lambda) = \det\big(\mathbf{A} - \lambda \mathbf{I}_4\big) = (-1-\lambda)(1-\lambda)^2(4-\lambda)$$

(verifique). Assim, os autovalores de $\mathbf{A}$ (com repetições) são $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 1$ e $\lambda_4 = 4$, e logo

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Autovetor associado à $\lambda_1 = -1$. Resolvendo $\mathbf{A}\mathbf{v} = -\mathbf{v}$, encontramos a solução particular

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Autovetores associados à $\lambda_2 = \lambda_3 = 1$. Resolvendo $\mathbf{A}\mathbf{v} = \mathbf{v}$, encontramos as soluções LI particulares

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Autovetor associado à $\lambda_4 = 4$. Resolvendo $\mathbf{Av} = 4\mathbf{v}$, encontramos a solução particular

$$\mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Assim, construímos **P seguindo a ordem dos autovalores/autovetores**:

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

Você pode verificar que, de fato,

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}.$$

□

7.4 Demonstrações

Demonstração do Teorema 7.1. Sejam $v_1, \ldots, v_r$ autovetores de T associados aos autovalores $\lambda_1, \ldots, \lambda_r$, distintos entre si. Considere a combinação linear

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathbf{0}. \tag{7.2}$$

Devemos mostrar que $a_i = 0$ para todo $i = 1, \ldots, r$. Para cada i, definimos a função $T - \lambda_i Id : V \to V$ por

$$(T - \lambda_i Id)(u) = T(u) - \lambda_i u.$$

É fácil mostrar que $T - \lambda_i Id$ é um operador linear sobre V. Veja que

$$(T - \lambda_i Id)(v_i) = T(v_i) - \lambda_i v_i = \mathbf{0}, \quad \forall i,$$

e que

$$(T - \lambda_i Id)(v_j) = T(v_j) - \lambda_i v_j = (\lambda_j - \lambda_i)v_j, \quad \forall i \neq j.$$

Assim, aplicando $T - \lambda_1 Id$ sobre (7.2) e usando sua linearidade, obtemos

$$\begin{aligned} a_1 \underbrace{(T - \lambda_1 Id)(v_1)}_{=\mathbf{0}} + a_2(T - \lambda_1 Id)(v_2) + \cdots + a_r(T - \lambda_1 Id)(v_r) &= \mathbf{0} \\ \Rightarrow \quad a_2(T - \lambda_1 Id)(v_2) + \cdots + a_r(T - \lambda_1 Id)(v_r) &= \mathbf{0} \\ \Rightarrow \quad a_2(\lambda_2 - \lambda_1)(v_2) + \cdots + a_r(\lambda_r - \lambda_1)(v_r) &= \mathbf{0} \end{aligned}$$

Da mesma forma, aplicando $T - \lambda_2 Id$ sobre a última equação acima obteremos

$$a_3(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)(v_3) + \cdots + a_r(\lambda_r - \lambda_2)(\lambda_r - \lambda_1)(v_r) = \mathbf{0}.$$

Aplicando $T - \lambda_3 Id, \ldots, T - \lambda_r Id$ sucessivamente obteremos

$$a_r\big[(\lambda_r - \lambda_{r-1}) \cdots (\lambda_r - \lambda_1)\big] = \mathbf{0}.$$

Como todos os autovalores são distintos entre si, concluímos que $a_r = 0$. Substituindo $a_r = 0$ na equação imediatamente anterior, obtida pela aplicação sucessiva dos operadores $T - \lambda_i Id$, concluímos também que $a_{r-1} = 0$. Continuando a substituir esses coeficientes nulos nas equações anteriores, concluiremos que $a_1 = a_2 = \cdots = a_r = 0$, e logo os vetores $v_1, \ldots, v_r$ são LI, como queríamos demonstrar. □

Demonstração do Corolário 7.1. Consequência direta do Teorema 7.1 com $r = n$. □

Demonstração do Teorema 7.2. A prova é demasiadamente sofisticada para este texto. Veja, por exemplo, o livro de Álgebra Linear de Hoffman e Kunze. □

Demonstração do Teorema 7.3. Consulte [2] para uma demonstração no caso 2×2. Uma demonstração geral requer conceitos sofisticados, e pode ser encontrada no livro de Álgebra Linear de Hoffman e Kunze. □

Demonstração do Teorema 7.4. Sejam $p(x)$ e $m(x)$ os polinômios característico e minimal de $[\,T\,]_\beta^\beta$, respectivamente. Devemos mostrar que $p(\lambda) = 0$ se, e somente se, $m(\lambda) = 0$.

Primeiro, suponha que $m(\lambda) = 0$. Então podemos colocar o termo $x - \lambda$, referente à raiz λ, em evidência, obtendo

$$m(x) = (x - \lambda)q(x),$$

onde $q(x)$ é um polinômio de grau menor do que o grau de $m(x)$. Note ainda que

$$m([\,T\,]_\beta^\beta) = ([\,T\,]_\beta^\beta - \lambda \mathbf{I})\, q([\,T\,]_\beta^\beta).$$

Pela definição de polinômio minimal, $m(x)$ é o polinômio de menor grau que anula $[\,T\,]_\beta^\beta$, e logo $q(x)$ não pode anular esta matriz, isto é,

$$q([\,T\,]_\beta^\beta) \neq \mathbf{0}.$$

Assim, podemos escolher uma matriz coluna $\mathbf{w}$ tal que $\mathbf{v} = q([\,T\,]_\beta^\beta)\,\mathbf{w} \neq \mathbf{0}$. Como $m([\,T\,]_\beta^\beta) = \mathbf{0}$, temos

$$([\,T\,]_\beta^\beta - \lambda \mathbf{I})\,\mathbf{v} = ([\,T\,]_\beta^\beta - \lambda \mathbf{I})\, q([\,T\,]_\beta^\beta)\,\mathbf{w} = \mathbf{0} = m([\,T\,]_\beta^\beta)\,\mathbf{w} = \mathbf{0},$$

ou seja, $\mathbf{v}$ é autovetor da matriz $[\,T\,]_\beta^\beta$ associado à λ. Portanto, λ é autovalor de $[\,T\,]_\beta^\beta$, e logo $p(\lambda) = 0$.

Reciprocamente, suponha que $p(\lambda) = 0$. Vamos mostrar que $m(\lambda) = 0$. Sendo λ autovalor de $[\,T\,]_\beta^\beta$, temos

$$[\,T\,]_\beta^\beta\,\mathbf{v} = \lambda \mathbf{v} \tag{7.3}$$

para alguma matriz coluna $\mathbf{v} \neq \mathbf{0}$. Escrevamos o polinômio minimal m na forma

$$m(x) = x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0.$$

Por simplicidade, chamemos $\mathbf{A} = [\,T\,]_\beta^\beta$. Multiplicando a equação (7.3) por $\mathbf{A} = [\,T\,]_\beta^\beta$ à esquerda sucessivas vezes, obtemos

$$\begin{aligned}
&\mathbf{A}\mathbf{v} = \lambda \mathbf{v} \\
\Rightarrow\ &\mathbf{A}^2\mathbf{v} = \lambda \mathbf{A}\mathbf{v} = \lambda^2 \mathbf{v} \\
\Rightarrow\ &\mathbf{A}^3\mathbf{v} = \lambda \mathbf{A}^2\mathbf{v} = \lambda^3 \mathbf{v} \\
&\quad\vdots \\
\Rightarrow\ &\mathbf{A}^r\mathbf{v} = \lambda^r \mathbf{v}.
\end{aligned}$$

Assim,

$$\begin{aligned} m(\mathbf{A})\,\mathbf{v} &= \big[\mathbf{A}^r + a_{r-1}\mathbf{A}^{r-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I}\big]\mathbf{v} \\ &= \mathbf{A}^r\mathbf{v} + a_{r-1}\mathbf{A}^{r-1}\mathbf{v} + \cdots + a_1\mathbf{A}\mathbf{v} + a_0\mathbf{I}\mathbf{v} \\ &= \lambda^r\mathbf{v} + a_{r-1}\lambda^{r-1}\mathbf{v} + \cdots + a_1\lambda\mathbf{v} + a_0\mathbf{v} \\ &= \big[\lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_1\lambda + a_0\big]\mathbf{v} = m(\lambda)\mathbf{v}. \end{aligned}$$

Isto é, $m([\,T\,]_\beta^\beta)\,\mathbf{v} = m(\lambda)\mathbf{v}$. Agora, como $m([\,T\,]_\beta^\beta) = \mathbf{0}$ e $\mathbf{v} \neq \mathbf{0}$, concluímos que $m(\lambda) = 0$, isto é, λ é raiz do polinômio minimal. Isso finaliza a demonstração. $\square$

Demonstração do Teorema 7.5. Olhando $\mathbf{A}$ como a matriz do operador linear $T : \mathbb{C}^n \to \mathbb{C}^n$ sobre o espaço vetorial das listas de números complexos $\mathbb{C}^n$, podemos interpretar as raízes de $p(\lambda)$ como autovalores associados à autovetores com componentes complexas $\mathbf{v} \in \mathbb{C}^n$. Assim, seja $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$, autovetor de $\mathbf{A}$ associado ao autovalor λ. Multiplicando $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ à esquerda pelo conjugado de $\mathbf{v}^t$, a matriz linha $\overline{\mathbf{v}}^t$, obtemos

$$\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v} = \lambda\,\overline{\mathbf{v}}^t\mathbf{v} \quad \Rightarrow \quad \lambda = \frac{\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v}}{\overline{\mathbf{v}}^t\mathbf{v}}.$$

Dado um número complexo qualquer $a + bi$, temos $(a + bi)(a - bi) = a^2 + b^2 \in \mathbb{R}$. Isto é, o produto de um número complexo com seu conjugado é um número real. Assim,

$$\overline{\mathbf{v}}^t\mathbf{v} = \begin{bmatrix} \overline{v}_1 & \cdots & \overline{v}_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \overline{v}_1 v_1 + \cdots + \overline{v}_n v_n \in \mathbb{R}.$$

Da mesma forma, é fácil verificar que dados dois números complexos quaisquer $a + bi$ e $c + di$, vale

$$\overline{(a+bi)}(c+di) + (a+bi)\overline{(c+di)} = 2(ac+bd) \in \mathbb{R}.$$

Pela simetria de $\mathbf{A}$, temos $a_{ij} = a_{ji}$ para todos i, j, e portanto podemos escrever

$$\begin{aligned} \overline{\mathbf{v}}^t\mathbf{A}\mathbf{v} &= \begin{bmatrix} \overline{v}_1 & \cdots & \overline{v}_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{v}_i v_j \\ &= 2\sum_{i=1}^{n} a_{ii}\overline{v}_i v_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}\big(\overline{v}_i v_j + v_i\overline{v}_j\big) \in \mathbb{R}. \end{aligned}$$

Assim, concluímos que $\lambda = \frac{\overline{\mathbf{v}}^t\mathbf{A}\mathbf{v}}{\overline{\mathbf{v}}^t\mathbf{v}} \in \mathbb{R}$, como queríamos demonstrar. $\square$

Demonstração do Teorema 7.6. A prova é feita por indução sobre n, a ordem da matriz simétrica $\mathbf{A}$. Para $n = 1$, a matriz $\mathbf{A}$ trivialmente admite 1 autovetor, que sozinho forma um conjunto LI. Suponha que $\mathbf{A}$ admita $n - 1$ ($n \geq 2$) autovetores LI e vamos mostrar que o teorema vale para n.

Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ autovetores LI associados à $\lambda_1, \ldots, \lambda_{n-1}$ respectivamente, e considere o conjunto

$$S^\perp = \{\mathbf{w} \in M(n,1) \mid \mathbf{v}_i^t \mathbf{w} = 0 \text{ para } i = 1, \ldots, n-1\}.$$

O sistema linear homogêneo

$$\begin{cases} \mathbf{v}_1^t \mathbf{w} &= & v_{11}w_1 &+ & \cdots &+ & v_{1n}w_n &= & 0 \\ \quad\vdots \\ \mathbf{v}_1^t \mathbf{w} &= & v_{(n-1)1}w_1 &+ & \cdots &+ & v_{(n-1)n}w_n &= & 0 \end{cases}$$

tem $n-1$ equações e n incógnitas (as entradas de $\mathbf{w}$). Logo, possui solução não trivial, isto é, $S^\perp \neq \{\mathbf{0}\}$.

Tomemos $\mathbf{v}_n \in S^\perp$ não nulo. Neste caso, afirmamos que $\mathbf{v}_n$ não é combinação linear dos outros $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. De fato, se fosse teríamos

$$\mathbf{v}_n = a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1} \tag{7.4}$$

e multiplicando a expressão acima à esquerda pela transposta de $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1}$, teríamos

$$\sum_{i=1}^{n} w_i^2 = \mathbf{w}^t\mathbf{w} = \big(a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1}\big)^t\mathbf{w} = a_1\mathbf{v}_1^t\mathbf{w} + \cdots + a_{n-1}\mathbf{v}_{n-1}^t\mathbf{w} = 0,$$

o que implica $\mathbf{w} = \mathbf{0}$. Mas como $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ são LI, isso implicaria $a_1 = \cdots = a_{n-1} = 0$, que por sua vez implicaria $\mathbf{v}_n = \mathbf{0}$, dado (7.4). Isso contraria o fato de $\mathbf{v}_n \neq \mathbf{0}$.

Agora, veja que

- o subespaço $[\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}]$ (subespaço gerado por $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$) tem dimensão $n-1$;
- o espaço das matrizes-coluna de tamanho n, $M(n,1)$, tem dimensão n ($M(n,1)$ é isomorfo à $\mathbb{R}^n$);
- pelo que foi dito anteriormente, $\mathbf{v}_n \notin [\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}]$, seja qual for $\mathbf{v}_n \in S^\perp$ não nulo. Logo,

  $$\dim S^\perp = \dim M(n,1) - \dim[\mathbf{v}_n] \cap [\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}] = n - (n-1) = 1,$$

  e assim $S^\perp = [\mathbf{v}_n]$ para um $\mathbf{v}_n \neq \mathbf{0}$ fixado.

Agora, veja que

$$(\mathbf{A}\mathbf{v}_n)^t\mathbf{v}_i = \mathbf{v}_n^t\mathbf{A}^t\mathbf{v}_i = \mathbf{v}_n^t\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_n^t\mathbf{v}_i = 0$$

para todo $i = 1, \ldots, n-1$. Isto é, $\mathbf{A}\mathbf{v}_n \in S^\perp$. Como $S^\perp = [\mathbf{v}_n]$, devemos ter $\mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n$ para algum $\lambda_n \in \mathbb{R}$. Assim, construímos n autovetores $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n$ de $\mathbf{A}$ que são LI. Pelo principio da indução, a demonstração está concluída. □

7.5 Exercícios

Veja a lista de exercícios 8.

Referências Bibliográficas

[1] Howard Anton e Chris Rorres. *Álgebra Linear com aplicações*. Bookman, 10 edition, 2010.

[2] José Luiz Boldrini e outros. *Álgebra Linear*. Harper & Row do Brasil, São Paulo, 3 edition, 1980.

[3] Alfredo Steinbruch e Paulo Winterle. *Álgebra Linear*. Pearson, São Paulo, 2 edition, 1987.

[4] David Lay. *Álgebra Linear*. LTC, Rio de Janeiro, 2 edition, 1999.

[5] David Poole. *Álgebra linear*. Thonsom Learning, São Paulo, 2006.