Curso de Física-Matemática II
Instituto de Física da USP
2000

João Carlos Alves Barata. Notas de aula[1] Versão de 15 de setembro de 2000

Tópicos de Álgebra Linear

Sumário

1 Rudimentos

Na presente versão destas notas uma certa familiaridade com o conceito de determinante de matrizes e suas propriedades é requerido.

- **Alguma Notação**

O conjunto de todas as matrizes $m \times n$ com entradas complexas será denotado aqui por $\text{Mat}(\mathbb{C},\ m,\ n)$. O conjunto de todas as matrizes quadradas $n \times n$ com entradas complexas será denotado simplesmente por $\text{Mat}(\mathbb{C},\ n)$.

Dado um conjunto de n números complexos $\alpha_1, \ \ldots, \ \alpha_n$, denotaremos por $\text{diag}(\alpha_1, \ \ldots, \ \alpha_n)$ a matriz $A \in \text{Mat}(\mathbb{C},\ n)$ cujos elementos A_{ij} são definidos da seguinte forma:

$$A_{ij} = \begin{cases} \alpha_i, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases} .$$

Uma tal matriz é dita ser *diagonal* pois apenas os elementos de sua diagonal principal são eventualmente não nulos. Na representação usual

$$A = \begin{pmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_n \end{pmatrix} .$$

[1] Estas notas podem ser copiadas do seguinte endereço WWW: http://www.fma.if.usp.br/~jbarata/notas_de_aula.html

Sejam $x^1, \ldots, x^n$ vetores, representados na base canônica por vetores-coluna

$$x^a = \begin{pmatrix} x_1^a \\ \vdots \\ x_n^a \end{pmatrix}.$$

Denotamos por $[\![x^1, \ldots, x^n]\!]$ a matriz $n \times n$ construída de forma que sua a-ésima coluna seja o vetor-coluna x^a, ou seja

$$[\![x^1, \ldots, x^n]\!] = \begin{pmatrix} x_1^1 & \cdots & x_1^n \\ \vdots & \ddots & \vdots \\ x_n^1 & \cdots & x_n^n \end{pmatrix}.$$

Essa notação é útil por permitir a seguinte observação. Seja B uma matriz qualquer. Então,

$$B[\![x^1, \ldots, x^n]\!] = [\![Bx^1, \ldots, Bx^n]\!]. \tag{1.1}$$

Essa relação é provada observando-se a regra de multiplicação de matrizes: a a-ésima coluna de $B[\![x^1, \ldots, x^n]\!]$ é

$$\begin{matrix} B_{11}x_1^a + \cdots + B_{1n}x_n^a \\ \vdots \\ B_{n1}x_1^a + \cdots + B_{nn}x_n^a \end{matrix}, \tag{1.2}$$

que vem a ser as componentes de Bx^a, representado como vetor-coluna na base canônica.

2 Noções Básicas sobre o Espectro de uma Matriz

• **O Espectro de uma Matriz**

Seja $A \in \mathrm{Mat}(\mathbb{C}, n)$ uma matriz $n \times n$ com entradas complexas. No estudo das propriedades de A é de grande importância saber para quais números complexos λ a matriz $\lambda\mathbb{1} - A$ é invertível e para quais não é.

Chegamos às seguintes importantes definições.

Definição. Um número complexo λ é dito ser um elemento do *espectro* de $A \in \mathrm{Mat}(\mathbb{C}, n)$ se a matriz $\lambda\mathbb{1} - A$ não possuir uma inversa.

Definição. Um número complexo λ é dito ser um elemento do *conjunto resolvente* de $A \in \mathrm{Mat}(\mathbb{C}, n)$ se a matriz $\lambda\mathbb{1} - A$ possuir uma inversa.

Em outras palavras, o *espectro* de $A \in \mathrm{Mat}(\mathbb{C}, n)$, denotado por $\sigma(A)$, é o conjunto de todos os $\lambda \in \mathbb{C}$ para os quais a matriz $\lambda\mathbb{1} - A$ não tem inversa.

O *conjunto resolvente* de $A \in \mathrm{Mat}(\mathbb{C}, n)$, denotado por $\rho(A)$, é o conjunto de todos os $\lambda \in \mathbb{C}$ para os quais a matriz $\lambda\mathbb{1} - A$ tem inversa.

É evidente que $\sigma(A)$ e $\rho(A)$ são conjuntos complementares, ou seja, $\sigma(A) \cap \rho(A) = \emptyset$ mas $\sigma(A) \cup \rho(A) = \mathbb{C}$.

Um fato importante é que $\lambda\mathbb{1} - A$ é não-invertível se e somente se $\det(\lambda\mathbb{1} - A) = 0$. Assim, um número complexo λ é um elemento do espectro de uma matriz A se e somente se for tal que $\det(\lambda\mathbb{1} - A) = 0$.

Chegamos ao importante conceito de polinômio característico de uma matriz.

• **O Polinômio Característico de uma Matriz**

Seja $A \in \mathrm{Mat}(\mathbb{C},\ n)$ uma matriz cujos elementos de matriz são A_{ij}. Para $\lambda \in \mathbb{C}$ a expressão

$$\det(\lambda\mathbb{1} - A) \;=\; \det\begin{pmatrix} \lambda - A_{11} & -A_{12} & \cdots & -A_{1n} \\ -A_{21} & \lambda - A_{22} & \cdots & -A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{n1} & \cdots & \cdots & \lambda - A_{nn} \end{pmatrix}$$

define, como facilmente se constata pelos métodos usuais e bem conhecidos de cálculo de determinantes, um polinômio de grau n na variável λ, com coeficientes complexos, os quais dependem dos elementos de matriz A_{ij} de A. Esse polinômio é denominado *polinômio característico* de A e desempenha um papel muito importante no estudo de propriedades de matrizes.

Denotaremos por vezes por p_A o polinômio característico de uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$. Como todo polinômio complexo de grau n, p_A possui n raízes, não necessariamente distintas no plano complexo (teorema fundamental da álgebra). As raízes do polinômio característico p_A são denominadas *autovalores* da matriz A. Assim o espectro de uma matriz A coincide com o conjunto de seus autovalores. O estudo de autovalores de matrizes é de grande importância na Álgebra Linear e em suas aplicações à Teoria das Equações Diferenciais, à Geometria, à Teoria dos Sistemas Dinâmicos e à Física, especialmente à Física Quântica.

Seja $A \in \mathrm{Mat}(\mathbb{C},\ n)$ uma matriz e sejam $\alpha_1,\ \ldots,\ \alpha_r$, $1 \leq r \leq n$, seus autovalores distintos, cada qual com multiplicidade $a_1,\ \ldots,\ a_r$, respectivamente, ou seja, cada α_i é uma raiz de ordem $a_i \in \mathbb{N}$ do polinômio característico de A:

$$q(\lambda) \;=\; \det(\lambda\mathbb{1} - A) \;=\; \prod_{i=1}^{r}(\lambda - \alpha_i)^{a_i}.$$

A quantidade a_i é um número inteiro positivo e é denominado *multiplicidade algébrica* do autovalor α_i.

Note-se que como o número de raízes de p_A (contando as multiplicidades) é exatamente igual a seu grau, segue facilmente que a seguinte relação é válida:

$$\sum_{i=1}^{r} a_i \;=\; n, \tag{2.1}$$

ou seja, a soma das multiplicidades algébricas dos autovalores de uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é n.

- **Matrizes Similares. Transformações de Similaridade**

Duas matrizes $A \in \mathrm{Mat}(\mathbb{C},\ n)$ e $B \in \mathrm{Mat}(\mathbb{C},\ n)$ são ditas *similares* se existir uma matriz invertível $P \in \mathrm{Mat}(\mathbb{C},\ n)$ tal que $P^{-1}AP = B$.

Para uma matriz invertível $P \in \mathrm{Mat}(\mathbb{C},\ n)$ fixa, a transformação que leva cada matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ à matriz $P^{-1}AP$ é denominada *transformação de similaridade.*

Sabemos que o determinante é invariante por transformações de similaridade, pois para toda matriz A vale $\det(A) = \det(P^{-1}AP)$.

O determinante não é o único objeto associado a uma matriz que é invariante por transformações de similaridade. O polinômio característico e, portanto, o conjunto de seus auto-valores (incluindo as multiplicidades), também o é. Isso pode ser visto da seguinte forma.

Sejam A e B duas matrizes similares com $B = P^{-1}AP$ para algum P. O polinômio característico de A é $p_A(\lambda) = \det(\lambda\mathbb{1} - A)$ e o de B é $p_B(\lambda) = \det(\lambda\mathbb{1} - B)$. Pela invariância do determinante vale

$$p_A(\lambda) \;=\; \det(\lambda\mathbb{1} - A) \;=\; \det(P^{-1}(\lambda\mathbb{1} - A)P) \;=\; \det(\lambda\mathbb{1} - P^{-1}AP) \;=\; \det(\lambda\mathbb{1} - B) \;=\; p_B(\lambda).$$

Assim, A e B têm o mesmo polinômio característico e, portanto, seus autovalores são iguais, incluindo suas multiplicidades.

- **Auto-vetores**

Seja λ_0 um autovalor de uma matriz A. Então $\lambda\mathbb{1} - A$ não tem inversa. Logo, como $V = \mathbb{C}^n$ é um espaço vetorial de dimensão finita, existe pelo menos um vetor não-nulo v tal que $(\lambda\mathbb{1} - A)v = 0$, ou seja,

$$Av \; = \; \lambda_0 v.$$

Chegamos a mais uma importante definição:

Definição. Um vetor não-nulo v é dito ser um *auto-vetor* de uma matriz A se houver $\lambda_0 \in \mathbb{C}$ tal que

$$Av \; = \; \lambda_0 v.$$

Note-se que se um tal λ_0 satisfaz a relação acima para algum $v \neq 0$ então $\lambda_0\mathbb{1} - A$ não tem inversa. λ_0 é então um elemento do espectro de A, ou seja, um autovalor. λ_0 é dito ser o autovalor associado ao auto-vetor v.

Uma observação importante é a seguinte. Sejam v_1 e v_2 dois autovetores aos quais está associado o <u>mesmo</u> autovalor, ou seja, $Av_1 \; = \; \lambda_0 v_1$ e $Av_2 \; = \; \lambda_0 v_2$. Então, para quaisquer números complexos c_1 e c_2 o vetor $v = c_1v_1 + c_2v_2$ também satisfaz $Av \; = \; \lambda_0 v$. De fato,

$$Av \; = \; A(c_1v_1 + c_2v_2) \; = \; c_1Av_1 + c_2Av_2 \; = \; c_1\lambda_0v_1 + c_2\lambda_0v_2 \; = \; \lambda_0(c_1v_1 + c_2v_2) \; = \; \lambda_0 v.$$

A conclusão a que se chega é que, para cada autovalor α_i de uma matriz A a coleção formada pelo vetor nulo e todos os autovetores de A com autovalor α_i é um subespaço vetorial. Vamos denotar esse subespaço por $\mathcal{E}(\alpha_i)$ ou simplesmente $\mathcal{E}_i$.

Se α_i e α_j são autovalores distintos de A então os sub-espaços de autovetores $\mathcal{E}(\alpha_i)$ e $\mathcal{E}(\alpha_j)$ têm em comum apenas o vetor nulo, ou seja, $\mathcal{E}(\alpha_i) \cap \mathcal{E}(\alpha_j) = \{0\}$. Isso é fácil de provar, pois se w é tal que $Aw = \alpha_i w$ e $Aw = \alpha_j w$ então, subtraindo-se uma relação da outra teríamos $0 = (\alpha_i - \alpha_j)w$, que implica $w = 0$, já que $\alpha_i \neq \alpha_j$.

Essas considerações nos levam a mais um conceito importante: o de multiplicidade geométrica de um autovalor.

• A Multiplicidade Geométrica de um Autovalor

Além do conceito de multiplicidade <u>algébrica</u> de um autovalor, há também o conceito de multiplicidade <u>geométrica</u> de um autovalor, do qual trataremos agora.

Como antes seja $A \in \mathrm{Mat}(\mathbb{C}, \; n)$ uma matriz e sejam $\alpha_1, \; \ldots \; , \; \alpha_r$, $1 \leq r \leq n$, seus autovalores <u>distintos</u>, cada qual com multiplicidade algébrica $a_1, \; \ldots \; , \; a_r$, respectivamente.

Acima introduzimos os sub-espaços $\mathcal{E}_i \; = \; \mathcal{E}(\alpha_i)$, definidos como sendo os sub-espaços gerados por todos os autovetores que têm α_i como autovalor. A *multiplicidade geométrica* de um autovalor α_i é definida como sendo a dimensão do subespaço $\mathcal{E}_i$, ou seja, como sendo o número máximo de autovetores linearmente independentes com autovalor α_i.

É importante advertir de imediato o leitor do fato que a multiplicidade algébrica e multiplicidade geométrica de autovalores nem sempre coincidem. Isso é bem ilustrado no seguinte exemplo simples. Seja

$$A \; = \; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Seu polinômio característico é

$$p_a(\lambda) \; = \; \det(\lambda\mathbb{1} - A) \; = \; \det\begin{pmatrix} \lambda & -1 \\ 0 & \lambda \end{pmatrix} \; = \; \lambda^2.$$

Assim seu (único) autovalor é 0 com multiplicidade algébrica 2. Quais os seus autovetores? São aqueles vetores que satisfazem $Av = 0$. Denotando v como um vetor coluna

$$v \; = \; \begin{pmatrix} a \\ b \end{pmatrix},$$

a relação $Av = 0$ significa

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} = 0.$$

Logo $b = 0$ e todos os autovetores são da forma

$$v = \begin{pmatrix} a \\ 0 \end{pmatrix},$$

$a \in \mathbb{C}$. É evidente que o subespaço gerado pelos autovetores com autovalor zero tem dimensão 1. Assim a multiplicidade algébrica do autovalor zero é 2 mas a sua multiplicidade geométrica é 1.

• A Multiplicidade Algébrica e a Multiplicidade Geométrica

Apesar de a multiplicidade algébrica e a multiplicidade geométrica de um autovalor nem sempre coincidirem, há uma relação de ordem entre eles. A saber, é possível mostrar que a multiplicidade geométrica de um autovalor é sempre menor ou igual à sua multiplicidade algébrica.

Isso segue das seguintes considerações. Seja λ_0 um autovalor de $A \in \mathrm{Mat}(\mathbb{C},\ n)$ e $\mathcal{E}(\lambda_0)$ o subespaço gerado pelos autovetores com autovalor λ_0, e cuja dimensão denotaremos por d. Vamos escolher uma base $v_1, \ldots, v_d, v_{d+1}, \ldots, v_n$ onde os primeiros d vetores são elementos de $\mathcal{E}(\lambda_0)$. Nessa base a matriz A tem a forma

$$\begin{pmatrix} D & Z \\ A_3 & A_4 \end{pmatrix},$$

onde D é uma matriz $d \times d$ diagonal $D = \mathrm{diag}\left(\underbrace{\lambda_0, \ldots, \lambda_0}_{d\,\text{vezes}}\right)$, A_4 é uma matriz $(n-d) \times (n-d)$, Z é uma matriz $d \times (n-d)$ cujos elementos de matriz são todos nulos e A_3 é uma matriz $(n-d) \times d$.

Fica claro que o polinômio característico de A é dado por

$$\det(\lambda\mathbb{1} - A) = (\lambda - \lambda_0)^d \det(\lambda\mathbb{1} - A_4).$$

Isso mostra que a multiplicidade algébrica de λ_0 é pelo menos igual a d, sua multiplicidade geométrica.

• Matrizes Simples

O que foi exposto acima leva-nos naturalmente ao conceito de matriz simples que, como veremos mais adiante, está intimamente ligado ao problema da diagonalizabilidade de matrizes.

Definição. Uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é dita ser *simples* se a multiplicidade algébrica de todos os seus autovalores for igual a sua multiplicidade geométrica.

Deixamos para o leitor provar o seguinte fato: toda matriz diagonal é simples.

Exercício. Prove isso.

Adiante faremos uso da seguinte proposição.

Proposição 2.1 *Se $A \in Mat(\mathbb{C},\ n)$ é uma matriz simples e $P \in Mat(\mathbb{C},\ n)$ é invertível então $P^{-1}AP$ é também simples.* □

Prova. Já vimos (página 3) que A e $P^{-1}AP$ têm o mesmo polinômio característico e, portanto, os mesmos autovalores, incluindo suas multiplicidades algébricas. Seja λ_0 um desses autovalores com multiplicidade algébrica d e sejam $v_1, \ldots, v_d$ um conjunto de d auto-vetores linearmente independentes de P. Os vetores $P^{-1}v_1, \ldots, P^{-1}v_d$ são auto-vetores de $P^{-1}AP$ com autovalor λ_0. De fato, $\left(P^{-1}AP\right) P^{-1}v_i = P^{-1}Av_i = \lambda_0 P^{-1}v_i$. Fora isso os d vetores

$P^{-1}v_1, \ldots, P^{-1} v_d$ são também linearmente independentes. Para ver isso, suponha houvesse constantes $c_1, \ldots, c_d$ tais que

$$c_1 P^{-1} v_1 + \cdots + c_d P^{-1} v_d \; = \; 0 .$$

Multiplicando-se à esquerda por P teríamos $c_1 v_1 + \cdots + c_d v_d = 0$. Como $v_1, \ldots, v_d$ são linearmente independentes as constantes c_i têm que ser todas nulas, provando que os vetores $P^{-1}v_1, \ldots, P^{-1} v_d$ são também linearmente independentes.

Isso prova que a multiplicidade geométrica do auto-valor λ_0 é pelo menos igual a d. Como ela não pode ser maior que d (página 5), conclui-se que é igual a d provando a proposição. ■

3 O Traço de uma Matriz

• **O Traço de uma Matriz**

Seja $A \in \mathrm{Mat}(\mathbb{C},\ n)$, cujos elementos de matriz são A_{ij}, $i, \ j = 1, \ldots n$. Sejam $\lambda_1, \ldots, \lambda_n$ seus n autovalores (não necessariamente distintos e repetidos conforme sua multiplicidade).

Definimos o traço de A como sendo a soma de seus n autovalores:

$$\mathrm{Tr}\,(A) \; := \; \sum_{a=1}^{n} \lambda_a .$$

Uma conclusão que se tira dessa definição é que se duas matrizes são similares, então ambas têm o mesmo traço, ou seja, para qualquer matriz invertível P e qualquer matriz A vale

$$\mathrm{Tr}\,(P^{-1}AP) \; = \; \mathrm{Tr}\,(A).$$

A razão reside na observação feita acima que duas matrizes similares têm o mesmo conjunto e autovalores e, portanto, o mesmo traço.

Temos a seguinte e importante proposição:

Proposição 3.1 *O traço de uma matriz $A \in Mat(\mathbb{C},\ n)$ é igual a soma dos elementos de sua diagonal principal, ou seja,*

$$Tr(A) \; := \; \sum_{a=1}^{n} \lambda_a \; = \; \sum_{a=1}^{n} A_{aa} .$$

□

Prova. A demonstração consistirá em se calcular o coeficiente de λ^{n-1} no polinômio característico $p(\lambda)$ de A de dois modos diferentes.

O polinômio característico de A é

$$p(\lambda) \; = \; \det(\lambda \mathbb{1} - A) \; = \; \det \begin{pmatrix} \lambda - A_{11} & -A_{12} & \cdots & -A_{1n} \\ -A_{21} & \lambda - A_{22} & \cdots & -A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{n1} & \cdots & \cdots & \lambda - A_{nn} \end{pmatrix} .$$

As técnicas de cálculo de determinantes nos dizem que o coeficiente de λ^{n-1} é $-\sum_{i=1}^{n} A_{ii}$. Por exemplo, para o caso $n = 2$

$$p(\lambda) \; = \; \det \begin{pmatrix} \lambda - A_{11} & -A_{12} \\ -A_{21} & \lambda - A_{22} \end{pmatrix} \; = \; \lambda^2 - \lambda(A_{11} + A_{22}) + A_{11}A_{22} - A_{12}A_{21} .$$

Exercício. Convença-se da veracidade da afirmativa acima para o caso de n arbitrário.

Por outro lado, os autovalores de A, $\lambda_1, \ldots, \lambda_n$, são por definição as raízes do polinômio característico. Logo,

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$$

Expandindo-se essa expressão, conclui-se que o coeficiente de λ^{n-1} é

$$-(\lambda_1 + \cdots + \lambda_n) = -\mathrm{Tr}\,(A).$$

Exercício. Certo?

Do exposto acima, conclui-se que o coeficiente de λ^{n-1} no polinômio característico de A é

$$-\sum_{i=1}^{n} A_{ii} = -(\lambda_1 + \cdots + \lambda_n) = -\mathrm{Tr}\,(A),$$

o que termina a prova. ■

Essa proposição leva a outra igualmente importante:

Proposição 3.2 (A Propriedade Cíclica do Traço) *Sejam $A, B \in Mat(\mathbb{C}, n)$. Então*

$$Tr(AB) = Tr(BA).$$

□

Prova. Pelo que vimos acima, tem-se

$$\mathrm{Tr}\,(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}B_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} B_{ji}A_{ij}\right) = \sum_{j=1}^{n}(BA)_{jj} = \mathrm{Tr}\,(BA).$$

Na segunda e quarta igualdades usamos a regra de produto de matrizes. Na terceira igualdade apenas trocamos a ordem das somas. ■

• **Sub-espaços Invariantes**

Um subespaço $\mathcal{E}$ de $\mathbb{C}^n$ é dito ser invariante pela ação de uma matriz A, se $Av \in \mathcal{E}$ para todo $v \in \mathcal{E}$.

4 Polinômios de Matrizes

• **Polinômios de Matrizes**

Seja p um polinômio de grau m:

$$p(x) = a_m x^m + \cdots + a_1 x + a_0$$

com $x \in \mathbb{C}$, $a_j \in \mathbb{C}$ e $a_m \neq 0$. Para uma matriz $A \in \mathrm{Mat}(\mathbb{C}, n)$ definimos o *polinômio matricial* $p(A)$ por

$$p(A) = a_m A^m + \cdots + a_1 A + a_0 \mathbb{1}.$$

Obviamente $p(A)$ é também uma matriz $n \times n$ com entradas complexas.

Se as raízes do polinômio p forem $\alpha_1, \ldots, \alpha_r$, com multiplicidades $m_1, \ldots, m_r$, respectivamente, então

$$p(x) \;=\; a_m \prod_{j=1}^{r} (x - \alpha_j)^{m_j},$$

para todo $x \in \mathbb{C}$. É fácil provar, então, que

$$p(A) \;=\; a_m \prod_{j=1}^{r} (A - \alpha_j \mathbb{1})^{m_j}.$$

Exercício. Justifique isso.

Exercício. Mostre que se $D = \text{diag}\,(d_1,\ \ldots,\ d_n)$ e q é um polinômio então

$$q(D) \;=\; \text{diag}\,(q(d_1),\ \ldots,\ q(d_n))\,.$$

Exercício. Suponha que $A = P^{-1}DP$, onde $D = \text{diag}\,(d_1,\ \ldots,\ d_n)$. Se q é um polinômio mostre que

$$q(A) \;=\; P^{-1}q(D)P \;=\; P^{-1}\text{diag}\,(q(d_1),\ \ldots,\ q(d_n))\,P.$$

• <u>O Polinômio Mínimo</u>

Vamos mostrar que para cada matriz $A \in \text{Mat}(\mathbb{C},\ n)$ sempre existe pelo menos um polinômio p com a propriedade que $p(A) = 0$.

Para tal notemos primeiramente que $\text{Mat}(\mathbb{C},\ n)$ é um espaço vetorial complexo de dimensão n^2. De fato toda a matriz $A \in \text{Mat}(\mathbb{C},\ n)$, cujos elementos de matriz são $A_{ij} \in \mathbb{C}$ pode ser trivialmente escrita na forma

$$A \;=\; \sum_{a=1}^{n}\sum_{b=1}^{n} A_{ab}\mathscr{E}^{ab}$$

onde $\mathscr{E}^{ab} \in \text{Mat}(\mathbb{C},\ n)$ são matrizes cujos elementos de matriz são $(\mathscr{E}^{ab})_{ij} = \delta_{i,a}\delta_{j,b}$, ou seja, todos os elementos de matriz de $\mathscr{E}^{ab}$ são nulos, exceto o elemento a, b, que vale 1.

Exercício. Certo?

Assim vemos que as matrizes $\{\mathscr{E}^{ab},\ a = 1, \ldots, n,\ b = 1, \ldots, n\}$ formam uma base em $\text{Mat}(\mathbb{C},\ n)$, mostrando que $\text{Mat}(\mathbb{C},\ n)$ é um espaço vetorial de dimensão n^2.

Isto posto, temos que concluir que qualquer conjunto de mais de n^2 matrizes é linearmente dependente. Assim, por exemplo, o conjunto $\mathbb{1},\ A,\ A^2,\ \ldots,\ A^{n^2}$ é linearmente dependente, pois possui $n^2 + 1$ elementos. Portanto, existem constantes $c_0,\ \ldots,\ c_{n^2}$, nem todas nulas, tais que

$$c_0\mathbb{1} + c_1A + c_2A^2 + \cdots + c_{n^2}A^{n^2} \;=\; 0.$$

Como o lado direito é um polinômio em A, fica provada nossa afirmação.

Chegamos à seguinte definição.

Definição. Dada uma matriz $A \in \text{Mat}(\mathbb{C},\ n)$, o polinômio mínimo de A é o polinômio de menor grau da forma

$$M(x) \;=\; x^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0$$

para o qual $M(A) = 0$.

As considerações acima mostram que um tal polinômio sempre existe e que tem grau no máximo igual a n^2. Essa é, no entanto, uma estimativa exagerada para o grau do polinômio mínimo de uma matriz $A \in \text{Mat}(\mathbb{C},\ n)$ pois, como veremos abaixo, o polinômio mínimo de uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ tem, na verdade, grau menor ou igual a n. Isso é um corolário de um teorema conhecido como Teorema de Hamilton-Cayley, que demonstraremos a seguir.

- **O Teorema de Hamilton-Cayley**

Teorema 4.1 (Teorema de Hamilton-Cayley) *Seja $A \in Mat(\mathbb{C},\ n)$ e seja $q(x) = \det(x\mathbb{1} - A)$ o polinômio característico de A (e que tem grau n). Então $q(A) = 0$.* $\square$

Prova. Desejamos mostrar que para todo vetor $y \in \mathbb{C}^n$ vale $q(A)y = 0$. Seja então um vetor y fixado, não nulo.

Como o espaço $\mathbb{C}^n$ tem dimensão n, nem todos os conjuntos de vetores da forma $\{y,\ Ay,\ A^2y,\ \ldots\ A^jy\}$ são formados por vetores linearmente independentes. Por exemplo, se $j \geq n$, o conjunto $\{y,\ Ay,\ A^2y,\ \ldots\ A^jy\}$ não é formado por vetores linearmente independentes pois seu número excede a dimensão do espaço.

Seja k o maior número tal que $\{y,\ Ay,\ A^2y,\ \ldots\ A^{k-1}y\}$ é um conjunto linearmente independente de vetores. É claro que $k \leq n$.

É claro também, pela definição, que

$$A^ky \ = \ h_ky + h_{k-1}Ay + \cdots + h_1A^{k-1}y, \tag{4.1}$$

para constantes $h_1, \ldots,\ h_k$.

Vamos denominar $z_1 = y$, $z_2 = Ay$, $\ldots$, $z_k = A^{k-1}y$. Caso $k < n$, escolhamos ainda vetores $z_{k+1}, \ldots, z_n$ de modo que o conjunto $\{z_1,\ \ldots\ ,\ z_n\}$ forme uma base em $\mathbb{C}^n$.

Coloquemo-nos agora a seguinte questão: qual é a forma da matriz A nessa base? No subespaço gerado pelos vetores $\{z_1,\ \ldots\ ,\ z_k\}$ tem-se o seguinte: para $i = 1, \ldots, k-1$ vale $Az_i = z_{i+1}$. Além disso, por (4.1), $Az_k = h_kz_1 + h_{k-1}z_2 + \cdots + h_1z_k$. Isso mostra que o subespaço gerado pelos vetores $\{z_1,\ \ldots\ ,\ z_k\}$ é invariante pela ação de A e a que matriz A, no mesmo subespaço, tem a forma

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ h_k & h_{k-1} & h_{k-2} & \ldots & h_2 & h_1 \end{pmatrix}. \tag{4.2}$$

Exercício. Justifique isso.

A matriz A na base $\{z_1,\ \ldots\ ,\ z_n\}$ tem portanto a seguinte forma:

$$A \ = \ \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix},$$

onde A_1 é a matriz $k \times k$ definida em (4.2), A_2 uma matriz $(n-k) \times k$ cujos elementos de matriz são todos nulos, A_3 é uma matriz $k \times (n-k)$ e A_4 é uma matriz $(n-k) \times (n-k)$. Não nos será necessário especificar os elementos de matriz das matrizes A_3 e A_4.

Alguns segundos (minutos?) de meditação nos levam a concluir que, como os elementos de matriz de A_2 são todos nulos, o polinômio característico q pode ser escrito como

$$q(x) \ = \ \det(x\mathbb{1} - A) \ = \ \det(x\mathbb{1} - A_1)\det(x\mathbb{1} - A_4).$$

Exercício. Realize essa meditação. Sugestão: demonstração por indução.

Vamos denominar $q_k(x) = \det(x\mathbb{1} - A_1)$ e $r_k(x) = \det(x\mathbb{1} - A_4)$. Claramente, $q(x) = q_k(x)r_k(x)$. Não será necessário no que segue calcular r_k, mas precisaremos calcular q_k.

Vamos mostrar por indução finita em k que vale a seguinte expressão:

$$q_k(x) \;=\; x^k - (h_1 x^{k-1} + \cdots + h_{k-1}x + h_k).$$

Temos que

$$q_k(x) \;=\; \det(x\mathbb{1} - A_1) \;=\; \det \begin{pmatrix} x & -1 & 0 & \ldots & 0 & 0 \\ 0 & x & -1 & \ldots & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \ldots & x & -1 \\ -h_k & -h_{k-1} & -h_{k-2} & \ldots & -h_2 & x - h_1 \end{pmatrix}. \tag{4.3}$$

Para $k = 2$ vale

$$q_2(x) \;=\; \det \begin{pmatrix} x & -1 \\ -h_2 & x - h_1 \end{pmatrix} \;=\; x^2 - h_1 x - h_2.$$

Para $k > 2$, tem-se, pelas bem conhecidas regras de cálculo de determinantes,

$$\begin{aligned} q_k(x) \;&=\; x \det \begin{pmatrix} x & -1 & \ldots & 0 & 0 \\ 0 & x & \ldots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \ldots & x & -1 \\ -h_{k-1} & -h_{k-1} & \ldots & -h_2 & x - h_1 \end{pmatrix} + 1 \det \begin{pmatrix} 0 & -1 & \ldots & 0 & 0 \\ 0 & x & \ldots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \ldots & x & -1 \\ -h_k & -h_{k-2} & \ldots & -h_2 & x - h_1 \end{pmatrix} \\ &=\; xq_{k-1}(x) + (-1)^{k+1} h_k \det \begin{pmatrix} -1 & 0 & \ldots & 0 & 0 \\ x & -1 & \ldots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \ldots & x & -1 \end{pmatrix} \\ &=\; xq_{k-1}(x) + (-1)^{k+1} h_k (-1)^{k-2} \\ &=\; xq_{k-1}(x) - h_k \end{aligned} \tag{4.4}$$

Exercício. Complete os detalhes.

Assim, se pela hipótese indutiva q_{k-1} é da forma

$$q_{k-1}(x) \;=\; x^{k-1} - (h_1 x^{k-2} + \cdots + h_{k-2}x + h_{k-1}).$$

seque de (4.4) que

$$\begin{aligned} q_k(x) \;&=\; x(x^{k-1} - (h_1 x^{k-2} + \cdots + h_{k-2}x + h_{k-1})) - h_k \\ &=\; x^k - (h_1 x^{k-1} + \cdots + h_{k-2}x^2 + h_{k-1}x + h_k) \end{aligned} \tag{4.5}$$

como queríamos provar.

Retomando, temos que $q(A)y = q_k(A)r_k(A)y = r_k(A)q_k(A)y$. Sucede, porém, que $q_k(A)y = 0$. De fato, pelo cômputo acima,

$$q_k(A)y \;=\; A^k y - h_1 A^{k-1} y - \cdots - h_{k-2}A^2 y - h_{k-1}Ay - h_k y$$

que é igual a zero por (4.1).

Logo $q(A)y = 0$. Como y foi escolhido arbitrário, segue que $q(A) = 0$, demonstrando o teorema. ∎

- **O Teorema de Hamilton-Cayley e a Inversa de Matrizes**

O Teorema de Hamilton-Cayley fornece-nos um método de calcular a inversa de matrizes não-singulares.

De fato, se $q(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ é o polinômio característico de uma matriz não-singular A, então o Teorema de Hamilton-Cayley afirma que

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 \mathbb{1} \;=\; 0,$$

ou seja,

$$A\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 \mathbb{1}\right) \;=\; -a_0 \mathbb{1}.$$

Isso tem por implicação

$$A^{-1} \;=\; -\frac{1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 \mathbb{1}\right).$$

Nota. Usando a definição de polinômio característico $q(x) = \det(x\mathbb{1} - A)$, é evidente (tomando-se $x = 0$) que $a_0 = (-1)^n \det(A)$. Assim $a_0 \neq 0$ se A for não-singular.

Em alguns casos é eficiente calcular A^{-1} usando essa fórmula.

Exercício. Use esse método para calcular a inversa das suas matrizes não-singulares favoritas.

5 Matrizes Diagonalizáveis e o Teorema Espectral

- **Matrizes Diagonalizáveis**

Vamos agora apresentar uma noção intimamente ligada a de matriz simples introduzida acima (página 5), mas de importância maior.

Definição. Uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é dita ser *diagonalizável* se existir uma matriz invertível $P \in \mathrm{Mat}(\mathbb{C},\ n)$ tal que $P^{-1}AP$ é uma matriz diagonal, ou seja,

$$P^{-1}AP \;=\; D \;=\; \mathrm{diag}(d_1,\ \ldots,\ d_n) \;=\; \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}.$$

É fácil de se ver que os elementos da diagonal de D são os autovalores de A. De fato, se A é diagonalizável por P, vale para seu polinômio característico

$$\begin{aligned} p(\lambda) \;&=\; \det(\lambda\mathbb{1} - A) \;=\; \det(P^{-1}(\lambda\mathbb{1} - A)P) \;=\; \det(\lambda\mathbb{1} - P^{-1}AP) \;=\; \det(\lambda\mathbb{1} - D) \\ &=\; \det\begin{pmatrix} \lambda - d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda - d_n \end{pmatrix} \;=\; (\lambda - d_1)\cdots(\lambda - d_n), \end{aligned}$$

o que mostra que os d_i são as raízes do polinômio característico de A e, portanto, seus autovalores.

Exercício. Justifique todas as passagens acima.

- **Diagonalização de Matrizes**

O próximo teorema é fundamental no estudo de matrizes diagonalizáveis.

Teorema 5.1 *Uma matriz $A \in Mat(\mathbb{C},\ n)$ é diagonalizável se e somente se possuir um conjunto de n autovetores linearmente independentes, ou seja, se e somente se o sub-espaço gerado pela coleção de todos os auto-vetores de A possuir dimensão n.* □

Prova. Vamos primeiro provar que se $A \in \mathrm{Mat}(\mathbb{C},\ n)$ possui um conjunto de n autovetores linearmente independentes então A é diagonalizável. Para tal vamos construir a matriz P que diagonaliza A.

Seja $\{v^1,\ \ldots,\ v^n\}$ um conjunto de n autovetores linearmente independentes de A, cujos autovalores são $\{d_1,\ \ldots,\ d_n\}$, respectivamente. Vamos denotar as componentes de v^i na base canônica por v^i_j, $j = 1, \ldots, n$. Seja a matriz P definida por $P = [\![v^1,\ \ldots,\ v^n]\!]$, ou seja,

$$P \ = \ \begin{pmatrix} v_1^1 & \cdots & v_1^n \\ \vdots & \ddots & \vdots \\ v_n^1 & \cdots & v_n^n \end{pmatrix}.$$

Como se vê pela construção, a a-ésima coluna de P é formada pelas componentes do vetor v^a. Por (1.1), segue que

$$AP \ = \ [\![Av^1,\ \ldots,\ Av^n]\!] \ = \ [\![d_1 v^1,\ \ldots,\ d_n v^n]\!].$$

É fácil de se ver, porém, que

$$[\![d_1 v^1,\ \ldots,\ d_n v^n]\!] \ = \ \begin{pmatrix} v_1^1 & \cdots & v_1^n \\ \vdots & \ddots & \vdots \\ v_n^1 & \cdots & v_n^n \end{pmatrix} \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix} \ = \ PD.$$

Exercício. Verifique.

Portanto, $AP = PD$. Como, por hipótese, as colunas de P são formadas por vetores linearmente independentes, tem-se que $\det(P) \neq 0$ (por que?). Logo, P é invertível e, portanto, $P^{-1}AP = D$, como queríamos demonstrar.

Vamos provar agora a afirmação recíproca que se A é diagonalizável, então possui n autovetores linearmente independentes. Suponha que exista P tal que

$$P^{-1}AP \ = \ D \ = \ \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}.$$

É evidente que os vetores da base canônica

$$e^1 \ = \ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e^2 \ = \ \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e^n \ = \ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

são autovetores de D com $De^a = d_a e^a$. Logo, $v^a = Pe^a$ são autovetores de A, pois

$$Av^a \ = \ APe^a \ = \ PDe^a \ = \ P(d_a e^a) \ = \ d_a Pe^a \ = \ d_a v^a.$$

Provar que os vetores v^a são linearmente independentes é fácil. Suponha que existam números complexos $\alpha_1,\ \ldots,\ \alpha_n$ tais que

$$\alpha_1 v^1 + \cdots + \alpha_n v^n \ = \ 0.$$

Multiplicando-se à esquerda por P^{-1} teríamos

$$\alpha_1 e^1 + \cdots + \alpha_n e^n \ = \ 0.$$

Como os e^a são obviamente linearmente independentes, segue que $\alpha_1 = \cdots = \alpha_n = 0$, provando que os v^a são linearmente independentes. ■

• Matrizes Diagonalizáveis e Matrizes Simples

Vamos agora discutir a relação entre os conceitos de matriz diagonalizável e o de matriz simples, conceito esse introduzido à página 5. Tem-se a saber o seguinte fato:

Proposição 5.2 *Uma matriz $A \in Mat(\mathbb{C},\ n)$ é diagonalizável se e somente se for simples, ou seja, se e somente se a multiplicidade algébrica de cada um dos seus autovalores coincidir com sua multiplicidade geométrica.* □

Prova. Se A é diagonalizável existe P tal que $P^{-1}AP = D$, diagonal. Como toda matriz diagonal, D é simples. Escrevamos D na forma

$$D \;=\; \mathrm{diag}\left(\underbrace{\alpha_1,\ \ldots\ ,\ \alpha_1}_{a_1 \text{ vezes}},\ \ldots\ ,\ \underbrace{\alpha_r,\ \ldots\ ,\ \alpha_r}_{a_r \text{ vezes}},\right).$$

Um conjunto de n-autovetores de D linearmente independentes é fornecido pelos vetores da base canônica:

$$e^1 \;=\; \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e^2 \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e^n \;=\; \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Os vetores $e^1, \ldots , e^{a_1}$ geram o subespaço de autovetores com autovalor α_1 de D etc.

Para a matriz A, os vetores $Pe^1, \ldots , Pe^{a_1}$ geram o subespaço de autovetores com autovalor α_1 etc. É claro que a dimensão desse subespaço é a_1, pois $Pe^1, \ldots , Pe^{a_1}$ são linearmente independentes, já que os vetores da base canônica $e^1, \ldots , e^{a_1}$ o são. Como isso também vale para os demais autovalores concluímos que A é simples.

Resta-nos agora mostrar que se $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é simples então A é diagonalizável. Como antes, sejam $\alpha_1,\ \ldots\ ,\ \alpha_r$, $1 \leq r \leq n$, seus autovalores <u>distintos</u>, cada qual com multiplicidade algébrica $a_1,\ \ldots\ ,\ a_r$, respectivamente, e seja $\mathcal{E}(\alpha_i)$ o subespaço gerado pelos autovetores com autovalor α_i. Como A é simples, tem-se que a dimensão de $\mathcal{E}(\alpha_i)$ é a_i. Já observamos (página 4) que sub-espaços $\mathcal{E}(\alpha_i)$ associados a autovalores distintos têm em comum apenas o vetor nulo. Assim, se em cada $\mathcal{E}(\alpha_i)$ escolhermos a_i vetores independentes, teremos ao todo um conjunto de $\sum_{i=1}^{r} a_i = n$ autovetores (vide (2.1)) linearmente independentes de A. Pelo Teorema 5.1, A é diagonalizável, completando a prova. ■

• Projetores

Uma matriz $E \in \mathrm{Mat}(\mathbb{C},\ n)$ é dita ser um *projetor* se satisfizer

$$E^2 \;=\; E.$$

Discutiremos várias propriedades importantes de projetores adiante, especialmente de uma classe especial de projetores denominados *projetores ortogonais*. Por ora vamos mostrar duas propriedades que usaremos logo abaixo quando discutirmos o teorema espectral.

A primeira propriedade é a afirmação que se λ é um autovalor de um projetor E então ou λ é igual a zero ou a um. De fato se v é um auto-vetor associado a um autovalor λ de E, tem-se que $Ev = \lambda v$ e $E^2 v = \lambda^2 v$. Como $E^2 = E$, segue que $\lambda^2 v = \lambda v$. Logo $\lambda(\lambda - 1)$ e, portanto, $\lambda = 0$ ou $\lambda = 1$.

A segunda propriedade é uma conseqüência da primeira: o traço de um projetor $E \in \mathrm{Mat}(\mathbb{C},\ n)$ é um número inteiro positivo ou nulo, mas menor ou igual a n. De fato, pela definição, o traço de um projetor E é a soma de seus autovalores. Como os mesmos valem zero ou um a soma é um inteiro positivo ou nulo. Como há no máximo n autovalores a soma não pode exceder n. Na verdade, o único projetor cujo traço vale exatamente n é a identidade $\mathbb{1}$ e o único projetor cujo traço vale exatamente 0 é a matriz nula (por que?).

Essas observações têm a seguinte conseqüência que usaremos adiante. Se $E_1, \ldots ,\ E_r$ são r projetores não-nulos com a propriedade que

$$\mathbb{1} \ = \ \sum_{a=1}^{r} E_a$$

então $r \leq n$. Para ver isso, basta tomar o traço de ambos os lados dessa expressão:

$$\mathrm{Tr}\,(\mathbb{1}) \ = \ \sum_{a=1}^{r} \mathrm{Tr}\,(E_a). \tag{5.1}$$

O lado esquerdo vale n enquanto que o lado direito é uma soma de r inteiros positivos. Obviamente isso só é possível se $r \leq n$.

- **O Teorema Espectral**

O teorema espectral é um dos mais importantes teoremas de toda a Álgebra Linear e, em verdade, de toda Análise Funcional, já que o mesmo possui generalizações para operadores limitados e não-limitados (auto-adjuntos) agindo em espaços de Hilbert. Nessa versão mais geral o teorema espectral é de importância fundamental para a própria interpretação da Física Quântica.

Teorema 5.3 (O Teorema Espectral para Matrizes) *Uma matriz $A \in Mat(\mathbb{C},\ n)$ é diagonalizável se e somente se existirem $r \in \mathbb{N}$, escalares $\alpha_1,\ \ldots ,\ \alpha_r$ e projetores $E_1,\ \ldots ,\ E_r \in Mat(\mathbb{C},\ n)$ tais que*

$$A \ = \ \sum_{a=1}^{r} \alpha_a E_a, \tag{5.2}$$

$$\mathbb{1} \ = \ \sum_{a=1}^{r} E_a$$

e

$$E_i E_j \ = \ \delta_{i,\,j} E_j.$$

Os escalares $\alpha_1,\ \ldots ,\ \alpha_r$ vem a ser os auto-valores de A. □

Vamos demonstrar esse teorema provando um outro um pouco mais geral (Teorema 5.4, abaixo).

Os projetores E_a que surgem em (5.2) são denominados *projetores espectrais* de A. A decomposição (5.2) é frequentemente denominada *decomposição espectral* de A.

Teorema 5.4 *Seja $A \in Mat(\mathbb{C},\ n)$. São equivalentes as seguintes afirmações:*

1. *A possui n autovalores linearmente independentes, ou seja, o sub-espaço gerado pelos autovalores de A tem dimensão n.*

2. *A é diagonalizável, ou seja, existe uma matriz $P \in Mat(\mathbb{C},\ n)$ invertível tal que $P^{-1}AP$ é uma matriz diagonal diag $(d_1,\ \ldots ,\ d_n)$, onde os d_i's são autovalores de A.*

3. *Para todo vetor $x \in \mathbb{C}^n$ e todo escalar $\lambda \in \mathbb{C}$ tais que $(A - \lambda \mathbb{1})^2 x = 0$, vale que $(A - \lambda \mathbb{1})x = 0$.*

4. *Se x é um vetor não-nulo tal que $(A - \lambda \mathbb{1})x = 0$ para algum $\lambda \in \mathbb{C}$ então <u>não</u> existe nenhum vetor y com a propriedade que $(A - \lambda \mathbb{1})y = x$.*

5. *Todas as raízes do polinômio mínimo de A têm multiplicidade 1.*

6. *Existem $r \in \mathbb{N}$, escalares $\alpha_1, \ldots, \alpha_r$ e projetores $E_1, \ldots, E_r \in Mat(\mathbb{C}, n)$ tais que*

$$A = \sum_{a=1}^{r} \alpha_a E_a.$$

Além disso, as matrizes E_a satisfazem

$$\mathbb{1} = \sum_{a=1}^{r} E_a$$

e

$$E_i E_j = \delta_{i,\,j} E_j.$$

□

Demonstração. A prova da equivalência será feita demonstrando-se sucessivamente as seguintes implicações: $1 \to 2$, $2 \to 3$, $3 \to 4$, $4 \to 5$, $5 \to 6$, $6 \to 1$. Que 1 implica 2 já foi demonstrado no teorema 5.1.

$2 \to 3$.

Seja $D = P^{-1}AP$ diagonal. $D = \text{diag}\,(d_1, \ldots, d_n)$. Seja $(A - \lambda \mathbb{1})^2 x = 0$. Segue que

$$P^{-1}(A - \lambda \mathbb{1})^2 P y = 0$$

onde $y = P^{-1}x$. Logo,

$$(D - \lambda \mathbb{1})^2 y = 0,$$

ou seja,

$$\begin{aligned} (d_1 - \lambda)^2 y_1 &= 0 \\ &\vdots \\ (d_n - \lambda)^2 y_n &= 0, \end{aligned}$$

onde y_j são as componentes de y:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Agora, é evidente que se $(d_a - \lambda)^2 y_a = 0$ então $(d_a - \lambda) y_a = 0$. Logo

$$(D - \lambda \mathbb{1}) y = 0.$$

Usando-se $y = P^{-1}x$ e multiplicando-se à direita por P, concluímos que

$$0 = P(D - \lambda \mathbb{1})P^{-1}x = (PDP^{-1} - \lambda \mathbb{1})x = (A - \lambda \mathbb{1})x,$$

que é o que queríamos provar.

$3 \to 4$.

A prova é feita por contradição. Vamos supor que para algum vetor $x \neq 0$ exista $\lambda \in \mathbb{C}$ tal que $(A - \lambda \mathbb{1})x = 0$. Suponhamos também que exista vetor y tal que $(A - \lambda \mathbb{1})y = x$. Teríamos

$$(A - \lambda \mathbb{1})^2 y = (A - \lambda \mathbb{1})x = 0.$$

Pelo item 3 isso implica $(A - \lambda \mathbb{1})y = 0$. Mas isso diz que $x = 0$, uma contradição.

$4 \to 5$.

Seja M o polinômio mínimo de A, ou seja, o polinômio de menor grau tal que $M(A) = 0$. Vamos mostrar que todas as raízes de M têm multiplicidade 1. Vamos, por contradição, supor que haja uma raiz, λ_0, com multiplicidade maior ou igual a 2. Teríamos, para $x \in \mathbb{C}$,

$$M(x) \ = \ p(x)(x - \lambda_0)^2 .$$

Assim $M(A) = p(A)(A - \lambda_0 \mathbb{1})^2 = 0$. Como M é, por definição, o polinômio de menor grau que zera em A, segue que

$$p(A)(A - \lambda_0 \mathbb{1}) \ \neq \ 0.$$

Assim, existe pelo menos um vetor z tal que $p(A)(A - \lambda_0 \mathbb{1})z \neq 0$. Vamos definir um vetor x por $x := p(A)(A - \lambda_0 \mathbb{1})z$. Então

$$(A - \lambda_0 \mathbb{1})x \ = \ (A - \lambda_0 \mathbb{1})p(A)(A - \lambda_0 \mathbb{1})z \ = \ p(A)(A - \lambda_0 \mathbb{1})^2 z \ = \ M(A)z \ = \ 0,$$

pois $M(A) = 0$. Agora, pela definição,

$$x \ = \ (A - \lambda_0 \mathbb{1})y$$

onde $y = p(A)z$. Pelo item 4, porém, isso é impossível.

$5 \to 6$.

Pelo item 5 temos para $x \in \mathbb{C}$,

$$M(x) \ = \ \prod_{j=1}^{r} (x - \alpha_j)$$

onde α_j são as raízes de M. Para $k = 1, \ \ldots, r$ defina-se os polinômios m_k por

$$M(x) \ = \ (x - \alpha_k) m_k(x),$$

ou seja,

$$m_k(x) \ = \ \prod_{\substack{j=1 \\ j \neq k}}^{r} (x - \alpha_j).$$

É claro que $m_k(\alpha_j) = 0 \iff j \neq k$ (por que?).

Vamos agora definir mais um polinômio, g, da seguinte forma:

$$g(x) \ = \ 1 - \sum_{k=1}^{r} \frac{1}{m_k(\alpha_k)} m_k(x).$$

Como os polinômios m_k têm grau $r - 1$, o polinômio g tem grau menor ou igual a $r - 1$. Porém, observe-se que, para todos os α_j, $j = 1, \ldots, r$, vale

$$g(\alpha_j) \ = \ 1 - \sum_{k=1}^{r} \frac{1}{m_k(\alpha_k)} m_k(\alpha_j) = 1 - \frac{m_j(\alpha_j)}{m_j(\alpha_j)} \ = \ 0$$

Assim g tem pelo menos r raízes distintas! O único polinômio de grau menor ou igual a $r - 1$ que tem r raízes distintas é o polinômio nulo. Logo, concluímos que

$$g(x) \ = \ 1 - \sum_{k=1}^{r} \frac{1}{m_k(\alpha_k)} m_k(x) \ \equiv \ 0$$

para todo $x \in \mathbb{C}$. Isso significa que todos os coeficientes de g são nulos. Assim, para qualquer matriz B tem-se $g(B) = 0$. Para a matriz A isso diz que

$$\mathbb{1} = \sum_{k=1}^{r} \frac{1}{m_k(\alpha_k)} m_k(A).$$

Definindo-se

$$E_k = \frac{1}{m_k(\alpha_k)} m_k(A), \tag{5.3}$$

concluímos que

$$\mathbb{1} = \sum_{k=1}^{r} E_k. \tag{5.4}$$

Para todo k vale $0 = M(A) = (A - \alpha_k \mathbb{1}) m_k(A)$, ou seja, $A m_k(A) = \alpha_k m_k(A)$. Pela definição de E_k isso significa

$$AE_k = \alpha_k E_k.$$

Assim, multiplicando-se ambos os lados de (5.4) por A, segue que

$$A = \sum_{k=1}^{r} \alpha_k E_k.$$

Para completar a demonstração de 6, resta-nos provar que $E_i E_j = \delta_{i,\,j} E_j$.

Para $i \neq i$ tem-se pela definição dos E_k's que

$$\begin{aligned}
E_i E_j &= \frac{1}{m_i(\alpha_i) m_j(\alpha_j)} m_i(A) m_j(A) \\
&= \frac{1}{m_i(\alpha_i) m_j(\alpha_j)} \left[\prod_{\substack{k=1 \\ k \neq i}}^{r} (A - \alpha_k \mathbb{1}) \right] \left[\prod_{\substack{l=1 \\ l \neq j}}^{r} (A - \alpha_l \mathbb{1}) \right] \\
&= \frac{1}{m_i(\alpha_i) m_j(\alpha_j)} \left[\prod_{\substack{k=1 \\ k \neq i,\ k \neq j}}^{r} (A - \alpha_k \mathbb{1}) \right] \left[\prod_{l=1}^{r} (A - \alpha_l \mathbb{1}) \right] \\
&= \frac{1}{m_i(\alpha_i) m_j(\alpha_j)} \left[\prod_{\substack{k=1 \\ k \neq i,\ k \neq j}}^{r} (A - \alpha_k \mathbb{1}) \right] M(A) \\
&= 0,
\end{aligned}$$

pois $M(A) = 0$. Resta-nos provar que $E_j^2 = E_j$ para todo j. Multiplicando-se ambos os lados de (5.4) por E_j teremos

$$E_j = \sum_{k=1}^{r} E_j E_k = E_j E_j,$$

já que $E_j E_k = 0$ quando $j \neq k$. Isso completa a demonstração do item 6.

$6 \to 1$.

Notemos primeiramente todo vetor x, os vetores $E_k x$ ou são nulos ou são autovetores de A. De fato, por 6,

$$AE_k x = \sum_{j=1}^{r} \alpha_j E_j E_k x = \alpha_k E_k x .$$

Logo ou $E_k x = 0$ ou $E_k x$ é auto-vetor de A.

Como há no máximo n autovetores, o espaço por eles gerado tem dimensão menor ou igual a n. Por (5.4), porém, vale para todo vetor x que

$$x = \mathbb{1}x = \sum_{k=1}^{r} \alpha_k E_k x .$$

Assim, todo vetor x pode ser escrito como uma combinação linear de autovetores de A, o que significa que o espaço gerado por esses autovetores tem dimensão exatamente igual a n.

Isso completa a demonstração. ■

• Diagonalizabilidade de Projetores

A proposição abaixo é uma aplicação simples do Teorema 5.4 a projetores. A mesma será usada abaixo quando falarmos de diagonalização simultânea de matrizes.

Proposição 5.5 *Seja $E \in Mat(\mathbb{C}, n)$ um projetor, ou seja, tal que $E^2 = E$. Então E é diagonalizável.* □

Prova. Seja $E \in \mathrm{Mat}(\mathbb{C}, n)$ um projetor. Definamos $E_1 = E$ e $E_2 = \mathbb{1} - E$. Então E_2 é também um projetor, pois

$$(E_2)^2 = (\mathbb{1} - E)^2 = \mathbb{1} - 2E + E^2 = \mathbb{1} - 2E + E = \mathbb{1} - E = E_2 .$$

Tem-se também que $E_1 E_2 = 0$, pois

$$E_1 E_2 = E(\mathbb{1} - E) = E - E^2 = E - E = 0 .$$

Fora isso, é óbvio que

$$\mathbb{1} = E_1 + E_2$$

e que

$$E = \alpha_1 E_1 + \alpha_2 E_2 ,$$

com $\alpha_1 = 1$ e $\alpha_0 = 0$. Ora, isso tudo diz que E satisfaz precisamente todas as condições do item 6 do Teorema 5.4. Portanto, pelo mesmo teorema, E é diagonalizável. ■

• O Cálculo Funcional para Matrizes Diagonalizáveis

O teorema espectral tem o seguinte corolário, muitas vezes conhecido como cálculo funcional.

Teorema 5.6 (Cálculo Funcional) *Seja $A \in Mat(\mathbb{C}, n)$ uma matriz diagonalizável e seja*

$$A = \sum_{a=1}^{r} \alpha_a E_a .$$

sua decomposição espectral, de acordo com o Teorema Espectral, o Teorema 5.3. Então para qualquer polinômio p vale

$$p(A) = \sum_{a=1}^{r} p(\alpha_a) E_a . \tag{5.5}$$

□

Prova. Tem-se, pelas propriedades dos E_a's,

$$A^2 \;=\; \sum_{a,\, b=1}^{r} \alpha_a \alpha_b E_a E_b \;=\; \sum_{a,\, b=1}^{r} \alpha_a \alpha_b \delta_{a,\, b} E_a \;=\; \sum_{a=1}^{r} (\alpha_a)^2 E_a .$$

Analogamente, mostra-se que

$$A^m \;=\; \sum_{a=1}^{r} (\alpha_a)^m E_a ,$$

para qualquer $m \in \mathbb{N}$. O resto da prova é trivial. ■

Exercício. Usando (5.5) demonstre novamente o teorema de Hamilton-Cayley, agora apenas para matrizes diagonalizáveis.

• **Diagonalização Simultânea de Matrizes**

Uma matriz $A \in \mathrm{Mat}(\mathbb{C},\ n)$ é dita ser diagonalizada por uma matriz $P \in \mathrm{Mat}(\mathbb{C},\ n)$ se $P^{-1}AP$ for uma matriz diagonal.

Uma questão muito importante é saber quando duas matrizes diagonalizáveis podem ser diagonalizadas por uma mesma matriz P. A resposta é fornecida no próximo teorema.

Teorema 5.7 (Diagonalização Simultânea de Matrizes) *Duas matrizes diagonalizáveis A e $B \in Mat(\mathbb{C},\ n)$ podem ser diagonalizadas pela mesma matriz $P \in Mat(\mathbb{C},\ n)$ se e somente se $AB = BA$, ou seja, se e somente se comutarem entre si.* □

Prova. A parte fácil da demonstração é provar que se A e B podem ser diagonalizadas pela mesma matriz P então A e B comutam entre si. De fato

$$P^{-1}(AB - BA)P \;=\; (P^{-1}AP)(P^{-1}BP) - (P^{-1}BP)(P^{-1}AP) \;=\; 0,$$

pois $P^{-1}AP$ e $P^{-1}BP$ são ambas diagonais e matrizes diagonais sempre comutam entre si (por que?). Assim, $P^{-1}(AB - BA)P = 0$ e, portanto, $AB = BA$.

Vamos agora passar a mostrar que se $AB = BA$ então ambas são diagonalizáveis por uma mesma matriz P.

Sejam $\alpha_1, \ \ldots \ , \alpha_r$ os r autovalores distintos de A e $\beta_1, \ \ldots \ , \beta_s$ os s autovalores distintos de B.

Evocando o teorema espectral, A e B podem ser escritos de acordo com suas decomposições espectrais como

$$A \;=\; \sum_{i=1}^{r} \alpha_i E_i^A$$

e

$$B \;=\; \sum_{j=1}^{s} \beta_j E_j^B ,$$

onde, de acordo com (5.3),

$$E_i^A \;=\; \left\{ \prod_{\substack{k=1 \\ k \neq i}}^{r} (\alpha_i - \alpha_k) \right\}^{-1} \left[\prod_{\substack{k=1 \\ k \neq i}}^{r} (A - \alpha_k \mathbb{1}) \right], \qquad i = 1, \ \ldots \ , r \tag{5.6}$$

e

$$E_j^B \;=\; \left\{\prod_{\substack{k=1\\k\neq j}}^{s}(\beta_j-\beta_k)\right\}^{-1}\left[\prod_{\substack{k=1\\k\neq j}}^{s}(B-\beta_k\mathbb{1})\right], \qquad j=1,\;\ldots\;,\;s. \tag{5.7}$$

Como A e B comutam entre si e como E_i^A e E_j^B, dados em (5.6)-(5.7), são polinômios em A e B, respectivamente, segue que E_i^A e E_j^B também comutam entre si para todo i e todo j.

Com isso, vamos definir

$$Q_{i,\,j} \;=\; E_i^A E_j^B \;=\; E_j^B E_i^A$$

para $i=1,\;\ldots\;,\;r$ e $j=1,\;\ldots\;,\;s$.

Note-se que os $Q_{i,\,j}$'s são projetores pois

$$Q_{i,\,j}^2 \;=\; (E_i^A E_j^B)(E_i^A E_j^B) \;=\; (E_i^A)^2(E_j^B)^2 \;=\; E_i^A E_j^B \;=\; Q_{i,\,j}.$$

Fora isso, é fácil ver que,

$$Q_{i,\,j}Q_{k,\,l} \;=\; \delta_{i,\,k}\delta_{j,\,l}\;Q_{i,\,j}. \tag{5.8}$$

Exercício. Mostre isso.

Note-se também que

$$\mathbb{1} \;=\; \sum_{i=1}^{r}\sum_{j=1}^{s}Q_{i,\,j}, \tag{5.9}$$

pois

$$\sum_{i=1}^{r}\sum_{j=1}^{s}Q_{i,\,j} \;=\; \sum_{i=1}^{r}\sum_{j=1}^{s}E_i^A E_j^B \;=\; \left(\sum_{i=1}^{r}E_i^A\right)\left(\sum_{j=1}^{s}E_j^B\right) \;=\; \mathbb{1}\mathbb{1} \;=\; \mathbb{1}.$$

Afirmamos que podemos escrever

$$A \;=\; \sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{i,\,j}^{A}Q_{i,\,j} \tag{5.10}$$

e

$$B \;=\; \sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{i,\,j}^{B}Q_{i,\,j}, \tag{5.11}$$

onde $\gamma_{i,\,j}^{A}=\alpha_i$ e $\gamma_{i,\,j}^{B}=\beta_j$. De fato, com essas definições,

$$\sum_{i=1}^{r}\sum_{j=1}^{s}\gamma_{i,\,j}^{A}Q_{i,\,j} \;=\; \sum_{i=1}^{r}\sum_{j=1}^{s}\alpha_i E_i^A E_j^B \;=\; \left(\sum_{i=1}^{r}\alpha_i E_i^A\right)\left(\sum_{j=1}^{s}E_j^B\right) \;=\; A\mathbb{1} \;=\; A.$$

Para B a demonstração é análoga.

Nas relações (5.10) e (5.11) é possível fazer simplificações em função do fato de que nem todos os projetores $Q_{i,\,j}$ são não-nulos. Seja $\mathcal{Q}_1\;\ldots\;,\;\mathcal{Q}_t$ a lista dos projetores $Q_{i,\,j}$ não-nulos, ou seja,

$$\{\mathcal{Q}_1\;\ldots\;,\;\mathcal{Q}_t\} \;=\; \{Q_{i,\,j}|\;Q_{i,\,j}\neq 0,\;i=1,\ldots,r \text{ e } j=1,\ldots,s\}.$$

É evidente por (5.8) que os $\mathcal{Q}_k$'s são projetores e que

$$\mathcal{Q}_k \mathcal{Q}_l = \delta_{k,\,l}\, \mathcal{Q}_k .$$

Por (5.9), tem-se

$$\mathbb{1} = \sum_{k=1}^{t} \mathcal{Q}_k \tag{5.12}$$

e por (5.10) e (5.11)

$$A = \sum_{k=1}^{t} \chi_k^A\, \mathcal{Q}_k \tag{5.13}$$

$$B = \sum_{k=1}^{t} \chi_k^B\, \mathcal{Q}_k \tag{5.14}$$

onde as constantes χ_k^A e χ_k^B estão relacionadas de modo óbvio com $\gamma_{i,\,j}^A$ e $\gamma_{i,\,j}^B$, respectivamente.

Em (5.13) e (5.14) vemos que A e B, por serem diagonalizáveis e por comutarem entre si, têm decomposições espectrais com os mesmos projetores espectrais. Note-se também que, pela observação feita no tópico **Projetores**, à página 13 (vide equação (5.1)), tem-se $1 \leq t \leq n$.

Vamos agora completar a demonstração que A e B podem ser diagonalizados por uma mesma matriz invertível P.

Seja $\mathcal{E}_k$ o subespaço dos autovalores de $\mathcal{Q}_k$ com autovalor 1. Sub-espaços $\mathcal{E}_k$'s diferentes têm em comum apenas o vetor nulo. De fato, se $k \neq l$ e w é um vetor tal que $\mathcal{Q}_k w = w$ e $\mathcal{Q}_l w = w$ então, como $\mathcal{Q}_k \mathcal{Q}_l = 0$ seque que

$$0 = (\mathcal{Q}_k \mathcal{Q}_l) w = \mathcal{Q}_k (\mathcal{Q}_l w) = \mathcal{Q}_k w = w .$$

Seja d_k a dimensão do subespaço $\mathcal{E}_k$ e seja

$$u_k^1, \ \ldots, \ u_k^{d_k}$$

um conjunto de d_k vetores linearmente independentes em $\mathcal{E}_k$. Notemos que d_k coincide com a multiplicidade algébrica do autovalor 1 de $\mathcal{Q}_k$, pois, conforme diz a Proposição 5.5, o projetor $\mathcal{Q}_k$ é diagonalizável e, portanto, é uma matriz simples (Proposição 5.2).

Como $\mathbb{1} = \sum_{k=1}^{t} \mathcal{Q}_k$, tem-se, tomando-se o traço, que $n = \sum_{k=1}^{t} d_k$.

Pelas definições, temos que

$$\mathcal{Q}_l u_k^a = \delta_{k,\,l}\, u_k^a , \tag{5.15}$$

pois $\mathcal{Q}_k u_k^a = u_k^a$ e, portanto, $\mathcal{Q}_l u_k^a = \mathcal{Q}_l(\mathcal{Q}_k u_k^a) = (\mathcal{Q}_l \mathcal{Q}_k) u_k^a = 0$ para $k \neq l$.

Afirmamos que o conjunto de vetores

$$u_1^1, \ \ldots, \ u_1^{d_1}, \ u_2^1, \ \ldots, \ u_2^{d_2}, \ \ldots u_t^1, \ \ldots, \ u_t^{d_t} \tag{5.16}$$

é um conjunto de n vetores linearmente independente. De fato, suponha que existam constantes $c_{k,\,j}$ tais que

$$\sum_{k=i}^{t} \sum_{j=1}^{d_k} c_{k,\,j}\, u_k^j = 0 .$$

Aplicando-se à direita $\mathcal{Q}_l$ teríamos

$$\sum_{j=1}^{d_l} c_{l,\,j}\; u_l^j \;=\; 0,$$

o que só é possível se $c_{l,\,j} = 0$ para todo j pois $u_l^1, \ldots, u_l^{d_l}$, foram escolhidos linearmente independentes. Como l é arbitrário, concluímos que $c_{l,\,j} = 0$ para todo l e todo j, o que mostra que o conjunto de vetores em (5.16) é linearmente independente.

Seja então a matriz $P \in \mathrm{Mat}\,(\mathbb{C},\; n)$ definida por

$$P \;=\; [\![u_1^1, \ldots, u_1^{d_1}, u_2^1, \ldots, u_2^{d_2}, \ldots u_t^1, \ldots, u_t^{d_t}]\!].$$

P é invertível pois o conjunto (5.16) é linearmente independente (e, portanto, $\det(P) \neq 0$).

Tem-se,

$$AP \;=\; [\![Au_1^1, \ldots, Au_1^{d_1}, Au_2^1, \ldots, Au_2^{d_2}, \ldots Au_t^1, \ldots, Au_t^{d_t}]\!]$$

Escrevendo $A = \sum_{l=1}^t \chi_l^A\; \mathcal{Q}_l$ (5.13) e usando (5.15), temos

$$Au_k^a \;=\; \sum_{l=1}^{t} \chi_l^A\; \mathcal{Q}_l u_k^a \;=\; \chi_k^A u_k^a.$$

Assim,

$$AP \;=\; [\![\chi_1^A u_1^1, \ldots, \chi_1^A u_1^{d_1}, \chi_2^A u_1^1, \ldots, \chi_2^A u_1^{d_1}, \ldots, \chi_t^A u_t^1, \ldots, \chi_t^A u_t^{d_t}]\!] \;=\; PD_A,$$

onde

$$D_A \;=\; \mathrm{diag}\left(\underbrace{\chi_1^A, \ldots, \chi_1^A}_{d_1 \text{ vezes}}, \underbrace{\chi_2^A, \ldots, \chi_2^A}_{d_2 \text{ vezes}}, \ldots, \underbrace{\chi_t^A, \ldots, \chi_t^A}_{d_t \text{ vezes}}\right).$$

Portanto,

$$P^{-1}AP \;=\; D_A.$$

Analogamente,

$$BP \;=\; [\![Bu_1^1, \ldots, Bu_1^{d_1}, Bu_2^1, \ldots, Bu_2^{d_2}, \ldots Bu_t^1, \ldots, Bu_t^{d_t}]\!].$$

Escrevendo $B = \sum_{l=1}^t \chi_l^B\; \mathcal{Q}_l$ (5.14) temos,

$$BP \;=\; [\![\chi_1^B u_1^1, \ldots, \chi_1^B u_1^{d_1}, \chi_2^B u_1^1, \ldots, \chi_2^B u_1^{d_1}, \ldots \;\; \chi_t^B u_t^1, \ldots, \chi_t^B u_t^{d_t}]\!] \;=\; PD_B,$$

onde

$$D_B \;=\; \mathrm{diag}\left(\underbrace{\chi_1^B, \ldots, \chi_1^B}_{d_1 \text{ vezes}}, \underbrace{\chi_2^B, \ldots, \chi_2^B}_{d_2 \text{ vezes}}, \ldots, \underbrace{\chi_t^B, \ldots, \chi_t^B}_{d_t \text{ vezes}}\right).$$

Portanto,

$$P^{-1}BP \;=\; D_B.$$

Isso provou que A e B são diagonalizáveis pela mesma matriz invertível P. A demonstração do Teorema 5.7 está completa. ■