Funções de Matrizes

(Versão Preliminar)

Hamilton Prado Bueno

Universidade Federal de Minas Gerais
Departamento de Matemática

A
Eliana Farias Bueno.
Inesquecível esposa,
eterna amiga.

Introdução

Funções $f : U \subset \mathbb{C} \to \mathbb{C}$ podem gerar funções de matrizes quadradas, chamadas **funções de matrizes**. Por exemplo, as matrizes A^k, A^{-1}, $\operatorname{sen} A$ e o fluxo e^{At}. Essas funções constituem um tópico essencial da Álgebra Linear, mas não estão presentes nas apresentações tradicionais do assunto.

Restringe-se a apresentação de funções de matrizes ao caso de polinômios de uma matriz quadrada ou então, quando a matriz A é simétrica e $A = P^{-1}DP$, com D diagonal, define-se $f(A)$ por $P^{-1}f(D)P$, em que $f(D)$ é obtida ao se avaliar f em cada uma das entradas diagonais de D.

Devido a sua enorme importância no estudo de equações diferenciais, e^{At} é usualmente definida por meio da expansão em série de potências, o que exige o conceito de convergência uniforme e, conseqüentemente, torna a exposição acessível apenas para alunos mais avançados. Além disso, e^{At} é obtida apenas em alguns casos muito simples (em geral, quando a matriz A está na Forma Canônica de Jordan) e o estudante fica com a impressão que e^{At} é uma solução "teórica" para o sistema $x' = Ax$. O fato que e^{At} é um polinômio (com coeficientes dependendo de t) na matriz A não é enfatizado.

Nossa exposição de funções de matrizes pode ser sintetizada como uma generalização da versão em dimensão finita do cálculo funcional de Dunford-Schwarz [9] e já era conhecida por Gantmacher [10]. Ela é simples e tem conseqüências notáveis: $f(A)$ é sempre um polinômio na matriz A (com coeficientes dependendo da função f), que pode ser facilmente obtido se são conhecidos os autovalores com multiplicidade de A. Essa abordagem, uma técnica corriqueira na Álgebra Linear Numérica, tem sido esquecida nos textos de Álgebra Linear. Livros bem reputados (veja [11], [12], [13], [20]) ou mesmo tratados mais avançados (veja [3] ou [14]) nem mesmo a mencionam. Com esse texto queremos contribuir para uma reavaliação do cálculo funcional na Álgebra Linear básica.

Descreveremos sucintamente o seu conteúdo. O primeiro capítulo apresenta resultados básicos sobre os polinômios característico e mínimo de um operador linear num espaço de dimensão finita. O capítulo 2 não passa de uma observação sobre a divisão de uma função por um polinômio. Mostraremos que, sob hipóteses naturais, tal divisão é euclidiana. (Esse enfoque, em última instância, conduz aos Teoremas de Preparação de Weierstrass e Malgrange, que estão muito além do propósito desse texto). O cálculo funcional, que é uma conseqüência desse fato, é estudado no terceiro capítulo. Lá é mostrado que toda função de uma matriz é um polinômio na matriz. O capítulo 4 apresenta vários exemplos de utilização do cálculo funcional, enquanto o capítulo 5 é dedicado às demonstrações (elementares) do Teoremas da Imagem do Espectro e Espectral por meio do cálculo funcional. (O Teorema Espectral é a versão no corpo $\mathbb{C}$ do Teorema da Decomposição Primária). O capítulo 6 mostra como se complexifica um espaço vetorial real e, com isso, apresenta o Teorema da Decomposição Primária.

A leitura desse texto pressupõe alguns conhecimentos simples de propriedades dos números complexos (manipulação de números complexos, fórmula de Euler e o teorema fundamental da Álgebra - alguns tópicos um pouco mais avançados são abordados em um exercício) e um curso introdutório de Álgebra Linear, no qual sejam abordadas as noções de espaço vetorial, aplicação linear e suas representações matriciais e somas diretas. Até mesmo a demonstração do Teorema de Cayley-Hamilton consta do texto. Esses tópicos básicos de Álgebra Linear tem uma excelente apresentação no livro "Geometria Analítica e Álgebra Linear", Partes I e II, de Reginaldo dos Santos, cuja leitura não podemos deixar de recomendar. Fazemos também uso de algumas noções básicas da Análise Matemática: norma, convergência uniforme (apenas na apresentação da definição tradicional do fluxo e^{At}) e, em uma seção, da definição da topologia C^k num compacto.

Agradecimentos. Esse texto é uma variação sobre um artigo aceito para publicação na revista Cubo. Ele difere daquele pela inclusão de pré-requisitos e por um maior detalhamento, além da eliminação de uma de suas seções. Esse texto, como aquele, contou com a participação de vários amigos: H. Rodrigues, M. Spira, E. Bueno, G. Svetlichny e C. Tomei. Não há como agradecer a contribuição dada.

Belo Horizonte, 20 de setembro de 2002.

Hamilton Prado Bueno

Sumário

Capítulo 1

Polinômios de matrizes

Denotaremos por V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$. Para simplificarmos a nossa exposição, $\mathbb{K}$ será o corpo dos complexos $\mathbb{C}$ ou o corpo dos reais $\mathbb{R}$.

Seja $T : V \to V$ uma aplicação linear. Escolhendo uma base $\mathcal{B}$ para V, obtemos a matriz $T_{\mathcal{B}}$, representação de T na base $\mathcal{B}$.

Exemplo 1.1 Seja $R : \mathbb{R}^2 \to \mathbb{R}^2$ a aplicação que roda em torno da origem por um ângulo $0 < \theta < 2\pi$ um ponto do $\mathbb{R}^2$, no sentido anti-horário. É claro que o único ponto fixo por R é a origem. A linearidade de R é geometricamente clara **(incluir figura)**. Escolhendo a base canônica $\mathcal{E} = \{e_1, e_2\}$, com $e_1 = (1\ 0)^T$ (estamos denotando assim a transposta) e $e_2 = (0\ 1)^T$, encontramos $T_{\mathcal{E}}$: se $Re_1 = P$, o ponto P tem coordenadas $(\cos\theta\ \operatorname{sen}\theta)^T$, de acordo com a própria definição das funções seno e cosseno. Do mesmo modo, se $Re_2 = Q$, as coordenadas de Q são $(\cos(\theta + \pi/2)\ \operatorname{sen}(\theta + \pi/2) = (-\operatorname{sen}\theta\ \ \cos\theta)^T$. Logo, a representação de R na base $\mathcal{E}$ é

$$A = R_{\mathcal{E}} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

$\diamond$

Seja $q(z) = a_k z^k + a_{k-1} z^{k-1} + a_1 z + a_0$ um polinômio com coeficientes no corpo $\mathbb{K}$. Claramente faz sentido calcular

$$q(T) := a_k T^k + a_{k-1} T^{k-1} + \cdots + a_1 T + a_0 I.$$

(Denotamos por I a aplicação identidade $I : V \to V$). Note que $q(T)$ é uma aplicação linear de V em V, que é representada por uma matriz $n \times n$ ao se escolher uma base de V.

1.1 O polinômio mínimo

Lembramos que um polinômio é **mônico** se o coeficiente de seu termo de maior grau for igual a 1.

Definição 1.2 *O* **polinômio mínimo** *m de uma aplicação $T : V \to V$ é o polinômio mônico de menor grau tal que $m(T) = 0$.*

Lema 1.3 *Se V é um espaço vetorial de dimensão n, toda aplicação linear $T : V \to V$ possui um polinômio mínimo.*

Demonstração: O espaço $\mathcal{L}(V, V)$ de todas as aplicações lineares $T : V \to V$ é um espaço vetorial de dimensão n^2. (Esse espaço é isomorfo ao espaço $\mathbb{M}_{n\times n}(\mathbb{K})$ de todas as matrizes $n \times n$ com entradas em $\mathbb{K}$). Assim, as aplicações lineares I, T, $T^2, \ldots, T^{n^2}$ são, necessariamente, linearmente dependentes. Quer dizer, existem escalares $a_0, a_1, \ldots, a_{n^2} \in \mathbb{K}$, nem todos nulos, tais que

$$a_0 I + a_1 T + \ldots + a_{n^2} T^{n^2} = 0.$$

Definindo $p(z) = a_0 + a_1 z + \ldots + a_{n^2} z^{n^2}$, temos $0 \neq p$ e $p(T) = 0$. Dividindo pelo coeficiente do termo de maior grau, obtemos um polinômio mônico p. O polinômio mínimo então existe, como decorrência da aplicação do Princípio da Boa Ordenação ao conjunto de todos os polinômios mônicos que anulam T. $\square$

Lema 1.4 *Se p é um polinômio tal que $p(T) = 0$, então p é um múltiplo de m.*

Demonstração: Se $\mathcal{I}$ denota o conjunto de todos os polinômios p (com coeficientes em $\mathbb{K}$) tais que $p(T) = 0$, claramente a soma de dois polinômios em $\mathcal{I}$, bem como a multiplicação de p por qualquer polinômio (com coeficientes em $\mathbb{K}$) estão em $\mathcal{I}$. (Quer dizer, $\mathcal{I}$ é um ideal.) A divisão euclidiana de p por m nos dá $p = qm + r$. Como $r = p - qm$ pertence a $\mathcal{I}$ e o grau de m é mínimo, concluímos que $r = 0$. $\square$

1.2 O teorema de Cayley-Hamilton

Recordamos a definição do polinômio característico:

Definição 1.5 *Seja $A = T_{\mathcal{B}}$. O* **polinômio característico** *da matriz A é o polinômio*

$$p(z) = \det(zI - A).$$

É fácil verificar que $p(z)$ é um polinômio mônico de grau n. Se $\mathcal{C}$ é uma outra base de V e a matriz $B = T_{\mathcal{C}}$ é a representação de T na base $\mathcal{C}$, então $A = P^{-1}BP$, sendo $P = P_{\mathcal{B}}^{\mathcal{C}}$ a matriz de mudança da base $\mathcal{B}$ para a base $\mathcal{C}$. Temos que

$$\begin{aligned}
\det(zI - A) &= \det(P^{-1}(zI - B)P) \\
&= \det P^{-1} \det(zI - B) \det P = \det(zI - B) \det(P^{-1}P) \\
&= \det(zI - B),
\end{aligned}$$

mostrando que qualquer representação de T numa base de V possui o mesmo polinômio característico. Em conseqüência, podemos definir:

Definição 1.6 *O polinômio característico da aplicação linear $T : V \to V$ é o polinômio característico de qualquer uma de suas representações matriciais.*

Teorema 1.7 (Cayley-Hamilton)
Se p é o polinômio característico de $T : V \to V$, então $p(T) = 0$.

Demonstração: Seja $0 \neq v \in V$ arbitrário. Queremos mostrar que $p(T)v = 0$. Seja m o maior natural tal que o conjunto

$$S = \{v, Tv, \ldots, T^{m-1}v\}$$

é linearmente independente. Então

$$T^m v = \alpha_0 v + \ldots + \alpha_{m-1} T^{m-1} v. \tag{1.1}$$

Seja $W = < S >$ o subespaço gerado por S. Então os elementos de S formam uma base de W. Afirmamos que $T(W) \subset W$. De fato, se $w \in W$, então $w = \beta_0 v + \beta_1 Tv + \ldots + \beta_{m-1} T^{m-1} v$, para escalares $\beta_0, \ldots, \beta_m \in \mathbb{K}$. Assim,

$$Tw = \beta_0 Tv + \beta_1 T^2 v + \ldots + \beta_{m-1} T^m v.$$

A igualdade (1.1) garante o afirmado.

Seja T_i a restrição de T ao subespaço W. A representação de T_i na base S é

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \cdots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{m-1} \end{pmatrix}.$$

Logo,

$$\begin{aligned}
\det(zI - A) &= \det \begin{pmatrix} z & 0 & \cdots & 0 & -\alpha_0 \\ -1 & z & \cdots & 0 & -\alpha_1 \\ 0 & -1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z - \alpha_{m-1} \end{pmatrix} \\
&= z \det \begin{pmatrix} z & 0 & \cdots & -\alpha_1 \\ -1 & z & \cdots & -\alpha_2 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & -1 & z - \alpha_{m-1} \end{pmatrix} + \\
& \qquad (-\alpha_0)(-1)^{m+1} \det \begin{pmatrix} -1 & z & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}.
\end{aligned}$$

Como o determinante da última matriz é $(-1)^{m-1}$, o último termo é justamente $-\alpha_0$. Procedendo do mesmo modo, obtemos

$$\det(zI - A) = z^m - \alpha_{m-1} z^{m-1} - \ldots - \alpha_0 = p_W(z),$$

sendo $p_W(z)$ o polinômio característico de T restrito a W. A equação (1.1) nos mostra então que $p_W(T)v = 0$.

Afirmamos agora que $p(z) = q(z)p_W(z)$, para algum polinômio $q(z)$. Daí decorre o resultado, pois $v \neq 0$ foi escolhido arbitrariamente e $p(T)v = q(T)p_W(T)v = 0$. Para provarmos a afirmação, notamos que ao completarmos S de forma a obter uma base $\mathcal{B}$ de V, a representação de T nessa base é

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}.$$

O resultado então decorre de propriedades do determinante, pois

$$\det(zI - T) = \det(zI - A)\det(zI - C) = p_W(z)q(z)$$

(em cada expressão, as ordens das matrizes I são diferentes). □

Relembramos:

Definição 1.8 *Seja $p(z)$ o polinômio característico da aplicação linear $T : V \to V$. As raízes $\lambda \in \mathbb{K}$ de $p(z)$ são chamadas* **autovalores** *de T, enquanto as soluções $v \neq 0$ de $Tv = \lambda v$ são chamados* **autovetores** *de T (associados a λ). O conjunto de todos os autovalores de T é chamado* **espectro** *de T e denotado por $\sigma(T)$.*

Observação 1.9 O polinômio característico é especialmente importante por causa de suas raízes. Por isso, também é comum chamar $\det(T - zI) = (-1)^n \det(zI - T)$ de polinômio característico de T. ◀

Exemplo 1.10 O polinômio característico da aplicação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (-y, x)$ é $p(z) = z^2 + 1$ e não possui raízes reais. Portanto, T não possui autovalores. Considerando $T : \mathbb{C}^2 \to \mathbb{C}^2$ definida da mesma maneira, $p(z) = z^2 + 1 = (z - i)(z + i)$ e T possui dois autovalores distintos. Isso mostra que a análise do espectro de uma aplicação linear $T : V \to V$ depende muito do corpo $\mathbb{K}$ sobre o qual V é espaço vetorial. Assim, estudar uma aplicação linear $T : V \to V$ é mais simples quando V é um espaço vetorial sobre $\mathbb{C}$ do que sobre $\mathbb{R}$: o Teorema Fundamental da Álgebra nos garante que o polinômio característico $p(z)$ possui n raízes (não necessariamente distintas) no corpo $\mathbb{C}$. ◇

Observação 1.11 Seja V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$. Dada uma aplicação linear $T : V \to V$, sejam p seu polinômio característico e m seu polinômio mínimo. Combinando o teorema de Cayley-Hamilton 1.7 com o lema 1.4, vemos que p é um múltiplo de m. Posteriormente mostraremos que p possui exatamente os mesmos fatores irredutíveis de m (veja a observação 5.4). Além disso, T é diagonalizável se, e somente se, $m(z)$ é produto de fatores lineares distintos (veja corolário 5.6). ◀

1.3 Decomposição de operadores

Seja V um espaço vetorial. Suponhamos que

$$V = W_1 \oplus \cdots \oplus W_\ell, \tag{1.2}$$

isto é, que cada ponto $x \in V$ tenha uma única representação

$$x = x_1 + \ldots + x_\ell, \quad x_j \in W_j, \quad j = 1, \ldots, \ell.$$

Para $j = 1, \ldots, \ell$, definimos as **projeções canônicas**

$$\begin{array}{rrl} \pi_j : & V \to & W_j \subset V \\ & x \mapsto & x_j. \end{array}$$

Claramente vale

$$\pi_j \pi_i = \delta_{ji} \pi_j, \tag{1.3}$$

em que $\delta_{ij} = 0$, se $i \neq j$, e $\delta_{ii} = 1$, com $i, j \in \{1, \ldots, \ell\}$. Além disso,

$$\sum_{j=1}^{\ell} \pi_j = I. \tag{1.4}$$

Reciprocamente, se os operadores lineares $\pi_1, \ldots, \pi_\ell$ satisfazem (1.3) e (1.4), definindo $W_j = \pi_j(V)$, temos que (1.2) se verifica e que π_j são as projeções canônicas dessa decomposição.

Definição 1.12 *Suponhamos que*

$$V = W_1 \oplus \cdots \oplus W_\ell$$

e que $T : V \to V$ satisfaça $T(W_j) \subset W_j$ para $j = 1, \ldots, \ell$. Dizemos então que os subespaços W_j são **invariantes** *pelo operador linear $T \in \mathcal{L}(V, V)$. Definimos então os* **blocos** *T_j de T por $T_j = T|_{W_j} : W_j \to W_j$.*

Proposição 1.13 *Suponhamos que $T \in \mathcal{L}(V, V)$ e $V = W_1 \oplus \cdots \oplus W_\ell$, com projeções correspondentes π_j, $j = 1, \ldots, \ell$. Então $T(W_j) \subset W_j$ se, e somente se,*

$$T\pi_j = \pi_j T.$$

Demonstração: Suponhamos que $T(W_j) \subset W_j$. Tome $x \in V$ arbitrário. Então $\pi_j x \in W_j$ e, conseqüentemente, $T\pi_j x \in W_j$. Logo $\pi_i T \pi_j x = \delta_{ij} T \pi_j x$ para todo $j = 1, \ldots, \ell$. Somando todos esses termos e utilizando (1.4), obtemos

$$T\pi_i x = \sum_{j=1}^{\ell} \delta_{ij} T \pi_j x = \sum_{j=1}^{\ell} \pi_i T \pi_j x = \pi_i T \left(\sum_{j=1}^{\ell} \pi_j x \right) = \pi_i T x.$$

Reciprocamente, se T comuta com todo π_j, para todo $x \in W_j$ vale

$$Tx = T\pi_j x = \pi_j T x \in W_j,$$

mostrando que $T(W_j) \subset W_j$.

□

1.4 Um homomorfismo de álgebras

Definição 1.14 *Uma* **álgebra** *$\mathcal{A}$ sobre o corpo $\mathbb{K}$ é um espaço vetorial sobre o corpo $\mathbb{K}$ que possui, adicionalmente, uma multiplicação satisfazendo às seguintes propriedades, para todos $u, v, w \in \mathcal{A}$ e $k \in \mathbb{K}$:*

(*i*) $(uv)w = u(vw)$ (*associatividade*);

(*ii*) $u(v + w) = uv + uw$ (*distributividade*);

(*iii*) $k(uv) = (ku)v = u(kv)$.

Se existir um elemento $e \in \mathcal{A}$ tal que $eu = ue = u$ para todo $u \in \mathcal{A}$, a álgebra $\mathcal{A}$ possui uma **unidade**. *Se $uv = vu$ para todos $u, v \in \mathcal{A}$, temos uma álgebra* **comutativa**.

Exemplo 1.15 O espaço vetorial $\mathcal{P}$ de todos os polinômios com coeficientes em $\mathbb{K}$ é uma álgebra comutativa com unidade. O espaço vetorial $\mathbb{M}_{n\times n}(\mathbb{K})$ é uma álgebra (não-comutativa) com unidade. O espaço $\mathcal{L}(V, V)$ é uma álgebra. Escolhida uma base de V, essa álgebra pode ser identificada com $\mathbb{M}_{n\times n}(\mathbb{K})$. Fixado $T \in \mathcal{L}(V, V)$, seja $\mathcal{P}(T)$ o conjunto de todas as aplicações lineares obtidas ao se avaliar o polinômio $p \in \mathcal{P}$ em $T \in \mathcal{L}(V, V)$:

$$T \mapsto p(T) \in \mathcal{L}(V, V).$$

É fácil verificar que $\mathcal{P}(T)$ é uma sub-álgebra de $\mathcal{L}(V, V)$. $\diamond$

Consideremos agora as álgebras $\mathcal{P}$ e $\mathcal{P}(T)$, definidas no exemplo anterior. A aplicação

$$\begin{array}{rccc} \phi: & \mathcal{P} & \to & \mathcal{P}(T) \\ & p & \mapsto & p(T) \end{array}$$

é uma aplicação **linear** que satisfaz, adicionalmente,

$$\phi(pq) = pq(T) = p(T)q(T) = \phi(p)\phi(q).$$

(A segunda igualdade é de verificação imediata). A aplicação ϕ é um **homomorfismo de álgebras**.

O núcleo de ϕ é o conjunto de múltiplos do polinômio mínimo m de T. A divisão euclidiana do polinômio p por m mostra que $\mathcal{P}(T)$ é constituída de polinômios em T com grau menor do que o do polinômio mínimo. (Estamos considerando que o grau do polinômio identicamente nulo é $-\infty$). Por definição, o homomorfismo ϕ é sobrejetivo.

1.5 Exercícios

Seja V um espaço vetorial de dimensão n sobre o corpo $\mathbb{K}$.

1. Seja $T : V \to V$ um operador linear. Mostre que o polinômio característico de T é um polinômio mônico de grau n com coeficientes em $\mathbb{K}$.

2. Suponha que o polinômio $p(z)$ seja da forma $(z-\lambda)^d q(z)$, com $q(\lambda) \neq 0$ e $d \in \{2, 3, \ldots\}$. Mostre que $p'(\lambda) = \ldots = p^{(d-1)}(\lambda) = 0$, mas $p^{(d)}(\lambda) \neq 0$. Dizemos então que a raiz λ de $p(z)$ tem **multiplicidade algébrica** d.

3. Sejam $T : V \to V$ uma aplicação linear, m o polinômio mínimo de T, A e B duas representações matriciais de T. Mostre que os polinômios mínimos de A e B são iguais a m.

4. Sejam $T : V \to V$ um operador linear e p um polinômio com coeficientes em $\mathbb{K}$. Mostre que se λ é um autovalor de T, então $p(\lambda)$ é um autovalor de $p(T)$.

5. Sejam $T : V \to V$ um operador linear e p um polinômio com coeficientes em $\mathbb{K}$. Mostre que se $\mathbb{K} = \mathbb{C}$ e μ é um autovalor de $p(T)$, então existe um autovalor λ de T tal que $\mu = p(\lambda)$. Dê um exemplo mostrando que esse resultado não é válido se $\mathbb{K} = \mathbb{R}$.

6. Compare a demonstração do Teorema de Cayley-Hamilton aqui apresentada com aquela em [6].

7. Seja $T : V \to V$ um operador linear. Se $V = W_1 \oplus \cdots \oplus W_\ell$ e T satisfaz $T(W_i) \subset W_i$, mostre que T pode ser representada por uma **matriz diagonal em blocos**

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell \end{pmatrix},$$

em que a A_i é uma matriz $j_i \times j_i$, sendo $j_i = \dim W_i$.

Reciprocamente, se a matriz diagonal em blocos A é a representação de um operador linear $T : V \to V$, mostre que existem espaços $W_i \subset V$ tais que $V = W_1 \oplus \cdots \oplus W_\ell$ e $T(W_i) \subset W_i$.

8. Seja A uma matriz diagonal em blocos:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell \end{pmatrix}.$$

Mostre que

$$A^k = \begin{pmatrix} A_1^k & 0 & \cdots & 0 \\ 0 & A_2^k & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell^k \end{pmatrix}.$$

9. Seja $V = W_1 \oplus W_2$, sendo W_1, W_2 invariantes pelo operador $T : V \to V$. Se r e s são os polinômios mínimos de $T|_{W_1}$ e $T|_{W_2}$, respectivamente, mostre que o polinômio mínimo de T é m.m.c.(r , s), o mínimo múltiplo comum dos polinômios r e s.

10. Uma aplicação linear $T : V \to V$ é **diagonalizável** T pode ser representada por uma matriz **diagonal**, isto é, por uma matriz diagonal em blocos cujos blocos diagonais são todos submatrizes 1×1. Mostre que um operador $T : V \to V$ é diagonalizável se, e somente se, V possui uma base formada por autovetores de T.

11. Sejam $p, q \in \mathcal{P}$ polinômios com coeficientes em $\mathbb{K}$ e $T : V \to V$ uma aplicação linear. Mostre que $(pq)(T) = p(T)q(T)$.

Capítulo 2

Divisão euclidiana

2.1 Funções euclidianas

Definição 2.1 *Uma função* $f : U \subset \mathbb{C} \to \mathbb{C}$ *(ou* $f : I \subset \mathbb{R} \to \mathbb{R}$*) é* **euclidiana** *com relação ao polinômio* p *se*

(i) todas as raízes de p *pertencem a* U *(resp., a* I*);*

(ii) se z_0 *é uma raiz de* p *com multiplicidade* k*, então* f *tem derivadas até a ordem* k *em* z_0*.*

Note que se U for um aberto e f analítica em U, a condição (ii) se verifica imediatamente.

A terminologia utilizada na definição acima é motivada pelo seguinte resultado, válido tanto para funções definidas em $I \subset \mathbb{R}$ como em $U \subset \mathbb{C}$. Como antes, definimos o grau do polinômio identicamente nulo como $-\infty$.

Proposição 2.2 *Seja* f *euclidiana com relação ao polinômio* p*. Então existem uma função* q*, contínua em cada uma das raízes do polinômio* p*, e um polinômio* r *tais que* $f = qp + r$*,* $\operatorname{gr} r < \operatorname{gr} p$*.*

Demonstração: Seja r um polinômio arbitrário. Consideremos a função q definida (nos pontos do domínio de f que não são raízes de p) por

$$q = \frac{f - r}{p}.$$

Queremos mostrar que podemos escolher r com grau menor do que o de p, de modo que q possua extensão contínua em cada uma das raízes de p. Notamos que q é tão suave quanto f em cada ponto z que não é uma raiz de p.

Seja z_0 uma raiz de multiplicidade k do polinômio p, isto é,

$$p(z) = (z - z_0)^k s(z),$$

sendo s um polinômio tal que $s(z_0) \neq 0$. Queremos achar r de modo que o quociente

$$\frac{f(z) - r(z)}{(z - z_0)^k}$$

possua extensão contínua em z_0. De acordo com a regra de L'Hospital, isso acontece quando

$$f(z_0) = r(z_0), \quad f'(z_0) = r'(z_0), \quad \ldots \quad , \quad f^{(k-1)}(z_0) = r^{(k-1)}(z_0).$$

Basta, portanto, mostrar que existe um polinômio $r(z)$ com grau menor do que o de p, satisfazendo relações como essas em cada raiz z_0 do polinômio p. A existência de tal polinômio será mostrada no lema abaixo. □

2.2 O polinômio interpolador

Denotaremos $f^{(0)} = f$.

Lema 2.3 *Suponhamos conhecidos os valores*

$$\begin{array}{ccccc} f(z_1) & f'(z_1) & \cdots & \cdots & f^{(d_1-1)}(z_1) \\ \vdots & \cdots & \cdots & \vdots & \\ f(z_\ell) & f'(z_\ell) & \cdots & f^{(d_\ell-1)}(z_\ell) & \end{array}$$

em que $z_1, \ldots, z_\ell$ são distintos. Denotemos $n := d_1 + d_2 + \ldots + d_\ell$. Então existe um único polinômio r, de grau menor ou igual a $n - 1$, satisfazendo

$$r^{(k)}(z_i) = f^{(k)}(z_i)$$

para todo $i = 1, \ldots, \ell$ e $k = 0, \ldots, d_i$.

Exemplo 2.4 Antes de passarmos ao caso geral, vejamos num exemplo a demonstração do lema 2.3. Suponhamos conhecidos os valores $f(z_1)$, $f(z_2)$ e $f'(z_2)$. Queremos encontrar um polinômio de grau 2 tal que $r(z_1) = f(z_1)$, $r(z_2) = f(z_2)$ e $r'(z_2) = f'(z_2)$. Seja $r(z) = az^2 + bz + c$. Então os coeficientes de r devem satisfazer ao sistema matricial:

$$\begin{pmatrix} z_0^2 & z_0 & 1 \\ z_1^2 & z_1 & 1 \\ 2z_1 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} f(z_0) \\ f(z_1) \\ f'(z_1) \end{pmatrix}. \tag{2.1}$$

Se os valores $f(z_0)$, $f(z_1)$ e $f'(z_1)$ são nulos, basta tomas $r \equiv 0$. A unicidade de r, nesse caso, é conseqüência do argumento apresentado a seguir.

Suponhamos que o sistema (2.1) não possua solução única. Então o sistema homogêneo associado possui uma solução não-trivial $(a_0\ b_0\ c_0)^T$. Consideremos o polinômio não-nulo

$$t(z) = a_0 z^2 + b_0 z + c_0.$$

Então $t(z)$ tem raízes z_0 e z_1, a segunda tendo multiplicidade 2 (já que é raiz da derivada de t). Mas isso implica que $t(z)$ é um múltiplo de $(z-z_0)(z-z_1)^2$ e tem grau maior ou igual a 3, o que é um absurdo. Logo (2.1) tem solução única para quaisquer valores $f(z_0)$, $f(z_1)$ e $f'(z_1)$. $\diamond$

Demonstração: Podemos supor que um dos valores dados seja não-nulo. O polinômio r procurado satisfaz a um sistema linear não-homogêneo, que pode ser escrito matricialmente como

$$Bz = b,$$

sendo z o vetor que tem como coordenadas os coeficientes procurados de r, b um vetor cujas n coordenadas são os valores conhecidos de f e B a matriz $n \times n$ do sistema linear assim formado.

Se B não tem inversa, o sistema homogêneo associado tem solução não trivial

$$z_0 = (a_0\ \ \ldots\ \ a_{n-1})^T.$$

Consideremos o polinômio

$$t(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1},$$

que é um polinômio de grau menor ou igual a $n-1$. Como z_0 satisfaz ao sistema homogêneo associado, temos que $t(z)$ deve ser um múltiplo de

$$(z - z_1)^{d_1} \cdots (z - z_\ell)^{d_\ell},$$

o que é um absurdo, pois o último polinômio tem grau n. Assim, B possui inversa e o sistema $Bz = b$ solução única. $\square$

O polinômio r é chamado **polinômio interpolador**.

Apresentamos agora uma conseqüência da Proposição 2.2 que está ausente de nossos cursos básicos de uma variável complexa: a álgebra $\mathcal{H}$ de todas as funções analíticas $f : \mathbb{C} \to \mathbb{C}$ é euclidiana com relação a todo polinômio p. Mais geralmente, temos

Proposição 2.5 *Na divisão euclidiana*

$$f = qp + r \qquad (\text{gr } r < \text{gr } p)$$

da função analítica $f : U \subset \mathbb{C} \to \mathbb{C}$ pelo polinômio p cujas raízes estão em U, o quociente q é analítico.

Demonstração: De acordo com a demonstração da Proposição 2.5, a função

$$q = \frac{f - r}{p}$$

é analítica, pois o numerador e o denominador se anulam exatamente nos mesmos pontos e os zeros do numerador possuem multiplicidade maior ou igual do que os do denominador. Assim, q possui uma expansão em série de potências em cada ponto de U. □

Esse resultado possui extensão para funções $f : I \subset \mathbb{R} \to \mathbb{R}$ de classe C^∞ e polinômios cujas raízes estão todas em I: a regra de L'Hospital implicará então que $q \in C^\infty$.

2.3 Exercícios

1. Sejam $U \subset \mathbb{C}$ um aberto e $f : U \to \mathbb{C}$ uma função. Dizemos que f é **analítica** se ela possui derivada em todos os pontos do aberto U e **holomorfa** se ela possui desenvolvimento em série de potências (com raio de convergência positivo) em todos os pontos do aberto U.

   Suponha que U seja um convexo e $f : U \to \mathbb{C}$ analítica em U. Se $z_0 \in U$, considere o quociente $q(z) = f(z)/(z - z_0)$. Seja γ um caminho contido em U. Mostre que $\int_\gamma q(z)dz = 0$. Conclua daí a fórmula integral de Cauchy para conjuntos convexos e então que toda função analítica é holomorfa.

2. Sejam $p(z) = (z - \lambda)^n$ e f uma função euclidiana com relação a p. Obtenha explicitamente os coeficientes do polinômio interpolador r tal que $f = qp + r$, gr $r <$ gr p.

Capítulo 3

O cálculo funcional em dimensão finita

3.1 Motivação

Algumas vezes escreveremos $f(z)$ para distinguir a função $f : U \subset \mathbb{C} \to \mathbb{C}$ (ou $f : I \subset \mathbb{R} \to \mathbb{R}$) da aplicação linear $f(T)$.

Seja $T : V \to V$ um operador linear e A uma representação matricial de T. Seja m o polinômio mínimo de A (que é o mesmo de T).

Funções de matrizes são usualmente definidas em duas situações: ou a função f é suave nos autovalores da matriz diagonalizável $A = P^{-1}DP$ (com D diagonal) e $f(A)$ é definida por $P^{-1}f(D)P$ ou a função f é analítica e $f(A)$ é definida por meio de uma expansão em série de potências de f (veja exemplos no capítulo 4). Em ambos os casos a função f é euclidiana com relação a m.

Assim, nos casos acima, se considerarmos a divisão euclidiana

$$f = qm + r, \tag{3.1}$$

ou q é contínua em cada autovalor da matriz diagonal D (de acordo com a proposição 2.2) e $q(A)$ é definida por $P^{-1}q(D)P$, ou q é analítica no espectro $\sigma(A)$ e $q(A)$ pode ser definida (veja definição 3.2). Resulta então que[1]

$$f(A) = r(A). \tag{3.2}$$

Esse é um dos resultados mais profundos da teoria espectral: $f(A)$ é um polinômio em A, cujos coeficientes são determinados pelos valores que f (e suas derivadas, conforme o caso) assume no espectro $\sigma(A)$ da matriz A.

[1]Para sermos precisos, é necessário mostrar que $(qm + r)(A) = q(A)m(A) + r(A)$, o que é uma generalização do homomorfismo mostrado na seção 1.4. Veja a seção 3.4.

Exemplo 3.1 Cálculo de potência de uma matriz (simétrica?)

$\diamond$

3.2 Definindo funções de matrizes

Consideremos agora o problema inverso. Dada uma matriz A e uma função $f(z)$, quando podemos definir $f(A)$?

Definição 3.2 *Seja $m(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_\ell)^{d_\ell}$ o polinômio mínimo de A. Se estão definidos os valores*

$$\begin{array}{ccccc} f(\lambda_1) & f'(\lambda_1) & \cdots & \cdots & f^{(d_1-1)}(\lambda_1) \\ \vdots & \cdots & \cdots & \vdots & \\ f(\lambda_\ell) & f'(\lambda_\ell) & \cdots & f^{(d_\ell-1)}(\lambda_\ell), & \end{array}$$

dizemos que f é **euclidiana com respeito a** *A e definimos*

$$f(A) = r(A),$$

sendo r o polinômio interpolador dado pelo lema 2.3.

A definição 3.2 tem uma conseqüência importante, que salientamos desde já: **a matriz $\mathbf{f(A)}$ sempre comuta com a matriz A!**

Observação 3.3 Se compararmos a definição acima com a definição de uma função euclidiana $f(z)$ com respeito a $m(z)$, vemos que as exigências sobre f são menos restritivas. Qual a razão dessa diferença?

A resposta é simples: ao considerarmos abstratamente a divisão

$$f(z) = q(z)m(z) + r(z),$$

precisamos impor condições em f que possibilitem definir uma função q que dê um sentido àquela divisão. Se essas exigências forem satisfeitas, podemos então concluir que $r(z)$ é dado pelo polinômio interpolador, que está definido sob condições menos exigentes. Por outro lado, ao considerarmos $f(A)$, uma vez mostrado que $(qm+r)(A) = q(A)m(A)+r(A)$, teremos que $f(A) = r(A)$ não importa como estiver definido $q(A)$. Assim, apenas o valor de A no polinômio r é importante. ◀

Entretanto, a definição 3.2 é, muitas vezes, pouco aplicável: é mais fácil obter o polinômio característico p da matriz A do que o polinômio mínimo m. Seria proveitoso se pudéssemos usar p ao invés de m para definir a função

$f(A)$. E isso pode ser feito. Podemos utilizar múltiplos de m enquanto a suavidade de f permitir. Ao mostrarmos esse resultado manteremos a notação $f = qm + r$ (sendo r o polinômio interpolador definido antes) para simbolizar que $f(A)$ foi definido como $r(A)$. Suponhamos que s seja outro polinômio que anula a matriz A e r_1 o polinômio interpolador gerado por s. Então teríamos $f = q_1 s + r_1$ (isto é, $f(A)$ seria definido como $r_1(A)$). Como s é múltiplo de m, a prova da proposição 2.2 garante que

$$r_1(z) = q_2(z)m(z) + r(z). \tag{3.3}$$

De fato, se λ é uma raiz de multiplicidade d de $m(z)$, notamos que

$$r_1^{(i)}(\lambda) = f^{(i)}(\lambda) = r^{(i)}(\lambda), \quad \text{for} \quad i = 0, \dots, d-1.$$

Uma vez que todos os termos da equação (3.3) são polinômios, a substituição de z por A faz sentido, de acordo com a seção 1.4. Assim,

$$r_1(A) = r(A),$$

o que autoriza a utilização de qualquer múltiplo $s(z)$ do polinômio mínimo $m(z)$ da matriz A ao invés de $m(z)$ na definição 3.2.

3.3 Justificando a definição

Precisamos mostrar que a definição 3.2 coincide com a definição usual em casos conhecidos. Para isso, começamos por mostrar o seguinte resultado auxiliar:

Lema 3.4 *Seja f uma função euclidiana com relação à matriz $n \times n$ em blocos*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell \end{pmatrix}.$$

Então

$$f(A) = \begin{pmatrix} f(A_1) & 0 & \cdots & 0 \\ 0 & f(A_2) & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & f(A_\ell) \end{pmatrix}.$$

Demonstração: Seja $r = a_0 + a_1 z + a_2 z^2 + \ldots + a_m z^m$ o polinômio interpolador da definição 3.2. De acordo com o exercício 8 do Capítulo 1 temos:

$$\begin{aligned} f(A) &= a_0 I + a_1 \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell \end{pmatrix} \\ & \quad + a_2 \begin{pmatrix} A_1^2 & 0 & \cdots & 0 \\ 0 & A_2^2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell^2 \end{pmatrix} + \ldots + a_m \begin{pmatrix} A_1^m & 0 & \cdots & 0 \\ 0 & A_2^m & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_\ell^m \end{pmatrix} \\ &= \begin{pmatrix} r(A_1) & 0 & \cdots & 0 \\ 0 & r(A_2) & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & r(A_\ell) \end{pmatrix} \end{aligned}$$

Assim, o resultado estará provado se tivermos

$$f(A_j) = r(A_j).$$

Para $j = 1, \ldots, \ell$, sejam m e m_j os polinômios mínimos de A e A_j, respectivamente. Como $m(A) = 0$, necessariamente cada bloco A_j é anulado por m. Pelo lema 1.4, temos que m é um múltiplo de m_j. Como vimos no final da seção 3.2, isso implica que $f(A_j) = r(A_j)$. (Veja, a esse respeito, o exercício 1.) □

Consideraremos aqui apenas o caso de matrizes diagonalizáveis. O caso geral, de matrizes na forma canônica de Jordan, será considerado no apêndice a esse capítulo. Notamos, entretanto, que se f for uma função analítica, o argumento apresentado na seção 3.1 implica a coincidência da definição 3.2 com a dada por meio de série de potências, uma vez mostrado que vale $(qm + r)(A) = q(A)m(A) + r(A)$ (o que faremos na seção 3.4).

Seja, portanto, f uma função definida nos autovalores da matriz diagonalizável $A = P^{-1}DP$ (sendo D matriz diagonal). A definição usual de $f(A)$ é $P^{-1}f(D)P$.

Consideremos a matriz diagonal D. De acordo com o lema 3.4, para calcularmos $f(D)$ segundo a definição 3.2, basta calcularmos f em cada um dos n blocos diagonais $D_1 = \lambda_1, \ldots, D_n = \lambda_n$ da matriz D. Como o polinômio mínimo do bloco D_j é $m_j = z - \lambda_j$, temos que $f(D_j) = r(D_j) = f(\lambda_j)$. Logo

$$f(D) = r(D) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}.$$

Mas então

$$r(A) = r(P^{-1}DP) = P^{-1}r(D)P = P^{-1}f(D)P,$$

mostrando que as duas definições coincidem.

3.4 Estendendo o homomorfismo de álgebras

Seja A uma matriz $n \times n$ e m o seu polinômio mínimo. Suponhamos que f seja euclidiana com relação a m, isto é, que seja válida a divisão euclidiana $f(z) = q(z)m(z) + r(z)$. Nosso objetivo nessa seção é mostrar que é válida a substituição de z por A: $f(A) = q(A)m(A)+r(A)$, o que produz $f(A) = r(A)$.

Assim, suponhamos que $m(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_\ell)^{d_\ell}$. Definimos $k = \max\{d_1 - 1, \ldots, d_\ell - 1\}$.

Como vimos na seção 1.4, existe um homomorfismo natural ϕ entre $\mathcal{P}$, a álgebra de polinômios com coeficientes em $\mathbb{K}$ e $\mathcal{P}(A)$, a álgebra de matrizes obtida ao se avaliar cada polinômio $p \in \mathcal{P}$ na matriz A.

Vamos agora introduzir uma topologia em $\mathcal{P}$, na qual esse homomorfismo será contínuo. Para isso, seja $K \subset \mathbb{K}$ um conjunto compacto tal que $\sigma(A) \subset K$. Definimos a norma

$$\|p\|_{C^k(K)} = \max_{z \in K}\{|p(z)|, \ldots, |p^{(k)}(z)|\}.$$

É de verificação imediata que a convergência nessa norma implica convergência na semi-norma

$$\|p\|_{\mathcal{P}} = \max\{|p(\lambda_1)|, \ldots, |p^{(d_1-1)}(\lambda_1)|, \ldots, |p(\lambda_\ell)|, \ldots, |p^{(d_\ell-1)}(\lambda_\ell)|\}.$$

Se considerarmos $\mathcal{P}(A)$ com a topologia de $\mathbb{K}^{n^2}$, o homomorfismo ϕ é contínuo. De fato, o polinômio (em A) $p(A) - q(A)$ tem coeficientes que dependem apenas dos valores assumidos pelos polinômios p e q (e, conforme o caso, suas derivadas até a ordem k) no espectro $\sigma(A) = \{\lambda_1, \ldots, \lambda_\ell\}$ da matriz A, de acordo com a definição3.2. Segue imediatamente que $p(A)$ estará perto de $q(A)$, se p e q estiverem suficientemente próximos na norma $\|\cdot\|_{C^k(K)}$.

Denotamos por $\mathcal{F}^k$ a álgebra de todas as funções f definidas e de classe C^k em todos os pontos do compacto K. As funções em $\mathcal{F}^k$ são euclidianas com respeito a A. Consideramos em $\mathcal{F}^k$ a mesma norma introduzida em $\mathcal{P}$. É claro que $\mathcal{P}$ é uma sub-álgebra de $\mathcal{F}^k$.

Definimos então $\Phi : \mathcal{F}^k \to \mathcal{P}(A)$ por $\Phi(f) = f(A)$, sendo $f(A)$ dado pela definição 3.2. Claramente Φ é uma aplicação linear. Vamos verificar que $\Phi(fg) = \Phi(f)\Phi(g)$. Para isso, escreveremos $f = qm + r$ para denotar o polinômio r tal que $f(A) = r(A)$. Se $f = q_1m + r$ e $g = q_2m + s$, claramente $\Phi(f)\Phi(g) = r(A)s(A) = (rs)(A)$. Suponhamos que $fg = q_3m + t$. Como vimos no final da seção 3.2, vale a divisão de polinômios $rs = q_4m + t$, o que implica $\Phi(fg) = t(A) = (rs)(A) = \Phi(f)\Phi(g)$. Isso mostra que Φ é um homomorfismo de álgebras, que estende o homomorfismo ϕ.

$$\begin{array}{ccc} & \phi & \\ \mathcal{P} & \to & \mathcal{P}(A) \\ \downarrow & \nearrow & \\ \mathcal{F}^k & \Phi & \end{array}$$

O mesmo argumento que prova a continuidade de ϕ continua válido. Assim, Φ é contínuo.

O núcleo de Φ é constituído pelas funções $f \in \mathcal{F}^k$ que possuem resto nulo quando divididas por m, isto é, pelas funções que se anulam no conjunto

$$\{|p(\lambda_1)|, \ldots, |p^{(d_1-1)}(\lambda_1)|, \ldots, |p(\lambda_\ell)|, \ldots, |p^{(d_\ell-1)}(\lambda_\ell)|\}.$$

3.5 Apêndice: Matrizes na Forma de Jordan

Como sabemos, a forma canônica de Jordan de uma matriz $n \times n$ complexa (ou que tenha n autovalores - não necessariamente distintos - no corpo $\mathbb{K}$) é da forma

$$J = \begin{pmatrix} J_{\lambda_1} & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_k} \end{pmatrix}$$

sendo que os blocos J_{λ_i} possuem a forma

$$J_{\lambda_i} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

Estamos denotando por λ_i um dos autovalores da matriz A. Ao mesmo autovalor λ_i podem estar associados diferentes blocos J_{λ_i}. Sabemos que existe pelo menos um bloco $r_i \times r_i$, sendo r_i a multiplicidade algébrica do autovalor λ_i (isto é, a multiplicidade de λ como fator do polinômio característico de A).
Consideremos um bloco J_λ de tamanho $k \times k$, com $k \leq d$. Suponhamos inicialmente que $k = d$. Nesse caso, como $(z-\lambda)^k$ é o polinômio mínimo de W_λ (o autoespaço associado ao autovalor λ), a função $f(J_i)$ é dada por um polinômio de grau no máximo igual a k_1, de acordo com o lema 3.2:

$$r(z) = a_0 + a_1(z-\lambda)^1 + \ldots + a_{k-1}(z-\lambda)^{k-1}.$$

Os coeficientes a_i são obtidos pela relações $f^{(i)}(\lambda) = r^{(i)}(\lambda)$. Assim,

$$\begin{aligned} f(J_\lambda) &= f(\lambda)I + f'(\lambda)(J_\lambda - \lambda I) + \ldots + \frac{f^{(k-1)}(\lambda)}{(k-1)!}(J_\lambda - \lambda I)^{(k-1)} \\ &= \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f"(\lambda)}{2!} & \cdots & \frac{f(k-1)(\lambda)}{(k-1)!} \\ 0 & f(\lambda) & \frac{f'(\lambda)}{1!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda) & \frac{f'(\lambda)}{1!} \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{pmatrix}. \end{aligned} \tag{3.4}$$

Comparando essa expressão, obtida através da definição 3.2, com a definição de função de matriz na forma de Jordan[2] vemos que elas coincidem.

No caso de blocos $k \times k$, com $1 \leq k < d$, basta então notarmos que o polinômio procurado sempre deverá ter grau $k-1$, pois o polinômio característico do bloco (que coincide com o polinômio mínimo) tem grau k. Assim, a expressão obtida acima continua válida para qualquer bloco $k \times k$.

Para passarmos dos blocos para a matriz na Forma Canônica de Jordan basta, como antes, notarmos que $r(A) = P^{-1}r(J)P$.

3.6 Exercícios

1. Ao mostrarmos o lema 3.4, tivemos que verificar que $f(A_j) = r(A_j)$. Qual a razão de não termos utilizado a divisão $f = qm + r$ e concluir daí que $f(A_j) = r(A_j)$?

[2]Em [19], o fluxo e^{Jt} de uma matriz J na Forma Canônica de Jordan é explicitamente calculado. Trocando-se a função $\exp zt$ por uma função f suficientemente suave, obtemos então uma expressão idêntica à equação (3.4). Veja, a esse respeito, [17].

2. Mostre que $\|p\|_{C^k(K)}$ é realmente uma norma em $\mathcal{P}$.

3. Uma aplicação $|\cdot| : \mathcal{P} \to \mathbb{R}$ é uma **semi-norma** se ela satisfaz

   (i) $|p| \geq 0$ para todo $p \in \mathcal{P}$;

   (ii) $|\alpha p| = |\alpha|\ |p|$ para todo $p \in \mathcal{P}$ e $\alpha \in \mathbb{K}$, sendo $|\alpha|$ o módulo de α.

   (iii) $|p+q| \leq |p| + |q|$ para todos $p, q \in \mathcal{P}$.

   Mostre que $|\cdot|_{\mathcal{P}}$, definida na seção 3.4, é uma semi-norma.

4. Mostre que não existe inteiro $j < k$ tal que a convergência na norma $\|\cdot\|_{C^j(K)}$ implica a convergência na semi-norma $\|\cdot\|_{\mathcal{P}}$.

5. Confira os detalhes na prova de que o homomorfismo de álgebras $\phi : \mathcal{P} \to \mathcal{P}(A)$ é contínuo.

6. Verifique que $\mathcal{F}^k$ é uma álgebra.

Capítulo 4

Exemplos

4.1 A exponencial

Começamos com a definição usual do fluxo e^{At}. Para isso, consideramos a função exponencial $\exp : \mathbb{C} \to \mathbb{C}$, cuja representação em série de potências

$$\exp(z\tau) = e^{z\tau} = 1 + \sum_{n=1}^{\infty} \frac{z^n \tau^n}{n!},$$

converge uniformemente em conjuntos compactos. Se $\|A\|$ denota a norma usual no espaço $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ das transformações lineares $A : \mathbb{C}^n \to \mathbb{C}^n$, afirmamos que

$$I + \sum_{n=1}^{\infty} \frac{A^n \tau^n}{n!}$$

define um operador linear. De fato, a norma em $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ tem a propriedade

$$\|AB\| \leq \|A\| \; \|B\|,$$

de onde decorre que $\|A^i\| \leq \|A\|^i$. Assim, para $k = 1, 2, \ldots$, segue

$$\left\| I + \sum_{n=1}^{k} \frac{A^n \tau^n}{n!} \right\| \leq 1 + \sum_{n=1}^{k} \frac{\|A\|^n |\tau|^n}{n!}. \tag{4.1}$$

Para cada valor de τ fixo, a série à direita converge. Como o espaço $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ é completo, provamos que

$$\exp(A\tau) = e^{A\tau} := I + \sum_{n=1}^{\infty} \frac{A^n \tau^n}{n!}$$

é um operador linear. Tomando $\tau = t \in \mathbb{R}$, definimos o fluxo e^{At}. Também notamos que (4.1) mostra que a convergência é uniforme se τ pertence a um conjunto compacto. Logo, diferenciação termo a termo produz sua derivada e

$$\frac{d}{dt}e^{At} = e^{At}A.$$

Além disso, quando $t = 0$, temos

$$e^{At}\Big|_{t=0} = e^0 = I.$$

Essas são as propriedades principais do fluxo e^{At}. Em particular, vemos que e^{At} é uma solução fundamental do sistema matricial $X' = AX$, $X(0) = I$.

Essa definição do fluxo e^{At} torna difícil o seu cálculo explícito: usualmente é necessário obter a Forma Canônica de Jordan $J = P^{-1}AP$ da matriz A, então e^{Jt} (veja o apêndice 3.5) e, finalmente, $e^{At} = Pe^{Jt}P^{-1}$. O cálculo funcional torna possível obter e^{At} facilmente.

Notamos também que as propriedades do fluxo e^{At} são conseqüências imediatas do cálculo funcional. Por exemplo, decorre das propriedades mostradas na seção 3.4 que

$$\frac{\partial}{\partial t}f(zt) = f'(zt)z \quad \Rightarrow \quad \frac{d}{dt}e^{At} = e^{At}A.$$

Observação 4.1 Embora a função $f(z) = e^z$ satisfaça à equação

$$e^{z+w} = e^z e^w,$$

não podemos deduzir que $e^{A+B} = e^A e^B$, uma vez que a substituição simultânea das variáveis z por A e w por B não é permitida pelo cálculo funcional. Contudo, se A e B comutam, o simples conhecimento de que e^A é um polinômio em A nos permite concluir que $e^A B = Be^A$. A prova de que $e^{A+B} = e^A e^B$ se, e somente se, $AB = BA$ então continua como usualmente (veja [1], [19]). ◀

Exemplo 4.2 Seja

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}.$$

O polinômio característico de A (e também o seu polinômio mínimo) é

$$p(z) = (z-1)(z+i)(z-i).$$

Para obtermos e^{At}, definimos a função $f(zt) = e^{zt}$. Basta então obter um polinômio, de grau no máximo igual a 2, tal que $r(1) = f(1t) = e^t$, $r(i) =$

$f(it) = \cos t + i \sin t$ e $r(-i) = f(-it) = \cos t - i \sin t$. Substituindo essas relações no polinômio $r(z) = az^2 + bz + c$, encontramos $a = (e^t/2) - (\cos t + \operatorname{sen} t)/2$, $b = \operatorname{sen} t$ e $c = (e^t/2) + (\cos t - \operatorname{sen} t)/2$. Assim,

$$e^{At} = \left[\frac{e^t}{2} - \frac{\cos t + \operatorname{sen} t}{2}\right] A^2 + (\operatorname{sen} t)A + \left[\frac{e^t}{2} + \frac{\cos t - \sin t}{2}\right] I,$$

que é uma "matriz" real (como se esperava), embora A tenha raízes complexas. ◇

Exemplo 4.3 Seja

$$A = \begin{pmatrix} 3 & -4 & -1 \\ -3 & 5 & 1 \\ 21 & -32 & -7 \end{pmatrix}.$$

O polinômio característico de A é

$$p(z) = (z-1)z^2.$$

Para calcularmos e^{At}, obtemos os coeficientes de $r(z) = az^2 + bz + c$ de modo que sejam satisfeitas as relações $r(1) = e^{1t} = e^t$, $r(0) = e^{0t} = 1$ e $r'(0) = te^{0t} = t$. Assim, $c = 1$, $b = t$ e $a = e^t - t - 1$. Concluímos que

$$e^{At} = (e^t - t - 1)A^2 + tA + 1I.$$

◇

Os exemplos acima mostram as vantagens práticas da obtenção do fluxo e^{At} através do cálculo funcional. Como conseqüência, deduzimos que o papel predominante dado à Forma Canônica de Jordan no estudo do sistema linear $x' = Ax$ não é intrínseca: toda a análise de sistemas hiperbólicos pode ser feita sem utilizá-la (veja [4]).

4.2 Funções trigonométricas

O estudo da seção anterior permanece válido para o caso da exponencial e^{iAt} (ou seja, o caso $\tau = it$, $t \in \mathbb{R}$, na seção anterior), o qual gera as funções trigonométrica $\operatorname{sen} At$ e $\cos At$. Essas funções são habitualmente definidas por meio das expansões em série de potências de the power series expansion of $\operatorname{sen} z$ and $\cos z$, mas também são fáceis de obter através do cálculo funcional.

As mesmas observações também se aplicam a outras funções trigonométricas.

4.3 Logaritmo

Um logaritmo da matrix A é usualmente definido através da Forma Canônica de Jordan. (Claro, a hipótese $\det A \neq 0$ é necessária). Como todos os autovalores de A não são nulos, habitualmente se toma um logaritmo dos blocos de Jordan, o que pode ser feito por meio da expansão em séries de $\log(1+z)$ (veja [3]). Contudo, como visto no apêndice 3.5, um logaritmo de um bloco de Jordan pode ser diretamente definido. Como antes, o principal inconveniente desse método é que a Forma de Jordan de uma matriz é necessária para se obter seu logaritmo.

O cálculo funcional permite a obtenção da matriz $B = \log A$, se $\det A \neq 0$. Apenas temos que escolher um ramo da função $f(z) = \log z$ que contém o espectro $\sigma(A)$ e então obter $B = \log A$ por meio do polinômio interpolador. Claro, a matriz B depende do ramo escolhido, mas a relação $e^B = A$ segue sempre de $e^{\log z} = z$.

Se todos os autovalores da matriz real A são positivos, podemos então considerar a função real $f(x) = \ln x$ e aplicar a mesma técnica. A matriz $B = \ln A$ assim obtida é a única solução real da equação $e^B = A$.

4.4 Raiz quadrada

Suponhamos que todos os autovalores da matriz real A sejam reais e não-negativos. Adicionalmente, se 0 for um autovalor de A, supomos que ele seja uma raiz simples do polinômio mínimo m de A. Nesse caso, podemos utilizar a função $f : (0, \infty) \to \mathbb{R}$, $f(x) = \sqrt{x}$ para definir $\sqrt{A}$. Nesse caso, o cálculo funcional é utilizado em uma função que é apenas contínua no autovalor simples $\lambda = 0$ da matriz A.

Contudo, podemos definir $\sqrt{A}$ mesmo que A e seus autovalores sejam complexos e não-nulos. Apenas precisamos escolher um ramo da função logaritmo $f(z) = \log z$ para o qual a raiz quadrada de todos os autovalores da matriz A esteja definida. Então aplicamos o cálculo funcional à função complexa $f(z) = \sqrt{z}$.

Observação 4.4 A definição de $\sqrt{A}$ não determina todas as soluções da equação $B^2 = A$. Se A é a matriz identidade 2×2,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

também são soluções de $B^2 = I$, além de $B = I$, a única solução que pode ser obtida através da função raiz quadrada real. Além disso, se $A = -I$, a equação $B^2 = A$ possui a solução real

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

que não vem de da função $\sqrt{A}$. ◄

4.5 A inversa

A maneira clássica de se obter a inversa por meio do polinômio característico p (ou mínimo) da matriz invertível A é a seguinte: se

$$p(z) = z^m + \ldots + a_1 z + a_0$$

temos

$$0 = A^m + a_{m-1}A^{m-1} + \ldots + a_1 A + a_0 I.$$

Multiplicando essa relação por A^{-1}, obtemos

$$a_0 A^{-1} = -[a_1 A + \ldots + a_m A^m].$$

Como A possui inversa, $a_0 \neq 0$. Obtemos A^{-1} dividindo o lado direito da igualdade acima por a_0.

Para uma matriz invertível arbitrária, esse procedimento não é vantajoso com relação ao cálculo da inversa por meio de eliminação gaussiana. Em geral, também o cálculo funcional não é vantajoso.

Mas, por exemplo, se a matriz invertível A é simétrica e possui poucos autovalores, o cálculo funcional é útil: veja [18] (ou [5]).

4.6 Exercícios

1. Seja A uma matriz real com todos os autovalores positivos. Mostre que $B = \ln A$ é a única solução real da equação $e^B = A$.

2. Seja A uma matriz real simétrica, com todos os autovalores não-negativos. Mostre que $B = \sqrt{A}$ é a única solução real de $B^2 = A$.

Capítulo 5

O teorema espectral

Nesta seção mostraremos a utilidade do cálculo funcional na demonstração de resultados abstratos.

5.1 Imagem do espectro

Começamos pelo

Teorema 5.1 (Teorema da Imagem do Espectro)
Seja f uma função euclidiana com relação à matriz $n \times n$ complexa A. Se λ é um autovalor de A, então $f(\lambda)$ é um autovalor de $f(A)$. Todo autovalor de $f(A)$ é da forma $f(\lambda)$, em que λ é um autovalor de A.

Demonstração: Como f é euclidiana com relação à A, $f(A) = r(A) = a_k A^k + \ldots + a_1 A + a_0 I$. Se v é um autovetor relacionado ao autovalor λ,

$$f(A)v = r(A)v = (a_k \lambda^k + \ldots + a_1 \lambda + a_0)v = r(\lambda)v = f(\lambda)v.$$

Suponhamos que μ seja um autovalor de $f(A) = r(A)$. Consideremos o polinômio $q(z) - \mu$, que pode ser fatorado em $\mathbb{C}$ como

$$q(z) - \mu = a_k \prod_{i=1}^{k} (z - \lambda_i).$$

Conseqüentemente,

$$q(A) - \mu I = a_k \prod_{i=1}^{k} (A - \lambda_i I).$$

Como o lado esquerdo da equação acima não possui inversa, ao menos um dos fatores $A - \lambda_i I$ não possui inversa. Assim, λ_i é, ao mesmo tempo, um autovalor de A e uma raiz de $q(z) - \mu$. Portanto,

$$f(\lambda_i) = r(\lambda_i) = \mu.$$

□

5.2 O teorema espectral

Definição 5.2 *Um operador $N : V \to V$ é* **nilpotente** *se existe $k \in \mathbb{N}$ tal que $N^k = 0$.*

Provaremos agora um dos resultados mais importantes da Álgebra Linear. Se V é um espaço vetorial complexo, ele é conhecido como Teorema Espectral. (No caso de V ser um espaço real, desse resultado pode ser obtido o Teorema da Decomposição Primária).

Teorema 5.3 (Teorema Espectral)
Sejam V um espaço vetorial complexo de dimensão n e $T : V \to V$ um operador linear com polinômio característico

$$p(z) = (z - \lambda_1)^{s_1} \cdots (z - \lambda_\ell)^{s_\ell},$$

em que os autovalores λ_i, $i = 1, \ldots, \ell$ são distintos.
Então existem subespaços $W_1, \ldots, W_\ell$ tais que

$$V = W_1 \oplus \cdots \oplus W_\ell$$

e $T(W_i) \subset W_i$. Então $\dim W_i = s_i$, o operador $T|_{W_i}$ tem polinômio mínimo $m_i = (z - \lambda_i)^{d_i}$ e $m(z) = m_1(z) \cdots m_\ell(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_\ell)^{d_\ell}$, sendo $1 \leq d_i \leq s_i$.
Além disso, o operador T se escreve como $D + N$, com D diagonalizável e N nilpotente, sendo que $DN = ND$.

Demonstração: Para cada λ_i consideramos um aberto $U_i \ni \lambda_i$, de modo que $U_i \cap U_k = \emptyset$, se $i \neq k$. Definimos $f_i(z) = 1$, se $z \in U_i$, e $f_i(z) = 0$, se $z \in U_j$, $j \neq i$. As funções $f_1, \ldots, f_\ell$ são euclidianas com relação à p e as relações

$$f_i^2 = f_i, \quad f_i f_j = 0, \text{ if } \quad i \neq j \quad \text{and} \quad \sum_{i=1}^{\ell} f_i = 1$$

são válidas em $\bigcup_{i=1}^{\ell} U_i \supset \sigma(T)$. Assim, denotando $f_i(T)$ por π_i, as relações

$$\pi_i^2 = \pi_i, \quad \pi_i\pi_j = 0, \text{ if } \quad i \neq j \quad \text{e} \quad \sum_{i=1}^{\ell} \pi_i = I, \tag{5.1}$$

continuam válidas (de acordo com a seção 3.4), mostrando assim que cada π_i é uma projeção.

Se W_i denota a imagem $\pi_i(V)$, obtemos

$$V = W_1 \oplus \cdots \oplus W_\ell.$$

Como π_i comuta com T, claramente vale $T(W_i) \subset W_i$ (veja a proposição 1.13).

Independente das bases $\mathcal{B}_1, \ldots, \mathcal{B}_\ell$ escolhidas para os espaços $W_1, \ldots, W_\ell$, respectivamente, T pode ser representado por uma matriz diagonal em blocos A com relação à base $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_\ell\}$ de $V = W_1 \oplus \cdots \oplus W_\ell$:

$$A = [T]_{\mathcal{B}} = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_\ell \end{pmatrix}.$$

Afirmamos que, para $i = 1, \ldots, \ell$, o polinômio característico de A_i é

$$\det(zI - A_i) = (z - \lambda_i)^{s_i},$$

o que implica que $\dim W_i = s_i$ e que o polinômio mínimo de A_i é $(z - \lambda_i)^{d_i}$, para $1 \leq d_i \leq s_i$ (de acordo com o lema 1.4). Daí decorre imediatamente que o polinômio m tem a forma dada pelo teorema.

Para provarmos nossa afirmação é suficiente mostrar que o único autovalor de A_i é λ_i pois, por um lado, temos a fatoração

$$p(z) = (z - \lambda_1)^{s_1} \cdots (z - \lambda_\ell)^{s_\ell}$$

e, por outro,

$$p(z) = \det(zI - A) = \det(zI - A_1) \cdots \det(zI - A_\ell).$$

Vamos considerar apenas $i = 1$, os casos restantes sendo análogos. Seja $\lambda \neq \lambda_1$ arbitrário. Definimos as funções

$$g(z) = \begin{cases} q_1(z) &= & z - \lambda & \text{se } z \in U_1 \\ q_j(z) &= & 1 & \text{se } z \in U_j, \end{cases} \quad \text{e} \quad h(z) = \begin{cases} 1/(z - \lambda) & \text{se } z \in U_1 \\ 1 & \text{se } z \in U_j, \end{cases}$$

em que $j = 2, \ldots, \ell$.

Notamos que, na construção das projeções $\pi_1, \ldots, \pi_\ell$, as vizinhanças disjuntas $U_1, \ldots, U_\ell$ foram escolhidas arbitrariamente. Reduzindo a vizinhança U_1 de λ_1, podemos supor que $\lambda \notin U_1$. Assim, temos que

$$g(z)h(z) = 1.$$

Isso garante que $g(A)$ possui inversa.

Agora calculamos $g(A)$. Para isso, notamos que

$$g(z) = q_1(z)f_1(z) + q_2(z)f_2(z) + \ldots + q_\ell(z)f_\ell(z).$$

Como $q_i(T)$ é um polinômio, vemos que

$$g(T) = (T - \lambda I)\pi_1 + \ldots + I\pi_\ell.$$

Representando o operador T na base $\mathcal{B}$ obtemos a expressão de $g(A)$:

$$g(A) = \begin{pmatrix} A_1 - \lambda I & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix}.$$

Como $g(A)$ tem inversa, $A_1 - \lambda I$ também possui inversa. Como $\lambda \neq \lambda_1$ foi tomado arbitrariamente, está provado que o único autovalor de A_1 é λ_1.

Consideramos agora o operador diagonalizável $D = \sum_{i=1}^{\ell} \lambda_i \pi_i$. (Em cada W_i temos $D_i := \lambda_i \pi_i = \lambda_i I$, em que I é o operador identidade em W_i, de acordo com a definição de f_i. Isso implica que D é diagonalizável. Veja o exemplo 5.5.)

Definimos $N = T - D$. Claramente, para $i = 1, \ldots, \ell$, vale $N = h(T)$, em que

$$h(z) = z - \lambda_i, \quad z \in U_i.$$

De acordo com o teorema de Cayley-Hamilton 1.7, $(A_i - \lambda_i I)^{s_i} = 0$, o que prova que $N^k = 0$, em que $k = \max\{s_1, \ldots, s_\ell\}$. Assim, $T = D + N$, sendo D diagonalizável e N nilpotente. (Na verdade, $(T_i - \lambda_i I)^{d_i} = 0$ para $d_i \in \{1, \ldots, s_i\}$. O inteiro d_i é o **índice** do autovalor λ_i).

Como $D = \sum_{i=1}^{\ell} \lambda_i \pi_i$ é uma soma de polinômios em T, D comuta com T. Assim, $ND = (T - D)D = D(T - D) = DN$. $\square$

Observação 5.4 1. A demonstração do Teorema Espectral 5.3 nos mostra como obter o espaço W_i (veja o exemplo 5.5). Contudo, o próprio teorema nos fornece uma outra caracterização desse espaço: uma vez que o polinômio característico de $T|_{W_i}$ é $(z-\lambda_i)^{r_i}$ (justifique!), todo elemento $w_i \in W_i$ satisfaz $(T-\lambda_i I)^{r_i} w_i = 0$ e, se $w_j \notin W_i$, $(T-\lambda_i I)^{r_i} w_j \neq 0$ (justifique). Assim, $W_i = \ker(T-\lambda_i I)^{r_i}$. Entretanto, o polinômio mínimo de $T|_{W_i}$ é $(z-\lambda_i)^{d_i}$. Do mesmo modo, $W_i = \ker(T-\lambda_i I)^{d_i}$. O índice d_i é encontrado por inspeção: obtemos $\ker(T-\lambda_i) \subset \ker(T-\lambda_i I)^2 \subset \ldots \subset \ker(T-\lambda_i I)^{d_i} = \ker(T-\lambda_i I)^{d_i+1} = \cdots = \ker(T-\lambda_i I)^{r_i}$. Quer dizer, d_i é encontrado quando os subespaços $\ker(T-\lambda_i I)^k$ passam a ser todos iguais.

2. Se V é um espaço vetorial de dimensão n sobre $\mathbb{R}$, o Teorema Espectral continua válido sempre que o operador $T : V \to V$ possui todos os seus n autovalores (contada a multiplicidade) no corpo $\mathbb{R}$.

Por outro lado, dado um operador linear $T : V \to V$ sobre um espaço vetorial real V, é possível definir a **complexificação** $V_{\mathbb{C}}$, espaço vetorial sobre $\mathbb{C}$ que contém V como subespaço, e uma extensão $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ de T - chamado **complexificação** de T. Os polinômios característico $p(z)$ e mínimo $m(z)$ de T e $T_{\mathbb{C}}$ coincidem, de modo que podemos concluir que os todos os fatores irredutíveis presentes na fatoração de $p(z)$ também estão presentes na fatoração de $m(z)$. Mostraremos essas afirmações no Capítulo 6. ◀

Exemplo 5.5 Seja $T : \mathbb{C}^4 \to \mathbb{C}^4$ definida por

$$T(x_1, x_2, x_3, x_4) = (2x_1 - x_2 + x_4, 3x_2 - x_3, x_2 + x_3, -x_2 + 3x_4).$$

O polinômio característico de T é $p(z) = (z-3)(z-2)^3$ e se verifica facilmente que $m(z) = (z-3)(z-2)^2$ é o polinômio mínimo de T.

Inicialmente exemplificaremos o teorema 5.3 com respeito à base canônica do $\mathbb{C}^4$. Denotaremos por A a matriz que representa T nessa base.

A projeção π_1 (associada ao autovalor 3) é obtida ao se resolver o sistema[1]

$$r(z) = az^2 + bz + c, \quad r(3) = 1,\ r(2) = 0,\ r'(2) = 0.$$

Assim, $a = 1$, $b = -4$, $c = 4$ e

$$\pi_1 = A^2 - 4A + 4I = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}.$$

[1] Para simplificar os cálculos, usamos o polinômio mínimo de T ao invés do polinômio característico.

Do mesmo modo,

$$\pi_2 = \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 \end{pmatrix}.$$

As relações (5.1) seguem imediatamente. Logo,

$$\begin{aligned} \left\{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right\} = \mathbb{R}^4 &= \left\{\begin{pmatrix} -2x_2 + x_3 + x_4 \\ 0 \\ 0 \\ -2x_2 + x_3 + x_4 \end{pmatrix}\right\} + \left\{\begin{pmatrix} x_1 + 2x_2 - x_3 - x_4 \\ x_2 \\ x_3 \\ 2x_2 - x_3 \end{pmatrix}\right\} \\ &= W_1 \oplus W_2. \end{aligned}$$

A matriz D é definida por

$$D = 3\pi_1 + 2\pi_2 = \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & -2 & 1 & 3 \end{pmatrix}$$

e a matriz nilpotente N por

$$N = A - D = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

É fácil verificar que $N^2 = 0$ e $ND = DN$.

Se escolhermos, por exemplo, bases

$$\mathcal{B}_1 = \{w_1 = (1, 0, 0, 1)\}$$

e

$$\mathcal{B}_2 = \{w_2 = (1, 0, 0, 0),\ w_3 = (0, 1, 0, 2),\ w_4 = (0, 0, 1, -1)\}$$

para os espaços W_1 e W_2, respectivamente, então T é representado pela matriz diagonal em blocos

$$B = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

na base $\{\mathcal{B}_1, \mathcal{B}_2\} = \{w_1, w_2, w_3, w_4\}$. Agora D é uma autêntica matriz diagonal

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

e

$$N = B - D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

também satisfaz $N^2 = 0$. ◇

Corolário 5.6 *Um operador linear $T : V \to V$ é diagonalizável se, e somente se, o seu polinômio mínimo é produto de fatores lineares distintos.*

Demonstração: Suponhamos que T seja diagonalizável. Sejam $\lambda_1, \ldots, \lambda_\ell$ os autovalores distintos de T. Então V possui uma base formada por autovetores de T, de acordo com o exercício 10 do Capítulo 1. Considere o polinômio

$$h(z) = (z - \lambda_1) \ldots (z - \lambda_k).$$

Se v é um autovetor de T associado ao autovalor λ_i, então $(T - \lambda_i I)v = 0$. Isso implica que $h(T)v = 0$ para qualquer autovetor de T. Como o Teorema Espectral 5.3 implica que o polinômio mínimo e característico possuem os mesmos fatores irredutíveis, mostramos que h é o polinômio mínimo de T.

Reciprocamente, se $p(z) = (z - \lambda_1) \ldots (z - \lambda_\ell)$ é o polinômio mínimo de T, então $W_i = \ker(T - \lambda_i I)$. Claramente todo elemento de W_i é um autovetor de T. Tomando bases $\mathcal{B}_i$ de cada espaço W_i, temos que $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ é uma base de V formada por autovetores de T. □

5.3 Exercícios

1. Na demonstração do Teorema Espectral 5.3, reduzimos as vizinhanças $U_i \ni \lambda_i$ para mostrar que o único autovalor da aplicação T restrita a W_i é λ_i. Essa redução não é necessária. Justifique.

2. Mostre que, na decomposição $T = D + N$, com $DN = ND$, sendo D diagonalizável e N nilpotente, então as aplicações lineares D e N são únicas.

Capítulo 6

O teorema da decomposição primária

Seja $T : V \to V$ um operador linear sobre o espaço real V de dimensão n. O teorema espectral 5.3 pode ser aplicado, desde que o polinômio característico p de T tenha suas n raízes em $\mathbb{R}$. Se esse não é o caso, aquele resultado não é imediatamente aplicável.

Notamos que, se todas as raízes de p estão em $\mathbb{K}$, então

$$p(z) = (z - \lambda_1)^{s_1} \cdots (z - \lambda_\ell)^{s_\ell}$$

é a decomposição de p em fatores irredutíveis, primos entre si dois a dois. O teorema da decomposição primária é o resultado análogo ao teorema espectral 5.3, no caso em que a decomposição de p possui fatores irredutíveis de grau 2.

6.1 A complexificação de um espaço vetorial

Definição 6.1 *Sejam $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ e $z \in \mathbb{K}^n$ um vetor qualquer. Definimos $\overline{A} \in \mathbb{M}_{n\times n}(\mathbb{K})$ como a matriz obtida ao se tomar o conjugado em cada uma das entradas de A e $\bar{z} \in \mathbb{K}^n$ como o vetor obtido ao se tomar o conjugado em cada uma das coordenadas de z.*

É de verificação imediata que $\overline{A + \lambda B} = \overline{A} + \bar{\lambda}\overline{B}$, $\overline{AB} = \overline{A}\,\overline{B}$ para quaisquer matrizes $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ e $\lambda \in \mathbb{K}$. Além disso, também vale $\overline{Az} = \overline{A}\bar{z}$ para qualquer $z \in \mathbb{K}^n$.

Definição 6.2 *Seja V um espaço vetorial real. Definimos a* **complexificação** *de V como sendo o conjunto*

$$V_{\mathbb{C}} = \{u + iv;\; u, v \in V\}.$$

Em $V_{\mathbb{C}}$ soma-se e multiplica-se por escalar (complexo) de maneira "natural". É fácil verificar que $V_{\mathbb{C}}$ torna-se, assim, um espaço vetorial sobre os complexos.

Seja $T : V \to V$ uma aplicação linear. Definimos a **complexificação** *de T como sendo a aplicação $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ definida por $T_{\mathbb{C}}(u+iv) = Tu + iTv$.*

Se identificarmos o vetor $v \in V$ com o vetor $v + i0 \in V_{\mathbb{C}}$, V passa a ser um subespaço de $V_{\mathbb{C}}$. Essa identificação será usada no próximo resultado:

Lema 6.3 *Sejam V um espaço vetorial real de dimensão finita e $T : V \to V$ uma aplicação linear. As seguintes afirmativas são válidas:*

(*i*) *toda base de V é base de $V_{\mathbb{C}}$;*

(*ii*) *os polinômios característicos de T e $T_{\mathbb{C}}$ são iguais;*

(*iii*) *se λ é um autovalor de $T_{\mathbb{C}}$, então $\bar{\lambda}$ é também um autovalor de $T_{\mathbb{C}}$; as multiplicidades algébricas dos autovalores λ e $\bar{\lambda}$ são iguais;*

(*v*) *seja $\tilde{W}$ um subespaço tal que $w = u + iv \in \tilde{W}$ implica que $\bar{w} = u - iv \in \tilde{W}$. Então $\tilde{W}$ possui uma base formada por vetores reais.*

Demonstração: (*i*) Basta notar que as partes real u e imaginária v de qualquer vetor $u + iv$ podem ser escritas como combinação linear dos elementos da base de V.

(*ii*) Decorre imediatamente de (*i*) com a identificação $V \ni v = v + i0 \in V_{\mathbb{C}}$, pois então as representações de T e $T_{\mathbb{C}}$ numa base de V são iguais.

(*iii*) Sejam λ um autovalor de $T_{\mathbb{C}}$ e $p(z)$ o polinômio característico de $T_{\mathbb{C}}$. Como $p(z)$ também é o polinômio característico de T, os coeficientes de $p(z)$ são reais. Tomando o conjugado na equação $p(\lambda) = 0$, obtemos $p(\bar{\lambda}) = 0$, o que mostra que $\bar{\lambda}$ também é uma raiz do polinômio característico de $T_{\mathbb{C}}$. Se $p'(\lambda) = \ldots = p^{(d-1)}(\lambda) = 0$ e $p^{(d)}(\lambda) \neq 0$ (isto é, se λ é raiz de multiplicidade d do polinômio característico[1]), tomando o conjugado em cada uma dessas equações obtemos $p'(\bar{\lambda}) = \ldots = p^{(d-1)}(\bar{\lambda}) = 0$ e $p^{(d)}(\bar{\lambda}) \neq 0$, mostrando que $\bar{\lambda}$ também tem multiplicidade d.

(*v*) Seja $\{w_1, \ldots, w_k\}$ uma base de $\tilde{W}$, com $w_j = u_j + iv_j$, $j = 1, \ldots, k$. Somando e subtraindo os vetores w_j e $\bar{w}_j$, obtemos que $u_j = u_j + i0$ e $v_j = v_j + i0$ estão em $\tilde{W}$. Assim, o conjunto $S = \{u_1, v_1, \ldots, u_k, v_k\}$ é um conjunto de vetores reais que gera $\tilde{W}$. Uma base formada de vetores reais é obtida ao se tomar um subconjunto de S com k elementos que seja linearmente independente em V. (Estamos então aplicando o item (*i*), acima.) □

[1] Veja exercício 2 do Capítulo 1.

Lema 6.4 *Sejam $T : V \to V$ um operador linear e $T_{\mathbb{C}}$ sua complexificação. Se o subespaço $\tilde{W} \subset V_{\mathbb{C}}$ possui uma base formada por vetores reais, então ele é a complexificação de um subespaço $W \subset V$. Se $W_{\mathbb{C}}$ é invariante por $T_{\mathbb{C}}$, então os polinômios mínimos de $T_{\mathbb{C}}|_{\tilde{W}}$ e de $T|_W$ são iguais.*

Demonstração: Todo vetor de $\tilde{W}$ é da forma $w = u + iv$, sendo u e v vetores reais. Escrevendo u e v em termos dos vetores da base real, segue imediatamente que $\tilde{W}$ é a complexificação do espaço real W gerado pelos vetores dessa base. Como a representação matricial de $T_{\mathbb{C}}|_{\tilde{W}}$ e de $T|_W$ em termos da base real é a mesma, seus polinômios mínimos coincidem. □

6.2 O teorema da decomposição primária

Definição 6.5 *Uma aplicação linear $T : V \to V$ definida no espaço real V é* **semi-simples** *se sua complexificação $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ for diagonalizável.*

Teorema 6.6 (Decomposição Primária)

Sejam V um espaço vetorial real de dimensão finita V e $T : V \to V$ uma aplicação linear. Seja $p \in \mathbb{R}[z]$ o polinômio característico de T. Se

$$p(z) = [p_1(z)]^{s_1} \cdots [p_\ell(z)]^{s_\ell}$$

é a decomposição de $p(z)$ em fatores irredutíveis, com $p_i \neq p_k$ para $i \neq k$. Então, o polinômio mínimo de T é

$$m(z) = [p_1(z)]^{d_1} \cdots [p_\ell(z)]^{d_\ell},$$

em que $0 < d_i \leq s_i$ para $i = 1, \ldots, \ell$. O espaço V se decompõe como soma direta de subespaços

$$V = W_1 \oplus \cdots \oplus W_\ell,$$

sendo $W_i = \ker[p_i(T)]^{d_i} = \ker[p_i(T)]^{s_i}$ invariante por T.

Demonstração: Suponhamos que

$$V_{\mathbb{C}} = \tilde{W}_1 \oplus \cdots \oplus \tilde{W}_\ell \tag{6.1}$$

seja a decomposição espectral de $T_{\mathbb{C}}$, de acordo com o Teorema Espectral 5.3 (ao espaço invariante $\tilde{W}_i$ está associado apenas o autovalor λ_i de $T_{\mathbb{C}}$).

De acordo com o lema 6.3 (i), escolhendo uma base $\mathcal{B}$ para V, obtemos uma matriz real A que representa tanto T quanto $T_{\mathbb{C}}$ nessa base.

Seja λ um autovalor **real** de $T_{\mathbb{C}}$ e $\tilde{W}_\lambda = \ker(T_{\mathbb{C}} - \lambda I)^d$ um dos subespaços da decomposição espectral (6.1) de $T_{\mathbb{C}}$. Sejam $w \in \tilde{W}_\lambda = \ker(T_{\mathbb{C}} - \lambda I)^d$ e x a representação de w na base $\mathcal{B}$. Então $(A - \lambda I)^d x = 0$. Tomando o conjugado nessa equação, obtemos $(A - \lambda I)^d \bar{x} = 0$. Assim, $\bar{w} \in \tilde{W}_\lambda$. De acordo com o lema 6.3 (iii), $\tilde{W}_\lambda$ possui uma base formada por vetores reais. Mas uma base formada por vetores reais para $\ker(T_{\mathbb{C}} - \lambda I)^d$ é uma base para $\ker(T - \lambda I)^d$.

Seja agora $\lambda \in \mathbb{C}\backslash\mathbb{R}$ um autovalor de $T_{\mathbb{C}}$. Então $\bar{\lambda}$ também é um autovalor de $T_{\mathbb{C}}$, de acordo com o item (iii) do lema 6.3. Assim, aos autovalores distintos λ e $\bar{\lambda}$, estão associados os subespaços $\tilde{W}_\lambda$ e $\tilde{W}_{\bar{\lambda}}$ da decomposição (6.1).

Suponhamos que $\tilde{W}_\lambda = \ker(T_{\mathbb{C}} - \lambda I)^d$ (ou seja, $(z - \lambda)^d$ é o polinômio mínimo de $T_{\mathbb{C}}$ restrito a $\tilde{W}_\lambda$). Temos que $\tilde{W}_{\bar{\lambda}} = \ker(T_{\mathbb{C}} - \bar{\lambda} I)^d$, pois os elementos de $\tilde{W}_{\bar{\lambda}}$ são os conjugados dos elementos de $\tilde{W}_\lambda$, o que pode ser verificado tomando-se o conjugado na equação $(A - \lambda I)^d x = 0$. Daí também segue que se $\{w_1, \ldots, w_k\}$ é uma base de $\tilde{W}_\lambda$, com $w_j = u_j + iv_j$, então $\{\bar{w}_1, \ldots, \bar{w}_k\}$ é uma base de $\tilde{W}_{\bar{\lambda}}$. Consideremos então o subespaço $\tilde{W}_\lambda \oplus \tilde{W}_{\bar{\lambda}}$. Uma vez que o conjunto de vetores reais

$$S = \{u_1, v_1, \ldots, u_k, v_k\}$$

gera esse espaço e possui $2k$ elementos, ele é uma base de $\tilde{W}_\lambda \oplus \tilde{W}_{\bar{\lambda}}$, pois esse subespaço de $V_{\mathbb{C}}$ tem dimensão $2k$. Seja $W_{\lambda\bar{\lambda}} =< S >$ o subespaço real gerado por S. Assim, o conjunto S é uma base real tanto para $W_{\lambda\bar{\lambda}}$ quanto para $\tilde{W}_\lambda \oplus W_{\bar{\lambda}}$. Claramente $W_{\lambda\bar{\lambda}}$ é invariante por T.

Procedendo dessa maneira, vemos que

$$V = W_{\lambda_1} \oplus \cdots W_{\lambda_s} \oplus W_{\lambda_1\bar{\lambda}_1} \oplus \cdots W_{\lambda_t\bar{\lambda}_t}$$

é a decomposição de T em subespaços invariantes, associada à decomposição espectral (6.1) de $T_{\mathbb{C}}$.

Como vimos, os polinômios mínimos de $T_{\mathbb{C}}$ restrito a $\tilde{W}_\lambda$ e $\tilde{W}_{\bar{\lambda}}$, são, respectivamente, $(z-\lambda)^d$ e $(z-\bar{\lambda})^d$. De acordo com o exercício 9 do Capítulo 1, o polinômio mínimo de T restrito a $\tilde{W}_\lambda \oplus \tilde{W}_{\bar{\lambda}}$ é o polinômio real

$$[(z - \lambda)(z - \bar{\lambda})]^d.$$

De acordo com o lema 6.4, o espaço $\tilde{W}_\lambda \oplus \tilde{W}_{\bar{\lambda}}$ é a complexificação do espaço real $W_{\lambda\bar{\lambda}}$ e seus polinômios mínimos coincidem. □

Observação 6.7 Os subespaços invariantes $W_{\lambda_1\bar{\lambda}_1}, \ldots, W_{\lambda_t\bar{\lambda}_t}$ não estão associados a autovalores reais, mas sim à fatores irredutíveis de grau 2 do polinômio característico de T. ◀

6.3 Exercícios

O objetivo dos próximos exercícios é mostrar que o operador T se escreve como $D+N$, com D semi-simples e N nilpotente, sendo que $DN = ND$.

1. Seja $V_{\mathbb{C}}$ a complexificação do espaço real V. Defina a **conjugação** $\sigma : V_{\mathbb{C}} \to V_{\mathbb{C}}$ por $\sigma(u+iv) = u - iv$. Então o conjunto dos pontos fixos por σ é justamente V.

2. Seja V um espaço real e $S : V_{\mathbb{C}} \to V_{\mathbb{C}}$ uma aplicação linear e σ a conjugação. Então S é a complexificação da aplicação linear $T : V \to V$ se, e somente se, $\sigma S = S\sigma$.

3. Seja $T_{\mathbb{C}} = D_0 + N_0$, com $D_0 N_0 = N_0 D_0$, sendo D_0 diagonalizável e N_0 nilpotente. Defina $D_1 = \sigma D_0 \sigma^{-1}$ e $N_1 = \sigma N_0 \sigma^{-1}$. Mostre que $T_{\mathbb{C}} = D_1 + N_1$, com $D_1 N_1 = N_1 D_1$, sendo D_1 diagonalizável e N_1 nilpotente. Aplique o exercício 2 do Capítulo 5 e conclua que D_0 e N_0 são complexificações de aplicações lineares $D : V \to V$ e $N : V \to V$. Mostre então que $DN = ND$.

Os próximos exercícios mostram como obter uma representação para as aplicações lineares semi-simples.

4. Sejam $u + iv$ um autovetor de $T_{\mathbb{C}}$ associado ao autovalor $\lambda = \alpha + i\beta$ e $S = \{u, v\}$. Então $W =< S >$ (o espaço gerado pelo conjunto S) é um subespaço invariante por T. A representação de $T|_W$ na base $\{u, v\}$ é
$$[T_W]_{\mathcal{B}} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

5. Conclua que todo operador semi-simples D pode ser representado na forma
$$\begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_r & & & & & \\ & & & a_1 & -b_1 & & & \\ & & & b_1 & a_1 & & & \\ & & & & & \ddots & & \\ & & & & & & a_s & -b_s \\ & & & & & & b_s & a_s \end{pmatrix},$$
sendo $\lambda_1, \cdots, \lambda_r$ os autovalores reais de D (podendo haver repetição) e $\lambda_1 = a_1 + ib_1, \cdots, a_s + ib_s$ os autovalores complexos com parte imaginária positiva (podendo haver repetição).

Referências Bibliográficas

[1] H. Amann: Ordinary Differential Equations - an Introduction to Nonlinear Analysis, Walter de Gruyter, Berlin, 1990.

[2] H. Anton and C. Rorres: Elementary Linear Algebra: Applications version, 6th. edition, Wiley, New York, 1991.

[3] R. Bellman: Introduction to Matrix Analysis, McGraw-Hill Book Company, New York, 1960. (Republished in the series "Classics in Applied Mathematics", vol. 12, SIAM, 1995).

[4] H. Bueno: Equações Diferenciais Ordinárias - 1a. parte. Lecture notes. Departamento de Matemática da UFMG, 2001.

[5] H. Bueno: Functions of Matrices, aceito para publicação na revista Cubo.

[6] B. Chisala: A quick Cayley-Hamilton, Amer. Math. Monthly 105 (1998), no. 9, 842-844.

[7] E. A. Coddington and N. Levinson: Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

[8] H. S. M. Coxeter: Regular Polytopes, 3rd. Edition, Dover, New York, 1973.

[9] N. Dunford and J. T. Schwarz: Linear operators I, Interscience, New York, 1968.

[10] Gantmacher, F. R.: The Theory of Matrices, vol. 1 and 2, Chelsea Publishing Co., New York, 1959.

[11] M. Hirsch and S. Smale: Differential Equations, Dynamical Systems and Linear Algebra, Academic Press, New York, 1974.

[12] K. Hoffman and R. Kunze: Linear Algebra, 2nd. edition, Prentice Hall, Englewoods Cliffs, 1971.

[13] S. Lang: Linear Algebra, 3rd. Edition, Springer-Verlag, New York, 1989.

[14] P. D. Lax: Linear Algebra, Wiley-Interscience Publication, New York, 1997.

[15] M. Reed and B. Simon: Methods of Modern Mathematical Physics, vol. I, Academic Press, New York, 1972.

[16] W. Rudin: Real and Complex Analysis, 3rd. Edition, McGraw-Hill, New York, 1987.

[17] R. J. Santos: Geometria Analítica e Álgebra Linear, Parte II, UFMG, 2000.

[18] N. C. Saldanha and C. Tomei: *Spectra of Regular Polytopes*, Discrete & Computational Geometry **7** (1992), 403-414.

[19] J. Sotomayor: Lições de Equações Diferenciais Ordinárias, IMPA, Rio de Janeiro, 1979.

[20] G. Strang: Linear Algebra and its Applications, 3rd. edition, Harcourt, Fort Worth, 1988.

[21] C. Tomei: Novos Cursos de Álgebra Linear para Alunos de Engenharia e Ciências Básicas, XXII CNMAC, 1999.

Índice Remissivo