Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Sobre matrizes circulantes

por

Cássio Nunes dos Anjos

Setembro/2015
João Pessoa - PB

Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Sobre matrizes circulantes

por

Cássio Nunes dos Anjos
sob orientação do

Prof. Dr. Antônio de Andrade e Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCEN - UFPB, como requisito parcial para obtenção do título de Mestre em Matemática.

Setembro/2015
João Pessoa - PB

Catalogação na Publicação
Setor de Catalogação e Classificação

A599s Anjos, Cássio Nunes dos.
Sobre matrizes e circulantes / Cássio Nunes dos Anjos. - João Pessoa, 2015.
55 f.

Orientador: Antônio de Andrade e Silva.
Dissertação (Mestrado) – UFPB/PPGM

1. Matrizes circulantes. 2. Raízes de polinômios. 3. Diagonalização de matrizes. I. Título.

UFPB/BC CDU – 519.612(043)

Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Sobre Matrizes Circulantes

por

Cássio Nunes dos Anjos

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCEN - UFPB, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Álgebra

Aprovada por:

Prof. Dr. Antônio de Andrade e Silva - UFPB

Prof^{a}. Dr^{a}. Jacqueline Fabíola Rojas Arancibia - UFPB

Prof. Dr. Orlando Stanley Juriaans - IME-USP

Setembro/2015

Dedicatória

"Quod in vita facimus, in aeternum resonat."

Maximus in Gladiator

A minha mãe Rita de Cássia que sempre trabalhou, lutou e sofreu para que eu pudesse ter tudo que ela não teve.

Agradecimentos

E chegamos ao começo de mais um fim para um novo recomeço. Agradeço a todos aqueles que torceram e torcem por mim.

Primeiramente aos meus pais pelo apoio de sempre e que apesar da vida simples que sempre levaram nunca deixaram se abalar e mesmo não tendo uma graduação sempre ensinram que a maior herança que podiam deixar para mim e minha irmã era o conhecimento.

A minha irmã de sangue Carla, essa que mesmo com todo as brigas, por todas as conversas e debates sobre a vida tentando me fazer alguém melhor. E a minha irmã por escolha Tathy pelos mesmo motivos e outros mais.

Ao meu orientador Andrade por me mostrar o verdadeiro sentido da palavra professor. E também por sua enorme paciência, dedicação, conselhos e conversas matemáticas e não-matemáticas tão úteis academicamente, profissionalmente e pessoalmente.

Aos professores do Programa de Pós-Graduação em Matemática da UFPB: Alberto Masayoshi, Cleto Brasileiro, Eduardo Gonçalves, Elisandra de Fátima, Lizandro Sanchez, Uberlandio Batista e Daniel Pellegrino.

A meus professores da UPE: Maria Cristina, Creuza Silva e Valdir Veneziani (in memorian) por suas orientações, conversas e por me acompanhar por toda a graduação e sempre ajudando no que podiam. Especialmente a Cleomacio Miguel pelos conselhos e por todo incentivo para fazer o mestrado e durante o mestrado.

Aos professores Jacqueline Fabiola e Orlando Stanley por aceitarem fazer parte banca. E também a Jacqueline por sua dedicação na disciplina de Estruturas Álgebricas com a nossa turma na difícil trasição da graduação para a pós-graduação. A Orlando pelo apoio incondicional desde do momento que me conheceu, sempre acreditar em mim e me incentivar a vir aqui tentar e ver que era possível, e principalmente por suas palavras de incetivo quando tudo parecia perdido.

Aos amigos da UPE Gerton Souza, Michel Galvão, Lizandra Islla, Karoline Torres, Renato Britto, Carlos Alberto, Kaliane Ribeiro e Marcelo Oliveira pela companhia e momentos de estudos dos 4 anos de matemática em Petrolina e pelo apoio mesmo distante nesses 2 anos de mais matemática em João pessoa.

Aos amigos de longa data autointitulados "irmandade". E aos amigos de poucos anos e longa convivência: Wasthenny Vasconcelos pelo apoio em basicamente tudo, desde do curso de verão até essa dissertação, Igor Laélio pelo desafio de estudar e dividir apartamento comigo e Isabelly Camila pelo apoio mútuo nas dificuldades ao longo do curso.

Aos amigos feitos ao longo dessa caminhada Manu, Tarcy, Marcius, Zeh, Tony, Pantoja, Sally, Camila, Caio, Raqueline, Jorge, Daniel. pelos momentos de "descontração".

À CAPES, pelo apoio financeiro.

Resumo

Neste trabalho introduzimos a teoria das matrizes circulantes e apresentamos 3 modelos para o seu espaço, sendo um o de uma álgebra comutiva finita. Além disso exibimos uma diagonalização para elas e calculamos os autovalores, autovetores e outros invariantes das matrizes circulantes. Em seguida usamos isso para calcular as raízes de equações polinomiais de grau ≤ 4.

Palavra-chave: matrizes circulantes, matrizes de Fourier, raízes de polinômios, diagonalização de matrizes.

Abstract

In this work we introduce the theory of circulat matrices and present 3 models for you space, being the one of a comutive finite algebra. Beyond this we show a diagonalization for they and calculate the the eigenvalues, eigenvectors and other invariants of circulating matrices. And then use it to calculate the roots of polynomial equations of degree ≤ 4.

Keywords: circulant matrices, Fourier matrices, polynomial roots, matrix diagonalization.

Notações

- $\mathbb{C}-$ Corpo dos números complexos
- $\mathbb{M}_{m\times n}(\mathbb{C})-$ O conjunto das matrizes $m \times n$ cujos elementos estão em $\mathbb{C}$
- $\mathbf{M}^t-$ Transposta de $\mathbf{M}$
- $\mathbf{M}^*-$ Conjugada de $\mathbf{M}$
- Se $\mathbf{M}$ é quadrada

  $\det \mathbf{M}-$ Determinante de $\mathbf{M}$

  $tr(\mathbf{M})-$ Traço de $\mathbf{M}$

  $\mathbf{M}^{-1}-$ Inversa de $\mathbf{M}$
- $Diag(d_1, d_2, ..., d_n) = Diag(d_1, d_2, ..., d_n)^t$

$$= \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

Matrizes quadradas especiais

- $\mathbf{0} = zero = circ(0, 0, ..., 0)$
- $\mathbf{I} = identidade = circ(1, 0, ..., 0)$
- $\mathbf{\Omega} = Diag(1, \epsilon, \epsilon^2, ..., \epsilon^{n-1})$, $\epsilon = e^{\frac{2\pi i}{n}}$, $\pi \approx 3,14...$
- $\mathbf{F} = Matriz\ de\ Fourier$
- $\mathbf{V} = Matriz\ de\ Vandermonde$
- $\mathcal{P}_{n-1}$ = Conjunto dos polinômios de grau $\leq n-1$ com coeficientes em $\mathbb{C}$
- $||\mathbf{M}||_F$ = Norma de Frobenius

Sumário

Introdução

A "Matemática", escreveu Alfred North Whitehead, "é a técnica mais poderosa para a compreensão do padrão e para a análise das relações de padrões." Em sua busca por padrão, no entanto, a própria matemática exibe um padrão. A matemática muitas vezes tem um apelo visual, arranjos espaciais incorporadas em fórmulas podem ser fonte de inspiração matemática e prazer estético.

A teoria de matrizes é interessante. Assim, as matrizes diagonais, matrizes simétricas, matrizes binárias, matrizes semelhantes são atraentes, independentemente das suas aplicações. Nessa mesma categoria se inserem as circulantes. A matriz circulante é aquela em que uma linha básica de números é repetida novamente e novamente, mas com uma mudança de posição.

Temos que teoria matricial para circulantes pode ser resolvida de forma fechada. Assim, as circulantes constituem um conjunto não trivial, mas simples de objetos que nos leva a um dos mótivos as quais se justificam o estudo da teoria circulante, o de aprofundar o conhecimento da teoria das matrizes.

Escritores sobre teoria das matrizes parecem não dão as circulantes a devida atenção. Dessa forma, este trabalho destina-se a servir como uma referência geral em português sobre circulantes, uma vez que a bibliografia nacional nessa área é pequena.

As circulantes são conhecidas pela humanidade desde pelo menos o início do século XIX, quando elas foram reveladas em sua manifestação original como determinantes circulantes. Séculos mais tarde, foram inventadas as matrizes e as circulantes foram reinterpretadas como matrizes. Muitos anos depois, as matrizes tornaram-se parte de uma nova, mais formal e abstrata álgebra, no século XX. Então as circulantes agora poderiam ser vistas como um tipo especial de álgebra, uma subálgebra da álgebra das matrizes.

Alguém pode desconhecer o nome "circulante", no entanto pode ser que as conheça por um nome alternativo o de "matriz cíclica".

As matrizes circulantes tem várias conexões com problemas em física, de processamento de imagem, de probabilidade e estatística, a análise numérica, a teoria dos números, a geometria. A periodicidade embutida em sua teoria significa que as circulantes combinam com análise Fourier e teoria de grupos.

Como as aplicações são principalmente em matemática pura e tecnologia, isso reflete misteriosamente uma dicotomia abstrata - concreta da teoria circulante. Por exemplo, as telecomunicações modernas seriam impossíveis sem a análise de frequência. Com o advento de computação digital, as técnicas de análise de frequência tornaram-se frequentes, principalmente

a da transformada de Fourier discreta. Esta transformação tem uma relação das circulantes em circulantes, tanto assim, que grande parte da teoria da circulantes pode ser considerada como a teoria da transformada de Fourier discreta. Circulantes são importantes na codificação digital; esta é uma tecnologia maravilhosa que permite que os dispositivos que vão desde computadores à leitores de música possam recuperar erros na transmissão e armazenamento de dados e restaurar os dados originais. No entanto, o ímpeto inicial para o estudo da circulantes não era tecnológica, mas sim resolução de problemas na matemática pura, particularmente a teoria dos números. Várias outras aplicações à matemática pura já foram descobertas. Prof. P. Davis por exemplo explora o ponto de vista geométrico da teoria. Seu livro é a base da maioria dos artigos que usam teoria circulante e é também a base do nosso trabalho para introduzir essa teoria.

Dada uma coleção tão imponente de aplicações, é com uma agradável surpresa que vamos descobrir que as matrizes circulante podem ser descritas de maneira muito simples. Além disso, Matrizes circulantes são sempre quadradas e aqui a matriz circulante geral é de tamanho n. No entanto vamos considerar os índice das entradas na matriz com os números 0 a $n-1$ ao invés do convencional de 1 a n.

0.1 Descrição do trabalho

Esta dissertação é constituída de três capítulos.

No Capítulo 1, apresentamos os conceitos básicos relativos a teoria matricial, fixando notação e mostrando resultados que serão usados no desenvolvimento da teoria circulante.

No Capítulo 2, temos nosso objetivo principal o de descrever as circulantes da teoria das matrizes. Porém sem desconsiderar resultados antigos e o contexto álgebrico a qual elas se inserem. Dessa forma oscilamos entre o ponto de vista das circulantes como uma álgebra comutativa e do ponto de vista das mesmas como matrizes enfatizando seus invariantes.

Finalmente no Capítulo 3, fornecemos uma das inúmeras aplicações da teoria circulante. Um fato interessante sobre as matrizes circulante é a facilidade do cálculo de seus autovalores e autovetores usaremos isso para resolver equações polinomias de grau ≤ 4, fazendo uma redescoberta da fórmula de Cardano e mostrando uma forma agradável de passar da resolução de cúbicas para quárticas e finalizamos com exemplos de resolução dessas equações polinomias.

Capítulo 1

Propriedades Básicas

Inicialmente vamos revisar conceitos de álgebra e algumas definições básicas da teoria matricial, além de fixar algumas notações e nomenclaturas. A maioria dos resultados são conhecidos e por isso são apresentados sem demonstração. Apresentamos ainda alguns operadores especiais, definições e teoremas que serão importantes no desenvolvimento do nosso trabalho. O leitor interessado em mais detalhes pode consultar [1] e [11].

1.1 Estruturas algébricas

Em tudo que segue $\mathbb{K}$ é o corpo dos números reais $\mathbb{R}$ ou o corpo dos números complexos $\mathbb{C}$.

Um *espaço vetorial* E sobre um corpo $\mathbb{K}$ (ou um $\mathbb{K}$-espaço vetorial) é um conjunto cujos elementos a qual chamaremos de vetores podem ser somados e multiplicados por escalares, isto é, os elementos do corpo $\mathbb{K}$. E satisfaz as seguintes propriedades para adição e multiplicação por escalar:

1. $v + w \in E$ e $av \in E$ (Fechamento)
2. $v + w = w + v$. (Comutatividade)
3. $(v + w) + x = v + (w + x)$ e $(ab)v = a\,(bv)$. (Associatividade)
4. $a(v + w) = av + aw$ e $(a + b)\,v = av + bv$. (Distributividade)
5. $\exists\ 0 \in E$ tal que $v + 0 = v$ (Elemento neutro)
6. $\exists\ -v \in E$ tal que $v + (-v) = 0$ (Inverso Aditivo)
7. $1v = v$ (Elemento neutro multiplicativo)

para todos a e $b \in \mathbb{K}$ e v, w e $x \in E$.

Seja $S \subset X$ um subconjunto qualquer de um espaço vetorial X. Uma combinação linear de elementos de S é uma soma (finita)

$$\sum_{i=1}^{k} \lambda_i x_i \ \text{ com } \ \lambda_i \in \mathbb{K} \ \text{ e } \ x_i \in S.$$

Dizemos que S é *linearmente dependente (L.D.)*, se $\exists$ um número finito de elementos $x_i's$ em S e $\lambda_i's$ em $\mathbb{K}$ não todos nulos, tais que

$$\sum_{i=1}^{k} \lambda_i x_i = 0,$$

caso contrário, S é dito *linearmente independente (L.I.)*. Diremos que S *gera* o espaço vetorial X se, todo $x \in X$, é combinação linear de elementos de S.

Denotaremos por $[S]$ o subespaço gerado pelo conjunto S. Uma *base* de X é um conjunto ordenado $\mathbf{b}$ que é $L.I.$ e gera X. X têm *dimensão finita*, se possuir uma base com um número finito n de elementos. E denotaremos por $\dim X$ a dimensão de X.

Sejam X e Y espaços vetoriais sobre o corpo $\mathbb{K}$. Uma aplicação satisfazendo

$$\begin{array}{ccccc} T & X & \to & Y \\ & T(x+\lambda y) & \mapsto & Tx + \lambda Ty \end{array}$$

para quaisquer $x, y \in X$ e $\lambda \in \mathbb{K}$ é chamada *transformação linear*. Se $X = Y$, chamaremos de *operador linear*. Se T for uma bijeção, dizemos que T é um *isomorfismo* e que X e Y são isomorformos o qual denotaremos por $X \cong Y$.

Observação 1.1 *Se quisermos enfatizar o corpo dizemos que é uma tranformãção $\mathbb{K}$-linear e um operador $\mathbb{K}$-linear.*

Definiremos a *imagem* de T, ImT e o *núcleo* de T, $\ker T$ como os conjuntos $ImT = \{y \in Y \mid y = Tx\}$ e $\ker T = \{x \in X \mid Tx = 0\}$, respectivamente. O núcleo e a imagem de T são subespaços vetoriais de X e Y, respectivamente.

Definimos o *posto* e a *nulidade* de T respectivamente por $posto(T) = \dim ImT$ e $null(T) = \dim \ker T$. Note que $\dim \ker T + \dim ImT = \dim X$.

Teorema 1.2 *Toda transformação linear $T : \mathbb{K}^n \to \mathbb{K}^m$ é da forma $y = T(x)$ sendo*

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \tag{1.1}$$

onde $x \in \mathbb{K}^n$, $y \in \mathbb{K}^m$ e $a_{ij} \in \mathbb{K}$, para $j = 1, ..., n$ e $i = 1, ..., m$.

Seja $(\mathbb{K}, +, \cdot)$ um corpo. Uma $\mathbb{K}-$álgebra é um conjunto $\mathcal{A}$ munido de três operações binárias:

$$\begin{array}{ccccc} + : & \mathcal{A}\times\mathcal{A} & \to & \mathcal{A} \\ & (\alpha,\beta) & \mapsto & \alpha+\beta \end{array} \quad \begin{array}{ccccc} * & \mathcal{A}\times\mathcal{A} & \to & \mathcal{A} \\ & (\alpha,\beta) & \mapsto & \alpha\beta \end{array} \quad \begin{array}{ccccc} \cdot & \mathbb{K}\times\mathcal{A} & \to & \mathcal{A} \\ & (a,\beta) & \mapsto & a\beta \end{array}$$

que têm as seguintes propriedades:

1. $(\mathcal{A}, +, \cdot)$ é um $\mathbb{K}$-espaço vetorial.

2. $(\mathcal{A}, +, *)$ é um anel.

3. $a(\alpha\beta) = (a\alpha)\,\beta = \alpha\,(a\beta)$ para quaisquer $a \in \mathbb{K}, a, \beta \in \mathcal{A}$.

Uma $\mathbb{K}$-álgebra é dita:

a Associativa se a operação $*$ for associativa.

b Comutativa se a operação $*$ for comutativa.

c Com unidade se $*$ tem um elemento neutro.

Exemplo 1.3 *O anel de polinômios $\mathbb{K}[x]$, munido com a adição e o produto usual de polinômios é uma $\mathbb{K}$-álgebra associativa, comutativa e com unidade e o anel da matrizes $\mathbb{M}_n(\mathbb{K})$ é uma $\mathbb{K}$-álgebra não comutativa com unidade.*

Sejam $\mathcal{A}$ e $\mathcal{B}$ duas $\mathbb{K}$-álgebras, um homomorfismo de $\mathbb{K}$-álgebras é uma função $\varphi : \mathcal{A} \to \mathcal{B}$ tal que:

1. φ é um homomorfismo de anéis.

2. φ é um homomorfismo de $\mathbb{K}$-espaços vetoriais.

Um homomorfismo é dito isomorfismo, quando φ também for uma bijecão. Definimos analogamente, o núcleo e a imagem de uma $\mathbb{K}$-álgebra.

Seja $(\mathcal{A}, +, \cdot, *)$ uma $\mathbb{K}$-álgebra. Um subconjunto $\mathcal{B} \subseteq \mathcal{A}$ é uma $\mathbb{K}$-*subálgebra* quando $\mathcal{B}$ for ao mesmo tempo um subanel de $(\mathcal{A}, +, \cdot)$ e um subespaço do espaço vetorial $(\mathcal{A}, +, *)$. E um subconjunto $I \subseteq \mathcal{A}$ é um *ideal de* $\mathcal{A}$ se é ao mesmo tempo um ideal do anel $(\mathcal{A}, +, \cdot)$ e um subespaço do espaço vetorial $(\mathcal{A}, +, *)$.

Dados uma $\mathbb{K}$-álgebra $\mathcal{A}$ e um ideal I, a álgebra quociente $\frac{\mathcal{A}}{I}$ é definida como o anel quociente $\frac{\mathcal{A}}{I}$, munido da estrutura de $\mathbb{K}$-espaço vetorial quociente $\frac{\mathcal{A}}{I}$.

Teorema 1.4 *Sejam $\mathcal{A}$ e $\mathcal{B}$ $\mathbb{K}$-álgebras e $\varphi : \mathcal{A} \to \mathcal{B}$ um homormorfismo de $\mathbb{K}-$álgebras. Então*

1. $\ker\varphi$ é um ideal de $\mathcal{A}$.

2. $Im\varphi$ é uma subálgebra de $\mathcal{B}$.

3. $\frac{\mathcal{A}}{\ker\varphi} \cong Im\varphi$.

Prova. A título de ilustração provaremos 1. De fato, é um subgrupo aditivo pois

$$\varphi(0) = 0 \text{ e } \varphi(a+b) = \varphi(a) + \varphi(b),$$

se $a, b \in \ker\varphi$. Além disso

$$b \in \mathcal{A}, a \in \ker\varphi \Rightarrow \varphi(ab) = \varphi(a)\varphi(b) = 0\varphi(b) = 0 \Rightarrow ab \in \ker\varphi.$$

■

Observação 1.5 *Para uma $\mathbb{K}$-álgebra $\mathcal{A}$. Definimos a base e dimensão de $\mathcal{A}$ como a base e a dimensão do $\mathbb{K}$-espaço vetorial.*

1.2 Matrizes e transformações lineares

É conveniente representar os coeficientes a_{ij} de (1.1) como um arranjo retangular

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ ou } \mathbf{A} = (a_{ij}).$$

denominamos tal arranjo *matriz* $m \times n$ sobre $\mathbb{K}$ sendo m o número de linhas e n o número de colunas com entradas em $\mathbb{K}$. O elemento a_{ij} é a entrada correspondente a linha i e a coluna j. Formalmente, uma matriz $m \times n$ é uma função

$$f : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{K}$$

definida como $f(i, j) = a_{ij}$. A matriz $\mathbf{A} = (a_{ij})$ chama-se uma matriz de *ordem* $m \times n$ e lê-se "m por n". Em particular, se $m = n$, dizemos que $\mathbf{A}$ é uma *matriz quadrada* de ordem n. Uma *submatriz* de $\mathbf{A}$ é uma matriz obtida de $\mathbf{A}$ ao se omitir algumas de suas linhas e/ou colunas.

O teorema 1.2 mostra que existe uma correspondência bijetiva entre o conjunto de matrizes mxn e o espaço das transformações lineares de $\mathbb{K}^n$ em $\mathbb{K}^m$. Denotaremos o elemento a_{ij} da matriz que representa T(com relação às bases canônicas de $\mathbb{K}^n$ e $\mathbb{K}^m$) por

$$T_{ij} = a_{ij} = (T(e_j))_i.$$

A associação acima entre matrizes e transformações lineares é válida para espaços gerais de dimensão finita. De fato, sejam X e Y espaços vetorias de dimensão n e m respectivamente e $T : X \to Y$ uma transformação linear. Então escolhendo uma base arbitrária $\mathbf{b} = \{x_1, ..., x_n\}$ do espaço X e escrevendo $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$, a aplicação linear $B : X \to \mathbb{K}^n$ definida por

$$Bx = (\lambda_1, ..., \lambda_n) = \lambda_1 e_1 + \cdots + \lambda_n e_n$$

é um isomorfismo entre X e $\mathbb{K}^n$. Da mesma forma, ao se escolher uma base $\mathbf{c} = \{y_1, ..., y_m\}$ no espaço Y, obtém-se um isomorfismo C entre Y e $\mathbb{K}^m$. Temos assim o diagrama abaixo, onde as setas verticais sempre indicam isomorfismos:

$$\begin{array}{ccc} & T & \\ & X \quad \to \quad Y & \\ B & \downarrow \qquad \downarrow & C \\ & \mathbb{K}^n \quad \to \quad \mathbb{K}^m & \end{array}$$

$$T_{\mathbb{K}} = C \circ T \circ B^{-1}$$

A aplicação linear $T_{\mathbb{K}}$ é representada por uma matriz $\mathbf{A}$**,** chamaremos de *representação da transformação linear* T com respeito as bases $\mathbf{b}$ e $\mathbf{c}$ e denotaremos $\mathbf{A} = T_{\mathbf{c}}^{\mathbf{b}}$. Dessa forma temos uma associação entre a transformação linear T e a matriz $\mathbf{A}$.

Denotaremos por $M_{m\times n}(\mathbb{K})$ o conjunto das matrizes mxn como entradas no corpo $\mathbb{K}$. Em particular, se $m = n$, dentoraremos $M_{m\times n}(\mathbb{K})$ simplesmente por $M_n(\mathbb{K})$ e sua dimensão é n^2. Note que que $M_{m\times n}(\mathbb{K})$ munido com as operações de adição

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$$

e multiplicação por escalar

$$c\mathbf{A} = (ca_{ij}),\ \ \forall\ c \in \mathbb{K},$$

é um espaço vetorial sobre $\mathbb{K}$. Neste caso, as *matriz unidade* $\mathbf{E}_{ij} = (e_{pq})$, com

$$e_{pq} = \delta_{pi}\delta_{qj} = \begin{cases} 1, & \text{se } (p,q) = (i,j) \\ 0, & \text{se } (p,q) \neq (i,j) \end{cases}$$

em que

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}$$

é o *símbolo de Kronecker*, isto é, $\mathbf{E}_{ij}$ são as matrizes cuja (i,j)-ésima entrada é igual a 1 e as demais zeros, formam uma base canônica para $M_{m\times n}(\mathbb{K})$,.

Observe que dado $\mathbf{A} = (a_{ij}) \in M_{m\times p}(\mathbb{K})$:

1. $$\mathbf{A} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathbf{E}_{ij}.$$

2. $\mathbf{E}_{ij} = \mathbf{E}_{pq}$ se, e somente se, $(p,q) = (i,j)$.

3. $\mathbf{E}_{ij}\mathbf{E}_{pq} = \delta_{jp}\mathbf{E}_{iq}$. Em particular, $\mathbf{E}_{ii}^2 = \mathbf{E}_{ii}$ e $\mathbf{E}_{ij}^2 = \mathbf{0}$ se $i \neq j$.

4. $\sum_{i=1}^{n} \mathbf{E}_{ii} = \mathbf{I}_n$, com $\mathbf{I}_n = (\delta_{ij})$ (ou simplesmente $\mathbf{I} = (\delta_{ij})$) é a *matriz identidade*.

5. $\mathbf{A}\mathbf{E}_{pq} = \sum_{i=1}^{m} a_{ip}\mathbf{E}_{iq}$, isto é, $\mathbf{A}\mathbf{E}_{pq}$ é a matriz cuja q-ésima coluna é igual a p-ésima coluna da matriz $\mathbf{A}$ e as demais zeros

6. $\mathbf{E}_{pq}\mathbf{A} = \sum_{j=1}^{n} a_{qj}\mathbf{E}_{pj}$, isto é, $\mathbf{E}_{pq}\mathbf{A}$ é a matriz cuja p-ésima linha é igual a q-ésima linha da matriz $\mathbf{A}$ e as demais zeros

7. $\mathbf{E}_{pq}\mathbf{A}\mathbf{E}_{rs} = a_{qr}\mathbf{E}_{ps}$, isto é, $\mathbf{E}_{pq}\mathbf{A}\mathbf{E}_{rs}$ é a matriz cuja (p,s)-ésima entrada é igual a a_{qr} e as demais zeros

Dado uma matriz $\mathbf{A} \in M_{mxn}(\mathbb{K})$ denotaremos por $row_i(\mathbf{A})$, a i-ésima linha de M e por $col_j(\mathbf{A})$ a j-ésima coluna de $\mathbf{A}$. Assim podemos particionar a matriz $\mathbf{A}$ tanto em linhas como em colunas e identificar $\mathbf{A}$ com:

$$(row_1(\mathbf{A}); row_2(\mathbf{A}); ...; row_m(\mathbf{A}))$$

e

$$(col_1(\mathbf{A})|col_2(\mathbf{A})|...|col_n(\mathbf{A}))$$

Sejam $\mathbf{A} = (a_{ij}) \in M_{m\times n}(\mathbb{K})$ e $\mathbf{B} = (b_{ij}) \in M_{n\times p}(\mathbb{K})$. Relembre que o *produto* de $\mathbf{A}$ por $\mathbf{B}$ é definido como

$$\mathbf{AB} = (c_{ij}),$$

em que

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \;\; i = 1, \ldots, m \;\text{ e }\; j = 1, \ldots, p.$$

Note que $\mathbf{AB} \in M_{m\times p}(\mathbb{K})$. Para fazer essa multiplicação é muitas vezes conveniente tanto para o trabalho teórico quanto computacional particionar uma matriz em submatrizes. Isto pode ser feito de várias maneiras. Um teorema clássico da teoria das matrizes é:

Teorema 1.6 *Sejam* $\mathbf{A} \in M_{m\times n}(\mathbb{K})$ *e,* $\mathbf{B} \in M_{n\times p}(\mathbb{K})$ *e*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \cdots & \mathbf{A}_{rk} \end{bmatrix} \;\; e \;\; \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{k1} & \mathbf{B}_{k2} & \cdots & \mathbf{B}_{kt} \end{bmatrix}$$

partições de $\mathbf{A}$ *e* $\mathbf{B}$ *em submatrizes tal que cada* $\mathbf{A}_{ij}$ *é uma matriz* $m_i \times n_j$ *e cada* $\mathbf{B}_{jl}$ *é uma matriz* $n_j \times p_l$. *Então* $m_1 + m_2 + \cdots + m_r = m$, $n_1 + n_2 + \cdots + n_k = n$ *e* $p_1 + p_2 + \cdots + p_t = p$. *Para cada* $i = 1, ..., r$ *e* $j = 1, ..., t$, *seja* $\mathbf{C}_{ij} = \sum_{q=1}^{k} \mathbf{A}_{iq}\mathbf{B}_{qj}$, *logo*

$$\mathbf{AB} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{r1} & \mathbf{C}_{r2} & \cdots & \mathbf{C}_{rt} \end{bmatrix}$$

Corolário 1.7 *Sejam* $\mathbf{A} \in M_{m\times n}(\mathbb{K})$, $\mathbf{B} \in M_{n\times p}(\mathbb{K})$, $\mathbf{X} = (x_1, x_2, ..., x_n) \in M_{n\times 1}(\mathbb{K})$ *e* $\mathbf{Y} = (y_1, y_2, ..., y_m) \in M_{1\times m}(\mathbb{K})$, *então*

1. $\mathbf{AX^t} = x_1 col_1(\mathbf{A}) + x_2 col_2(\mathbf{A}) + \cdots + x_n col_n(\mathbf{A})$
2. $\mathbf{YA} = y_1 row_1(M) + y_2 row(M) + \cdots + y_m row_m(M)$
3. $\mathbf{AB} = (\mathbf{A}col_1(\mathbf{B})|\mathbf{A}col_2(\mathbf{B})|...|\mathbf{A}col_p(\mathbf{B}))$
4. $\mathbf{AB} = (row_1(\mathbf{A})\mathbf{B}; row_2(\mathbf{A})\mathbf{B}; ...; row_m(\mathbf{A})\mathbf{B})$

Seja $\mathbf{A} \in M_n(\mathbb{K})$. O *determinante* de $\mathbf{A}$ é um elemento de $\mathbb{K}$ associado a $\mathbf{A}$. Existem várias maneiras de definir o determinante. A mais concisa e às vezes conveniente é dada em termos de permutações é. Para isto, vamos apresentar alguns conceitos e resultados sobre permutações.

Uma *permutação* dos inteiros 1,2,3,...,n é uma bijeção $\sigma : S \to S$, onde $S = \{1, 2, ..., n\}$. Vamos denotar por $S_n = \{\sigma : S \to S \;/\; \sigma \text{ é bijetiva}\}$ o grupo de todas as permutações de S e é

fácil verificar que S_n possui $n!$ permutações distintas. Um exemplo particular de permutação, quando $n > 1$, é dado pelas transposições. Uma *transposição* $\tau : S \to S$ é definida fixando-se dois elementos $i \neq j$ em S, pondo

$$\tau(i) = j, \tau(j) = i \text{ e } \tau(k) = k \text{ se } k \notin \{i, j\}.$$

Se σ,$\tau : S \to S$ são permutações então a função composta $\sigma \circ \tau : S \to S$ também é uma permutação, chamada o *produto* das permutações σ e τ e a indicaremos pela notação $\sigma\tau$. Definimos a *permutação inversa* $\sigma^{-1} : S \to S$ como a aplicação $\sigma^{-1}(i) = j \in S$ tal que $\sigma(j) = i$, caracterizada pelo fato de que $\sigma\sigma^{-1} = \sigma^{-1}\sigma = id$ onde id é a permutação identidade, isto é, a permutação tal que $id(i) = i \ \forall \ i = 1, , , n$.

Existem varias notações para uma permutação $\sigma \in S_n$, as vezes ela é representada pela lista ordenada $(\sigma(1), ..., \sigma(n))$ dos valores que ela assume nos números $1, 2, ..., n$, respectivamente. Como também pode ser escrita sob a forma

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Por exemplo, para $n = 3$, temos que os seis elementos de S_3 são:

$$\begin{aligned} I &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \ \sigma^2 = \sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \\ \tau &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \ \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \ \sigma^2 \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \end{aligned}$$

Ou ainda podemos representar uma permutação por uma matriz $A = (a_{ij})$, sendo

$$a_{ij} = \begin{cases} 1, & j = \sigma(i) \\ 0, & j \neq \sigma(i) \end{cases}$$

chamada representação matricial da permutação.

Exemplo 1.8 *Considere a permutação*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$$

a permutação σ é representada pela matriz

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

Seja $\sigma \in S_n$ uma permutação. Dizemos que σ é um *k-ciclo* se existirem elementos distintos

$$i_1, \ldots, i_k \in S$$

tais que

$$\sigma(i_1) = i_2, \;\; \sigma(i_2) = i_3, \ldots, \sigma(i_k) = i_1$$

e

$$\sigma(x) = x, \;\; \forall \;\; x \in S - \{i_1, \ldots, i_k\},$$

o qual será denota por $\sigma = (i_1 \ldots i_n)$. O número k chama-se o *comprimento* ou a *ordem* do ciclo. Por exemplo,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} = (123)(4)(5) = (123)$$

é um 3-ciclo. As transposições são os ciclos de comprimento dois, também chamados de inversões.

Teorema 1.9 *Qualquer permutação $\sigma \in S_n$, com $\sigma \neq I$, pode ser escrita como um produto de transposições.*

Prova. Vamos usar indução sobre n. Se $n = 1$, nada há para ser provado. Suponhamos que o resultado seja válido para todo k, com $1 \leq k \leq n-1$. Seja $\sigma \in S_n$, com $\sigma \neq I$, tal que $\sigma(n) = k$. Consideremos a transposição $\tau \in S_n$ tal que $\tau(n) = k$ e $\tau(k) = n$. Então $\tau\sigma \in S_n$ e

$$(\tau\sigma)(n) = \tau(k) = n.$$

Assim, podemos ver $\tau\sigma$ como um elemento de S_{n-1}. Logo, por hipótese de indução, existem transposições $\tau_1, \ldots . \tau_m \in S_n$ tais que

$$\tau\sigma = \tau_1 \cdots \tau_m,$$

de modo que

$$\sigma = \tau^{-1}\tau_1 \cdots \tau_m = \tau\tau_1 \cdots \tau_m,$$

que é o resultado desejado. ■

Para cada permutação $\sigma \in S_n$, consideremos o número

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Por exemplo, se $n = 3$ e $\sigma = (123)$ é um 3-ciclo, então

$$\prod_{1 \leq i < j \leq 3} \frac{\sigma(i) - \sigma(j)}{i - j} = \frac{\sigma(1) - \sigma(2)}{1 - 2} \cdot \frac{\sigma(1) - \sigma(3)}{1 - 3} \cdot \frac{\sigma(2) - \sigma(3)}{2 - 3} = 1.$$

Mais geralmente,

$$\prod_{1\leq i<j\leq n}\frac{\sigma(i)-\sigma(j)}{i-j}=\prod_{i=1}^{n-1}\prod_{j=i+1}^{n}\frac{\sigma(i)-\sigma(j)}{i-j}=\pm 1.$$

De fato, como σ é bijetora temos que existem únicos k e l tais que $\sigma(k)=i$ e $\sigma(l)=j$. Se $k<l$, então o fator

$$\sigma(k)-\sigma(l)=i-j$$

aparece no numerador do produto. Se $k>l$, então o fator

$$\sigma(l)-\sigma(k)=j-i=-(i-j)$$

aparece no numerador do produto. Portanto,

$$\frac{\sigma(k)-\sigma(l)}{i-j}=1 \text{ ou } \frac{\sigma(l)-\sigma(k)}{i-j}=-1,$$

e fatores distintos no denominador dão origem a fatores distintos no numerador, ou seja, se $\sigma(k)=\sigma(l)$ e $\sigma(k')=\sigma(l')$, então $k=l$ e $k'=l'$.

Seja $\sigma\in S_n$ uma permutação. Definimos o *sinal* de σ como

$$sgn\sigma=\prod_{1\leq i<j\leq n}\frac{\sigma(i)-\sigma(j)}{i-j}=\prod_{i=1}^{n-1}\prod_{j=i+1}^{n}\frac{\sigma(i)-\sigma(j)}{i-j}.$$

Assim, $sgn\sigma=-1$ ou $sgn\sigma=1$. Dizemos que σ é uma *permutação par* se $sgn\sigma=1$ e uma *permutação ímpar* se $sgn\sigma=-1$. Observe que transposições são sempre permutações ímpares, pois se

$$\tau=\begin{pmatrix}1&2&\cdots&i&\cdots&j&\cdots&n-1&n\\1&2&\cdots&j&\cdots&i&\cdots&n-1&n\end{pmatrix}\in S_n,\ \text{com } i<j,$$

então o número de inversões (cruzamentos) é $1+2(j-i-1)$, o qual é um número ímpar. Portanto, $sgn\tau=-1$. Consequentemente, pelo teorema 1.9,

$$sgn\sigma=(-1)^N,$$

em que N é o número de transposições na decomposição de σ. Neste caso, σ é uma permutação par se, e somente se, N é um número par.

Seja $\mathbf{A}\in M_n(\mathbb{F})$. O determinante de $\mathbf{A}$ é definido como

$$\det(\mathbf{A})=\sum_{\sigma\in S_n}(sgn\sigma)\prod_{i=1}^{n}a_{i\sigma(i)},$$

Assim, $\det(\mathbf{A})$ é a soma de $n!$ termos, em que o sinal está bem definido, e qualquer termo tem n elementos, um e somente um, de cada linha e coluna de $\mathbf{A}$. Por exemplo, se $n=3$, então

$$\begin{aligned}\det\mathbf{A} &= (-1)^0a_{11}a_{22}a_{33}+(-1)^2a_{12}a_{23}a_{31}+(-1)^2a_{13}a_{21}a_{32}\\ &\quad +(-1)^1a_{11}a_{23}a_{32}+(-1)^1a_{12}a_{21}a_{33}+(-1)^1a_{13}a_{22}a_{31}\\ &= (a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32})\\ &\quad -(a_{13}a_{22}a_{31}+a_{11}a_{23}a_{32}+a_{12}a_{21}a_{33})\end{aligned}$$

que é exatamente a *Regra de Sarrus*. Observe que podemos agrupar o $\det \mathbf{A}$ sob a forma:

$$\det \mathbf{A} = a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Mais geralmente, pode ser provado a *expansão de Laplace* em relação i-ésima linha (j-ésima coluna) de $\mathbf{A}$:

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}), \;\; i = 1, \ldots, n.$$

com $\mathbf{A}_{ij}$ a matriz obtida de $\mathbf{A}$ eliminando-se a i-ésima linha e j-ésima coluna da matriz $\mathbf{A}$. O escalar $c_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ij})$ chama-se o *cofator* do termo a_{ij} no $\det \mathbf{A}$ e a matriz $\mathbf{C} = (c_{ij}) \in M_n(\mathbb{K})$ chama-se a *matriz dos cofatores* da matriz $\mathbf{A}$.

Seja $\mathbf{A} = (a_{ij}) \in M_{n\times n}(\mathbb{K})$. Dizemos que $\mathbf{A}$ é *invertível* ou *não-singular*, se existir uma matriz $\mathbf{B} = (b_{ij}) \in M_{n\times n}(\mathbb{K})$ tal que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

Caso contrário, $\mathbf{A}$ é *não-invertível* ou *singular*. A matriz $\mathbf{B}$ é chamada de inversa de $\mathbf{A}$ e é denotada por $\mathbf{A}^{-1}$. Se $T : X \to Y$ e $S : Y \to X$ forem aplicações lineares invertíveis então valem

$$(ST)^{-1} = T^{-1}S^{-1} \;\text{ e }\; \left(S^t\right)^{-1} = \left(S^{-1}\right)^t.$$

O *traço* de uma matriz quadrada $\mathbf{A} \in M_n(\mathbb{K})$ é definido como a soma dos elementos da sua diagonal

$$tr\mathbf{A} = \sum_{j=1}^{n} a_{jj}.$$

As principais propriedades do traço são:

1. $tr\,(a\mathbf{A}+b\mathbf{B}) = atr\mathbf{A}+btr\mathbf{B}$.
2. $tr\,(\mathbf{AB}) = tr\,(\mathbf{BA})$.
3. $tr\,(\mathbf{A}) = tr\,(\mathbf{S}^{-1}\mathbf{AS})$, com $\mathbf{S}$ *não-singular*.

Seja $\mathbf{A} = (a_{ij}) \in M_{m\times n}(\mathbb{K})$. A *matriz transposta* de $\mathbf{A}$ é a matriz obtida escrevendo-se as linhas da matriz $\mathbf{A}$ como colunas, ou seja,

$$\mathbf{A}^t = (a_{ji}), \;\; i = 1, \ldots, m \;\text{ e }\; j = 1, \ldots, n.$$

A *matriz conjugada* de $\mathbf{A}$ é a matriz obtida substituindo as entradas a_{ij} de $\mathbf{A}$ por $\overline{a}_{ij}$ e será denotada por $\overline{\mathbf{A}}$. Além disso, a matriz transconjugada de $\overline{\mathbf{A}}^t$ será denotada por $\mathbf{A}^*$.

Seja $\mathbf{A} \in M_n(\mathbb{K})$. Dizemos que $\mathbf{A}$ é:

- *diagonal* se $a_{ij} = 0$, quando $i \neq j$, e denotada por $\mathbf{D} = diag(a_{11}, \ldots, a_{nn})$.

- *triangular superior* $a_{ij} = 0$, quando $i > j$.

- *simétrica* se $\mathbf{A}^t = \mathbf{A}$.

- *antissimétrica* se $\mathbf{A}^t = -\mathbf{A}$.

- *Hermitiana* se $\mathbf{A}^* = \mathbf{A}$.

- *anti-Hermitiana* se $\mathbf{A}^* = -\mathbf{A}$.

- *normal* se $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$.

- *unitária* se $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^* = \mathbf{I}$.

- *ortogonal* se $\mathbf{A}^t\mathbf{A} = \mathbf{A}\mathbf{A}^t = \mathbf{I}$.

Observação 1.10 *Os operadores lineares tem a mesma denominação que as matrizes que representam tais operadores com relação a uma base ortogonal. E claramente operadores unitários e hermitianos são normais.*

1.3 Operadores e Polinômios

Sejam $\mathbf{A} \in M_n(\mathbb{K})$ e $\lambda \in \mathbb{K}$. A matriz

$$\lambda\mathbf{I} - \mathbf{A} \in M_n(\mathbb{K})$$

chama-se *matriz característica* de $\mathbf{A}$, a qual é uma função do escalar $\lambda \in \mathbb{K}$.

Seja $\mathbf{A} \in M_n(\mathbb{K})$. Um escalar $\lambda \in \mathbb{F}$ é um *autovalor* de $\mathbf{A}$ se existir $\mathbf{X}$ em $\mathbb{F}^n$, com $\mathbf{X} \neq \mathbf{O}$, tal que

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{O}.$$

O vetor $\mathbf{X}$ é chamado um *autovetor* de $\mathbf{A}$ associado a λ.

Seja $\mathbf{A} \in M_n(\mathbb{K})$. Lembremos que

$$\det(\mathbf{A}) = \sum_{\sigma \in S_n} (sgn\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Assim,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \prod_{i=1}^{n} (a_{ii} - \lambda) + f_{n-2}(\lambda),$$

é um polinômio de grau n em λ, em que $f_{n-2}(\lambda)$ é um polinômio de grau no máximo $n-2$ em λ, pois se $i \neq j$, então os termos

$$\pm \prod_{i=1}^{n} a_{i\sigma(i)}$$

contêm no máximo $n-2$ entradas da diagonal principal de $\mathbf{A}$. Como

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \Leftrightarrow \det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

temos que $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$ é uma equação polinomial de grau n em λ, a saber,

$$\lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_{n-1}\lambda + b_n = 0,$$

em que

$$\begin{aligned}
b_1 &= (-1)^1 tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}, \\
b_2 &= (-1)^2 \sum_{i<j} \det\left(\begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}\right) \\
b_3 &= (-1)^3 \sum_{i<j<k} \det\left(\begin{bmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{bmatrix}\right) \\
&\vdots \\
b_n &= (-1)^n \det(\mathbf{A}).
\end{aligned}$$

O polinômio

$$p_{\mathbf{A}}(x) = \det(x\mathbf{I} - \mathbf{A})$$

será chamado o *polinômio característico* de $\mathbf{A}$. A equação polinomial

$$\det(x\mathbf{I} - \mathbf{A}) = 0$$

será chamada a *equação característica* de $\mathbf{A}$ e as raízes dessa equação são os autovalores de $\mathbf{A}$.

Para cada $\lambda \in \mathbb{K}$, definimos o autoespaço associado a λ, por

$$E_\lambda = ker(A - \lambda I).$$

Teorema 1.11 *Seja $\mathbf{A} \in M_n(\mathbb{K})$ e $\lambda \in \mathbb{F}$. Então as seguintes condições são equivalentes*:

1. *λ é um autovalor de $\mathbf{A}$;*
2. *$\mathbf{A} - \lambda I$ é uma matriz singular, isto é, $E_\lambda \neq \{\mathbf{O}\}$;*
3. *$\det(\mathbf{A} - \lambda I) = 0$, isto é, λ é raiz característica.*

Note, pelo teorema 1.11, que o problema de determinar os autovalores e autovetores de uma matriz $\mathbf{A}$ é equivalente a resolver o sistema homogêneo

$$\sum_{j=1}^{n} (\lambda\delta_{ij} - a_{ij})x_j = 0, \;\; i = 1, \ldots, n.$$

Assim, o sistema possui uma solução não nula se, e somente se, $\det(\lambda I - \mathbf{A}) = 0$ se, e somente se, a matriz característica $\lambda I - \mathbf{A}$ é singular ou não invertível.

Exemplo 1.12 *Determine os autovalores e autovetores da matriz*

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{O} & \mathbf{I}_2 \\ \mathbf{I}_1 & \mathbf{O} \end{pmatrix}.$$

Solução. O polinômio característico de $\mathbf{W}$ é

$$p_{\mathbf{W}}(x) = \det(x\mathbf{I} - \mathbf{A}) = x^3 - 1.$$

Assim, os autovalores de $\mathbf{A}$ em $\mathbb{C}$ são

$$1, \omega = -\frac{1}{2} + \frac{1}{2}\sqrt{3}i \text{ e } \omega^2 = \overline{\omega} = -\frac{1}{2} - \frac{1}{2}\sqrt{3}i.$$

Note que os autovalores dependem do corpo. Para obter os autovetores de $\mathbf{W}$ devemos encontrar $\mathbf{X} \in \mathbb{C}^3$ tal que

$$(\lambda\mathbf{I} - \mathbf{W})\mathbf{X} = \mathbf{O},$$

isto é, resolver o sistema homogêneo

$$\begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -1 & 0 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

É fácil verificar que

$$\mathbf{x}_1 = (1, 1, 1), \ \mathbf{x}_2 = (1, \omega, \omega^2) \text{ e } \mathbf{x}_3 = (1, \omega^2, \omega)$$

são os autovetores de $\mathbf{W}$ associados aos autovalores 1, ω e ω^2, respectivamente. Note que o conjunto

$$\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$$

forma uma base de $\mathbb{C}^3$ sobre $\mathbb{C}$. ■

Sejam E um espaço vetorial de dimensão finita sobre $\mathbb{K}$ e $T : E \to E$ um operador linear. Os autovalores, autovetores e polinômio característico de T são os autovalores, autovetores e polinômio característico de qualquer representação matricial de T em relação a alguma base ordenada de E.

Teorema 1.13 *Sejam $T : E \to E$ um operador linear com $\dim V = n$, cuja representação matricial em relação a alguma base ordenada $\mathbf{a}$ de E é $\mathbf{A} = [T]^{\mathbf{a}}_{\mathbf{a}}$, e*

$$\mathbf{X}_j = [v]_{\mathbf{a}} = (x_{1j}, x_{2j}, ..., x_{nj})^t \in \mathbb{R}^{n\times 1}$$

as coordenadas de um autovetor v de T associado ao autovalor $\lambda_j, j = 1, ..., n$. Se os vetores $\mathbf{X}_1, ..., \mathbf{X}_n$ geram $\mathbb{R}^{n\times 1}$, então a matriz $\mathbf{P} = [x_{ij}]$ é tal que

$$\mathbf{PAP}^{-1} = \mathbf{D} = diag(\lambda_1, \ldots, \lambda_n)$$

Se $T : E \to E$ é um operador linear com $\dim E = n$. Então esse teorema motiva a seguinte definição se existir uma base $\mathbf{b}$ de E tal que a matriz $\mathbf{A} = [\mathbf{T}]_{\mathbf{b}}$ é uma matriz diagonal dizemos que T é *diagonalizável*. E segue o teorema abaixo que caracteriza os operadores lineares diagonálizaveis.

Teorema 1.14 *Sejam $T : E \to E$ um operador linear com* $\dim E = n$. *Um operador linear $T : E \to E$ é* diagonalizável *se, e somente se, existir uma base $\mathbf{b}$ formada por autovetores de T.*

Seja

$$f = c_m x^m + \cdots + c_1 x + c_0 \in \mathbb{K}[x]$$

um polinômio de grau m sobre $\mathbb{K}$ e $\mathbf{A} \in \mathbb{M}_n(\mathbb{K})$. Então $f(\mathbf{A})$ é uma matriz de ordem n sobre $\mathbb{K}$ definida como

$$f(\mathbf{A}) = c_m \mathbf{A}^m + \cdots + c_1 \mathbf{A} + c_0 \mathbf{I}.$$

Note que $f(\mathbf{A})$ é obtida de f substituindo-se a variável x pela matriz A e o escalar c_0 pela matriz escalar $c_0 I$.

Proposição 1.15 *Seja $T : E \to E$ um operador linear tal que $T(u) = \lambda u$, com $u \neq 0$. Então*

$$f(T)(u) = f(\lambda)u, \forall f \in \mathbb{K}[x].$$

A função $f_{\mathbf{A}} : \mathbb{K}[x] \to \mathbb{M}_n(\mathbb{K})$ definida por $f_{\mathbf{A}}(c_m x^m + \cdots + c_1 x + c_0) = c_m A^m + \cdots + c_1 A + c_0 I$ é claramente uma transformação linear.

Teorema 1.16 *Sejam $f, g \in \mathbb{K}[x]$. Seja $\mathbf{A} \in \mathbb{M}_n(\mathbb{K})$ então*

$$(f + g)(\mathbf{A}) = f(\mathbf{A}) + g(\mathbf{A}),$$

$$\mathbf{A} f(\mathbf{A}) = f(\mathbf{A})\mathbf{A}, \ \forall \ f \in \mathbb{K}[x]$$

e

$$f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A}), \ \forall \ f, g \in \mathbb{K}[x].$$

Um *polinômio mínimo* $m_T \in \mathbb{K}[x]$ de uma tranformação linear $T : E \to E$ é o polinômio mônico de menor grau tal que $m_T(T) = 0$.

Observação 1.17 *Se $\mathbf{A}$ é uma matriz sobre um corpo $\mathbb{K}$. E se I é o conjunto dos polinômios f de $\mathbb{K}[x]$ tais que $f(\mathbf{A}) = 0$, então I é um ideal. O polinômio unitário gerador de I é o polinômio mínimo de $\mathbf{A}$ sobre $\mathbb{K}$.*

Proposição 1.18 *Todo transformação linear $T : E \to E$, definido em um espaço E com* $\dim E = n$, *possui um polinómio mínimo.*

Prova. O espaço $\mathcal{L}(E,E)$ de todas as aplicações lineares $T : E \to E$ tem dimensão n^2 (esse espaço é isomorfo ao espaço $\mathbb{M}_n(\mathbb{K})$ de todas as matrizes $n \times n$ com entradas em $\mathbb{K}$. Logo, os operadores lineares $I, T, T^2, ..., T^{n^2}$ são L.D., ou seja. existem $a_0, a_1, ..., a_{n^2} \in \mathbb{K}$ não todos nulos tais que

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

Definindo $p(x) = a_0 + a_1 x + \cdots + a_{n^2} x^{n^2}$, temos que $0 \neq p$ e $p(T) = 0$. dividindo pelo coeficiente do termo de maior grau, temos um polinômio mônico q. Portanto, o polinômio mínimo existe, em decorrência da aplicação do Príncipio da Boa Ordenação ao conjunto de todos os polinômios mônicos que anulam T.

Proposição 1.19 *Se $p(T) = 0$ para um polinômio $p \in \mathbb{K}[x]$ e m_T é um polinômio mínimo de T, então p é um múltiplo de m_T.*

Em particular isso garante a unicidade do polinômio mínimo de T.

Teorema 1.20 (Cayley-Hamilton) *Seja E um espaço de dimensão finita. Se $p_T \in \mathbb{K}[x]$ for o polinômio característico de $T : E \to E$, então $p_T(T) = 0$.*

A prova do teorema acima se encontra em [1].

Corolário 1.21 *Seja $T : E \to E$ um operador do espaço complexo de dimensão finita E. O polinômio mínimo de T é um divisor do polinômio característico de T.*

O espaço $\mathcal{L}(E,E)$ é uma $\mathbb{K}-$álgebra. Se E for um espaço de dimensão finita n, essa $\mathbb{K}-$álgebra pode ser identificada com $\mathbb{M}_n(\mathbb{K})$, escolhendo uma base em E.

Fixado $T \in \mathcal{L}(E,E)$, seja $\mathbb{K}[T]$ o conjunto de todas as aplicações lineares obtidas ao se avaliar o polinômio $p \in \mathbb{K}[x]$ em T, então $\mathbb{K}[T]$ é uma $\mathbb{K}-$subálgebra comutativa de $\mathcal{L}(E,E)$.

Defina

$$\begin{array}{rccc} \varphi: & \mathbb{K}[x] & \to & \mathbb{K}[T] \\ & p & \mapsto & p(T) \end{array}$$

temos que a φ é um homomorfismo de $\mathbb{K}-$álgebras, onde o núcleo de φ é o conjunto dos múltiplos do polinômio mínimo m_T de T.

Seja E um espaço vetorial sobre $\mathbb{K}$. A função $\langle\ ,\ \rangle : E \times E \to \mathbb{K}$ é dita um produto interno se as seguintes condições são satisfeitas:

1. $\langle ax, y\rangle = a\langle x, y\rangle\ \forall\ x, y \in E$ e $a \in \mathbb{K}$. (Linearidade)
2. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle\ \forall\ x, y, z \in E$. (Distributividade)
3. $\langle x, y\rangle = \overline{\langle y, x\rangle}\ \forall\ x, y \in E$. (Simetria conjugada)

4. $\langle x, x\rangle \geq 0$, $\forall\ x \in E$ e $\langle x, x\rangle = 0 \Leftrightarrow x = 0$. (Positividade)

   Se $\mathbb{K} = \mathbb{R}$ dizemos que E é um espaço *euclidiano*. Se $\mathbb{K} = \mathbb{C}$ chamaremos o espaço E e o produto interno de espaço *hermitiano* e *produto hermitiano*.

Para $E = \mathbb{R}^n$ e $E = \mathbb{C}^n$ definimos produto interno canônico respectivamente por

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

e

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

Sejam $x, y \in E$. Dizemos que x e y são *ortogonais* se $\langle x, y\rangle = 0$ e denotamos por $x \perp y$.

Dado um espaço vetorial E sobre o corpo $\mathbb{K}$. Uma *norma* em E é uma função $||\cdot|| : E \to [0, \infty)$ que satisfaz as seguintes propriedades:

1. $||ax|| = |a|\,||x||$, $\forall\ a \in \mathbb{K}$ e $\forall x \in E$. (homogeneidade absoluta)
2. $||x + y|| \leq ||x|| + ||y||$. (Desigualdade de Minkowski)
3. $||x|| > 0$ para $x \neq 0$, se $||x|| = 0 \Leftrightarrow x = 0$. (Positividade)

Se E admite uma norma $||\cdot||$, dizemos que E é um espaço *normado*.

Teorema 1.22 (Pitágoras) *Seja E um espaço com produto interno e $||x|| = \sqrt{\langle x, x\rangle}$. Então, se $x \perp y$, temos*

$$||x + y||^2 = ||x||^2 + ||y||^2 .$$

Proposição 1.23 (Desigualdade de Cauchy-Schwarz) *Seja E um espaço com produto interno. Então, se $||x|| = \sqrt{\langle x, x\rangle}$, $\forall\ x, y \in E$ vale:*

$$|\langle x, y\rangle| \leq ||x||\,||y|| .$$

Prova. Se x e y são linearmente dependentes, então $x + ay = 0$ e a igualdade é facilmente satisfeita. Considere x e y linearmente independentes, isto é, $x + ay \neq 0$ para todo $a \in \mathbb{K}$. Faremos a demonstração para o caso $E = \mathbb{C}$, temos que

$$\begin{aligned} 0 &< \langle x + ay, x + ay\rangle \\ &= \langle x, x\rangle + \langle x, ay\rangle + \langle ay, x\rangle + |a|^2 \langle y, y\rangle \\ &= \langle x, x\rangle + \overline{a}\,\langle x, y\rangle + a\,\langle y, x\rangle + |a|^2 \langle y, y\rangle \end{aligned}$$

Note que

$$\langle x, y\rangle = \exp(i\theta)\,|\langle x, y\rangle| \ \text{ com } \ \theta \in [0, 2\pi) \Rightarrow \overline{\langle x, y\rangle} = \exp(-i\theta)$$

Assim, temos que

$$\langle x, x\rangle + 2Re(a\exp(i\theta))\,|\langle x, y\rangle| + |a|^2 \langle y, y\rangle > 0\ \forall a \in \mathbb{K}$$

Fazendo $\chi = a\exp(i\theta) \in \mathbb{C} \Rightarrow |\chi|^2 = |a|^2$, observe que $|Re(\chi)| \leq |\chi|$ logo temos a inequação do segundo grau em $|\chi|$ que verifica

$$\langle x,x\rangle + 2\,|\chi|\,|\langle x,y\rangle| + |\chi|^2\,\langle y,y\rangle > 0 \;\forall\; |\chi| \in \mathbb{R}$$

portanto

$$4\,|\langle x,y\rangle|^2 - 4\,\langle x,x\rangle\,\langle y,y\rangle < 0 \Rightarrow |\langle x,y\rangle|^2 < \langle x,x\rangle\,\langle y,y\rangle\,.$$

Como $||x|| = \sqrt{\langle x,x\rangle}$ concluímos a demonstração ■

Proposição 1.24 *Todo espaço com produto interno E tem uma norma definida por $||x|| = \sqrt{\langle x,x\rangle}$. Dizemos que essa norma é gerada pelo produto interno $\langle\cdot,\cdot\rangle$.*

Prova. A homogeneidade absoluta e a positividade decorrem respectivamente da homogeneidade e positividade do produto interno.

Mostremos que satisfaz a Desigualdade de Minkowski. Para o caso real basta desenvolver $||x+y||^2$ e aplicar a desigualdade de Cauchy-Schwarz.

Provemos para o caso complexo

$$\begin{aligned} ||x+y||^2 &= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle y,x\rangle} + \langle y,y\rangle \\ &= \langle x,x\rangle + 2Re(\langle x,y\rangle) + \langle y,y\rangle \\ &\leq \langle x,x\rangle + 2\,|Re(\langle x,y\rangle)| + \langle y,y\rangle \\ &\leq \langle x,x\rangle + 2\,|\langle x,y\rangle| + \langle y,y\rangle \end{aligned}$$

aplicando a desigualdade de Cauchy-Schwarz, segue que

$$\begin{aligned} ||x+y||^2 &\leq \langle x,x\rangle + 2\,||x||\,||y|| + \langle y,y\rangle \\ &= ||x||^2 + 2\,||x||\,||y|| + ||y||^2 \\ &= (||x|| + ||y||)^2 \end{aligned}$$

que concluí a demonstração ■

Seja E um espaço vetorial com produto interno. Um subcojunto $X \subset E$ é *ortogonal*, se $x \perp y$ para quaisquer $x, y \in X$. Se além disso, todos os seus vetores forem unitários, então X é *ortonormal*. Dizemos que uma base $\mathbf{b} = \{x_1, ..., x_n\}$ de E é uma *base ortogonal* se $\mathbf{b}$ é um conjunto ortogonal, da mesma forma se $\mathbf{b}$ for ortonomal dizemos que é uma *base ortonormal*.

Teorema 1.25 (Schur) *Todo operador $\mathbb{C}$-linear é triangularizável, ou seja, para toda matriz $\mathbf{A} \in \mathbb{M}_n(\mathbb{C})$ existe uma matriz unitária $\mathbf{U} \in \mathbb{M}_n(\mathbb{C})$ e uma matriz triangular superior $\mathbf{T} \in \mathbb{M}_n(\mathbb{C})$ tais que*

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{T}.$$

Prova. Vamos demostrar por indução sobre n. Para $n = 1$ é claramente é válido. Suponhamos que vale para toda matriz de orden $(n-1) \times (n-1)$. Seja A uma matriz de ordem $n \times n$. Seja λ_1 um autovalor de A e x_1 um de seus autovetores associados, isto é, $\mathbf{A}x_1 = \lambda_1 x_1$. Assuma x_1 unitario, isto é $|x_1| = 1$. Existem $n-1$ vetores $x_2, ..., x_n$ que completam $\{x_1\}$ até formar uma base ortonormal de $\mathbb{C}^n$. A matriz $\mathbf{V} = [x_1|...|x_n]$ onde $x_i^t = col_i(\mathbf{V})$ é unitária, ou seja, $\mathbf{V}^*\mathbf{V} = I$. Então

$$\mathbf{V}^*\mathbf{A}\mathbf{V}e_1 = \mathbf{V}^*\mathbf{A}_1 = \lambda_1^*\mathbf{V}x_1 = \lambda_1 e_1,$$

a matriz $\mathbf{U}^*\mathbf{A}\mathbf{U}$ têm a forma

$$\begin{bmatrix} \lambda_1 & a \\ 0 & \mathbf{A}_1 \end{bmatrix}$$

onde $\mathbf{A}_1$ é uma matriz de ordem n-1 e $a^t \in \mathbb{C}^{n-1}$. Por hipótese de indução existe uma matriz unitaria $\mathbf{V}_1$ de orden $n-1$ tal que $\mathbf{V}_1^*\mathbf{A}\mathbf{V}_1$ é uma matriz triangular superior. Finalmente, a matriz

$$\mathbf{U} = \mathbf{V}\begin{bmatrix} 1 & 0 \\ 0 & \mathbf{V}_1 \end{bmatrix}$$

é uma matriz unitária de ordem n que satisfaz

$$\begin{aligned} \mathbf{U}^*\mathbf{A}\mathbf{U} &= \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{V}_1^* \end{bmatrix}\mathbf{V}^*\mathbf{A}\mathbf{V}\begin{bmatrix} 1 & 0 \\ 0 & \mathbf{V}_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{V}_1^* \end{bmatrix}\begin{bmatrix} \lambda_1 & a \\ 0 & \mathbf{A}_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \mathbf{V}_1 \end{bmatrix}. \end{aligned}$$

Portanto

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

E claramente a matriz acima é triangular superior, logo o teorema está provado. ■

Um operador $U : E \to E$ é *unitário* quando $U^* = U^{-1}$, ou seja,

$$\langle Ux, Uy\rangle = \langle x, y\rangle$$

para todo $x, y \in E$.

Teorema 1.26 *Um operador U é unitário se, e somente se, $||Ux|| = ||x||$, $\forall x \in E$.*

Teorema 1.27 *Seja $\mathbf{U} \in \mathbb{M}_n(\mathbb{C})$ uma matriz unitária. Então, $|\det \mathbf{U}| = 1$.*

Prova. Temos que $\mathbf{U}^* = \overline{\mathbf{U}^t}$ o que implica que $\det \mathbf{U}^* = \det \overline{\mathbf{U}^t} = \det \overline{\mathbf{U}} = \overline{\det \mathbf{U}}$ assim $\det \mathbf{U}\overline{\det \mathbf{U}} = \det \mathbf{U}\mathbf{U}^* = \det \mathbf{I} = 1$, assim $|\det \mathbf{U}| = 1$. ■

Teorema 1.28 (Espectral para operadores complexos) *Seja E um espaço complexo, munido de um produto interno hermitiano. Um operador $\mathbb{C}$-linear $N : E \to E$ é normal se, e somente se, existe em E uma base ornormal formada por autovetores de N.*

Equivalentemente provaremos a versão matricial desse teorema

Teorema 1.29 *Seja $\mathbf{A} \in \mathbb{M}_n(\mathbb{K})$. A fim de que exista uma matriz unitária $\mathbf{U}$ tal que $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$ é diagonal é necessário e suficiente que $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$.*

Prova. Seja $\mathbf{U}$ uma matriz unitária tal que

$$\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{U}$$

seja uma matriz triangular. Então

$$\mathbf{T}^* = (\mathbf{U}^*\mathbf{A}\mathbf{U})^* = \mathbf{U}^*\mathbf{A}^*\mathbf{U}$$

multiplicando as duas matrizes

$$\mathbf{T}\mathbf{T}^* = \mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{U}^*\mathbf{A}^*\mathbf{U} = \mathbf{U}^*\mathbf{A}\mathbf{A}^*\mathbf{U}$$

e

$$\mathbf{T}^*\mathbf{T} = \mathbf{U}^*\mathbf{A}^*\mathbf{U}\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{U}^*\mathbf{A}^*\mathbf{A}\mathbf{U}$$

como $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$, segue que $\mathbf{T}^*\mathbf{T} = \mathbf{T}\mathbf{T}^*$. E uma vez que $\mathbf{T}$ é triangular e normal então é diagonal. Assim $\mathbf{U}^*\mathbf{A}\mathbf{U}$ é diagonal. Reciprocamente, se $\mathbf{D} = \mathbf{U}^*\mathbf{A}\mathbf{U}$ é diagonal então $\mathbf{D}\mathbf{D}^* = \mathbf{D}^*\mathbf{D}$ o que implica que

$$\mathbf{U}^*\mathbf{A}\mathbf{A}^*\mathbf{U} = \mathbf{U}^*\mathbf{A}^*\mathbf{A}\mathbf{U}$$

multiplicando a esquerda por $\mathbf{U}$ e a direita por $\mathbf{U}^*$ obtemos $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$, logo é normal. ■

Corolário 1.30 *Se $H : E \to E$ é hermitiano, então existe uma base ortonormal de E formada por autovetores H. Em particular, os autovalores de H são todos reais*

Corolário 1.31 *Se $U : E \to E$ é unitário, então existe uma base ortonormal de E formada por autovetores U. Em particular, os autovalores de U são números complexos de módulo* 1.

Prova. Como $U^*U = 1$, U tem inversa $U^* = U^{-1}$. Isso implica que U é normal, possuindo assim uma base ortonormal formada por seus autovetores. Se λ for um autovalor de U associado ao autovetor v, então $||Uv|| = ||\lambda v|| = |\lambda|\,||v||$ pelo teorema 1.26 $|\lambda| = 1$. ■

Vamos definir no espaço vetorial $M_n(\mathbb{C})$ uma norma e transforma-lo em um espaço normado. A *norma de Frobenius* (ou do traço). Essa norma em $M_n(\mathbb{C})$ é definida através de um produto interno neste espaço vetorial.

A aplicação que, a cada duas matrizes $\mathbf{A}$ e $\mathbf{B}$ em $M_n(\mathbb{C})$, faz corresponder o número complexo

$$\langle \mathbf{A}, \mathbf{B} \rangle = tr(\mathbf{A}^* \mathbf{B}),$$

é um produto interno complexo.

De fato, a homogeneidade e a distributividade valem

$$\langle \mathbf{A}, \mathbf{B} + aC \rangle = tr(\mathbf{A}^*(\mathbf{B} + aC)) = tr(\mathbf{A}^*\mathbf{B}) + atr(\mathbf{A}^*C) = \langle \mathbf{A}, \mathbf{B} \rangle + a \langle \mathbf{A}, C \rangle$$

é simetrica conjugada

$$\langle \mathbf{A}, \mathbf{B} \rangle = tr(\mathbf{A}^*\mathbf{B}) = tr\left((\mathbf{B}^*\mathbf{A})^*\right) = tr\left(\overline{(\mathbf{B}^*\mathbf{A})^t}\right) = \overline{tr(\mathbf{B}^* A)} = \langle \mathbf{B}, \mathbf{A} \rangle$$

e por fim é positiva, temos que:

$$\mathbf{A}^*\mathbf{B} = \left(\sum_{i=1}^{n} \overline{a}_{ik} b_{ij} \right) = (c_{kj}) = C$$

logo,

$$tr(\mathbf{A}^*\mathbf{B}) = tr\,(C) = \sum_{j=1}^{n} \sum_{i=1}^{n} \overline{a}_{ij} b_{ij}. \tag{1.2}$$

Assim

$$\langle \mathbf{A}, \mathbf{A} \rangle = \sum_{i,,j}^{n} |a_{ij}|^2 \geq 0 \text{ e } \sum_{i,,j}^{n} |a_{ij}|^2 = 0 \Leftrightarrow a_{ij} = 0 \quad \forall\ i, j. \tag{1.3}$$

A norma de Frobenius é a norma (matricial) associada a este produto interno, ou seja, é a aplicação que, a cada matriz em $\mathbb{M}_n(\mathbb{C})$, faz corresponder o número real

$$||\mathbf{A}||_F = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{tr(\mathbf{A}^*\mathbf{A})} = \sqrt{\sum_{i,,j}^{n} |a_{ij}|^2}.$$

Mostremos que $||\mathbf{A}||_F$ cumpre os axiomas de uma norma. A primeira e a segunda condição seguem das equações (1.2) e (1.3) respectivamente.

Provemos a desigualdade triangular:

$$\begin{aligned} ||\mathbf{A} + \mathbf{B}||_F^2 &= \sum_{i,,j}^{n} |a_{ij} + b_{ij}|^2 \\ &\leq \sum_{i,,j}^{n} (|a_{ij}|^2 + |b_{ij}|^2 + 2\,|a_{ij}|\,|b_{ij}|) \\ &= \sum_{i,,j}^{n} |a_{ij}|^2 + \sum_{i,,j}^{n} |b_{ij}|^2 + 2 \sum_{i,,j}^{n} |a_{ij}|\,|b_{ij}| \end{aligned}$$

aplicando duas vezes a desigualdade de Cauchy-Schwarz a norma $|\cdot|$ uma com respeito ao índices i e outra a j temos

$$\sum_{i,,j}^{n} |a_{ij}|\,|b_{ij}| \leq \sqrt{\sum_{i,,j}^{n} |a_{ij}|^2}\sqrt{\sum_{i,,j}^{n} |b_{ij}|^2}$$

isto é,

$$||\mathbf{A}+\mathbf{B}||_F^2 \leq ||\mathbf{A}||_F^2 + ||\mathbf{B}||_F^2 + 2\,||\mathbf{A}||_F\,||\mathbf{B}||_F = ||\mathbf{A}||_F^2 + ||\mathbf{B}||_F^2\,.$$

Capítulo 2

Matrizes Circulantes

Neste capítulo vamos introduzir as definições básicas relativas a teoria das matrizes circulantes e apresentar três modelos para o espaço dessas matrizes, incluindo o de uma álgebra comutativa finita. Além disso, vamos estudar outros tipos de matrizes como matrizes de permutação que são um tipo especial de matrizes circulantes e a matriz diagonalizante de uma matriz circulante que é um múltiplo de uma matriz designada por matriz da transformada discreta de Fourier. Essas matrizes como veremos mais adiante aparecem naturalmente no desenvolvimento da teoria circulante.

2.1 Raízes da unidade

Seja $z \in \mathbb{C}$, uma raiz *n-ésima da unidade*, para $n \in \mathbb{N}$, é uma solução da equação $z^n = 1$. Note que

$$\sqrt[n]{1} = \sqrt[n]{1\,(\cos 0 + i \sin 0)} = \sqrt[n]{1}(\cos \frac{0 + 2k\pi}{n} + i \sin \frac{0 + 2k\pi}{n}) = (\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}),$$

para $k = 0, 1, ..., n-1$. Então fixando n, o conjunto das raízes da n-ésimas da unidade é dado por

$$\epsilon_k = \exp\left(\frac{2k\pi i}{n}\right) = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$$

As raízes n-ésimas da unidade desempenham um papel importante em vários ramos da matemática e particularmente na teoria das equações algébricas.

Teorema 2.1 (Fórmula de De Moivre) *Para todo inteiro n e todo número complexo $z = \rho(\cos\theta + i \sin\theta) \neq 0$, temos que*

$$z^n = [\rho(\cos\theta + i \sin\theta)]^n = \rho^n(\cos n\theta + i \sin n\theta). \tag{2.1}$$

Corolário 2.2 *Um número complexo $z \neq 0$ tem m raízes distintas.*

Prova. Tomando $n = 1/m$ em (2.1), temos

$$z^{1/m} = \sqrt[m]{z} = \sqrt[m]{\rho}\left(\cos \frac{\theta + 2k\pi}{m} + i \sin \frac{\theta + 2k\pi}{m}\right).$$

Mostremos que k varia de 0 até $m-1$. Suponha que $k > m$ então existe $p \in \mathbb{N}$ tal que $k = p + m$. Então:

$$\sqrt[m]{z} = \sqrt[m]{\rho}\left(\cos\frac{\theta + 2(p+m)\pi}{m} + i\sin\frac{\theta + 2(p+m)\pi}{m}\right)$$

como

$$\begin{aligned}\cos(a+b) &= \cos a\cos b - \sin a\sin b\\ \sin(a+b) &= \sin a\cos b + \sin b\cos a\end{aligned}$$

segue que

$$\sqrt[m]{z} = \sqrt[m]{\rho}\left(\cos\frac{\theta + 2p\pi}{m} + i\sin\frac{\theta + 2p\pi}{m}\right)$$

se $p > m$ ainda, reduzimos novamente fazendo $p = m + p_1$ até chegarmos a $p_i \leq m$. Ou seja, $0 \leq k \leq m-1$, que nos da m possibilidades para as raízes m-ésimas de z. ■

Proposição 2.3 *Sejam ϵ_l e ϵ_k raízes n-ésimas da unidade em $\mathbb{C}$, onde os índices são tomados módulo n, então elas gozam das seguintes propriedades:*

1. $\epsilon_l\epsilon_k = \epsilon_{l+k}$.
2. $\epsilon_l^k = \epsilon_{kl}$, *para todo* $k \in \mathbb{Z}$.
3. $\epsilon_l^n = 1$.
4. $\epsilon_l^{-1} = \overline{\epsilon_l} = \epsilon_{n-l}$.
5. $\epsilon_l\overline{\epsilon_l} = 1$.
6. $\overline{\epsilon_l}^k = \epsilon_l^{-k} = \epsilon_l^{n-k}$.
7. $1 + \epsilon_l + \epsilon_l^2 + ... + \epsilon_l^{n-1} = 0$ *para todo* $\epsilon_l \neq 1$.

Prova. A título de ilustração provaremos o item 1. Temos que:

$$\epsilon_l\epsilon_k = \cos(\frac{2l\pi}{n} + \frac{2k\pi}{n}) + i\sin(\frac{2l\pi}{n} + \frac{2k\pi}{n}) = \cos\frac{2(l+k)\pi}{n} + i\sin\frac{2(l+k)\pi}{n}$$

se $l + k \equiv m mod n$, isso significa que m é o resto da divisão de $l+k$ por n ou seja, existe um inteiro p tal que $l + k = pn + m$ e $m \in \{0, 1, 2, ..., n-1\}$. Então,

$$\begin{aligned}\epsilon_l\epsilon_k &= \cos\frac{2(l+k)\pi}{n} + i\sin\frac{2(l+k)\pi}{n}\\ &= \cos\frac{2(pn+m)\pi}{n} + i\sin\frac{2(pn+m)\pi}{n}\\ &= \cos(2pn + \frac{2m\pi}{n}) + i\sin(2pn + \frac{2m\pi}{n})\\ &= \cos\frac{2m\pi}{n} + i\sin\frac{2m\pi}{n} = \epsilon_m = \epsilon_{l+k}\end{aligned}$$

■

Algumas raízes n-ésimas da unidade tem um destaque maior sobre as outras, vejamos um exemplo que nos será útil tanto para classificar essas raízes como mais adiante na resolução de equações.

Exemplo 2.4 *As raízes cúbicas da unidade são*

$$\begin{aligned}\epsilon^0 &= \cos 0 + i\sin 0 = 1\\ \epsilon^1 &= \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}\\ \epsilon^2 &= \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}\end{aligned}$$

note que $(\epsilon^1)^0 = 1, (\epsilon^1)^{-1} = \epsilon^2, (\epsilon^2)^0 = 1$ *e* $(\epsilon^2)^2 = \epsilon^1$.

Chamaremos *raiz n-ésima primitiva da unidade* aquela cujas potências geram todas as raízes n-ésimas da unidade. Pela proposição $\epsilon_k^n = \epsilon_{nk} = 1 = \epsilon_k^0$, assim $\epsilon_k^{n+j} = \epsilon_{(n+j)k} = \epsilon_{nk+jk} = \epsilon_{nk}\epsilon_{jk} = \epsilon_{jk}$ logo, temos n potências a serem consideradas que variam de 0 a $n-1$.

Proposição 2.5 *A raiz n-ésima da unidade*

$$\epsilon_l = \cos\frac{2l\pi}{n} + i\sin\frac{2l\pi}{n} \quad (l = 0, 1, ..., n-1)$$

é primitiva se e somente se l *e* n *são relativamente primos.*

Prova. Suponha, por absurdo, que l e n não primos relativos, então podemos simplificar a fração $\frac{l}{n}$ ficando $\frac{p}{q}$, com $q < n$. Logo a raiz n-ésima da unidade $\epsilon_l = \cos\frac{2p\pi}{q} + i\sin\frac{2p\pi}{q}$ também é uma raiz q-ésima da unidade ϵ_p. Dessa forma suas potências, uma vez que são raízes q-ésimas da unidade, podem ter no máximo q valores distintos e como $q < n$ não podem dar origem as n raízes n-ésimas da unidade.

Reciprocamente, considere l e n primos relativos e as m potências de ϵ_l com $0 \leq m \leq n-1$

$$\epsilon_l^0, \epsilon_l^1, \epsilon_l^2, ..., \epsilon_l^{n-1}.$$

Potências inteiras de raizes n-ésima da unidade são raízes n-ésima da unidade. Se essas potências forem todas distintas, teremos todas as raízes n-ésimas da unidade. Provemos então que essas potências são todas distintas.

De fato, dados p e $q \in \mathbb{Z}$, com $p < q$ e tome as potências ϵ_l^p e ϵ_l^q.

Suponha que $\epsilon_l^p = \epsilon_l^q$, então

$$\begin{aligned}\epsilon_l^{p-q} &= 1\\ \left(\cos\frac{2l\pi}{n} + i\sin\frac{2l\pi}{n}\right)^{p-q} &= 1\\ \cos\frac{2l(p-q)\pi}{n} + i\sin\frac{2l(p-q)\pi}{n} &= 1\end{aligned}$$

Logo

$$\frac{2l(p-q)\pi}{n} = 2k\pi,\ k \in \mathbb{Z}$$

ou seja, $\frac{l(p-q)}{n} \in \mathbb{Z}$. Uma vez que l e n são primos relativos, $p - q = jn$ onde $j \in \mathbb{Z}$. Absurdo já que $p, q \in \{0, 1, ..., n-1\}$ e como $p < q$ temos $p - q \in \{1, ..., n-1\}$. ■

2.2 Matriz de Permurtação

Fixamos daqui em diante um inteiro positivo $n \geq 2$. Nossas principais estruturas são o espaço vetorial complexo $\mathbb{C}^n$ com o produto hermitiano e o anel das $n \times n$ matrizes complexas $\mathbb{M}_n$. Iremos estudar a multiplicação $\mathbf{M}v$ de matrizes $\mathbf{M} \in \mathbb{M}_n$ por vetores $v \in \mathbb{C}^n$. A este respeito, podemos visualizar v como um vetor coluna. No entanto, por vezes, é útil matematicamente e mais conveniente tipograficamente considerar

$$v = (v_0, v_1, ..., v_{n-1}) \in \mathbb{C}^n$$

como um vetor linha.

Daqui em diante representaremos a base ortonomal usual de $\mathbb{C}^n$, salvo menção explícita em contrário, como:

$$e_i = (\delta_{i,0}, ..., \delta_{i,n-1}), i = 0, ..., n-1, \tag{2.2}$$

onde $\delta_{i,j}$ é o símbolo de Kronecker ($= 1$ para $i = j$ e 0 para $i \neq j$). Denotaremos essa base por $\mathbf{e}$.

Uma matriz de permutação é uma matriz binária a qual tem exatamente um elemento 1 em cada linha e coluna e 0 em todas as outras entradas. Cada uma dessas matrizes representa uma permutação específica de n elementos e, ao ser usada para multiplicar uma outra matriz, produzira permutações nas linhas ou colunas da outra matriz.

Seja σ uma permutação do conjunto $N = \{1, 2, .., n\}$. A *matriz de permutação* P de ordem n, é da forma

$$\mathbf{P}_\sigma = (a_{ij}) \text{ onde } \begin{cases} a_{ij} = 1, & \text{se } j = \sigma(i) \\ a_{ij} = 0, & \text{se } j \neq \sigma(i). \end{cases}$$

A matriz do Exemplo 1.8 é uma matriz de permutação, note as linhas da matriz A são formadas pelos vetores e_i da base canônica do $\mathbb{C}^n$ onde a $\sigma(i)$-ésima coordenada é igual a 1. Seja $\mathbf{E}_j$ designando o *vetor linha unidade* de n componentes o qual tem 1 em sua j-ésima posição e zeros em todas as outras:

$$\mathbf{E}_j = (0, ..., 0, 1, 0, ..., 0).$$

Logo, podemos escrever as matrizes de permutação de ordem n da seguinte forma

$$\mathbf{P}_\sigma = \begin{pmatrix} \mathbf{E}_{\sigma(i_1)} \\ \mathbf{E}_{\sigma(i_2)} \\ \vdots \\ \mathbf{E}_{\sigma(i_n)} \end{pmatrix}.$$

Entre as matrizes de permutação, a matriz $\mathbf{W}$ que iremos definir tem um papel fundamental na teoria das circulantes. Ela corresponde a permutação *deslocamento para frente* $\sigma(1) = 2$, $\sigma(2) = 3, ..., \sigma(n-1) = n, \sigma(n) = 1$, isto é, ao ciclo $\sigma = (123...n)$ gerando um grupo cíclico de ordem n. Assim dado o ciclo $\sigma = (12...n)$, vamos definir a *matriz de permutação fundamental (MPF)* $\mathbf{W}$.

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

que é justamente a matriz identidade com sua linha superior movida para a parte inferior , ela tem muitas propriedades interessantes, vamos analisar algumas delas.

O efeito de multiplicar um vetor, à esquerda por $\mathbf{W}$, consiste em deslocar, inferiormente, os elementos do vetor, passando o último a tomar a posição do primeiro. De fato, considere $\mathbf{X}$ o vetor $(x_1, ..., x_n)^t$. Particionando $\mathbf{W}$ em colunas

$$\mathbf{W} = (e_n|e_1|...|e_{n-1})$$

e aplicando o Corolário 1.7 temos:

$$\mathbf{WX} = x_1e_n + x_2e_1 + ... + x_ne_{n-1} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ x_1 \end{pmatrix} = \begin{pmatrix} x_{\sigma(1)} \\ \vdots \\ x_{\sigma(n-1)} \\ x_{\sigma(n)} \end{pmatrix}. \tag{2.3}$$

Então se $\mathbf{A} = (a_{ij})$ é uma matriz $n \times r$,

$$\mathbf{WA} = (a_{\sigma(i),\, j}), \tag{2.4}$$

isto é, $\mathbf{WA}$ é a matriz $\mathbf{A}$ com as suas linhas permutadas por σ.

Exemplo 2.6 *Seja* $\mathbf{W}$ *a MPF do ciclo* (123) *e uma matriz* 3×3 *A então*

$$\mathbf{WA} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = \begin{pmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_1 & y_1 & z_1 \end{pmatrix}.$$

Analogamente multiplicar um vetor, à direita por $\mathbf{W}$, consiste em deslocar, lateralmente, os elementos do vetor, ou seja,

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \mathbf{W} = \begin{pmatrix} x_n & x_1 & \cdots & x_{n-1} \end{pmatrix} = \begin{pmatrix} x_{\sigma^{-1}(1)} & x_{\sigma^{-1}(2)} & \cdots & x_{\sigma^{-1}(n)} \end{pmatrix}$$

de modo que se $\mathbf{A} = (a_{ij})$ é uma matriz $r \times n$,

$$\mathbf{AW} = (a_{i,\, \sigma^{-1}(j)}), \tag{2.5}$$

isto é, $\mathbf{AW}$ é matriz $\mathbf{A}$ com as suas colunas permutadas por σ^{-1}.

Proposição 2.7 *A composta de duas permutações do conjunto* $\{1, ..., n\}$ *equivale à multiplicação das matrizes de suas permutações.*

Prova. Seja o espaço $\mathbb{C}^n$ e sua base canônica. Podemos identificar o conjunto $\{1, ..., n\}$ com $\{e_1, ..., e_n\}$. Uma permutação do conjunto $\mathbf{e} = \{e_1, ..., e_n\}$ induz uma aplicação linear $T : \mathbb{C}^n \to \mathbb{C}^n$ definida por $T(e_j) = e_{\sigma(j)}$. A matriz de T na base canônica $\mathbf{e}$ é

$$(e_{\sigma(1)}|e_{\sigma(2)}|...|e_{\sigma(1)})$$

que corresponde a matriz da permutação σ

$$\begin{pmatrix} i \\ \sigma(i) \end{pmatrix}.$$

Essa aplicação linear é um isomorfismo linear pois leva base em base. Note que a composição de permutações equivale à composta dessas aplicações lineares. De fato,

$$T_{\sigma\tau}(e_j) = e_{\sigma\tau(j)} = e_{\sigma(\tau(j))} = T_\sigma\left(e_{\tau(j)}\right) = T_\sigma \circ T_\tau(e_j) \ \forall \ j \in \{1, ..., n\}.$$

Sabemos que a composta de aplicações lineares equivale a multiplicação das matrizes que as representam. Daí, segue o resultado. ■

Corolário 2.8 *Dadas duas permutações* σ *e* τ *e suas matrizes de permutação* P_σ *e* P_τ *respectivamente, então o produto dessas matrizes de permutação é uma matriz de permutação correspondente a matriz da composta dessas permutações ,ou seja,* $P_\sigma P_\tau = P_{\sigma\tau}$.

Prova. Segue da Proposição 2.7. ■

Proposição 2.9 *A transposta de uma matriz de permutação é também uma matriz de permutação, e mais é dada pela matriz da permutação inversa.*

Prova. Como

$$j = \sigma(i) \text{ segue que } \sigma^{-1}(j) = i.$$

Assim,

$$(P_\sigma)^* = (a_{ji}) = (a_{j\sigma^{-1}(j)}) = P_{\sigma^{-1}}.$$

Então

$$(P_\sigma)^* P_\sigma = P_{\sigma^{-1}} P_\sigma = P_{id} = I.$$

Portanto,

$$(P_\sigma)^* = (P_\sigma)^{-1} = P_{\sigma^{-1}}.$$

E temos o resultado. ■

Observação 2.10 *De (2.5), (2.4) e da Proposição 2.9, segue que* $P_\sigma A(P_\sigma)^{-1} = P_\sigma A P_{\sigma^{-1}} = (a_{\sigma(i),\, \sigma(j)})$.

Voltemos a matriz $\mathbf{W}$ e vejamos mais algumas de suas interessantes propriedades. Após calcular $\mathbf{W}^2$, obtemos

$$\mathbf{W}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

vemos que $\mathbf{W}^2$ corresponde á permutação σ^2 para o qual

$$\sigma^2(1) = 3, \sigma^2(2) = 4, ..., \sigma^2(n) = 2.$$

Da mesma forma a matriz W^k corresponde a permutação σ^k. Uma vez que $\sigma^n = id$ segue que

$$\mathbf{W}^n = \mathbf{I} \Rightarrow \mathbf{W}^{n-1} = \mathbf{W}^{-1}.$$

Observe também que, uma vez que $\mathbf{W}$ é uma matriz real, $\mathbf{W}^* = \mathbf{W}^t$ e como $\mathbf{W}$ é uma matriz de permutação $\mathbf{W}^* = \mathbf{W}^{-1}$, logo

$$\mathbf{W}^t = \mathbf{W}^{-1} = \mathbf{W}^* = \mathbf{W}^{n-1}.$$

Então da primeira igualdade temos que $\mathbf{W}$ é uma matriz ortogonal, da segunda igualdade que $\mathbf{W}$ é unitária, e portanto $\mathbf{W}$ é uma matriz normal.

Teorema 2.11 *O polinômio característico da MPF* $\mathbf{W} \in \mathbb{M}_n$ *é*

$$p(X) = X^n - 1$$

Além disso, esse também é seu polinômio minimal.

Prova. O polinômio característico de $\mathbf{W}$ é dado pelo $\det(X\mathbf{I} - \mathbf{W})$, ou seja,

$$\det(X\mathbf{I} - \mathbf{W}) = \det \begin{bmatrix} X & -1 & \cdots & 0 & 0 \\ 0 & X & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & X & -1 \\ -1 & 0 & \cdots & 0 & X \end{bmatrix}$$

fazendo a expansão de laplace pela última linha

$$\det(X\mathbf{I} - \mathbf{W}) = (-1)^{n+1}(-1)\det \begin{bmatrix} -1 & 0 & 0 & 0 \\ X & -1 & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & X & -1 \end{bmatrix} + X(-1)^{n+n} \begin{bmatrix} X & -1 & 0 & 0 \\ 0 & X & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ 0 & 0 & \cdots & X \end{bmatrix}$$

note que as duas matrizes do lado direito da equação são triangular superior e triangular inferior e seus respectivos determinantes são $(-1)^{n-1}$ e X^{n-1} logo,

$$\begin{aligned}\det(X\mathbf{I}-\mathbf{W}) &= (-1)^{n+1}(-1)\,(-1)^{n-1}+X(-1)^{n+n}X^{n-1}\\ &= (-1)^{n}(-1)(-1)^{n}+(-1)^{2n}X^{n}\\ &= (-1)^{2n}(X^{n}-1)=X^{n}-1.\end{aligned}$$

Substituindo $\mathbf{W}$ em seu polinômio característico e usando o fato que $\mathbf{W}^n=\mathbf{I}$ segue o resultado. ■

2.3 A matriz de Fourier

Seja S_N o conjunto das funções de $\{0,\ldots,N-1\}$ em $\mathbb{C}$ ou sequências em $\mathbb{C}$ de comprimento N. Seja ϵ uma raiz N-ésima da unidade. A Transformada Discreta de Fourier (DFT) $\mathcal{F}: S_N \to S_N$ é definida por

$$\mathcal{F}(x_\bullet)_k=\sum_{m=0}^{N-1}x_m\epsilon^{-mk},\ \ k=0,...,N-1 \text{ e } x(i)=x_i \tag{2.6}$$

A DFT é uma aproximação da transformada contínua de Fourier de uma função. A relação entre a transformada de Fourier discreta e contínua é bem conhecida. Na nossa discussão da DFT, vamos restringir a nossa atenção para algumas das propriedades que serão usadas ao longo deste trabalho.

Uma propriedade importante da DFT é a unicidade do par x e $\mathcal{F}(x)$, usando a DFT como um operador linear com a transformação direta definida por (2.6) e a transformação inversa (IDFT) definida por $y\in S$, definida por

$$y_l=\frac{1}{N}\sum_{k=0}^{N-1}\mathcal{F}(x)_k\epsilon^{lk},\ \ l=0,...,N-1 \tag{2.7}$$

De fato, substituindo (2.6) em (2.7) temos

$$y_l=\frac{1}{N}\sum_{k=0}^{N-1}\sum_{m=0}^{N-1}x_m\epsilon^{-mk}\epsilon^{lk}$$

e pelo item 3 da Proposição 2.3**,** o caso não nulo nos dá $y=x$.

Observação 2.12 *O fator multiplicando a DFT e a IDFT (aqui* 1*,* $\frac{1}{N}$ *e os sinais dos expoentes são apenas convenções, e diferem em alguns tratamentos. As únicas exigências dessas convenções são que a DFT e a IDFT tenha expoentes de sinais opostos e que o produto de seus fatores de multiplicação seja* $\frac{1}{N}$*.*

Observação 2.13 *Observe que* S_n *é um espaço vetorial sobre* $\mathbb{C}$ *tendo por base canônica* $\{e_i\}_{i=0}^{N}$ *sendo que* $e_i(k)=\delta_{ik}$*.*

Para o nosso estudo das matrizes circulantes iremos utilizar a transformada discreta de fourier normalizada.

Teorema 2.14 *Seja* $\mathbf{F}$ *a matriz da transformada de Fourier na base canônica e* $\mathbf{F}^*$ *sua adjunta. Então* $||\mathbf{F}||_F = \sqrt{n}$.

Prova. Primeiro calculemos a matriz da DFT na base canônica $\{e_i\}_{i=0}^{N-1}$, a DFT de e_j é a sequência $\mathcal{F}(e_j)$ dada por

$$\mathcal{F}(e_j)_k = \sum_{l=0}^{n-1} \delta_{j,l}\epsilon^{-lk} = \delta_{j,j}\epsilon^{-jk}$$

logo temos

$$\begin{aligned}
\mathbf{F}(e_0) &= (\epsilon^{-0\times 0}, \epsilon^{-1\times 0}, ..., \epsilon^{-(n-1)\times 0}) = (1, 1, ..., 1) \\
\mathbf{F}(e_1) &= (\epsilon^{-0\times 1}, \epsilon^{-1\times 1}, ..., \epsilon^{-(n-1)\times 1}) = (1, \epsilon^{-1}, ..., \epsilon^{-(n-1)}) \\
&\vdots \\
\mathbf{F}(e_{n-1}) &= (\epsilon^{-0\times n-1}, \epsilon^{-1\times (n-1)}, ..., \epsilon^{-(n-1)\times (n-1)}) = (1, \epsilon^{-(n-1)}, ..., \epsilon^{-(n-1)(n-1)}).
\end{aligned}$$

Note que $\mathbf{F}^* = \overline{\mathbf{F}}$ então pela Proposição 2.3

$$\begin{aligned}
\mathbf{F}\mathbf{F}^* &= \sum_{k=0}^{n-1} \epsilon^{-k(i-1)}\overline{\epsilon^{-k(j-1)}} \\
&= \sum_{k=0}^{n-1} \epsilon^{-(i-1)k}\overline{\epsilon}^{-k(j-1)} \\
&= \sum_{k=0}^{n-1} \epsilon^{-k(i-1)}\epsilon^{k(j-1)} \\
&= \sum_{k=0}^{n-1} \epsilon^{k(j-i)}
\end{aligned}$$

Se $i = j$ então

$$\sum_{k=0}^{n-1} \epsilon^0 = \sum_{k=0}^{n-1} 1 = n.$$

Se $i \neq j$ note que $\sum_{k=0}^{n-1} \epsilon^{k(j-i)}$ é uma P.G. com n termos, cujo primeiro termo é igual a 1 e razão ϵ^{j-i} logo

$$\sum_{k=0}^{n-1} \epsilon^{k(j-i)} = \frac{\epsilon^{(j-i)n} - 1}{\epsilon^{j-i} - 1} = \frac{1^{j-i} - 1}{\epsilon^{j-i} - 1} = 0$$

assim $\mathbf{F}\mathbf{F}^* = n\mathbf{I}$, e portanto $||F||_F = \sqrt{tr(n\mathbf{I})} = \sqrt{n}$. ■

A *transformada discreta de Fourier normalizada* (NDFT) de x é o vetor $\mathbf{X} = \mathcal{F}(x)$ de $\mathbb{C}^n$ cujas as coordenadas são dadas por

$$\mathbf{X}_k = \frac{1}{\sqrt{n}} \sum_{l=0}^{n-1} x_l \epsilon^{-kl}.$$

A *matriz de Fourier normalizada* de ordem n, é a matriz $\mathbf{E}$ onde

$$\mathbf{E}^* = \frac{1}{\sqrt{n}}(\epsilon^{(i-1)(j-1)}) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \epsilon & \epsilon^2 & \cdots & \epsilon^{n-1} \\ 1 & \epsilon^2 & \epsilon^4 & \cdots & \epsilon^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \epsilon^{n-1} & \epsilon^{2(n-1)} & \cdots & \epsilon^{(n-1)(n-1)} \end{pmatrix} = \frac{1}{\sqrt{n}} \mathbf{F}^*$$

com $0 \leq i, j \leq n-1$.

A sequência $\epsilon^k, k = 0, 1, ...$ é periódica com periódo n, istó é, $\epsilon^{k+n} = \epsilon^k \ \forall \ k \in \mathbb{Z}_+$; dai há somente n elementos distintos em $\mathbf{E}$. Álem disso, elevando ϵ^j a $n-1$ para $1 \leq j \leq n-1$.

$$\epsilon^{j(n-1)} = \epsilon^{nj-j} = \epsilon^{nj}\epsilon^{n-j} = \epsilon^{n(j+1)-j} = \epsilon^{-j} = \epsilon^{n-j}.$$

Portanto, $\mathbf{E}^*$ pode ser escrita alternativamente como

$$\mathbf{E}^* = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \epsilon & \epsilon^2 & \cdots & \epsilon^{n-1} \\ 1 & \epsilon^2 & \epsilon^4 & \cdots & \epsilon^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \epsilon^{n-1} & \epsilon^{(n-2)} & \cdots & \epsilon \end{pmatrix}$$

Teorema 2.15 *A matriz* $\mathbf{E}$ *é simétrica e unitária.*

Prova. A simetria é evidente. Escreva $\mathbf{E}\mathbf{E}^* = \frac{1}{\sqrt{n}}\mathbf{F}\frac{1}{\sqrt{n}}\mathbf{F}^* = (a_{ij})_{i \leq i,j \leq n}$, então

$$\begin{aligned} a_{ij} &= \frac{1}{n} \sum_{k=0}^{n-1} \epsilon^{(i-1)k} \overline{\epsilon^{k(j-1)}} \\ &= \frac{1}{n} \sum_{k=0}^{n-1} \epsilon^{k(i-j)}. \end{aligned}$$

Então se $i = j$

$$\frac{1}{n} \sum_{k=0}^{n-1} \epsilon^{k(i-j)} = \frac{n}{n} = 1,$$

e se $i \neq j$

$$\frac{1}{n} \sum_{k=0}^{n-1} \epsilon^{k(i-j)} = 0.$$

Assim, $\mathbf{E}\mathbf{E}^* = \mathbf{I}$ ■

Corolário 2.16 *Os autovalores de* $\mathbf{E}$ *são* ± 1, $\pm i$, *com as multiplicidades apropriadas.*

Prova. Isto se deduz do fato que $\mathbf{F}$ é uma matriz unitaria e todo autovalor λ de uma matriz unitaria satisfaz $|\lambda| = 1$.

Por uma *matriz de Vandermonde* $V(z_0, z_1, ..., z_{n-1})$ entendemos a matriz de ordem n da forma

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_0 & z_1 & \cdots & z_{n-1} \\ z_0^2 & z_1^2 & \cdots & z_{n-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_0^{n-1} & z_1^{n-1} & \cdots & z_{n-1}^{n-1} \end{pmatrix}$$

Então $V(1, \epsilon, \epsilon^2, ..., \epsilon^{n-1}) = \sqrt{n}\mathbf{F}^*$ e $V(1, \overline{\epsilon}, \overline{\epsilon}^2, ..., \overline{\epsilon}^{n-1}) = \sqrt{n}\overline{\mathbf{F}}^* = \sqrt{n}\mathbf{F}$.

Proposição 2.17 *Seja* $\mathbf{V}$ *a matriz de Vandermonde. Então* $\det \mathbf{V} = \prod_{i>j}(z_i - z_j)$ *em que o produtório é tomado sobre todos os termos* $z_i - z_j$ *com* $i > j$. *Esse determinante é chamado o* determinante de Vandermonde.

Observação 2.18 *A matriz de Fourier* $\mathbf{F}$ *é a matriz de Vandermonde associada ao vetor* $(1, \epsilon, ..., \epsilon^{n-1})$. *Então se nós associamos ao vetor* v *o polinômio* $v(x) = \sum_{i=0}^{n-1} v_i x^i$, *denotando por* $\hat{v}$ *a transformada discreta de Fourier de um vetor, então* $\hat{v}$ *é simplesmente o vetor cujos componentes correspondem à avaliação de* $v(x)$ *das raízes enésima da unidade.*

Agora que sabemos que a matriz $\mathbf{E}$ é uma matriz de Vandermonde podemos estabelecer rapidamente seu determinante:

$$\det \mathbf{E} = \frac{1}{\sqrt{n^n}} \prod_{0 \le i,j \le n-1} (\epsilon^j - \epsilon^i) \neq 0$$

então $\mathbf{E}$ é não singular. Logo possui inversa $\mathbf{E}^{-1} = \overline{\mathbf{E}}$ uma vez que $\mathbf{E}$ é unitária e simétrica. Fixaremos uma raíz n-ésima primitiva da unidade

$$\epsilon = e^{\frac{2\pi i}{n}}$$

defina para $l = 0, 1..., n-1$,

$$x_l = \frac{1}{\sqrt{n}}(1, \epsilon^l, \epsilon^{2l}, ..., \epsilon^{(n-1)l}) \in \mathbb{C}^n, \tag{2.8}$$

assim as colunas e linhas de $\mathbf{E}$ são dadas pelo vetor $\{x_l\}$.

Olharemos para $\mathbf{E}$ como um operador linear de $\mathbb{C}^n$, e temos que $\mathbf{E}(e_i) = x_i$. Uma vez que $\mathbf{E}$ é não singular, segue que os vetores $\{x_l\}$ são outra base ortonormal de $\mathbb{C}^n$, a qual denotaremos por $\mathbf{x}$. O operador linear de $\mathbb{C}^n$ definido pela matriz $\mathbf{E}$ depende, é claro, das bases definidas para o domínio e do contradomínio; para mostrar essa dependência a aplicação deve ser denotada como $\mathbf{E}_{\mathbf{e},\mathbf{e}}$. Note que como aplicações lineares, $\mathbf{E}_{\mathbf{e},\mathbf{e}} = \mathbf{I}_{\mathbf{e},\mathbf{x}}$, onde $\mathbf{I}$ é a matriz identidade $n \times n$.

2.4 Matrizes Circulantes

Vamos definir o *operador deslocamento*

$$\begin{aligned} T &: \mathbb{C}^n \to \mathbb{C}^n \\ T(v_0, v_1, ..., v_{n-1}) &= (v_{n-1}, v_0, ..., v_{n-2}). \end{aligned}$$

Começaremos com uma definição básica e fundamental. Por uma *matriz circulante* de ordem n, ou *circulante* para ser mais breve, queremos dizer uma matriz quadrada de forma

$$\mathbf{V} = \begin{pmatrix} v_0 & v_1 & \cdots & v_{n-2} & v_{n-1} \\ v_{n-1} & v_0 & \cdots & v_{n-3} & v_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_2 & v_3 & \cdots & v_0 & v_1 \\ v_1 & v_2 & \cdots & v_{n-1} & v_0 \end{pmatrix}$$

Note que os elementos de todas as linhas de $\mathbf{V}$ são idênticos aos da linha anterior diferindo apenas por um deslocamento a direita. Dessa forma, $\mathbf{V}$ pode ser definida inteiramente por sua primeira linha e o operador deslocamento T.

Assim a matriz circulante $\mathbf{V}$ associada ao vetor $v \in \mathbb{C}^n$, chamado *vetor circulante*, é a matriz $n \times n$ cujas linhas são dadas por iterações do operador deslocamento agindo em v em que sua k-ésima linha é dada por $T^{k-1}v$, com $k = 1, ..., n$ e denotaremos $\mathbf{V}$ por $circ\{(v)\}$ ou $circ(v_0, v_1, ..., v_{n-1})$. Poderemos ainda escrever uma circulante na forma $\mathbf{C} = (c_{ij}) = (c_{(i-j) mod\ n})$.

Exemplo 2.19 *O primeiro exemplo de matriz circulante é qualquer matriz da forma* $c\mathbf{I}$ *onde* c *é um escalar e* $\mathbf{I}$ *é a matriz identidade. Em particular, a identidade e a matriz nula são matrizes circulantes.*

Exemplo 2.20 *A matriz de permutação* $\mathbf{W} = circ\,(0, 1, 0, ..., 0)$ *é uma circulante chamada de matriz circulante básica.*

Exemplo 2.21 *Tome* $n = 4$ *e* $v = (1, 2, 3, 4)$ *geramos a matriz circulante* 4×4

$$\mathbf{C} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

Matrizes circulantes tem valores constantes em cada diagonal descendente, isto é, ao longo das linhas de entrada paralelas à diagonal principal.

No exemplo abaixo, o produto de duas matrizes circulantes é, em si, uma matriz circulante. Esta é uma propriedade geral das circulantes como será provado em breve. Desde que uma

circulante é determinada pela sua primeira linha, uma boa parte da aritmética normalmente usada na multiplicação de matrizes circulantes é reduntante. Para circulantes de ordem baixa, a multiplicação pode ser feita com lápis e papel, usando o esquema abreviado esboçado abaixo.

Exemplo 2.22 *Tome $n = 3$ e os vetores circulantes $v = (1, 2, 4)$ e $w = (4, 5, 6)$*

$$\begin{array}{ccc} 1 & 2 & 4 \\ 4 & 5 & 6 \\ \hline 4 & 8 & 16 \\ 20 & 5 & 10 \\ 12 & 24 & 6 \\ \hline 36 & 37 & 32 \end{array}$$

Acima multiplicamos cada coordenada do vetor v por w gerando 3 vetores que chamaremos $w_1 = (x_1, y_1, z_1)$, $w_2 = (x_2, y_2, z_2)$, $w_3 = (x_3, y_3, z_3)$ e distribuimos as coordenadas dos w_i na "matriz" como se segue

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ y_3 & y_1 & y_2 \\ z_2 & z_3 & z_1 \end{bmatrix}.$$

Após, somamos o valores das colunas da "matriz"e geramos um vetor linha que será o vetor circulante da matriz que é o produto das matrizes $\mathbf{VW}$,onde $\mathbf{V}$ e $\mathbf{W}$ são as matrizes geradas pelos vetores circulantes v e w respectivamente.

Denotaremos por $Circ(n) \subset \mathbb{M}_n$ o conjunto de todas as matrizes circulantes complexas $n \times n$, note que $circ(v)+circ(w) = circ(v+w)$ e $\alpha circ(v) = circ(\alpha v)$, logo $Circ(n)$ é um espaço vetorial complexo de acordo com as operações usuais de adição de matrizes e multiplicação de matrizes por escalares, daí o nosso primeiro modelo de matrizes circulantes é fornecido pelo isomorfismo $\mathbb{C}$-linear

$$\mathcal{J} : Circ(n) \to \mathbb{C}^n$$

onde $\mathcal{J}$ manda a matriz em sua primeira linha. Matrizes podem, naturalmente, ser multiplicadas e para o caso das matrizes circulante temos os seguinte resultados

Teorema 2.23 *Seja $\mathbf{A}$ uma matriz de tamanho $n \times n$ e $\mathbf{W} = circ(0, 1, 0, \ldots, 0)$. Então $\mathbf{A}$ é circulante se, e somente se, $\mathbf{WA} = \mathbf{AW}$.*

Prova. Dados $\mathbf{A} = (a_{ij})$ e σ a permutação do ciclo $\sigma = (1, 2, ..., n)$, temos que $\mathbf{P}_\sigma \mathbf{A} \mathbf{P}_{\sigma^*} = (a_{\sigma(i),\sigma(j)})$ no nosso caso $\mathbf{P}_\sigma = \mathbf{W}$ então

$$\mathbf{WAW}^* = \mathbf{WAW}^{-1} = (a_{(i+1)(j+1)})$$

(onde os índices são tomados $mod n$) e essa matriz é igual a $\mathbf{A} = (a_{ij})$ se, e somente se, $\mathbf{A}$ é circulante. Equivalentemente se, e somente se, $\mathbf{WAW}^* = \mathbf{A}$ ■

Corolário 2.24 *Sejam* $\mathbf{A}$ *e* $\mathbf{B}$ *matrizes circulantes de mesma ordem. Então valem as seguintes propriedades:*

1- $\mathbf{A}^*$ *é circulante*

2- $\mathbf{AB}$ *é circulante*

Prova. Seja $\mathbf{A}$ uma matriz circulante, então pelo Teorema 2.23

$$\mathbf{A} = \mathbf{WAW}^* \Rightarrow \mathbf{A}^* = (\mathbf{WAW}^*)^* = \mathbf{WA}^*\mathbf{W}^*,$$

segue que $\mathbf{A}^*$ é circulante.

Se $\mathbf{A}$ e $\mathbf{B}$ são circulantes novamente temos

$$\mathbf{A} = \mathbf{WAW}^* \text{ e } \mathbf{B} = \mathbf{WBW}^*$$

aqui

$$\mathbf{AB} = \mathbf{WAW}^*\mathbf{WBW}^* = \mathbf{WABW}^*,$$

uma vez que $\mathbf{W}^* = \mathbf{W}^{-1}$. ■

Observação 2.25 *Pelos resultados acima temos outras caracterizações para as matrizes circulantes como:* $\mathbf{A}$ *é circulante se, e somente se,* $\mathbf{A}^*$ *é circulante, e que as circulantes compreendem todas as matrizes (quadradas) que comutam com* $\mathbf{W}$.

Veremos agora uma segunda representação para as circulantes. Tendo em vista a estrutura das matrizes de permutação $\mathbf{W}^k$, $k = 0, 1, ..., n-1$, temos a seguinte proposição.

Proposição 2.26 $\mathbf{V}$ *é circulante se, e somente se,* $\mathbf{V} = p(\mathbf{W})$ *para algum polinômio* $p(x)$.

Prova. Seja $\mathbf{V}$ uma circulante definida pelo vetor $v = (v_0, v_1, ..., v_{n-1})$. Considere agora um vetor e_i da base canônica (como definido no capítulo 1) calcularemos a multiplicação do vetor e_i por $\mathbf{W}$ a direita , então por (2.3). $\mathbf{W}e_1 = e_n, \mathbf{W}e_2 = e_1, ..., \mathbf{W}e_n = e_{n-1}$ assim:

$$\begin{aligned}
\mathbf{W}e_i &= e_{\sigma^{-1}(i)} \\
\mathbf{W}^2e_i &= \mathbf{W}e_{\sigma^{-1}(i)} = e_{\sigma^{-2}(i)} \\
&\vdots \\
\mathbf{W}^ke_i &= \mathbf{W}^{k-1}e_{\sigma^{-1}(i)} = e_{\sigma^{-k}(i)}
\end{aligned}$$

usando isso em

$$(v_0\mathbf{I} + v_1\mathbf{W} + \cdots + v_{n-1}\mathbf{W}^{n-1})e_i = v_0e_i + v_1e_{\sigma^{-1}(i)} + \cdots + v_{n-1}e_{\sigma^{-n+1}(i)} = \mathbf{V}e_i.$$

Logo $\mathbf{V} = v_0\mathbf{I} + v_1\mathbf{W} + \cdots + v_{n-1}\mathbf{W}^{n-1} = p(\mathbf{W})$. ■

Defina,

$$\mathbf{W}_i = circ\{e_i\},$$

onde e_i é dado como em (2.2), logo pela Proposição 2.26 temos uma representação usual ou forma para as matrizes circulantes:

$$circ(v_0, v_1, ..., v_{n-1}) = \sum_{i=0}^{n-1} v_i \mathbf{W}^i.$$

Como $\mathbf{W}$ é uma matriz de permutação, segue que $\mathbf{W}_i\mathbf{W}_j = \mathbf{W}_{i+j}$, onde todos os índices são interpretados *modn*. Convencionaremos que $\mathbf{W}^0 = \mathbf{I}$ e $\mathbf{W}^1 = \mathbf{W}$. Note que $\mathbf{W}^i = \mathbf{W}_i$.

Observação 2.27 *Com respeito a base canônica de $\mathbb{C}^n$, o operador deslocamento é representado pela transposta da matriz $\mathbf{W}$, isto é, por $circ\{(0, 0, ..., 0, 1)\}$.*

O *representante* P_V da matriz circulante $\mathbf{V} = circ\{(v_0, v_1, ..., v_{n-1})\}$ é

$$P_V(X) = \sum_{i=0}^{n-1} v_i X^i.$$

Então,

$$\mathbf{V} = circ(v) = P_V(\mathbf{W}).$$

A função

$$\phi(\theta) = \phi_v(\theta) = v_0 + v_2 e^{i\theta} + \cdots + v_n e^{i(n-1)\theta}$$

também é útil como representante.

Proposição 2.28 *Todas as matrizes circulantes são normais, isto é, se $\mathbf{V}$ é circulante então $\mathbf{V}\mathbf{V}^* = \mathbf{V}^*\mathbf{V}$.*

Prova. Uma vez que polinômios em uma mesma matriz comutam, segue que todas as circulantes de mesma ordem comutam, e se $\mathbf{V}$ é circulante $\mathbf{V}^*$ também é, logo

$$\mathbf{V}\mathbf{V}^* = P_V(\mathbf{W})P_V(\mathbf{W}^*) = P_V(\mathbf{W}^*)P_V(\mathbf{W}) = \mathbf{V}^*\mathbf{V}.$$

Então $\mathbf{V}$ e $\mathbf{V}^*$ comutam e portanto todas as circulantes são matrizes normais. ■

Corolário 2.29 *Se $\mathbf{V}$ é circulante então $\mathbf{V}^k$ para $k \in \mathbb{Z}_+$ é uma matriz circulante.*

Teorema 2.30 *$Circ(n)$ é uma álgebra comutativa com as operações usuais de adição e multiplicação de matrizes que é gerada (sobre $\mathbb{C}$) pela matriz $\mathbf{W}$. A aplicação que manda $\mathbf{W}$ em sua indeterminada X estende por linearidade e multiplicatividade um a isomorfismo de $\mathbb{C}$-algebras*

$$\mathcal{J} : Circ(n) \to \frac{\mathbb{C}[x]}{(X^n - 1)}.$$

A aplicação que manda a matriz circulante $\mathbf{V}$ em sua transposta $\mathbf{V}^t$ é uma involução de $Circ(n)$ e corresponde sobre $\mathcal{J}$ a um automorfismo de $\frac{\mathbb{C}[x]}{(X^n-1)}$ induzido por $X \to X^{n-1}$.

Prova. Inicialmente, queremos mostrar que $Circ(n)$ é uma álgebra comutativa gerada, sobre $\mathbb{C}$, por potências de $\mathbf{W}$. Pelo primeiro modelo para as matrizes circulantes sabemos que sua $\mathbb{C}$-álgebra têm dimensão n. Assim, uma base para sua $\mathbb{C}$-álgebra deve conter n elementos. Pela Proposição 2.26 temos que se $\mathbf{V}$ é uma matriz circulante definida por $v = (v_0, \ldots, v_{n-1})$ então

$$\mathbf{V} = v_0\mathbf{I} + v_1\mathbf{W} + \cdots + v_{n-1}\mathbf{W}^{n-1}$$

Logo, Toda matriz circulante é uma combinação linear dos elementos do conjunto $\{\mathbf{I}, \mathbf{W}, ..., \mathbf{W}^{n-1}\}$, além disso esse conjunto é L.I. uma vez que a matriz nula é uma circulante e $p(\mathbf{W})$ é igual a matriz nula se, e somente se, os coeficientes do polinômio p são todos nulos. Portanto o conjunto

$$\{\mathbf{I}, \mathbf{W}, ..., \mathbf{W}^{n-1}\}$$

é uma base para a $\mathbb{C}$-álgebra das matrizes circulantes.

Agora, provaremos que $Circ(n) \cong \frac{\mathbb{C}[x]}{(X^n-1)}$ é um isomorfismo de $\mathbb{C}$-álgebras. De fato, como as circulantes são uma $\mathbb{C}$-álgebra com gerador $\mathbf{W}$. E sabemos que cada elemento pode ser expresso por um poliômio de grau $\leq n-1$. Assim, podemos construir o seguinte anel.

Considere $\mathbb{C}[\mathbf{W}]$ o anel do polinômios complexos em $\mathbf{W}$ de grau $\leq n$ e com coeficientes complexos, vamos manter em $\mathbb{C}[\mathbf{W}]$ as regras usuais de adição e multiplicação de polinômios, mas as potências mais altas vão ser substituidas por potências mais baixas usando o fato de que $\mathbf{W}^n = 1$. Note que a aplicação

$$\begin{array}{ccc} Circ(n) & \to & \mathbb{C}[\mathbf{W}] \\ \mathbf{V} & \mapsto & P_V(\mathbf{W}) \end{array}$$

é um isomorfismo de aneis.

Equivalentemente, esta construção nos permite ver as matrizes circulantes como o quociente de uma álgebra polinomial quocientada pelo ideal gerado pelo polinômio mínimal $X^n - 1$ do gerador $\mathbf{W}$, ou seja, $\frac{\mathbb{C}[x]}{(X^n-1)}$. Além disso, visto como $\mathbb{C}$-espaço vetorial, claramente o espaço das matrizes circulantes é isormorfo ao espaço $\mathcal{P}_{n-1}$ dos polinômios complexos de grau $\leq n-1$. Que pelo Teorema do Algoritmo da divisão é isomorfo a $\frac{\mathbb{C}[x]}{(X^n-1)}$.

Temos então um isomorfismo de aneis e de $\mathbb{C}$-espaços vetorias, logo temos um isomorfismo de $\mathbb{C}$-álgebras. Então, as matrizes circulante sobre $\mathbb{C}$ são isomorfas a $\frac{\mathbb{C}[x]}{(X^n-1)}$.

Finalmente, mostraremos a última afirmação. A demontração é exemplificada no diagrama abaixo

$$\begin{array}{cccc} & \mathcal{I} & & \mathcal{A} \\ \mathcal{J} & Circ(n) & \to & \frac{\mathbb{C}[x]}{(X^n-1)} \\ & \downarrow & & \downarrow \\ \mathcal{J}_t & Circ(n) & \to & \frac{\mathbb{C}[x]}{(X^n-1)} \end{array}$$

Seja $\mathcal{I}$ a involução de $Cir(n)$ que manda $\mathbf{V}$ em $\mathbf{V}^t$, essa função é claramente um isomorfismo e assim leva base em base. Como vimos a $\mathbb{K}$-álgebra $Circ(n)$ é gerada por $\mathbf{W}$. Então $\mathcal{I}$ manda $\mathbf{W}$ em $\mathbf{W}^t$. Note que $(\mathbf{W}^i)^t = \mathbf{W}^{n-i}$.

Defina $\mathcal{J}_t$ por

$$\begin{array}{rcl} Circ(n) & \to & \frac{\mathbb{C}[x]}{(X^n-1)} \\ \mathbf{V}^t & \mapsto & P_{\mathbf{V}}(\mathbf{W}^t) \end{array}$$

temos então que $\mathcal{J}_t$ é um isomorfismo. E mais, $\mathcal{J}_t$ manda $\mathbf{W}^t$ em X^{n-1} e uma vez que $\mathcal{J}$ manda $\mathbf{W}$ em X, segue o resultado. ■

Teorema 2.31 *Seja $v = (v_0, v_1, , ..., v_{n-1})$ um vetor em $\mathbb{C}^n$ e $\mathbf{V} = circ\{v\}$. Se ϵ é uma raiz n-ésima primitiva da unidade, então*

$$\det \mathbf{V} = \prod_{l=0}^{n-1} \left(\sum_{j=0}^{n-1} \epsilon^{jl} v_j \right) = \prod_{l=0}^{n-1} P_V(\epsilon^l).$$

Prova. Vemos que a matriz $\mathbf{V}$ é um operador linear $V_{e,e}$ de $\mathbb{C}^n$. Para cada inteiro l, $0 \leq l \leq n-1$, seja x_l como em (2.8). Vamos calcular os autovalores de $\mathbf{V}$. Assim,

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{n-2} & v_{n-1} \\ v_{n-1} & v_0 & \cdots & v_{n-3} & v_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_2 & v_3 & \cdots & v_0 & v_1 \\ v_1 & v_2 & \cdots & v_{n-1} & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon^l \\ \vdots \\ \epsilon^{(n-2)l} \\ \epsilon^{(n-1)l} \end{bmatrix} = \lambda_l \begin{bmatrix} 1 \\ \epsilon^l \\ \vdots \\ \epsilon^{(n-2)l} \\ \epsilon^{(n-1)l} \end{bmatrix}.$$

a equação acima é equivalente a n equações

$$\sum_{k=0}^{m-1} v_{n-m+k} \epsilon^{lk} + \sum_{k=m}^{n-1} v_{k-m} \epsilon^{lk} = \lambda_l \epsilon^m,$$

para $m = 0, 1, ..., n-1$.

Trocando os índices mudo da soma temos

$$\sum_{k=0}^{n-1-m} v_k \epsilon^{l(k+m)} + \sum_{k=n-m}^{n-1} v_k \epsilon^{l(k-n-m)} = \lambda_l \epsilon^{lm}$$

dividindo ambos os lados por ϵ^m ficamos com:

$$\sum_{k=0}^{n-1-m} v_k \epsilon^k + \epsilon^{-n} \sum_{k=n-m}^{n-1} v_k \epsilon^k = \lambda_l$$

pela Proposição 2.3 temos o autovalor de $\mathbf{V}$

$$\sum_{k=0}^{n-1} v_k \epsilon^{lk} = \lambda_l.$$

com autovetor normalizado correspondente x_l. Uma vez que $\{x_0, x_1, ..., x_{n-1}\}$ é um base ortonormal de $\mathbb{C}$, então $\mathbf{V}$ é diagonálizavel. Uma vez que o determinante de uma matriz diagonal é produtório de seus autovalores, defina

$$\lambda_l = v_0 + \epsilon^l v_1 + \cdots + \epsilon^{(n-1)l} v_{n-1} = P_V(\epsilon^l),$$

e concluímos que $\det \mathbf{V} = \prod_{l=0}^{n-1} P_V(\epsilon^l)$.

Corolário 2.32 *Todas as matrizes circulantes têm o mesmo conjunto ordenado de autovetores ortonormais* $\{x_l\}$.

Corolário 2.33 *O polinômio característico* p_V *de* V *é dado por*

$$p_v(X) = \det(X\mathbf{I} - \mathbf{V}) = \prod_{l=0}^{n-1}(X - \lambda_l) = X^n + \sum_{i=n-1}^{0} b_i X^i$$

(aqui, deixamos na última igualdade os b_i*'s definidos como funções dos* λ_l*'s. Eles são as funções simétricas elementares dos* λ_l*'s.*

Corolário 2.34 *A nulidade de* $\mathbf{V} \in Circ(n)$ *é o número de autovalores* λ_l *nulos.*

Corolário 2.35 *o traço da circulante* $\mathbf{V}$ *é o somátorio de seus autovalores, ou seja,*

$$\sum_{l=0}^{n-1} \lambda_l = nv_0.$$

Prova. Temos que

$$\sum_{l=0}^{n-1} \lambda_l = \sum_{l=0}^{n-1}\sum_{j=0}^{n-1} \epsilon^{jl} v_j = \sum_{j=0}^{n-1} \left(\sum_{l=0}^{n-1} \epsilon^{jl} \right) v_j$$

e uma vez que

$$\sum_{l=0}^{n-1} \epsilon^{jl} = \begin{cases} n, & \text{se } i = 0 \\ 0, & \text{se } i = 1, ..., n-1 \end{cases},$$

e assim, segue o resultado. ■

Se denotarmos por $v(\mathbf{V})$ a nulidade de $\mathbf{V}$, então para todo $\mathbf{V} \in Circ(n)$, $v(\mathbf{V}) = \deg mdc(p_V(X), X^n)$. De fato, seja k a cardinalidade do conjunto dos autovalores nulos, então o polinômio característico é da forma $X^k q(X)$, onde $\deg q(x) \leq n$, note que o mdc entre o polinômio característico e X^n é igual a k. Logo $\deg mdc(p_V(X),\ X^n) = \deg mdc(X^k q(X), X^n) = k$. Pelo Corolário (2.34), $v(\mathbf{V}) = \deg mdc(p_V(X),\ X^n)$.

Corolário 2.36 *Seja* V *uma matriz circulante com representante* P_V*. Então as afirmações abaixo são equivalentes:*

(a) *A matriz* $\mathbf{V}$ *é singular.*

(b) $P_V(\epsilon^l) = 0$ *para algum* $l \in \mathbb{Z}$.

(c) *Os polinômios* $P_V(X)$ *e* $X^n - 1$ *não são primos relativos.*

Prova. $(a) \Leftrightarrow (b)$ Se $\mathbf{V}$ é singular então $\det \mathbf{V} = 0$, mas pelo Teorema 2.31

$$\det \mathbf{V} = \prod_{l=0}^{n-1} P_V(\epsilon^l) = 0,$$

portanto

$$\prod_{l=0}^{n-1} P_V(\epsilon^l) = 0 \Leftrightarrow P_V(\epsilon^l) = 0$$

para algum $l \in \mathbb{Z}$.

$(b) \Rightarrow (c)$ Se $P_V(\epsilon^l) = 0$ para algum $l \in \mathbb{Z}$, então ϵ^l é raiz de $P_V(X)$. Logo, $X - \epsilon^l$ divide $P_V(X)$ e pelo Proposição 2.3 $\left(\epsilon^l\right)^n = 1$, assim ϵ^l também é raíz de $X^n - 1$. Portanto, $P_V(X)$ e $X^n - 1$ não são primos relativos.

$(c) \Rightarrow (b)$ Se $P_V(X)$ e $X^n - 1$ não são primos relativos. Então o polinômio $(x - \alpha)$, onde α é raíz de $X^n - 1$, divide $P_V(X)$. Uma vez que as raízes de $X^n - 1$ são as raízes da unidades, então $P_V(\epsilon^l) = 0$ para algum $l \in \mathbb{Z}$. ■

Novamente, temos uma reformulação de parte do último corolário como

• Para todo $\mathbf{V} \in Circ(n)$, $v(\mathbf{V}) = \deg mdc(P_V(X), X^n - 1)$.

Para obter o nosso *terceiro modelo* para $Circ(n)$, vamos começar definindo $\mathbb{D}_n$ como o espaço das matrizes diagonais $n \times n$. Este espaço é linearmente isomorfo a $\mathbb{C}^n$.

A diagonalização das matrizes circulantes seguirá imediatamente a partir da diagonalizacao da circulante básica $\mathbf{W}$.

Teorema 2.37 *A matriz* $\mathbf{W}$ *é unitariamente diagonalizada por* $\mathbf{E}$. *Então,*

$$\mathbf{W} = \mathbf{E\Omega E}^*.$$

onde $\mathbf{\Omega}$ $=Diag\,(1, \epsilon, ..., \epsilon^{n-1})$.

Prova. A k-ésima linha de $\mathbf{E}$ é

$$\left(\frac{1}{\sqrt{n}}\right)(\overline{\epsilon}^{(k-1)0}, \overline{\epsilon}^{(k-1)1}, ..., \overline{\epsilon}^{(k-1)n-1})$$

Então a k-ésima linha de $\mathbf{E\Omega}$ é

$$\left(\frac{1}{\sqrt{n}}\right)(\overline{\epsilon}^{(k-1)r}\epsilon^r) = \left(\frac{1}{\sqrt{n}}\right)\epsilon^{2r-kr}$$

para $r = 0, 1, ..., n-1$.

A j-ésima coluna de $\mathbf{E}^*$ é

$$\left(\frac{1}{\sqrt{n}}\right)(\epsilon^{(j-1)0}, \epsilon^{(j-1)1}, ..., \epsilon^{(j-1)n-1}).$$

Assim, o (k, j)-ésimo elemento de $\mathbf{E\Omega E}^*$ é

$$\begin{aligned} \frac{1}{n}\sum_{r=0}^{n-1} \epsilon^{2r-kr}\epsilon^{(j-1)r} &= \frac{1}{n}\sum_{r=0}^{n-1} \epsilon^{r(j-k+1)} \\ &= \begin{cases} 1, & \text{se } j = k-1, \\ 0, & \text{se } j \neq k-1, \end{cases} \end{aligned}$$

onde os índices são tomados *modn* e na primeira igualdade usamos a Proposição 2.3. Dessa forma, temos $\mathbf{W} = \mathbf{E\Omega E}^*$. ■

Teorema 2.38 *Todos os elementos de $Circ(n)$ são simultaneamente diagonalizaveis pela matriz unitária* $\mathbf{E}$*, isto é, para todo* $\mathbf{V} \in Circ(n)$*,*

$$\mathbf{E}^{-1}\mathbf{V}\mathbf{E} = \mathbf{D}_{\mathbf{V}}$$

onde $\mathbf{D}_{\mathbf{V}}$ *é a matriz* $Diag(P_{\mathbf{V}}(1), P_{\mathbf{V}}(\epsilon), ..., P_{\mathbf{V}}(\epsilon^{n-1}))$ *e a aplicação resultante*

$$\mathcal{D} : Circ(n) \to \mathbb{D}_n$$

é um isomorfismo de $\mathbb{C}$*-algebras.*

Prova. Se $\mathbf{V}$ é uma matriz circulante, então

$$\begin{aligned} V &= circ\{v\} = P_V(W) \\ &= \sum_{i=0}^{n-1} v_i \mathbf{W}^i = \sum_{i=0}^{n-1} v_i \mathbf{E}\mathbf{\Omega}^i\mathbf{E}^* \\ &= \mathbf{E}\left(\sum_{i=0}^{n-1} v_i \mathbf{\Omega}^i\right)\mathbf{E}^* \\ &= \mathbf{E}Diag\left(P_{\mathbf{V}}\left(\epsilon^l\right)\right)\mathbf{E}^* \end{aligned}$$

multiplicando por $\mathbf{E}^*$ a esquerda, e a direita por $\mathbf{E}$ e usando o teorema acima temos que $\mathbf{E}^{-1}\mathbf{V}\mathbf{E} = \mathbf{D}_{\mathbf{V}}$. ■

Corolário 2.39 *A inversa de um elemento invertível de $Circ(n)$ também pertence a $Circ(n)$.*

Prova. Se $\mathbf{V}$ é uma matriz circulante não singular, então $\mathbf{V}$ é invertível. Assim

$$\mathbf{V}^{-1} = \left(\mathbf{E}^{-1}\mathbf{D}_{\mathbf{V}}\mathbf{E}\right)^{-1} = \mathbf{E}\mathbf{D}_{\mathbf{V}}^{-1}\mathbf{E}^{-1}$$

logo, $\mathbf{D}_V^{-1} = \mathbf{E}^{-1}\mathbf{V}^{-1}\mathbf{E} = \mathbf{D}_{\mathbf{V}^{-1}}$.

Corolário 2.40 *O polinômio característico de* $\mathbf{V} \in Circ(n)$ *é dado por*

$$p_{\mathbf{V}}(X) = \det(X\mathbf{I} - \mathbf{V}) = \det(X\mathbf{I} - \mathbf{D}_{\mathbf{V}}).$$

Capítulo 3

Resolução de equações polinomiais

Faremos agora uma pequena discursão sobre o método de Cardano para as cúbicas e uma aplicação da teoria circulante. Como aplicação vamos discutir a utilização das matrizes circulantes na resolução de equações polinomiais com grau $n \leq 4$. Para isso, construiremos uma matriz circulante com polinômio caracteristico específico e então as raízes do polinômio serão os autovalores da matriz circulantes. Por fim, aplicaremos o metodo circulante na resolução de cúbicas e quárticas.

3.1 A fórmula de Cardano

O problema de resolver equações cúbicas surgiu inicialmente de um problema géometrico na grécia antiga, o de triseccionar um ângulo usando apenas régua e compasso, com o desenvolvimento posterior da trigonometria fica claro que esse problema se reduz a resolver uma equação cúbica.

Para encontrar um terço de um ângulo θ dado, podemos começar pensando que θ é três vezes o ângulo procurado α, isto é, $3\alpha = \theta$ e aplicando a fórmula do cosseno de 3α, ficamos com

$$\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha \tag{3.1}$$

uma vez que sabemos quem é θ, consequentemente quem é $\cos\theta$ que chamaremos de a. Para construir $3^{-1}\theta$, basta encontrar seu cosseno. Fazendo $x = \cos(3^{-1}\theta)$ (e lembrando que $\alpha = 3^{-1}\theta$) em (3.1), obtemos

$$4x^3 - 3x - a = 0.$$

Portanto, podemos encontrar x resolvendo a equação.

O matemático Girolamo Cardano encontrou uma fórmula para resolver uma equação cúbica da forma

$$x^3 + px + q = 0. \tag{3.2}$$

Consideremos o binômio de Newton $(u+v)^3$, desenvolvendo teremos

$$(u+v)^3 = u^3 + 3u^2v + 3uv^2 + v^3,$$

fazendo $x = u + v$ obtemos

$$x^3 - 3uvx - u^3 - v^3 = 0. \tag{3.3}$$

Igualando as equações (3.2) e (3.3),fícamos com o sistema

$$\begin{aligned} -u^3 - v^3 &= q \\ -3uv &= p, \end{aligned}$$

vamos resolver esse sistema em uma próxima seção e veremos que

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

e

$$u = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

Daí segue a fórmula de Cardano para uma solução da cúbica (3.2),

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \tag{3.4}$$

Isto dava uma solução para muitas cúbicas, mas em certos casos havia uma ilógica para a época. Suponhamos, por exemplo, que a equação seja

$$x^3 - 15x - 4 = 0$$

e aplicamos a fórmula para obter:

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}. \tag{3.5}$$

Com base em nossa experiência com quadráticas, a conclusão correta pareceria ser que não há solução (real). Como veremos em uma seção posterior essa equação possui uma raiz real e mais as outras duas raízes também são reais.

Cardano notou esse problema, mas parece não saber o que fazer a respeito, lembrando que naquela epóca ainda não se trabalhava com os números complexos, no entanto, mesmo hoje com números complexos não é fácil sair de (3.5) para $x = 4$.

Ao mencionar esse problema em seu livro duas vezes, primeiro ele diz que esse caso precisa ser resolvido usando um método diferente a ser descrito em outro livro. Na segunda vez, ele escreveu que resolvendo

$$y^3 - 8y + 3 = 0$$

obtem 3 como resposta, o que é intrigante pois ao aplicar (3.4) aparece

$$\sqrt{\frac{3^2}{4} + \frac{(-8)^3}{27}} = \sqrt{\frac{-1805}{108}}.$$

3.2 Raízes de polinômios no caso geral

Usaremos a teoria circulante para explorar conexões entre a álgebra matricial e as raízes de polinômios. Estas conexões vão nos guiar para a solução de equações polinomiais.

O que mais nos chama atenção sobre as matrizes circulantes é o simples cálculo de seus autovalores e autovetores utilizando as raízes da unidade. Vamos agora usar alguns resultados sobre matrizes circulantes e ver como eles nos fornecem uma forma de construir polinômios com raízes conhecidas.

Exemplo 3.1 *Calcular os autovalores, autovetores e o polinômio característico da matriz*

$$\mathbf{V} = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 3 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 2 & 1 & 3 & 1 \end{bmatrix}$$

Solução. Temos que o representante de $\mathbf{V}$ é

$$P_{\mathbf{V}}(X) = 1 + 2X + X^2 + 3X^3.$$

Logo, seus autovalores são $P_{\mathbf{V}}(1) = 7$, $P_{\mathbf{V}}(-1) = -3$, $P_{\mathbf{V}}(i) = -i$ e $P_{\mathbf{V}}(-i) = i$, com autovetores correpondentes $(1,1,1,1), (1,-1,1,-1), (1,i,-1,-i), (1,-i,-1,i)$. Seu polinômio caracterítisco é

$$p_{\mathbf{V}}(X) = X^4 - 4X^3 - 20X^2 - 4X - 21.$$

A cada matriz circulante $\mathbf{V}$ de ordem $n \times n$ temos dois polinômios naturalmente associados a ela, a saber, o seu representante $P_{\mathbf{V}}$ e o polinômio característico $p_{\mathbf{V}}$. Os quais são ambos descritos explicitamente em termos dos autovalores λ_l de $\mathbf{V}$. O polinômio característico $p_{\mathbf{V}}$ é o único polinômio mônico de grau n que se anula em cada λ_l e o representante $P_{\mathbf{V}}$ é o único polinômio grau $n-1$, cujo valor em ϵ^l é λ_l.

As raízes do polinômio característico de uma matriz arbitrária $n \times n$ (estes são os autovalores da matriz $\mathbf{V}$) são obtidos através da resolução de uma equação polinomial mônica de grau n. No caso de matrizes circulantes, as raízes do $p_{\mathbf{V}}$ são facilmente calculadas usando o polinômio representante $P_{\mathbf{V}}$.

Assim, se um dado polinômio p é conhecido por ser o polinômio característico de uma matriz circulante dada $\mathbf{V}$, suas raízes podem ser facilmente determinadas.

Dessa forma é muito interessante determinar quais polinômios mônicos seriam polinômios característicos das matrizes circulantes e, assim, um problema muito natural aparece: se $p = p_{\mathbf{V}}$ para uma coleção de matrizes circulantes $\mathbf{V}$, podemos determinar uma dessas $\mathbf{V}$ ou, equivalente seu representante $P_{\mathbf{V}}$ diretamente a partir de p?

Proposição 3.2 *Todo polinômio mônico p é o polinômio característico de alguma matriz circulante $\mathbf{V}$.*

Prova. Se p é um polinômio mônico de grau n com raízes $z_1, ..., z_n$, a proposição é equivalente a encontrar o polinômio representante de $\mathbf{V}$, ou seja, encontrar um polinômio q de grau $n-1$ tal que, para cada $1 \leq l \leq n$,

$$q\left(\epsilon^l\right) = z_l.$$

Dessa forma, tome q igual ao representante P_V de $\mathbf{V}$ que é igual a

$$a_0 + a_1 X + \cdots + a_{n-1} X^{n-1},$$

então podemos definir $\mathbf{V}$ pelo vetor circulante

$$\left(a_0, a_1, ..., a_{n-1}\right).$$

Uma vez que os autovalores de $\mathbf{V}$ são $q\left(\epsilon^l\right) = z_l$, o polinômio característico de $\mathbf{V}$ é igual a p.■

Pela Proposição 3.2 temos que as n raízes de um dado polinômio mônico p são os valores de P_V nas n raízes n-ésimas da unidade. No entanto, o argumento não nos da uma solução para o nosso problema de encontrar as raízes de um polinômio p, uma vez que para construir $\mathbf{V}$ a partir de p supomos conhecidas todas as raízes de p. Dessa forma, a questão mais díficil levantada foi a construção de $\mathbf{V}$, ou seja, de seu representante P_V em função dos coeficientes do polinômio p.

O espaço vetorial $\mathcal{P}_{n-1}$ dos polinômios de grau $\leq n-1$ é canonicamente isomorfo a $Circ(n)$ uma vez que ambos são canonicamente isomorfos ao espaço vetorial $\mathbb{C}^n$. De fato, já vimos que o espaço $Circ(n)$ é isomorfo a $\mathbb{C}^n$ atráves do primeiro modelo para as matrizes circulantes. Provaremos agora que a aplicação linear

$$\begin{array}{rccc} \Phi: & \mathcal{P}_{n-1} & \to & \mathbb{C}^n \\ & q & \mapsto & \left(q\left(1\right), q\left(\epsilon^l\right), ..., q\left(\epsilon^{(n-1)l}\right)\right) \end{array} \tag{3.6}$$

é um isomorfismo. Note que se $\Phi\left(q\right) = (0, ..., 0)$, então $q \in \mathcal{P}_{n-1}$ se anula em todas as raízes da unidade, portanto, q é o polinômio nulo. Logo, $\ker \Phi = 0$ e, portanto, Φ é injetora. Uma vez que ambos os espaços têm a mesma dimensão, pelo teorema do núcleo e da imagem $\Phi\left(\mathcal{P}_{n-1}\right)$ é igual a $\mathbb{C}^n$ e, portanto, Φ é bijetora. Além disso, a caracterização acima prova a existência e unicidade do polinômio represente P_V de $\mathbf{V}$.

Seja $\mathcal{M}$ o espaço dos polinômios mônicos de grau n, novamente esse espaço também pode ser identificado com $\mathbb{C}^n$. De fato, pela Proposição 3.2 e pela aplicação (3.6) temos o resultado. Portanto, segue que

$$\mathcal{P}_{n-1} \cong Circ(n) \cong \mathcal{M}.$$

Vamos agora trabalhar com um desses isomorfimos. Defina

$$\mathbf{\Lambda}: \mathcal{P}_{n-1} \to \mathcal{M}$$

da seguinte forma. Para cada $p \in \mathcal{P}_{n-1}$, existe uma única matriz $\mathbf{V}$ em $Circ\left(n\right)$ tal que $p = P_V$. Assim sendo, enviamos p em p_V. Em cada um desses três espaços vamos definir um subespaço:

1. $\mathcal{P}^0_{n-1}$constituído por aqueles $\{p \in \mathcal{P}_{n-1}$ com $p(\epsilon^i) \neq p(\epsilon^j)$ para todos os inteiros $0 \leq i < j \leq n-1\}$.

2. $Circ^0(n)$ constituído por aqueles $\{\mathbf{V} \in Circ(n)$ com autovalores distintos$\}$.

3. $\mathcal{M}^0$ constituído por aqueles $\{p \in \mathcal{M}$ com raízes distintas$\}$.

E restrigir a aplicação $\mathbf{\Lambda}$ aos seus subespaços

$$\mathbf{\Lambda} : \mathcal{P}^0_{n-1} \to \mathcal{M}^0.$$

Portanto, uma forma explícita para a inversa desta aplicação iria fornecer um algoritmo para resolver equações de todos os graus.

O problema encontrado acima é de fundamental importância e bastante difícil em geral. Vamos agora voltar nossa atenção apenas para os casos de grau baixo.

Observação 3.3 *Sabemos que*

$$\mathcal{P}^0_{n-1} \cong Circ^0(n) \cong \mathcal{M}^0.$$

Cada um destes espaços é definido analiticamente. No entanto, o último tem uma caracterização algébrica alternativa. Seja p' a derivada de p. O conjunto $\mathcal{M}^0$ pode ser descrito como

$$\{p \in \mathcal{M} : \deg \ mdc\,(p, p') = 0\}.$$

Assim, a resolução de equações gerais podem ser reduzidos através de um procedimento para a resolução de equações algébricas com raízes distintas, o cálculo de p' é bastante algébrico e, assim, é o cálculo de $d = mdc\,(p, p')$ através do algoritmo de divisão.

3.3 Raízes de polinômios de grau ≤ 4

Matrizes circulantes fornecem uma abordagem unificada para resolver equações de grau ≤ 4. Vamos ilustrar isso para os graus 3 e 4.

Com tudo que vimos ao longo do trabalho, vemos que a teoria circulante fornecem uma novo método para a construação de polinômios com raízes conhecidas. Dado um conjunto de raízes $z_1, ..., z_n$ multiplicando os fatores da forma $(x - z_l)$ determinamos os coeficientes do polinomio com essas raízes. Para encontrar as raízes desse polinômio aplicamos o processo inverso, temos os coeficientes e tentamos extrair as raízes.

Com as circulantes temos uma nova perspectiva para esse problema. Começamos com uma matriz circulante $\mathbf{V}$ e obtemos ao mesmo tempo os coeficientes e as raízes de um polinômio p. Onde como já sabemos o polinômio p é o polinômio caracteristico e suas raízes são os autovalores de $\mathbf{V}$ que são obtidas avaliando as raízes da unidade no representante q da matriz $\mathbf{V}$. Dessa forma a abordagem circulante nos fornece não só um método para resolver equações polinomias como cria uma conexão ainda mais forte entre matrizes e polinômios.

Vamos agora a algumas definições rápidas. Dado um polinômio mônico p de grau n, uma matriz circulante $\mathbf{V}$ é dita *aderente* a p se o polinômio característico p_V de $\mathbf{V}$ é igual a p. Uma matriz *traceless* é uma matriz cujo traço é zero.

Dados um polinômio geral

$$p(X) = X^n + \alpha_{n-1}X^{n-1} + \alpha_{n-2}X^{n-2} + \cdots + \alpha_1 X + \alpha_0 = 0,$$

e a circulante geral $n \times n$

$$\mathbf{V} = \begin{pmatrix} v_0 & v_1 & \cdots & v_{n-2} & v_{n-1} \\ v_{n-1} & v_0 & \cdots & v_{n-3} & v_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_2 & v_3 & \cdots & v_0 & v_1 \\ v_1 & v_2 & \cdots & v_{n-1} & v_0 \end{pmatrix}.$$

Vamos analisar alguns conexões entre essa matriz e esse polinômio do ponto de vista da teoria das circulantes.

Sabemos pelo método de newton que $-\alpha_{n-1}$ é a soma das raízes de $p(X)$. E assim, se p for polinômio característico da circulante $\mathbf{V}$ então temos

$$\sum_{i=0}^{n-1} \lambda_l = -\alpha_{n-1}.$$

E note que pelo Corolário 2.10 2.35

$$nv_0 = -\alpha_{n-1},$$

o que nos dá $v_0 = -\alpha_{n-1}/n$ que é um dos paramêtros de uma matriz circulante que desejamos.

Calculando $\det \mathbf{V}$ sem fazer a expansão do polinômio, aparece no resultado o seguinte termo $(X - v_0)^3$, note que a expansão desse determinante é uma tarefa demasiadamente trabalhosa. Então, para termos uma circulante com um determinante mais simples de se calcular seria conveniente a mudança de varíavel

$$Y = X - v_0 = X + \frac{\alpha_{n-1}}{n}.$$

Assim, fazendo a mudança de varíavel $X = Y - \frac{\alpha_{n-1}}{n}$ na equação

$$p(X) = X^n + \alpha_{n-1}X^{n-1} + \alpha_{n-2}X^{n-2} + \cdots + \alpha_1 X + \alpha_0 = 0,$$

elimina o termo de grau $n-1$ do polinômio p e conduz a equação

$$q(Y) = Y^n + \gamma_{n-2}Y^{n-2} + ... + \gamma_1 X + \gamma_0 = 0$$

a ser resolvida.

E a equação resultante não tem o termo quadrático em Y. Álem disso a teoria circulante motiva de forma muito natural a etapa preliminar existente na solução tradicional das cúbicas

de fazer uma mudança linear de variáveis afim de eliminar o termo quadrático. Daqui em diante para resolver as equações vamos assumir que a mudança de variáveis já foi feita.

Portanto, se $\mathbf{V} = circ\{(v_0, v_1, ..., v_{n-1})\}$ a circulante aderente a p, então a matriz traceless $\mathbf{V} - v_0\mathbf{I}$ é aderente à q pelo Corolário 2.35.

Dessa forma um método razoável para resolver equações polinomiais p de grau n pode, portanto, consistir em uma mudança de variáveis afim de reduzir p para uma equação q com coeficiente zero no monômio de grau $n-1$, e em seguida, encontrar uma matriz circulante traceless $\mathbf{V}$ que é aderente à q. Os autovalores de $\mathbf{V}$ são as raízes de $p_V = q$ e podem ser facilmente calculados utilizando o representante P_V de $\mathbf{V}$. Neste método parece que estamos substituindo o difícil problema de resolver uma equação polinomial mônica de grau n pelo problema mais difícil de resolver $n-1$ equações não-lineares em $n-1$ variáveis. No entanto, por causa das simetrias presentes no último conjunto de equações, elas podem ser mais fáceis de manusear.

3.4 Resolução de Cúbicas e Quárticas

Vamos agora resolver uma cúbica geral. Nosso objetivo é encontrar qualquer matriz circulante traceless 3×3 $\mathbf{V}$

$$\mathbf{V} = \begin{bmatrix} 0 & a & b \\ b & 0 & a \\ a & b & 0 \end{bmatrix}$$

que é aderente a cúbica geral

$$q(Y) = Y^3 + \alpha Y + \beta.$$

e avaliando o representante $P_{\mathbf{V}}(Y) = aY + bY^2$ em $Y = e^{j\frac{2\pi l}{n}}$, $j = 0, 1, 2$, vamos então produzir as raízes de q.

Calculemos o polinômio característico $circ\{(0, a, b)\}$, vemos que

$$p_V(Y) = \det(Y\mathbf{I} - \mathbf{V}) = \det \begin{bmatrix} Y & -a & -b \\ -b & Y & -a \\ -a & -b & Y \end{bmatrix} = Y^3 - 3abY - (a^3 + b^3),$$

agora igualamos p_V a q e assim

$$\begin{aligned} 3ab &= -\alpha \\ a^3 + b^3 &= -\beta. \end{aligned} \tag{3.7}$$

Para completar a solução da equação temos que encontra os valores a e b que satisfazem a equação acima.

Vamos resolver o sistema (3.7) definindo como incognítas b^3 e c^3. Primeiro dividindo a primeira equação por 3 e depois elevando a equação resultante a cubo ficamos com o seguinte

sistema equivalente

$$\begin{aligned} a^3b^3 &= -\frac{\alpha^3}{27} \\ a^3+b^3 &= -\beta. \end{aligned}$$

Note que a^3 e b^3 são as raízes da equação quadrática

$$r\left(X\right) = X^2 + \beta X - \frac{\alpha^3}{27}, \tag{3.8}$$

e uma vez que as raízes da equação acima são dadas por

$$-\frac{1}{2}\beta \pm \frac{1}{2}\sqrt{\frac{4}{27}\alpha^3+\beta^2}.$$

Assim,

$$a = \left(-\frac{1}{2}\beta \pm \frac{1}{2}\sqrt{\frac{4}{27}\alpha^3+\beta^2}\right)^{\frac{1}{3}}$$

e

$$b = \left(-\frac{1}{2}\beta \mp \frac{1}{2}\sqrt{\frac{4}{27}\alpha^3+\beta^2}\right)^{\frac{1}{3}},$$

podemos escolher qualquer conjunto consistente de valores, uma vez que necessitamos somente de um representante, e as raízes de q são dadas pelos valores de P_V nas três raízes cúbicas da unidade

$$\begin{aligned} r_1 &= a+b, \\ r_2 &= ae^{\frac{2\pi i}{3}} + be^{2\frac{2\pi i}{3}}, \\ r_3 &= ae^{2\frac{2\pi i}{3}} + be^{\frac{2\pi i}{3}}. \end{aligned}$$

Temos que a solução acima não é exatamente uma nova fórmula para a solução de uma cúbica. Existem outros métodos que fornecem sistemas de equações idênticos, inclusive a solução dada por Cardano em 1545 que utiliza essencialmente as mesmas equações como já vimos. A vantagem da nossa abordagem consiste no fato de se extender imediatamente para o caso das quárticas.

Vamos agora extender o resultado acima e procurar uma solução circulante para a equação quártica. Considere o polinômio geral

$$q\left(Y\right) = Y^4 + \beta Y^2 + \gamma Y + \delta,$$

com β, γ e δ não nulos. Nós procuramos por uma matriz circulante traceless $\mathbf{V}$ tal que

$$p_{\mathbf{V}}(Y) = det(Y\mathbf{I} - \mathbf{V}) = q(Y). \tag{3.9}$$

Dessa forma temos $\mathbf{V} = circ\{(0,b,c,d)\}$ e seu polinômio característico é dado por

$$p_{\mathbf{V}}(Y) = Y^4 - (4bd + 2c^2)Y^2 - 4c(b^2 + d^2)Y + (c^4 - b^4 - d^4 - 4bc^2d + 2b^2d^2),$$

e de (3.9) temos então

$$\begin{aligned} 4bd + 2c^2 &= -\beta, \\ 4c(b^2 + d^2) &= -\gamma, \\ c^4 - b^4 - d^4 - 4bc^2d + 2b^2d^2 &= \delta, \end{aligned} \tag{3.10}$$

um sistema nas incógnitas b, c, d.

Vamos resolver a primeira e a segunda equação em função de c

$$\begin{aligned} bd &= \frac{-\beta - 2c^2}{4}, \\ b^2 + d^2 &= \frac{-\gamma}{4c}. \end{aligned} \tag{3.11}$$

E reescrevemos a terceira equação de modo a deixar em evidência os termos bd e $b^2 + d^2$, usando o fato de que

$$-\left(b^2 + d^2\right)^2 = -b^4 - 2b^2d^2 - d^4$$

umas vez que o termo $-2b^2d^2$ não aparece na equação 3 vamos somá-lo e subtraí-lo da equação, e a terceira equação é reescrita com

$$c^4 - \left(b^2 + d^2\right)^2 + 4b^2d^2 - 4bdc^2 = \delta. \tag{3.12}$$

Agora, substituindo (3.11) em (3.12) vamos obter uma equação só em c como se segue

$$\begin{aligned} c^4 + \frac{\gamma^2}{16c^2} + \frac{\left(\beta + 2c^2\right)^2}{4} + \left(\beta + 2c^2\right)c^2 &= \delta \\ 16c^6 + \gamma^2 + 4c^2\left(\beta^2 + 4\beta c^2 + 4c^4\right) + 16\beta c^4 + 32c^6 - 16c^2\delta &= 0 \\ 16c^6 + \gamma^2 + 4\beta^2c^2 + 16\beta c^4 + 16c^6 + 16\beta c^4 + 32c^6 - 16\delta c^2 &= 0 \\ 64c^6 + 32\beta c^4 + \left(4\beta^2 - 16\delta\right)c^2 + \gamma^2 &= 0. \end{aligned}$$

a qual é uma equação cúbica se fizermos $c^2 = x$,

$$x^3 + \frac{1}{2}\beta x^2 + \frac{(4\beta^2 - 16\delta)}{64}x - \frac{\gamma^2}{64} \tag{3.13}$$

e podemos resolvê-la usando o método anterior para cúbicas atráves da mudança de variável $x = y - \frac{\beta}{6}$.

Encontramos o valor de c, em seguida, encontramos valores correspondentes para b e d satisfazendo (3.11). Note que b^2 e d^2 são raízes da quadrática

$$\begin{aligned} s\left(X\right) &= X^2 + \frac{\gamma}{4c}X + \left(\frac{-\left(\beta + 2c^2\right)}{4}\right)^2 \\ s\left(X\right) &= 4cX^2 + \gamma X + c^5 + c^3\beta + \frac{c}{4}\beta^2 \end{aligned}$$

que tem como raízes,

$$X = -\frac{1}{8c}\left(\gamma \pm \sqrt{-4c^2\beta^2 + \gamma^2 - 16c^4\beta - 16c^6}\right).$$

Logo,

$$\begin{aligned} b &= \sqrt{-\frac{1}{8c}\left(\gamma + \sqrt{-4c^2\beta^2 + \gamma^2 - 16c^4\beta - 16c^6}\right)} \\ d &= \sqrt{-\frac{1}{8c}\left(\gamma - \sqrt{-4c^2\beta^2 + \gamma^2 - 16c^4\beta - 16c^6}\right)} \end{aligned}$$

Finalmente construimos o representante

$$P_{\mathbf{V}}(Y) = bY + cY^2 + dY^3$$

de $\mathbf{V}$ e o avaliamos em 1,-1,i e $-i$. Assim, as raízes de q são

$$\begin{aligned} q(1) &= b + c + d \\ q(-1) &= -b + c - d \\ q(i) &= -c + (b - d)\, i \\ q(-i) &= -c - (b - d)\, i \end{aligned}$$

Vimos assim que a arbodagem circulante explora as conexões entre álgebra matricial e as raizes dos polinômios. Essas conexões não apenas nos guiaram para as soluções de equações de grau $n \leq 4$, elas também lançaram uma nova luz sobre outras propriedades de polinômio, como por exemplo o resultado familiar sobre a eliminação do termo de grau $n-1$ de um polinômio de grau n. Vamos agora apresentar um outro resultado muito atraente que caracteriza polinômios reais com todas as raízes reais.

Proposição 3.4 *O polinômio $p \in \mathbb{R}[x]$ possui todas as raízes reais se, e somente se, é associado a alguma matriz circulante hermitiana $\mathbf{V} = P_{\mathbf{V}}(\mathbf{W})$.*

Prova. Suponhamos que $\mathbf{V}$ é uma circulante hermitiana. Então pelo Corolário 1.64 $\mathbf{V}$ tem todos os autovalores reais, logo p aderente a $\mathbf{V}$ tem todas as raízes reais.

Por outro lado suponah que p tem todas as raízes reais , e seja $\mathbf{V}$ alguma matriz circulante aderente a p, assim $\mathbf{V}$ tem todas as raízes reais. E temos pelo Teorema 2.35 que

$$\mathbf{V} = \mathbf{E}\mathbf{D}_{\mathbf{V}}\mathbf{E}^*$$

onde aqui $\mathbf{D}_{\mathbf{V}}$ é uma matriz real e assim $\mathbf{D}_{\mathbf{V}} = \mathbf{D}_{\mathbf{V}}^*$. Portanto

$$\mathbf{V}^* = (\mathbf{E}\mathbf{D}_{\mathbf{V}}\mathbf{E}^*)^* = \mathbf{E}\mathbf{D}_{\mathbf{V}}^*\mathbf{E}^* = \mathbf{V}.$$

■

Vamos resumir a ideia geral do método a um algoritmo para a resolução de equações polinomiais.

1. Dado um polinômio p de grau n fazemos a mudança de variável $X = Y - \frac{\alpha_{n-1}}{n}$ e obtemos um polinômio q sem o termo de grau $n-1$.

2. Escrevemos a matriz traceless $\mathbf{V}$ geral $n \times n$ aderente a q.

3. Calculamos o polinômio característico $p_{\mathbf{V}}$ de $\mathbf{V}$.

4. Igualamos os coeficiêntes de $p_{\mathbf{V}}$ aos de q gerando um sistema de $n-1$ equações.

5. Resolvemos o sistema.

6. Calculamos o representante $P_{\mathbf{V}}$ de $\mathbf{V}$.

7. Avalimos $P_{\mathbf{V}}$ nas raízes enésimas da unidade

Exemplo 3.5 *Resolver a cúbica* $p(X) = X^3 - 3X^2 - 3X - 1$.

Solução. Fazendo a mudança de varíavel $X = Y + 1$

$$p(Y+1) = (Y+1)^3 - 3(Y+1)^2 - 3(Y+1) - 1 = 0$$

$$q(Y) = Y^3 - 6Y - 6,$$

que é aderente a matriz circulante traceless $\mathbf{V} = circ\{(0, a, b)\}$.

Por (3.8) a^3 e b^3 são as raízes quadrática

$$r(Y) = Y^2 - 6Y + 8$$

logo $a = \sqrt[3]{2}$ e $b = \sqrt[3]{4}$.

E o representante de $\mathbf{V}$ é

$$P_V(Y) = \sqrt[3]{2}Y + \sqrt[3]{4}Y^2.$$

avaliando $P_V(x)$ nas raízes cúbicas da unidade temos

$$\begin{aligned}
P_V(1) &= \sqrt[3]{2} + \sqrt[3]{4} = r_1 \\
P_V\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right) &= \sqrt[3]{2}\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right) + \sqrt[3]{4}\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right)^2 = r_2 \\
P_V\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right) &= \sqrt[3]{2}\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right) + \sqrt[3]{4}\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right)^2 = r_3.
\end{aligned}$$

Logo as são as raízes de p são $1 + r_1$,$1 + r_2$ e $1 + r_3$.

Exemplo 3.6 *Resolver a cúbica* $p(X) = X^3 - 15X - 4$.

Solução. Devemos primeiro encontar uma matriz circulante $\mathbf{V} = circ\{(0, a, b)\}$ aderente a p. Por (3.8) a^3 e b^3 são as raízes quadrática

$$r(Y) = Y^2 - 4Y + 125$$

logo $a = \sqrt[3]{2 + 11i}$ e $b = \sqrt[3]{2 - 11i}$.

Assim a matriz circulante aderente a p é

$$\mathbf{V} = \begin{bmatrix} 0 & \sqrt[3]{2 + 11i} & \sqrt[3]{2 - 11i} \\ \sqrt[3]{2 - 11i} & 0 & \sqrt[3]{2 + 11i} \\ \sqrt[3]{2 + 11i} & \sqrt[3]{2 - 11i} & 0 \end{bmatrix}.$$

e uma vez que a hermitiana de $\mathbf{V}$ é

$$\begin{bmatrix} 0 & \sqrt[3]{(2 + 11i)} & \sqrt[3]{(2 - 11i)} \\ \sqrt[3]{(2 - 11i)} & 0 & \sqrt[3]{(2 + 11i)} \\ \sqrt[3]{(2 + 11i)} & \sqrt[3]{(2 - 11i)} & 0 \end{bmatrix},$$

Então pela Proposição 3.4 p tem todas as raízes reais.

Vamos agora calcular as raízes de p, o representante de $\mathbf{V}$ é

$$P_{\mathbf{V}}(X) = \sqrt[3]{2 + 11i}X + \sqrt[3]{2 - 11i}X^2,$$

agora avaliando $P_{\mathbf{V}}(X)$ nas raízes cúbicas da unidade temos

$$\begin{aligned} P_{\mathbf{V}}(1) &= \sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i} = r_1 \\ P_{\mathbf{V}}\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right) &= \sqrt[3]{2 + 11i}\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right) + \sqrt[3]{2 - 11i}\left(-\frac{1}{2} + \frac{1}{2}i\sqrt{3}\right)^2 = r_2 \\ P_{\mathbf{V}}\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right) &= \sqrt[3]{2 + 11i}\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right) + \sqrt[3]{2 - 11i}\left(-\frac{1}{2} - \frac{1}{2}i\sqrt{3}\right)^2 = r_3. \end{aligned}$$

Ná fórmula $r_1 = \sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i}$, cada radical tem 3 valores. Dessa forma parece que obteremos mais de 3 raízes para a a cúbica $p(X) = X^3 - 15X - 4$. No entanto como $r_1 = a + b$, com a e b satisfazendo (3.8), assim escolhendo um valor para a entre os três valores possíveis de $\sqrt[3]{2 + 11i}$ o valor b fica imediatamente determinado.

Note que $r_1 = 4$, uma vez que $2 \pm 11i = (2 \pm i)^3$.

Exemplo 3.7 *Resolver a quártica* $p(X) = Z^4 - 2Z^2 + 8Z - 3$

Solução Uma vez que esse polinômio não tem o termo cúbico podemos aplicar diretamente o método circulante sem a necessidade de uma mudança de variável. Devemos encontrar a circulante traceless $\mathbf{V} = circ\{(0, b, c, d)\}$ que seja aderente a p, onde b, c e d satisfazem (3.11) e mais c^2 é raíz de (3.13)**.**

Assim temos que

$$\begin{aligned} bd &= \frac{2-2c^2}{4} \\ b^2+d^2 &= \frac{-8}{4c} \end{aligned}$$

e

$$q(Z) = Z^3 - Z^2 + Z + 1.$$

Fazendo a mudança de variável $Z = Y + \frac{1}{3}$

$$\begin{aligned} q\left(Y+\frac{1}{3}\right) &= \left(Y+\frac{1}{3}\right)^3 - \left(Y+\frac{1}{3}\right)^2 + \left(Y+\frac{1}{3}\right) + 1. \\ q(Y) &= Y^3 + \frac{2}{3}Y + \frac{34}{27} \end{aligned}$$

e temos que q é aderente a circulante $\mathbf{U} = circ\{(0, r, s)\}$.

Por (3.8) r^3 e s^3 raízes da quadrática

$$r(X) = X^2 + \frac{34}{27}X - \frac{8}{729}$$

logo $r = \sqrt[3]{\frac{1}{9}\sqrt{3}\sqrt{11} - \frac{17}{27}}$ e $s = \sqrt[3]{-\frac{1}{9}\sqrt{3}\sqrt{11} - \frac{17}{27}}$ e segue que

$$P_{\mathbf{U}}(X) = \sqrt[3]{\frac{1}{9}\sqrt{3}\sqrt{11} - \frac{17}{27}}X^2 + \sqrt[3]{-\frac{1}{9}\sqrt{3}\sqrt{11} - \frac{17}{27}}X$$

é o representante de $\mathbf{U}$.

Referências Bibliográficas

[1] H. P. Bueno, **Álgebra linear. um segundo curso**. coleção textos universitários, sociedade brasileira de matemática, Rio de Janeiro, 2006.

[2] P. J. Davis, **Circulant Matrices**, AMS Chelsea Publishing, 1994.

[3] D. Geller, I. Kra, S. Popescu and S. Simanca, **On circulant matrices**, preprint 2002. Disponível em < www.math.sunysb.edu/˜sorin/eprints/circulant.pdf >.

[4] R. M. Gray, **Toeplitz and Circulant Matrices: A review** (Foundations and Trends in Communications and Information Theory), NOW, 2005.

[5] D. Kalman and J. E. White, **Polynomial equations and circulant matrices**, Amer. Math. Monthly 108 (2001), p. 821 - 840.

[6] Houssam Khalil. **Matrices structurées et matrices de Toeplitz par blocs de Toeplitz en calcul numérique et formel**. Mathematics. Université Claude Bernard - Lyon I, 2008.

[7] I. Kra and R. Simanca, **On Circulant Matrices**, Notices of AMS Vol. 59 No. 3(2012), pp. 368-377.

[8] Teodoro Lara. **Matrices circulantes.** Divulgaciones Matemáticas Vol. 9 No. 1(2001), pp. 85–102.

[9] Lima, E. L., **Meu Professor de Matemática e Outras Histórias**, Editora da SBM, Rio de Janeiro, 1991.

[10] H. J. Nussbaumer. **Fast Fourier Transform and Convolution Algorithms**. Springer-Verlag, 1982.

[11] Silva, A. de A. e, **Introdução à Álgebra Linear**, Editora Universitária-UFPB, 2007.

[12] A. Wyn-Jones, **Cirulants**, New York, 2013. Disponível em < http://www.circulants.org/circ/circall.pdf >.