UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS

Introdução aos métodos de iteração de subespaço

TRABALHO DE CONCLUSÃO DE CURSO

Débora Zichtl Campos Mariani Pichetti

Florianópolis, 08 de março de 2013

Débora Zichtl Campos Mariani Pichetti

Introdução aos métodos de iteração de subespaço

Trabalho de Conclusão de Curso para obter o grau de Licenciado em Matemática do Centro de Ciências Físicas e Matemáticas da Universidade Federal de Santa Catarina

Orientador:
Licio Hernanes Bezerra

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Florianópolis

08 de março de 2013

Esta monografia foi julgada adequada como TRABALHO DE CONCLUSÃO DE CURSO no Curso de Matemática - Habilitação Licenciatura, e aprovada em sua forma final pela Banca Examinadora designada pela Portaria n13/CCM/2013.

Prof. Nereu Estanislau Burin
Professor da disciplina

Banca Examinadora:

Prof. Licio Hernanes Bezerra
Orientador

Prof. Juliano de Bem Francisco

Prof. Flávia Tereza Giordani

Sumário

Resumo

Neste trabalho, cujo objetivo principal é estudar certos métodos iterativos para a computação de alguns autovalores de uma matriz, veremos três importantes processos iterativos: o Método de Potência, a Iteração Ortogonal e o Método LOPSI. Os dois últimos métodos, que na verdade consistem em uma generalização do Método de Potência, são utilizados para encontrarmos simultaneamente mais de um autovalor de uma dada matriz, recebendo a denominação Iterações de Subespaço. Para finalizar, faremos testes com matrizes de autovalores sensíveis, utilizando os dois métodos de iteração de subespaço.

Palavras-chave: Autovalor, Método de Potência, Iteração ortogonal, Método LOPSI, Iteração de Subespaço.

Abstract

In this work, whose goal is to study certain iterative methods for computing some eigenvalues of a matrix, we see three major iterative processes: the Power Method, Orthogonal Iteration and LOPSI Method. The latter two methods, which are actually a generalization of the power method, are used to find more than one eigenvalue of a given matrix and are called Subspace Iterations. Finally, we test matrices with sensitive eigenvalues, using both methods of subspace iteration.

Keywords: Eigenvalue, Power Method, Orthogonal Iteration, LOPSI Method, Subspace Iteration.

Agradecimentos

Dedico esta, bem como todas as minhas conquistas, aos meu grandes mentores e heróis, os quais tenho enorme prazer em chamar de pais, Antonio Pichetti Junior e Geanne Zichtl Campos. Não posso deixar de incluir porém, aqueles que estiveram sempre na minha companhia, que me enchem de orgulho e me inspiram vontade de ser alguém melhor; meus irmãos Douglas Pichetti e Daniella Pichetti.

Agradeço à minha família em geral, por todo o carinho e atenção que sempre me deram e, em especial, durante os cinco anos de caminhada nesse curso, pois em nenhum momento deixaram de acreditar em mim e de me darem forças, mesmo à distância. Agradeço também aos meus amigos, que compreenderam muitas vezes minha ausência, porém mesmo assim continuaram ao meu lado, transformando minhas angústias em sorrisos.

Agradeço a cada um de meus professores, dos quais nenhum passou sem deixar em mim alguma marca. Obrigada pelos desafios que propuseram, pelo conhecimento que conosco dividiram e por mais esta vitória que me proporcionaram alcançar. Um agradecimento especial aos professores Nereu Estanislau Burin, Aldrovando Luiz Azeredo e Licio Hernanes Bezerra, que por vezes foram muito mais que professores; foram grandes amigos que espero poder ter sempre, independente do rumo que a vida nos trilhar.

Acima de tudo agradeço à Deus, força maior que permite que nossa existência aconteça da maneira mais brilhante e surpreendente que possamos imaginar.

Introdução

Autovalores e autovetores são objetos matemáticos amplamente utilizados na matemática, física, engenharia, entre diversas outras áreas. No contexto de Controle, são essenciais nos processos de observação de estabilidade, frequências naturais e modos de vibração, por exemplo. A própria resposta temporal de um sistema linear invariante no tempo é uma exponencial que depende do autovalor.

O problema de autovalores é um dos problemas centrais da Algebra Linear. Jacobi foi pioneiro na formulação de um método iterativo para calcular o espectro de uma matriz simétrica, isso em 1846. Hoje em dia, há várias maneiras práticas de obtermos os autovalores de uma dada matriz. O sistema iterativo MATLAB, por exemplo, tem a função *eig* para calcular o espectro de uma matriz genérica via método **QR**, método surgido no início da segunda metade do século XX. A convergência deste método para matrizes não normais pode ser problemática, por exemplo, se alguns dos autovalores estiverem agrupados, muito próximos um do outro. Há métodos que computam apenas um autovalor de uma matriz, como por exemplo o Método de Potência, que abordaremos neste trabalho. Contudo, nosso foco será tratar de métodos que computam simultaneamente vários autovalores de uma dada matriz, mais especificamente, os métodos de iteração de subespaço, Iteração Ortogonal e LOPSI.

O trabalho foi organizado de forma que pudéssemos trazer grande parte dos resultados importantes da Álgebra Linear relacionados com nosso tema de pesquisa. Por isso os Capítulos 1, 2 e 3, consistem em uma grande revisão de conceitos básicos envolvendo espaços vetoriais e transformações lineares, bem como trazem algumas notações importantes que serão utilizadas no decorrer da pesquisa. Os Capítulos 4, 5 e 6 trazem conceitos mais específicos a respeito de autovalores e operadores lineares, bem como algumas decomposições matriciais importantes como a fatoração **QR** e a decompsição de Schur, que serão utilizadas nos métodos iterativos que trabalharemos. Finalmente, no Capítulo 7 introduzimos o Método de Potência e em seguida os Métodos de Iteração de Subespaço: Iteração Ortogonal e Método LOPSI, foco desta pesquisa. Ainda neste Capítulo, apresentaremos os resultados de experimentos que foram feitos utilizando tais métodos para a obtenção dos autovalores de algumas matrizes especiais.

1 Conceitos Básicos da Álgebra Linear

Neste primeiro tópico do trabalho, revisaremos alguns conceitos básicos da Álgebra Linear bem como definiremos notações que serão utilizados no decorrer desta pesquisa.

1.1 Matrizes

Seja $\mathbb{K}$ um corpo de característica zero. Vamos denotar por $\mathbb{K}^{m\times n}$ o conjunto das matrizes $m \times n$ cujas entradas estão em $\mathbb{K}$. Por exemplo, $\mathbb{R}^{m\times n}$ denota o conjunto das matrizes reais $m \times n$. Como é usual, denotaremos por a_{ij} o elemento de uma matriz $\mathbf{A}$ que está na linha i e coluna j.

$$\begin{pmatrix} a_{11} & & \cdots & & a_{1n} \\ & \ddots & & & \\ \vdots & & a_{ij} & & \vdots \\ & & & \ddots & \\ a_{m1} & & \cdots & & a_{mn} \end{pmatrix}$$

Para denotar o vetor-linha i da matriz A escreveremos $\mathbf{A}(i,:)$, que é a notação utilizada pelo sistema interativo MATLAB (entre outros softwares, como os sistemas interativos lives FreeMat, GNU Octave e Scilab). Da mesma forma, para denotarmos o vetor coluna j , utilizaremos $\mathbf{A}(:,j)$. A submatriz de A formada pelos elementos que estão ao mesmo tempo nas linhas $i_1 < i_2, < \ldots < i_r$ e nas colunas $j_1 < j_2 < \ldots < j_s$ será denotada por $\mathbf{A}([i_1,\ldots,i_r],[j_1,\ldots,j_s])$. Se as linhas e as colunas forem consecutivas, vamos denotar essa mesma submatriz por

$$\mathbf{A}(i_1 : i_r, j_1 : j_s).$$

Submatrizes desse tipo são chamadas de blocos. Por exemplo, na matriz:

$$\mathbf{A} = \left(\begin{array}{c|cc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right)$$

as submatrizes

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}, etc.,$$

são blocos. Podemos reescrever a matriz $\mathbf{A}$ como sendo

$$\begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \mathbf{P}_{13} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \mathbf{P}_{23} \end{pmatrix}$$

em que cada $\mathbf{P}_{ij}$ é um dos blocos indicados na partição da matriz $\mathbf{A}$. Podemos perceber que se dois elementos dessa última matriz têm o primeiro índice igual, então os blocos que representam têm a mesma quantidade de linhas e se dois elementos têm o segundo índice igual, então os blocos que representam têm a mesma quantidade de colunas.

Uma propriedade importante de matrizes em blocos é que estas podem ser tratadas, até certo ponto, como se cada bloco fosse um só elemento. É claro que é necessário preservar a ordem desses elementos durante os cálculos, para que os blocos tenham dimensões adequadas para que todas as somas e produtos envolvidos estejam bem definidos. O teorema que segue mostra como pode-se fazer isto.

Teorema 1. *Sejam*

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1s} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2s} \\ \vdots & \vdots & & \vdots \\ \mathbf{P}_{r1} & \mathbf{P}_{r2} & \cdots & \mathbf{P}_{rs} \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \cdots & \mathbf{Q}_{1t} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \cdots & \mathbf{Q}_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{Q}_{s1} & \mathbf{Q}_{s2} & \cdots & \mathbf{Q}_{st} \end{pmatrix}$$

matrizes em blocos tais que para cada j, $j = 1,\ldots,r$ e para cada l, $l = 1,\ldots,t$, o número de colunas de $\mathbf{P}_{jk}$ é igual ao número de linhas em $\mathbf{Q}_{kl}$, $k = 1,\ldots,s$.

Então o produto $\mathbf{PQ}$ pode ser particionado nos blocos $\mathbf{R}_{ik}$, $i = 1,\ldots,r$, $k = 1\ldots,t$, em que

$$\mathbf{R}_{ik} = \mathbf{P}_{i1}\mathbf{Q}_{1k} + \mathbf{P}_{i2}\mathbf{Q}_{2k} + \ldots + \mathbf{P}_{is}\mathbf{Q}_{sk}$$

Definição 1. *Seja $\mathbf{A} = (a_{ij})$ uma matriz $m \times n$. Então a matriz $\mathbf{C} = (c_{ij})$, de ordem $n \times m$ tal que $a_{ij} = c_{ji}$, $i = 1,\ldots,m$, $j = 1,\ldots n$, é chamada de transposta de $\mathbf{A}$ e é denotada por $\mathbf{A}^T$.*

A operação transposição tem duas importantes propriedades:

i) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$, sempre que o produto $\mathbf{AB}$ está bem definido, e

ii) $(\mathbf{A}^T)^T = \mathbf{A}$, para toda matriz $\mathbf{A}$.

Esta notação será utilizada no decorrer deste trabalho, especialmente com vetores coluna. Um vetor coluna $\mathbf{u}$ com entradas $u_1, u_2, \ldots, u_n$ será escrito como $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$.

Definição 2. *Uma matriz* $\mathbf{A}$ *que satisfaz* $\mathbf{A} = \mathbf{A}^T$ *é dita uma matriz simétrica.*

Definição 3. *O conjugado transposto ou transposto Hermitiano, ou matriz adjunta de uma matriz* $m \times n$ $\mathbf{A}$ *com entradas complexas é a matriz* $n \times m$ $\mathbf{A}^H$ *obtida de* $\mathbf{A}$ *por tomar a transposta e então tomar o conjugado complexo de cada entrada. Denotamos o transposto conjugado de uma matriz* $\mathbf{A} = (a_{ij})$ *é denotado por* $\mathbf{A}^H = (\overline{a_{ji}})$.

Definição 4. *Dada uma matriz* $\mathbf{A} = (a_{ij})$ *de ordem* $m \times n$, *então ela é dita hermitiana (ou auto-adjunta) se for igual à sua transposta conjungada. Simbolicamente,* $\mathbf{A} = a_{ij} = \overline{a_{ji}} = \mathbf{A}^H$. *Dizemos que* $\mathbf{A}$ *é anti-hermitiana se* $\mathbf{A} = -\mathbf{A}^H$. *Se* $\mathbf{A}^{-1} = \mathbf{A}^H$ *então* $\mathbf{A}$ *é denominada unitária. Uma matriz* $\mathbf{A}$ *que satisfaz* $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ *é dita uma matriz normal.*

Vamos definir agora a adjunta clássica de uma matriz, que não deve ser confundida com a matriz adjunta, ou conjugado transposto, definida em (3).

Definição 5. *A adjunta (ou adjunta clássica) de uma matriz quadrada* $\mathbf{A} = (a_{ij})$ *é a transposta da matriz que se obtém substituindo cada termo* (a_{ij}) *pelo determinante da matriz resultante de retirar de* $\mathbf{A}$ *a linha* i *e a coluna* j *(isso é, o determinante menor) multiplicado por* $(-1)^{i+j}$. *A adjunta de uma matriz é denotada por* $adj(\mathbf{A})$.

Nossa atenção especial será voltada às matrizes quadradas. Neste âmbito, muitas classes especias de matrizes são importantes.

Em uma matriz quadrada $\mathbf{A} = (a_{ij})$ de ordem n, os elementos a_{ij} que satisfazem $i = j,\ i = 1, \ldots, n,\ j = 1, \ldots, n$, compõem a diagonal principal de $\mathbf{A}$, também chamada simplismente de diagonal de $\mathbf{A}$.

Definição 6. *Uma matriz* $\mathbf{A} = (a_{ij})$ *de ordem* n *é dita triangular inferior se* $a_{ij} = 0$ *para* $i < j,\ \ i, j = 1, \ldots, n$ *e é dita triangular superior se* $a_{ij} = 0$ *para* $i > j,\ \ i, j = 1, \ldots, n$.

Definição 7. *Uma matriz quadrada* $\mathbf{A} = (a_{ij})$ *de ordem* n *é dita diagonal se for simultaneamente triangular superior e inferior, ou seja,* $a_{ij} = 0$ *para* $i \neq j$, $i, j = 1, \ldots, n$. *Dados* $d_1, \ldots, d_k \in \mathbb{K}$, *denotamos a matriz diagonal* $\mathbf{D}$ *tal que* $(\forall i) d_{ii} = d_i$, *por* $diag(d_1, \ldots, d_k)$.

A matriz diagonal $\mathbf{I}_n$ definida por

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

é chamada de matriz identidade, que em MATLAB é definida pelo comando $eye(n)$. Sua i-ésima coluna é o i-ésimo vetor cânonico $\mathbf{e_i} \in \mathbb{K}^n$. Simbolicamente,

$$\mathbf{I}_n = (\mathbf{e_1} \;\; \mathbf{e_2} \;\; \cdots \;\; \mathbf{e_n})$$

Quando no contexto estiver clara qual a dimensão da matriz identidade, eliminaremos o índice n e escreveremos apenas I.

Exemplo 1. *As matrizes*

$$\begin{pmatrix} 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

podem ser classificadas respectivamente como triangular inferior, triangular inferior, triangular superior e diagonal. A quarta matriz também pode ser escrita como $diag(2,-1,3)$.

Definição 8. *Chamamos de superdiagonal de uma matriz a qualquer diagonal paralela à diagonal principal que esteja na parte triangular superior da matriz e de subdiagonal de uma matriz a qualquer diagonal paralela à diagonal principal que esteja na parte triangular inferior da matriz.*

Teorema 2. *O produto de duas matrizes triangulares inferiores de mesma ordem é ainda uma matriz triangular inferior. O produto de duas matrizes triangulares superiores é ainda uma matriz triangular superior. O produto de duas matrizes diagonais é ainda uma matriz diagonal.*

Definição 9. *Uma matriz* $\mathbf{A} = (a_{ij})$ *de ordem n é dita uma matriz de Hessenberg superior se* $a_{ij} = 0$ *para* $i > j+1$ *e Hessenberg inferior se* $a_{ij} = 0$ *para* $i < j+1$.

Definição 10. *Uma matriz* $\mathbf{A} = (aij)$ *de ordem n é dita uma matriz de banda se possuir p superdiagonais e q subdiagonais,* $p, q > 1$.

Um outro conceito importante é o de singularidade de matrizes.

Definição 11. *Uma matriz quadrada* $\mathbf{A}$ *de ordem n é dita não singular se existir uma matriz* $\mathbf{B}$ *tal que* $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$*, onde* $\mathbf{I}$ *representa a matriz identidade de ordem n. A matriz* $\mathbf{B}$ *é definida como sendo a matriz inversa de* $\mathbf{A}$ *e é denotada por* $\mathbf{A}^{-1}$.

Teorema 3. *Dada uma matriz não singular* $\mathbf{A}$*, sua matriz inversa,* $\mathbf{A}^{-1}$*, é determinada de forma única. Além disso, temos que* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

Teorema 4. *Se* $\mathbf{A}$ *e* $\mathbf{B}$ *são duas matrizes não singulares de mesma ordem, então o produto* $\mathbf{AB}$ *é também não singular e*

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Em outras palavras, o inverso do produto é o produto dos inversos em ordem trocada.

Teorema 5. *Se* $\mathbf{A}$ *é uma matriz não singular então* $\mathbf{A}^T$ *é também uma matriz não singular e*

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$$

Em outras palavras, a inversa da transposta é a transposta da inversa.

A partir do conceito de singularidade de matrizes, surgem muitas outras definições e resultados, como por exemplo o conceito de semelhança de matrizes.

Definição 12. *Duas matrizes* $\mathbf{A}$ *e* $\mathbf{B}$ *de ordem n são ditas semelhantes se existir uma matriz não singular* $\mathbf{C}$ *tal que* $\mathbf{A} = \mathbf{C}^{-1}\mathbf{BC}$.

A noção de semelhança define uma relação de equivalência no conjunto das matrizes quadradas no sentido que

i) $\mathbf{A}$ é semelhante consigo mesma;

ii) Se $\mathbf{A}$ é semelhante a $\mathbf{B}$, então $\mathbf{B}$ é semelhante a $\mathbf{A}$; e

iii) Se $\mathbf{A}$ é semelhante a $\mathbf{B}$ e $\mathbf{B}$ é semelhante a $\mathbf{C}$, então $\mathbf{A}$ é semelhante a $\mathbf{C}$.

O estudo das funções traço e determinante de matrizes quadradas é importante para um estudo posterior a respeito de matrizes não singulares e inversas.

Definição 13. *Seja* $\mathbf{A} = (a_{ij})$ *uma matriz de ordem n. O traço de* $\mathbf{A}$ *é o número*

$$tr(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Definição 14. *Seja* $\mathbf{A} = (a_{ij})$ *uma* $m \times n$ *matriz. O determinante de* $\mathbf{A}$*,* $det(\mathbf{A})$ *é o número*

$$det(\mathbf{A}) = \sum_{P=k_1,\ldots,k_n} \sigma(P) a_1 k_1 a_2 k_2 \cdots a_n k_n \tag{1.1}$$

O somatório é sobre todas as permutações P dos números inteiros $1,\ldots,n$ *e o fator* $\sigma(P)$ *é o sinal da permutação P. É igual a* $+1$ *se P permutação é par, e* -1 *se P é ímpar. A permutação*

$$\begin{pmatrix} 1, & 2, & \ldots, & n \\ k_1, & k_2, & \ldots, & k_n \end{pmatrix}$$

é dita par se tiver um número par de inversões e ímpar se tiver um número ímpar de inversões.

Vamos agora citar duas consequências da definição de (1.1). Claramente, se $\mathbf{I}$ é a matriz identidade, então

$$det(\mathbf{I}) = 1 \tag{1.2}$$

As propriedades da permutação implicam que para uma matriz quadrada $\mathbf{A}$ qualquer,

i) $det(\mathbf{A}^T) = det(\mathbf{A})$

ii) $det(\mathbf{A}^H) = \overline{det(\mathbf{A})}$

iii) $det(\mathbf{AB}) = det(\mathbf{A})det(\mathbf{B})$

iv) Se $\mathbf{A}$ tem ordem n e α é um escalar qualquer, então, $det(\alpha\mathbf{A}) = \alpha^n det(\mathbf{A})$

v) $det(\mathbf{A}^{-1}) = det(\mathbf{A})^{-1}$

Há várias matrizes especiais que podem ser definidas a partir de comandos no MATLAB. Abaixo alguns exemplos.

- A matriz nula $m \times n$ que é toda formada por zeros, é definida pelo comando $zeros(m,n)$;
- A matriz $m \times n$ formada toda por uns é definida pelo comando $ones(m,n)$;
- A matriz companheira fica melhor definida a partir do conceito de autovalor, objeto que ainda abordaremos neste trabalho. Seja $p(x) = a_n x^n + \ldots + a_1 x + a_0$, $a_n \neq 0$ um polinômio em $\mathbb{K}$. A matriz companheira associada a p é a matriz $\mathbf{C} = compan([a_n,\ldots,a_0])$, que é a matriz cuja primeira linha é $-a_{n-1}/a_n \ \ldots \ -a_1/a_n \ -a_0/a_n$, a segunda linha é o primeiro vetor canônico de $\mathbb{K}^n$, $\mathbf{e_1}$, a terceira linha é $\mathbf{e_2}$, ..., a enésima linha é $\mathbf{e_{n-1}}$. Exemplo: se $p(x) = x^4 - 5x^3 + 3x^2 + x - 1$, então

$$\begin{pmatrix} 5 & -3 & -1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Veremos em um capítulo adiante que o polinômio dado é o polinômio característico de $\mathbf{C}$. Isso é um fato geral: dado um vetor $\mathbf{v} = (a_n, \ldots, a_0) \in \mathbb{K}^n$, em que $a_n \neq 0$, o polinômio $x^n - \frac{a_{n-1}}{a_n}x^{n-1} - \cdots - \frac{a_1}{a_n}x - \frac{a_0}{a_n}$ é o polinômio característico de $\mathbf{C} = compan(v)$.

- Matrizes de Hankel são matrizes retangulares $m \times n$ tais que para todo k, $0 \leq k \leq m+n-2$, $h_{ij} = h_k$, em que $1 \leq i \leq m$ e $1 \leq j \leq n$ são tais que $i-1+j-1 = k$. Ou seja, são matrizes cujos elementos de cada diagonal paralela à diagonal secundária são todos iguais. Para gerar tais matrizes, basta sabermos sua primeira coluna e sua última linha. Em MATLAB, define-se uma matriz de Hankel pelo comando

$$hankel([h_0, h_1, \ldots, h_{m-1}], [h_{m-1}, \ldots, h_{m+n-2}]).$$

- Um exemplo de matriz de Hankel é a matriz de Hilbert, que é uma matriz $n \times n$, $\mathbf{H} = hilb(n)$, cujas entradas são definidas por $H_{i,j} = 1/(i+j-1)$.

- Um quadrado mágico simples de ordem n é uma matriz cujas entradas vão de 1 a n^2 e estão dispostas de tal modo que a soma dos elementos de cada linha, de cada coluna e de cada diagonal (principal e secundária) é sempre igual a um mesmo valor. No MATLAB, o comando para gerar tal matriz é dado por $magic(n)$.

- O comando $\mathbf{P} = pascal(n)$ gera uma matriz simétrica que satisfaz a propriedade similar a que define um triângulo de Pascal: $p(i,1) = p(1,j)$ e $p(i+1,j+1) = p(i+1,j) + p(i,j+1)$ para $i,j \geq 1$. A matriz de Pascal triangular inferior de ordem n, denotada por $\mathbf{P}_n$, é a matriz cujo elemento da linha i e coluna j é definida por

$$(\mathbf{P}_n)_{ij} = \begin{cases} \begin{pmatrix} i-1 \\ j-1 \end{pmatrix}, & \text{se } i \geq j, \\ 0, & \text{se } i < j. \end{cases}$$

Por exemplo,

$$\mathbf{P}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{pmatrix}.$$

Essa matriz é gerada pelo comando $abs(pascal(4,1))$.

- Matrizes de Toeplitz também são matrizes retangulares de ordem $m \times n$ tais que para todo k $1-n \leq k \leq m-1$, $h_{ij} = h_k$, em que $1 \leq i \leq m$ e $1 \leq j \leq n$ satisfazem $i - j = k$. Ou seja, são matrizes cujas entradas de cada diagonal paralela à diagonal principal são todas iguais. Uma matriz de Toeplitz é determinada se soubermos sua primeira coluna e sua primeira linha. Em MATLAB, uma matriz de Toeplitz é construída a partir do comando

$$toeplitz([h_0, h_1, \ldots, h_{m-1}], [h_{1-n-(1-n)}, h_{1-n-(2-n)}, \ldots, h_{1-n}]).$$

- Dado um vetor $x = (x_1, \ldots, x_n)$, em MATLAB o comando $\mathbf{V} = vander(x)$ gera a matriz de Vandermonde $n \times n$ cujos elementos são definidos por $v_{ij} = x_{n-j+1}^{i-1}$:

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_n & x_{n-1} & \cdots & x_1 \\ x_n^2 & x_{n-1}^2 & \cdots & x_1^2 \\ \vdots & \vdots & \cdots & \vdots \\ x_n^{n-1} & x_{n-1}^{n-1} & \cdots & x_1^{n-1} \end{pmatrix}$$

2 Espaços Vetoriais

2.1 Espaços Vetoriais e Subespaços

Vamos formalizar o conceito de espaço vetorial e demonstrar alguns resultados importantes que serão necessários no decorrer desta pesquisa.

Este primeiro teorema será de fundamental importância quando discutirmos raízes de certos polinômios no Capítulo 5. Não daremos sua demonstração aqui, mas ela pode ser encontrada facilmente em livros de Álgebra.

Teorema 6. *(Teorema Fundamental da Álgebra) Todo polinômio com coeficientes complexos possui raízes complexas.*

Um conjunto que satistfaz a propriedade acima é dito algebricamente fechado.

Definição 15. *Um conjunto não vazio $\mathbb{K}$ é um corpo se em $\mathbb{K}$ pudermos definir duas operações, denotadas por $+$ (adição) e $\cdot$ (multiplicação), satisfazendo as seguintes prorpiedades:*

A1) $\alpha+\beta=\beta+\alpha$, $\forall\ \alpha,\beta\in\mathbb{K}$ *(comutatividade da adição)*

A2) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$, $\forall\ \alpha,\beta,\gamma\in\mathbb{K}$ *(associatividade da adição)*

A3) *Existe um único elemento* $0\in\mathbb{K}$ *tal que* $\alpha+0=0+\alpha=\alpha$, $\forall\ \alpha\in\mathbb{K}$ *(exitência do elemento neutro da adição)*

A4) *Para cada* $\alpha\in\mathbb{K}$ *existe um ínico elemento* $(-\alpha)$ *tal que* $\alpha+(-\alpha)=0$ *(existência do elemento oposto da adição)*

M1) $\alpha\cdot\beta=\beta\cdot\alpha$, $\forall\ \alpha,\beta\in\mathbb{K}$ *(propriedade comutativa da multiplicação)*

M2) $\alpha\cdot(\beta\cdot\gamma)=(\alpha\cdot\beta)\cdot\gamma$, $\forall\ \alpha,\beta,\gamma\in\mathbb{K}$ *(propriedade associativa da multiplicação)*

M3) *Existe um único* $1\in\mathbb{K}$ *tal que* $\alpha\cdot 1=1\cdot\alpha=\alpha$, $\forall\ \alpha\in\mathbb{K}$ *(existência do elemento neutro da multiplicação)*

M4) *Para cada* α *não nulo em* $\mathbb{K}$ *existe um único* α^{-1} *tal que* $\alpha \cdot \alpha^{-1} = 1$ *(existência do elemento inverso da multiplicação)*

D) $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$, $\forall\ \alpha, \beta, \gamma \in \mathbb{K}$ *(propriedade distributiva)*

Definição 16. *Um conjunto não vazio* V *é um espaço vetorial sobre (um corpo)* $\mathbb{K}$ *se em seus elementos, denominados vetores, estiverem definidas as seguintes duas operações:*

A) *A cada par* $\mathbf{u}, \mathbf{v}$ *de vetores de* V*, corresponde a um vetor* $\mathbf{u} + \mathbf{v} \in V$ *chamado de soma de* $\mathbf{u}$ *e* $\mathbf{v}$ *de modo que sejam satisfeitas as propriedades da adição para corpos.*

Observação 1. *O elemento neutro da adição em* V *é chamado vetor nulo e é denotado por* $\mathbf{0}$*.*

M) *A cada par* $\alpha \in \mathbb{K}$ *e* $v \in V$*, corresponde a um vetor* $\alpha \cdot \mathbf{v} \in V$*, denominado produto por escalar de* α *por* $\mathbf{v}$*, de modo que*

M1) $(\alpha\beta) \cdot \mathbf{v} = \alpha(\beta \cdot \mathbf{v})$, $\forall\ \alpha, \beta \in \mathbb{K}$ *e* $\mathbf{v} \in V$ *(propriedade associativa)*

M2) $1 \cdot \mathbf{v} = \mathbf{v}$, $\forall\ \mathbf{v} \in V$ *(no qual 1 é o elemento neutro da multiplicação em* $\mathbb{K}$*)*

D1) $\alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$, $\forall\ \alpha \in \mathbb{K}$ *e* $\forall\ \mathbf{u}, \mathbf{v} \in V$

D2) $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \beta + \beta \cdot \mathbf{v}$, $\forall\ \alpha, \beta \in \mathbb{K}$ *e* $\forall\ \mathbf{v} \in V$

Observação 2. *Podemos denotar o espaço vetorial simplesmente por* V *ou quando for desejável especificar o corpo, usaremos a expressão* $\mathbb{K}$*-espaço vetorial* V*.*

Observação 3. *Seja* V *um* $\mathbb{K}$*-espaço vetorial. O conjunto* V *com a operação de soma de vetores é um grupo abeliano. Portanto cada vetor nulo é único assim como cada vetor tem um único vetor oposto.*

Definição 17. *Seja* V *um espaço vetorial e* W *um subconjunto não vazio de* V*. Se* W *é um espaço vetorial em relação às operações de* V*, dizemos que* W *é um subespaço de* V*.*

Exemplo 2. *Todo espaço vetorial tem pelo menos dois subespaços; ele próprio e o subespaço que contém apenas o vetor nulo. Este último é conhecido como subespaço nulo.*

Teorema 7. *Um subconjunto não vazio* W *de* V *é um subespaço de* V *se, e somente se, para cada par de vetores* $\mathbf{u}, \mathbf{v} \in W$ *e cada escalar* $\alpha \in \mathbb{K}$*, o vetor* $\alpha\mathbf{u} + \mathbf{v}$ *pertence a* W*.*

Demonstração: Suponhamos que W seja um subconjunto não vazio de V tal que $\alpha\mathbf{u}+\mathbf{v}$ pertença a W para todos os vetores $\mathbf{u},\mathbf{v} \in W$ e todos os escalares $\alpha \in \mathbb{K}$. Como W é não vazio, existe um vetor $\mathbf{p} \in W$. Logo, $(-1)\mathbf{p}+\mathbf{p} = \mathbf{0}$ está em W. Então se $\mathbf{u}$ é um vetor arbitrário em W e α um escalar arbitrário, o vetor $\alpha\mathbf{u} = \alpha\mathbf{u}\mathbf{0}$ está em W. Em particular, $(-1)\alpha = -\alpha$ está em W. Finalmente se $\mathbf{u}$ e $\mathbf{v}$ estão em W, então $\mathbf{u}+\mathbf{v} = 1\mathbf{u}+\mathbf{v}$ está em W. assim, W é um subespaço de V.

Reciprocamente, se W é um subespaço de V, $\mathbf{u}$ e $\mathbf{v}$ estão em W e α é um escalar, certamente $\alpha\mathbf{u}+\mathbf{v}$ está em W. ■

Exemplo 3. *Seja* $\mathbf{A}$ *uma* $m \times n$ *matriz sobre* $\mathbb{K}$. *Então o conjunto de todas as* $n \times 1$ *matrizes (colunas)* $\mathbf{X}$ *sobre* $\mathbb{K}$ *tais que* $\mathbf{A}X = \mathbf{0}$*, é um subespaço do espaço de todas as* $n \times 1$ *matrizes sobre* $\mathbb{K}$ *Para demonstrar isso, é preciso provas que* $\mathbf{A}(\alpha\mathbf{X}+\mathbf{Y}) = \mathbf{0}$ *para* $\mathbf{AX} = \mathbf{0}$*,* $\mathbf{AY} = \mathbf{0}$ *e* α *um escalar arbtrário de* $\mathbb{K}$. *Isto decorre imediatamente do seguinte lema:*

Lema 1. *Se* $\mathbf{A}$ *é uma* $m \times n$ *matriz sobre* $\mathbb{K}$ *e* $\mathbf{B},\mathbf{C}$ *são* $n \times p$ *matrizes sobre* $\mathbb{K}$*, então*

$$\mathbf{A}(\alpha\mathbf{B}+\mathbf{C}) = \alpha(\mathbf{AB})+\mathbf{AC}$$

para todo escalar $\alpha \in \mathbb{K}$.

Demonstração:

$$\begin{aligned}
[\mathbf{A}(\alpha\mathbf{B}+\mathbf{C})]_{ij} &= \sum_{k=1}^{n} \mathbf{A}_{ik}(\alpha\mathbf{B}+\mathbf{C})_{kj} \\
&= \sum_{k=1}^{n} (\alpha\mathbf{A}_{ik}\mathbf{B}_{kj} + \mathbf{A}_{ik}\mathbf{C}_{kj}) \\
&= \alpha\sum_{k=1}^{n} \mathbf{A}_{ik}\mathbf{B}_{kj} + \sum_{k=1}^{n} \mathbf{A}_{ik}\mathbf{C}_{kj} \\
&= \alpha(\mathbf{AB})_{ij} + (\mathbf{AC})_{ij} \\
&= [\alpha(\mathbf{AB})+\mathbf{AC}]_{ij}.
\end{aligned}$$

■

Analogamente, pode-se mostrar que $(\alpha\mathbf{B}+\mathbf{C})\mathbf{A} = \alpha(\mathbf{BA})+\mathbf{CA}$, se as somas e produtos de matrizes estão definidos.

Teorema 8. *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{K}$. *A intersecção de uma coleção arbitrária de subespaços de* V *é um subespaço de* V.

Demonstração: Seja $\{W_a\}$ uma coleção de subespaços de V e seja $W = \bigcap_a W_a$ a sua intersecção. Recordemos que W é definido como sendo o conjunto dos elementos pertencentes simultaneamente a W_a. Como cada W_a é um subespaço, todos contêm o vetor nulo. Assim, o vetor nulo está na intersecção W o que implica que W é não vazio. Sejam $\mathbf{u}$ e $\mathbf{v}$ vetores em W e seja α um escalar. Pela definição de W, tanto $\mathbf{u}$ como $\mathbf{v}$ pertencem a cada W_a e, como cada W_a é um subespaço, o vetor $(\alpha\mathbf{u}+\mathbf{v})$ está em todo W_a. Assim $(\alpha\mathbf{u}+\mathbf{v})$ está em W. Pelo Teorema (7), W é um subespaço de V. ■

Este teorema mostra que se S é uma coleção arbitrária de vetores em V, então existe um menor subespaço de V que contém S, isto é, um subespaço que contém S e que está contido em todos os outros subespaços que contêm S.

Definição 18. *Seja S um conjunto de vetores num espaço vetorial V. O subespaço gerado por S é definido como sendo a intersecção W de todos os subespaços de V que contêm S. Quando S é um conjunto finito de vetores, $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, denominamos W simplesmente o subespaço gerado pelos vetores $\mathbf{v}_1, \dots, \mathbf{v}_n$, que será denotado por $[\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_n]$.*

Definição 19. *Seja V um -espaço vetorial. Dizemos que um vetor $\mathbf{v} \in V$ é uma combinaçao linear dos vetores $\mathbf{v}_1 \cdots \mathbf{v}_n \in V$ se existirem os escalares $\alpha_1 \cdots \alpha_n$ em $\mathbb{K}$ tais que $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \sum_{i=1}^{n} \alpha_i\mathbf{v}_i$.*

Teorema 9. *O subespaço gerado por um subconjunto não vazio S de um espaço vetorial V é o conjunto de todas as combinações lineares de vetores em S.*

Observação 4. *Por definição temos que o conjunto vazio gera o espaço vetorial $\{\mathbf{0}\}$.*

Demonstração: Seja W o subespaço gerado por S. Então, cada combinação linear

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_m$$

de vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ em S evidentemente está em W. Assim, W contém o conjunto L de todas as combinações lineares de vetores em S. O conjunto L, por outro lado, contém S e é não vazio. Se $\mathbf{v}, \mathbf{u}$ pertencem a L, então $\mathbf{v}$ é uma combinação linear

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_m$$

de vetores $\mathbf{v}_i$ em S e $\mathbf{u}$ é uma combinação linear

$$\mathbf{u} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \ldots + \beta_n\mathbf{u}_n$$

de vetores $\mathbf{u}_i$ em S. Para cada escalar γ,

$$\gamma\mathbf{v} + \mathbf{u} = \sum_{i=1}^{m} (\gamma\alpha_i)\mathbf{v}_i + \sum_{j=1}^{n} \beta_j\mathbf{u}_j$$

Logo, $\gamma\mathbf{v}+\mathbf{u}$ pertence a L. Assim, L é um subespaço de V. ∎

Exemplo 4. *Seja*

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

uma $m\times n$ matriz sobre o corpo $\mathbb{K}$. As linhas $\mathbf{l}_i$ de $\mathbf{A}$, $i=1,\ldots m$,

$$\begin{aligned} \mathbf{l}_1 &= (a_{11},a_{12},\ldots,a_{1n}) \\ \mathbf{l}_2 &= (a_{21},a_{22},\ldots,a_{2n}) \\ &\vdots \\ \mathbf{l}_m &= (a_{m1},a_{m2},\ldots.a_{mn}) \end{aligned}$$

consideradas vetores em $\mathbb{K}^n$, geram um subespaço de $\mathbb{K}^n$ chamado espaço linha de $\mathbf{A}$. Analogamente, as colunas $\mathbf{c}_j$ de $\mathbf{A}$, $j=1,\ldots,n$,

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \cdots, \quad \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

consideradas vetores em $\mathbb{K}^n$, geram um subespaço de $\mathbb{K}^n$ chamado espaço coluna de $\mathbf{A}$, que denotaremos nesta pesquisa por $C(\mathbf{A})$.

Mostramos acima que L é um subespaço de V que contém S e também que todo subespaço que contém S contém L. Decorre que L é a intersecção de todos os subespaços que contêm S, isto é, que L é o subespaço gerado pelo conjunto S.

2.2 Bases

Definição 20. *Sejam V um $\mathbb{K}$-espaço vetorial. Diz-se que um conjunto $X\subseteq V$ é linearmente independente se $\alpha_i\mathbf{v}_i+\cdots+\alpha_n\mathbf{v}_n=0$, para $\mathbf{v}_i\in X$ e $\alpha_i\in\mathbb{K}$, $i=1\cdots n$, implica que $\alpha_i=\cdots=\alpha_n=0$. Um subconjunto de V é dito linearmente dependente se ele não é linearmente independente.*

Observação 5. *Como um resultado dessa definição, podemos também classificar um conjunto $X \subseteq \mathbb{K}$ como sendo linearmente independente quando nenhum vetor $\mathbf{v} \in X$ é combinação linear de outros elementos de X. No caso em que $X = \{\mathbf{v}\}$ (ou seja, X é um conjunto unitário), diz-se se X é linearmente independente se e só se $\mathbf{v}$ é diferente do vetor nulo.*

Definição 21. *Dizemos que $W_1, \ldots, W_k$ são independentes se todo conjunto $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$, com $\mathbf{w}_i \in W_i$, é linearmente independente.*

Definição 22. *Uma base de um $\mathbb{K}$-espaço vetorial V é um conjunto $\mathfrak{B} \subseteq V$, lineamente independente que gera V.*

Exemplo 5. *Seja $\mathbb{K}$ um corpo e, em $\mathbb{K}^n$, seja $\mathfrak{B}$ o subconjunto constituído dos vetores $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ definidos por*

$$\begin{array}{rcl} \mathbf{e}_1 & = & (1,0,0,\ldots,0) \\ \mathbf{e}_2 & = & (0,1,0,\ldots,0) \\ \vdots & & \vdots \\ \mathbf{e}_n & = & (0,0,0,\ldots,1) \end{array}$$

Sejam $\alpha_1, \alpha_2, \ldots, \alpha_n$ escalares em $\mathbb{K}$ e coloquemos $\mathbf{u} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 \ldots + \alpha_n\mathbf{e}_n$. Então

$$\mathbf{u} = (\alpha_1, \alpha_2, \ldots, \alpha_n). \tag{2.1}$$

Isso mostra que $\mathbf{e}_1, \ldots, \mathbf{e}_n$ geram $\mathbb{K}^n$. Como $\mathbf{u} = \mathbf{0}$ se, e somente se $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, os vetores $\mathbf{e}_1, \ldots, \mathbf{e}_n$ são linearmente independentes. O conjunto $\mathfrak{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ é portanto uma base de $\mathbb{K}^n$. Denominamos essa base particular de base canônica de $\mathbb{K}^n$.

Teorema 10. *Se $\mathfrak{B} = \{v_1, v_2, \ldots, v_n\}$ constitui uma base para o $\mathbb{K}$-espaço vetorial V e se $T = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ é um conjunto linearmente independente de vetores em V, então $r \leq n$.*

Demonstração: Seja $T_1 = \{\mathbf{w}_1, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Como $\mathfrak{B}$ gera V, T_1 também gera V. Como $\mathbf{w}_1$ é uma combinação linear de vetores em $\mathfrak{B}$, T_1 é linearmente dependente. Logo algum $\mathbf{v}_j$ é uma combinação linear dos vetores precedentes em T_1. Remova esse vetor particular $\mathbf{v}_j$.

Seja $\mathfrak{B}_1 = \{\mathbf{w}_1, \mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\}$. Note que $\mathfrak{B}_1$ gera V. Considere, agora, $T_2 = \{\mathbf{w}_2, \mathbf{w}_1, \mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\}$. Então T_2 é linearmente dependente e algum vetor em T_2 é uma combinação linear dos vetores precedentes em T_2. Como T é linearmente independente, esse vetor não pode ser $\mathbf{w}_1$, logo tem que ser algum $\mathbf{v}_i$, com $i \neq j$. Repita esse processo quantas vezes forem necessárias. Se todos os vetores V forem eliminados antes de acabarem os vetores $\mathbf{w}$, então o conjunto resultante de vetores $\mathbf{w}$, um subconjunto de T, é linearmente dependente, o que implica que T também é linearmente dependente, uma contradição. Podemos então concluir que o número r de vetores de T não pode ser maior que o número n de vetores de $\mathfrak{B}$, isto é, $r \leq n$. ■

Corolário 1. *Se $\mathfrak{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ e $T = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ são bases para um espaço vetorial, então $n = m$.*

Demonstração: Como T é um conjunto linearmente independente de vetores, o Teorema (10) implica que $m \leq n$. Analogamente, $n \leq m$, pois $\mathfrak{B}$ é linearmente independente. Portanto, $n = m$. ■

Definição 23. *Um $\mathbb{K}$-espaço vetorial que possui um conjunto gerador finito é denominado um espaço finitamente gerado.*

Observação 6. *Observe que se V não for finitamente gerado, então qualquer base de V possui infinitos elementos. Neste caso é possível mostrar que as bases são equivalentes como conjuntos, isto é, podemos mostrar que duas bases de V têm sempre a mesma cardinalidade. No entanto, não faremos aqui esta distinção. Os resultados acima justificam a seguinte definição.*

Definição 24. *Se V é um $\mathbb{K}$-espaço vetorial de dimensão finita, a dimensão de V é definida como sendo o número de elementos de uma base de V. Caso contrário dizemos que a dimensão de V é infinita. Denotamos a dimensão de um espaço V por $dim(V)$.*

Definição 25. *Um espaço vetorial V possui dimensão finita se ele possui uma base finita.*

Proposição 1. *Seja V um espaço vetorial sobre $\mathbb{K}$ e considere $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ um conjunto linearmente independente em V. Se existir $\mathbf{v} \in V$ que não seja combinação linear dos elementos de $\mathfrak{B}$, então $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}\}$ é linearmente independente.*

Demonstração: Sejam $\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}$ escalares tais que

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_m \mathbf{v}_m + \alpha_{m+1} \mathbf{v} = 0$$

Se $\alpha_{m+1} \neq 0$, então podemos escrever

$$\mathbf{v} = -\frac{\alpha_1}{\alpha_{m+1}} \mathbf{v}_1 - \ldots - \frac{\alpha_m}{\alpha_{m+1}} \mathbf{v}_m$$

o que é uma contradição com a nossa hipótese de $\mathbf{v}$ não ser uma combinação linear de elementos de $\mathfrak{B}$. Então $\alpha_{m+1} = 0$ e, portanto, $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$. Como o conjunto $\mathfrak{B}$ é linearmente independente, segue então que $\alpha_1 = \ldots = \alpha_m = 0$, uma contradição com a hipótese sobre os α_i's. Portanto $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}\}$ é linearmente independente. ■

Teorema 11. *Todo espaço vetorial finitamente gerado não nulo possui uma base.*

Demonstração: Seja V um espaço vetorial finitamente gerado não nulo sobre $\mathbb{K}$. Então V possui um conjunto gerador finito, digamos com m elementos, $m \geq 1$. Seja agora $\mathbf{v}_1 \in V$ um vetor não nulo. Então $\mathfrak{B}_1 = \{\mathbf{v}_1\}$ é linearmente independente. Se $\mathfrak{B}_1$ gerar V, então $\mathfrak{B}_1$ é uma base de V. Caso contrário, existe $\mathbf{v}_2 \in V$ que não é um múltiplo de $\mathbf{v}_1$. Pela Proposição (1), $\mathfrak{B}_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ é linearmente independente. De novo, se $\mathfrak{B}_2$ gerar todo o espaço V, então será uma base de V. Caso contrário, existe $\mathbf{v}_3 \in V$ tal que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é linearmente independente. Repetindo este procedimento, chegaremos ou a uma base de V ou construiremos conjuntos linearmente independentes em V arbitrariamente grandes. O segundo caso não é possível, pois como mostramos no teorema (10), todo conjunto linearmente independente neste espaço deve possuir no máximo m elementos. ∎

Teorema 12. *Se W é um subespaço de um espaço vetorial V de dimensão finita, todo subconjunto de W que é linearmente independente é finito e é parte de uma base (finita) de W.*

Demonstração: Suponhamos que $\mathfrak{B}_0$ seja um subconjunto de W linearmente independente. Se $\mathfrak{B}$ é um subconjunto de W linearmente independente contendo $\mathfrak{B}_0$, então $\mathfrak{B}$ também é um subconjunto de W linearmente independente; como V é de dimensão finita, S contém no máximo $dim(V)$ elementos. Portanto, existe um subconjunto $\mathfrak{B}$ de W linearmente independente que é maximal e contém $\mathfrak{B}_0$. Como $\mathfrak{B}$ é um subconjunto de W linearmente independente e maximal contendo $\mathfrak{B}_0$, a Proposição (1) mostra que W é o subespaço gerado por $\mathfrak{B}$. Logo, $\mathfrak{B}$ é uma base de W e o conjunto original $\mathfrak{B}_0$ é parte de uma base de W. ∎

Corolário 2. *Se W é um subespaço próprio de um espaço vetorial V de dimensão finita, então W é de dimensão finita e $dim(W) < dim(V)$.*

Demonstração: Podemos supor que W contém um vetor $\mathbf{v} \neq \mathbf{0}$. Pelo Teorema (12) e sua demonstração, existe uma base de W que contém $\mathbf{v}$ e no máximo $dim(V)$ elementos. Logo W é de dimensão finita e $dim(W) \leq dim(V)$. Como W é subespaço próprio, existe um vetor $\mathbf{u}$ em V que não está em W. Acrescentando $\mathbf{u}$ a uma base arbitrária de W obtemos um subconjunto de V linearmente independente. Portanto $dim(W) < dim(V)$. ∎

Corolário 3. *Num espaço vetorial V de dimensão finita todo conjunto não vazio de vetores linearmente independentes é parte de uma base.*

Corolário 4. *Seja $\mathbf{A}$ uma matriz $n \times n$ sobre um corpo $\mathbb{K}$ e suponhamos que os vetores-linha de $\mathbf{A}$ formem um conjunto de vetores de $\mathbb{K}^n$ linearmente independentes. Então $\mathbf{A}$ é inversível.*

Demonstração: Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ os vetores-linha de $\mathbf{A}$ e suponhamos que W seja o subespaço de $\mathbb{K}^n$ gerado por $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n$. Como $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são linearmente independentes, a dimensão de W é n. O Corolário (2) mostra agora que $W = \mathbb{K}^n$. Logo, existem escalares b_{ij} em $\mathbb{K}$ tais que

$$\mathbf{e}_i = \sum_{j=1}^{n} b_{ij}\mathbf{v}_j, \quad 1 \leq i \leq n$$

em que $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ é a base canônica de $\mathbb{K}^n$. Portanto, para a matriz $\mathbf{B}$ com elementos b_{ij}, temos

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

■

Proposição 2. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão n e seja $\mathfrak{B} \subseteq V$. As seguintes afirmações são equivalentes:*

i) *$\mathfrak{B}$ é uma base de V*

ii) *Cada elemento de V se escreve de maneira única como combinação linear dos elementos de $\mathfrak{B}$.*

Demonstração:

$(i) \Rightarrow (ii)$ Vamos supor que $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ seja uma base de V. Em particular, $\mathfrak{B}$ gera V e, portanto, todo elemento $\mathbf{v} \in V$ se escreve como combinação linear de $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Para mostrar a unicidade, suponha que $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ e $\mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$. Então $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$ ou $\mathbf{v} = \sum_{i=1}^{n} (\alpha_i - \beta_i)\mathbf{v}_i = 0$. Como $\mathfrak{B}$ é linearmente independente, segue então que $\alpha_i - \beta_i = 0$ para todo $i = 1, \ldots, n$. Logo, $\alpha_i = \beta_i$, para todo i, de onde segue a unicidade requerida.

$(ii) \Rightarrow (i)$ Assuma agora que cada elemento de V se escreve de maneira única como combinação linear de elementos de $\mathfrak{B}$. Em particular, $\mathfrak{B}$ gera V. Para mostrarmos que S é uma base, falta verificar que $\mathfrak{B}$ é linearmente independente. Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ e $\gamma_1, \ldots, \gamma_n \in \mathbb{K}$ tais que $\sum_{i=1}^{n} \gamma_i \mathbf{v}_i = \mathbf{0}$ Como $\mathbf{0} = \sum_{i=1}^{n} \mathbf{0}\mathbf{v}_i$, segue da condição de unicidade dada no item (ii) que $\gamma_i = 0$ para cada $i = 1, \ldots, n$. Portanto $\mathfrak{B}$ é uma base.

■

2.3 Coordenadas

Seja $\mathbf{v}$ um vetor do $\mathbb{K}$-espaço vetorial V e $\mathfrak{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ uma base ordenada de V. Segue da Proposição (2) demonstrada na seção anterior, que existem escalares $\alpha_1, ..., \alpha_n \in \mathbb{K}$ tais que $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$. Assim sendo, denotamos por

$$[\mathbf{v}]_{\mathfrak{B}}$$

a matriz do vetor v com relação à base ordenada $\mathfrak{B}$. Essa notação será particularmente útil ao passarmos agora a descrver o que acontece com as coordenadas de um vetor $\mathbf{v}$ quando passamos de uma base ordenada à outra.

Suponhamos então que V seja n-dimensional e que $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $\mathfrak{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ sejam duas bases ordenadas de V. Existem escalares a_{ij}, bem determinados, tais que

$$\mathbf{v}'_j = \sum_{i=1}^{n} a_{ij} v_i, \quad 1 \leq j \leq n. \tag{2.2}$$

Sejam $x'_1, \ldots x'_n$ as coordenadas de um dado vetor $\mathbf{v}$ em relação à base ordenada $\mathfrak{B}'$. Então

$$\begin{aligned}
\mathbf{v} &= x'_1 \mathbf{v}'_1 + \ldots + x'_n \mathbf{v}'_n \\
&= \sum_{j=1}^{n} x'_j \mathbf{v}'_j \\
&= \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} a_{ij} \mathbf{v}_i \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} (a_{ij} x'_j) \mathbf{v}_i \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x'_j \right) \mathbf{v}_i.
\end{aligned}$$

Portanto, obtemos a relação

$$\mathbf{v} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x'_j \right) \mathbf{v}_i \tag{2.3}$$

Como as coordenadas $x_1, \ldots, x_n$ de v em relação à base ordenadas $\mathfrak{B}$ são determinadas de modo único, decorre de (2.3) que

$$x_i = \sum_{j=1}^{n} a_{ij} x'_j, \quad 1 \leq i \leq n. \tag{2.4}$$

Seja $\mathbf{A}$ a $n \times n$ matriz cujo elemento i, j é o escalar a_{ij} e sejam $\mathbf{X}$ e $\mathbf{X}'$ as matrizes das coordenadas do vetor $\mathbf{x}$ em relação às bases $\mathfrak{B}$ e $\mathfrak{B}'$. Podemos então reformular (2.4) como

$$\mathbf{X} = \mathbf{A}\mathbf{X}'. \tag{2.5}$$

Como $\mathfrak{B}$ e $\mathfrak{B}'$ são conjuntos linearmente independentes, $\mathbf{X} = \mathbf{0}$ se, e somente se, $\mathbf{X}' = 0$. Assim, temos que $\mathbf{A}$ é inversível e assim

$$\mathbf{X}' = \mathbf{A}^{-1}\mathbf{X}. \tag{2.6}$$

Se usarmos a notação acima introduzida para a matriz das coordenadas de um vetor em relação a uma base ordenada, então (2.5) e (2.6) afirmam que

$$\begin{aligned} [\mathbf{v}]_{\mathfrak{B}} &= \mathbf{A}[\mathbf{v}]_{\mathfrak{B}'} \\ [\mathbf{v}]_{\mathfrak{B}'} &= \mathbf{A}^{-1}[\mathbf{v}]_{\mathfrak{B}} \end{aligned}$$

Portanto, a discussão acima pode ser resumida como segue.

Teorema 13. *Seja V um $\mathbb{K}$-espaço vetorial n-dimensional e sejam $\mathfrak{B}$ e $\mathfrak{B}'$ duas bases ordenadas de V. Então existe uma única matriz $\mathbf{A}$ em $\mathbb{K}^{n\times n}$, necessariamente inversível, tal que para todo vetor $\mathbf{v} \in V$,*

i) $[\mathbf{v}]_{\mathfrak{B}} = \mathbf{A}[\mathbf{v}]_{\mathfrak{B}'}$

ii) $[\mathbf{v}]_{\mathfrak{B}'} = \mathbf{A}^{-1}[\mathbf{v}]_{\mathfrak{B}}$

Para completar a análise acima, demonstraremos também o resultado que segue.

Teorema 14. *Suponhamos que $\mathbf{A}$ seja uma $n \times n$ matriz inversível sobre $\mathbb{K}$. Seja V um espaço vetorial n-dimensional sobre $\mathbb{K}$ e seja $\mathfrak{B}$ uma base ordenada de V. Então, existe uma única base ordenada $\mathfrak{B}'$ de V tal que*

i) $[\mathbf{v}]_{\mathfrak{B}} = \mathbf{A}[\mathbf{v}]_{\mathfrak{B}'}$

ii) $[\mathbf{v}]_{\mathfrak{B}'} = \mathbf{A}^{-1}[\mathbf{v}]_{\mathfrak{B}}$

para todo vetor $\mathbf{v} \in V$.

Demonstração: Seja $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Se $\mathfrak{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ é uma base ordenada de V para a qual o item (i) é válido, é claro que

$$\mathbf{v}'_j = \sum_{i=1}^{n} \mathbf{A}_{ij}\mathbf{v}_i.$$

Assim, basta mostrar que os vetores $\mathbf{v}'_j$, definidos por estas equações, formam uma base. Seja $\mathbf{Q} = \mathbf{A}^{-1}$. Então

$$\begin{aligned}
\sum_j \mathbf{Q}_{jk}\mathbf{v}'_j &= \sum_j \mathbf{Q}_{jk} \sum_i \mathbf{A}_{ij}\mathbf{v}_i \\
&= \sum_j \left(\sum_i \mathbf{A}_{ij}\mathbf{Q}_{jk} \right) \mathbf{v}_i \\
&= \sum_i \left(\sum_j \mathbf{A}_{ij}\mathbf{Q}_{jk} \right) \mathbf{v}_i \\
&= \mathbf{v}_k.
\end{aligned}$$

Portanto, o subespaço gerado pelo conjunto

$$\mathfrak{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$$

contém $\mathfrak{B}$, logo, é igual a V. Assim, $\mathfrak{B}'$ é uma base e, de sua definição e do Teorema (13), é evidente que (i) é válido, logo (ii) também o é. ■

2.4 Somas Diretas

Seja V um espaço vetorial sobre $\mathbb{K}$. Às vezes, é conveniente escrever seus elementos como soma de elementos de dois (ou mais) subespaços.

Definição 26. *Sejam W_1 e W_2 subespaços vetoriais de um espaço vetorial V. Diremos que a soma $W_1 + W_2$ é direta se $W_1 \bigcap W_2 = \{\mathbf{0}\}$ e, neste caso, escrevemos $W_1 \oplus W_2$.*

Exemplo 6. *Sejam W_1 e W_2 dois subespaços de $\mathbb{C}^4$ com bases iguais a $\{(1,2,0,i),(i,0,0,1)\}$ e $\{(0,0,3,1)\}$, respectivamente. A soma de W_1 e W_2 é direta, pois $W_1 \bigcap W_2 = \{0\}$. De fato, se $(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3,\mathbf{z}_4) \in W_1 \bigcap W_2$, então*

$$(z_1,z_2,z_3,z_4) = a(1,2,0,i) + b(i,0,0,1) = c(0,0,3,1)$$

com $a,b,c \in \mathbb{C}$. Não é difícil ver então que $a = b = c = 0$ e, portanto, $(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3,\mathbf{z}_4) = (0,0,0,0)$, como queríamos.

Definição 27. *Seja V um espaço vetorial sobre um corpo $\mathbb{K}$ e sejam W_1 e W_2 dois subespaços de V. Dizemos que V é a soma direta de W_1 e W_2 se $V = W_1 \oplus W_2$.*

Os próximos resultados serão muito importantes em nossas considerações futuras.

Proposição 3. *Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$ e W_1, W_2 dois subespaços de V. Então, $V = W_1 \oplus W_2$ se, e só se cada elemento $\mathbf{v} \in V$ se escreve de maneira única como uma soma $\mathbf{w}_1 + \mathbf{w}_2$ com $\mathbf{w}_i \in W_i$, $i = 1, 2$.*

Demonstração:

($\Rightarrow$) Vamos supor que $V = W_1 \oplus W_2$. Segue então que cada elemento $\mathbf{v} \in V$ se escreve como soma de um elemento de W_1 e um elemento de W_2. Suponha agora que $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w'}_1 + \mathbf{w'}_2$, com $\mathbf{w}_1, \mathbf{w'}_1 \in W_1$ e $\mathbf{w}_2, \mathbf{w'}_2 \in W_2$. Daí segue que $\mathbf{w}_1 - \mathbf{w'}_1 = -\mathbf{w}_2 + \mathbf{w'}_2 \in W_1 \bigcap W_2$, pois $\mathbf{w}_1 - \mathbf{w'}_1 \in W_1$ e $-\mathbf{w}_2 + \mathbf{w'}_2 \in W_2$. Como $W_1 \bigcap W_2 = \{\mathbf{0}\}$, teremos $\mathbf{w}_1 = \mathbf{w'}_1$ e $\mathbf{w}_2 = \mathbf{w'}_2$, como queríamos.

($\Leftarrow$) Como cada elemento se escreve como soma de elementos de W_1 e W_2, então $V = W_1 + W_2$. Suponha agora que $W_1 \bigcap W_2$ tenha um elemento não nulo $\mathbf{w}$. Observe então que $\mathbf{w}$ pode ser escrito como $\mathbf{w} = 0 + \mathbf{w}$ se considerarmos $\mathbf{0} \in W_1$ e $\mathbf{w} \in W_2$ e também como $\mathbf{w} = \mathbf{w} + \mathbf{0}$ se considerarmos $\mathbf{w} \in W_1$ e $\mathbf{0} \in W_2$, o que contradiz a nossa hipótese da unicidade. Logo, $W_1 \bigcap W_2 = \{\mathbf{0}\}$ e o resultado está provado.

■

Proposição 4. *Sejam V um espaço vetorial finitamente gerado não nulo e W_1 um subespaço de V. Então existe um subespaço W_2 de V tal que $V = W_1 \oplus W_2$.*

Demonstração: Se $W_1 = V$, não há nada a fazer, pois bastaria escolher $W_2 = 0$. Suponha $W_1 \neq V$. Seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ uma base de W_1 e estenda-a a uma base $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ de V. O subespaço vetorial W_2 gerado pelos vetores $\{\mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ satisfaz as propriedades desejadas. De fato, é claro que $V = W_1 + W_2$ pois o conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ é um conjunto gerador de V. Por outro lado, como $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ é linearmente independente, segue que $W_1 \bigcap W_2 = \{\mathbf{0}\}$, como queríamos. ■

O subespaço W_2 como no teorema acima é chamado de complemento de W_1 em V. O complemento de um subespaço vetorial nem sempre é único.

Discutimos acima a soma de dois subespaços. isso pode ser generalizado para a soma direta de vários subespaços da seguinte maneira. Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Para subespaços $W_1, W_2, \ldots, W_t$, definimos

$$W_1 + \ldots + W_t = \{\mathbf{v}_1 + \ldots + \mathbf{v}_t ; \mathbf{v}_i \in W_i, i = 1, \ldots, t\}.$$

Se $W_i \bigcap (W_1 + \ldots + W_{i-1} + W_{i+1} + \ldots + W_t) = \{\mathbf{0}\}$, para cada $i = 1, \ldots t$, então a soma $W_1 + \ldots W_t$ é chamada de soma direta de $W_1, \ldots, W_t$ e será indicada por $W_1 \oplus \ldots \oplus W_t$. Também diremos que o espaço V é a soma direta dos subespaços $W_1, \ldots, W_t$ se $V = W_1 \oplus \ldots \oplus W_t$.

3 Transformações Lineares

Neste capítulo vamos estudar funções entre espaços vetoriais que preservam as operações destes espaços, chamadas transformações lineares, importantes na continuação de nosso estudo.

3.1 Conceitos Básicos

Definição 28. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). Uma transformação linear de V em W é uma função T de V em W tal que*

$$T(\alpha\mathbf{u}+\mathbf{v}) = \alpha\cdot T(\mathbf{u})+T(\mathbf{v})$$

para todos $\mathbf{u},\mathbf{v}\in V$ e para todos os escalares $\alpha\in\mathbb{K}$.

Exemplo 7. *Se V é um espaço vetorial arbitrário, a transformação identica I, definida por $I(\mathbf{u}) = \mathbf{u}$, é uma transformação linear de V em V. A transformação nula 0, definida por $0(\mathbf{u}) = \mathbf{0}$, é uma transformação linear de V em V.*

Exemplo 8. *Seja V o espaço das matrizes de ordem $m\times n$ sobre um corpo $\mathbb{K}$, seja $\mathbf{P}$ uma matriz $m\times n$ fixa sobre $\mathbb{K}$ e seja $\mathbf{Q}$ uma matriz $n\times n$ fixa sobre $\mathbb{K}$. Então a função $T: V\to V$ definida por $T(\mathbf{A}) = \mathbf{PAQ}$ é uma transformação linear, pois para quaisquer $\mathbf{A},\mathbf{B}\in V$ e $\alpha\in\mathbb{K}$,*

$$\begin{aligned} T(\alpha\mathbf{A}+\mathbf{B}) &= \mathbf{P}(\alpha\mathbf{A}+\mathbf{B})\mathbf{Q} \\ &= (\alpha\mathbf{PA}+\mathbf{PB})\mathbf{Q} \\ &= \alpha\mathbf{PAQ}+\mathbf{PBQ} \\ &= \alpha T(\mathbf{A})+T(\mathbf{B}) \end{aligned} \tag{3.1}$$

É importante notar que se T é uma transformação linear de V em W, então $T(\mathbf{0}) = \mathbf{0}$, pois

$$T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0})+T(\mathbf{0}).$$

Outra propriedade importante a respeito de transformações lineares é que estas conservam combinações lineares, isto é, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são vetores em V e $\alpha_1, \alpha_2, \ldots, \alpha_n$ são escalares em $\mathbb{K}$, então decorre da definição que

$$T(\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n) = \alpha_1 T(\mathbf{v}_1) + \ldots + \alpha_n T(\mathbf{v}_n).$$

Teorema 15. *Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ordenada de V. Seja W um espaço vetorial sobre o mesmo corpo $\mathbb{K}$ e sejam $\mathbf{u}_1, \ldots, \mathbf{u}_n$ vetores arbitrários em W. Então existe exatamente uma transformação linear $T : V \to V$ tal que*

$$T(\mathbf{v}_i) = \mathbf{u}_i, \;\; i = 1, \ldots, n$$

Demonstração: Para demonstrar que existe pelo menos uma transformação linear T com $T(\mathbf{v}_i) = \mathbf{u}_i$, procedemos como segue. Dado $\mathbf{v}$, existe uma única n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ tal que

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n.$$

Para este vetor $\mathbf{v}$, definamos

$$T(\mathbf{v}) = \alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n$$

Então T é uma regra bem definida para se associar a cada vetor $\mathbf{v}$ em V um vetor $T(\mathbf{v})$ em W. Pela definição, é evidente que $T(\mathbf{v}_j) = \mathbf{u}_j$ para todo j. Para ver que T é linear, seja

$$\mathbf{u} = \beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n$$

em V e γ escalar arbitrário. Ora,

$$\gamma\mathbf{v} + \mathbf{u} = (\gamma\alpha_1 + \beta_1)\mathbf{v}_1 + \ldots + (\gamma\alpha_n + \beta_n)\mathbf{v}_n$$

portanto, pela definição

$$T(\gamma\mathbf{v} + \mathbf{u}) = (\gamma\alpha_1 + \beta_1)\mathbf{u}_1 + \ldots + (\gamma\alpha_n + \beta_n)\mathbf{u}_n.$$

Por outro lado,

$$\begin{aligned} \gamma(T(\mathbf{v})) + T(\mathbf{u}) &= \gamma\sum_{i=1}^{n}\alpha_i\mathbf{u}_i + \sum_{i=1}^{n}\beta_i\mathbf{u}_i \\ &= \sum_{i=1}^{n}(\gamma\alpha_i + \beta_i)\mathbf{u}_i \end{aligned}$$

e assim

$$T(\gamma\mathbf{v} + \mathbf{u}) = \gamma(T(\mathbf{v})) + T(\mathbf{u}).$$

Se U é uma transformação linear de V em W com $U(\mathbf{v}_j) = \mathbf{u}_j$, $j = 1, \ldots, n$, então, para o vetor $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ temos

$$\begin{aligned} U(\mathbf{v}) &= U\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) \\ &= \sum_{i=1}^{n} \alpha_i (U(\mathbf{v}_i)) \\ &= \sum_{i=1}^{n} \alpha_i \mathbf{u}_i \end{aligned}$$

de modo que U é exatamente a regra T que definimos acima. Isto mostra que a transformação linear T com $T(\mathbf{v}_i) = \mathbf{u}_i$ é única. ■

Teorema 16. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ e seja T uma transformação linear $T : V \to W$. Então, a imagem de T é um subespaço de W. O conjunto de vetores $\mathbf{v} \in V$ tais que $T(\mathbf{v}) = 0$ também é u subespaço de V, dito o núcleo de T e denotado por $N(T)$.*

Demonstração: Indiquemos por Im_T a imagem de T, isto é, o conjunto de vetores $\mathbf{u}$ em W tais que $\mathbf{u} = T(\mathbf{v})$ para algum $\mathbf{v}$ em V. Sejam $\mathbf{u}_1, \mathbf{u}_2 \in R\Im_T$ e seja $\alpha \in \mathbb{K}$. Existem vetores $\mathbf{v}_1, \mathbf{v}_2 \in V$ tais que $T(\mathbf{v}_i) = \mathbf{u}_i$, $i = 1, 2$. Como T é linear

$$T(\alpha \mathbf{v}_1 + \mathbf{v}_2) = \alpha \mathbf{u}_1 + \mathbf{u}_2$$

o que mostra que $\alpha \mathbf{u}_1 + \mathbf{u}_2$ também está na imagem de T. Portanto, Im_T é um subespaço de W.

Considere o núcleo de T, $N(T)$. Se $\mathbf{v}_1$ e $\mathbf{v}_2$ estão em $N(T)$ e α é um escalar arbitrário, então

$$\begin{aligned} T(\alpha \mathbf{v}_1 + \mathbf{v}_2) &= \alpha(T(\mathbf{v}_1)) + T(\mathbf{v}_2) \\ &= +\alpha(\mathbf{0}) + \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

de modo que $\alpha \mathbf{v}_1 + \mathbf{v}_2$ está novamente em $N(T)$. logo, $N(T)$ é um subespaço. ■

No exemplo 7, a imagem da transformação idêntica é todo o espaço V, e seu núcleo é o subespaço nulo. A imagem da transformação nula é o subespaço nulo e seu núcleo é todo o espaço V. No exemplo 8, a imagem e o núcleo de T são um tanto difíceis de escrever, exceto pela repetição de suas definições.

Se T é uma transformação linear de V em W e $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ são vetores que geram V, então é evidente que os vetores $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ geram a imagem de T. Em particular, se V posuir dimensão finita, a imagem de T será um subespaço de W de dimensão finita.

Definição 29. *Seja $T : V \to W$ uma transformação linear, sendo V de dimensão finita. Então, temos que*

i) *O posto de T é a dimensão da imagem de T;*

ii) *A nulidade de T é a dimensão do núcleo de T.*

Teorema 17. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ e seja $T : V \to W$ uma transformação linear. Suponha que V possua dimensão finita. Então*

$$posto(T) + nulidade(T) = dim(V)$$

Demonstração: Seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ uma base de $N(T)$. Existem vetores $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n \in V$ tais que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ seja uma base de V. Demonstraremos agora que $\{T(\mathbf{v}k+1), \ldots, T(\mathbf{v}_n)\}$ é uma base da imagem de T. Os vetores $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ certamente geram a imagem de T e, como $T(\mathbf{v}_j) = \mathbf{0}$ para $j \leq k$, vemos que $T(\mathbf{v}_{k+1}, \ldots, T(\mathbf{v}_n)$ geram a imagem. Para ver que esses vetores são independentes, suponhamos que existam escalares α_i tais que

$$\sum_{i=k+1}^{n} \alpha_i(T(\mathbf{v}_i)) = \mathbf{0}.$$

Isto diz que

$$T\left(\sum_{i=k+1}^{n} \alpha_i \mathbf{v}_i\right) = \mathbf{0}$$

e, consequentemente, o vetor $\mathbf{v} = \sum_{i=k+1}^{n} \alpha_i \mathbf{v}_i$ está no núcleo de T. Como $\mathbf{v}_1, \ldots, \mathbf{v}_k$ formam uma base de $N(T)$, existem necessariamente escalares $\beta_1, \ldots, \beta_k$ tais que

$$\mathbf{v} = \sum_{i=1}^{k} \beta_i \mathbf{v}_i.$$

Assim

$$\sum_{i=1}^{k} \beta_i \mathbf{v}_i - \sum_{j=k+1}^{n} \alpha_j \mathbf{v}_j = 0$$

e, como $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são linearmente independentes, devemos ter

$$\beta_1 = \ldots = \beta_k = \alpha_{k+1} = \ldots = \alpha_n = 0.$$

Se r é o posto de T, o fato de $T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)$ formarem uma base da imagem de T nos diz que $r = n - k$. Como k é a nulidade de T e n é a dimensão de V, está completa a demonstração. ■

Teorema 18. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$. Sejam T e U transformações* *lineares de V em W. A função $(T+U)$ definida por*

$$(T+U)(\mathbf{v}) = T(\mathbf{v}) + U(\mathbf{v})$$

é uma transformação linear de V em W. Se $\alpha \in \mathbb{K}$ é um escalar arbitrário, a função (αT) definida por

$$(\alpha T)(\mathbf{v}) = \alpha T(\mathbf{v})$$

é uma transformação linear de V em W. O conjunto das transformações lineares de V em W, munido da adição e multiplicação por escalar acima definida, é um espaço vetorial sobre o corpo $\mathbb{K}$.

Demonstração: Suponhamos que T e U sejam transformações lineares de V em W e definamos $(T+U)$ como acima. Então

$$\begin{aligned}
(T+U)(\alpha\mathbf{v}+\mathbf{u}) &= T(\alpha\mathbf{v}+\mathbf{u}) + U(\alpha\mathbf{v}+\mathbf{u}) \\
&= \alpha(T(\mathbf{v})) + T(\mathbf{u}) + \alpha(U(\mathbf{v})) + U(\mathbf{u}) \\
&= \alpha(T(\mathbf{v}) + U(\mathbf{v})) + (T(\mathbf{u}) + U(\mathbf{u})) \\
&= \alpha(T+U)(\mathbf{v}) + (T+U)(\mathbf{u})
\end{aligned}$$

o que mostra que $(T+U)$ é uma transformação linear. Analogamente,

$$\begin{aligned}
(\alpha T(\mathbf{v}))(\beta\mathbf{v}+\mathbf{u}) &= \alpha[T(\beta\mathbf{v}+\mathbf{u})] \\
&= \alpha[\beta T(\mathbf{v}) + T(\mathbf{u})] \\
&= \alpha\beta T(\mathbf{v}) + \alpha T(\mathbf{u}) \\
&= \beta[\alpha T(\mathbf{v})] + \alpha T(\mathbf{u})
\end{aligned}$$

mostrando que (αT) é uma transformação linear. ■

Indicaremos o espaço das transformações lineares de V em W por $L(V,W)$. É claro que $L(V,W)$ está definido somente para V e W espaços vetoriais sobre o mesmo corpo.

Teorema 19. *Seja V um espaço vetorial n-dimensional sobre o corpo $\mathbb{K}$ e seja W um espaço vetorial m-dimensional sobre $\mathbb{K}$. Então o espaço $L(V,W)$ é de dimensão finita igual a mn.*

Demonstração: Sejam $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $\mathfrak{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ bases ordenadas de V e W, respectivamente. Para cada par de inteiros (p,q) com $1 \leq p \leq m$ e $1 \leq q \leq n$, definamos uma transformação linear $E^{p,q}$ de V em W por

$$\begin{aligned} E^{p,q}(\mathbf{v}_i) &= \begin{cases} 0, & se\ i \neq q \\ \mathbf{u}_p, & se\ i = q \end{cases} \\ &= \delta_{iq}\mathbf{u}_p. \end{aligned}$$

De acordo com o Teorema (15), existe uma única transformação linear de V em W que satisfaz estas condições. Afirmamos que as mn transformações $E^{p,q}$ formam uma base de $L(V,W)$. Seja T uma transformação linear de V em W. Para cada j, $1 \leq j \leq n$, sejam $a_{ij}, \ldots, a_{mj}$ as coordenadas do vetor $T(\mathbf{v}_j)$ em relação à base ordenada $\mathfrak{B}'$, isto é,

$$T(\mathbf{v}_j) = \sum_{p=1}^{m} a_{pj}\mathbf{u}_p. \tag{3.2}$$

Desejamos mostrar que

$$T = \sum_{p=1}^{m}\sum_{q=1}^{m} a_{pq}E^{p,q} \tag{3.3}$$

Seja U a transformação linear no segundo membro de (3.3). Então para cada j

$$\begin{aligned} U(\mathbf{v}_j) &= \sum_p\sum_q a_{pq}E^{p,q}(\mathbf{v}_j) \\ &= \sum_p\sum_q a_{pq}\delta_{jq}\mathbf{u}_p \\ &= \sum_{p=1}^{m} a_{pj}\mathbf{u}_p \\ &= T(\mathbf{v}_j) \end{aligned}$$

e, consequentemente, $U = T$. Agora (3.3) mostra que as $E^{p,q}$ geram $L(V,W)$; precisamos demonstrar que elas são independentes. mas isto é evidente pelo que fizemos acima, pois, se a transformação

$$U = \sum_p\sum_q a_{pq}E^{p,q}$$

é a transformação nula, então $U(v_j) = 0$ para cada j, portanto

$$\sum_{p=1}^{m} a_{pj}\mathbf{u}_p = 0$$

e a independência dos $\mathbf{u}_p$ implica que $a_{pj} = 0$ para todos p e j. ■

Teorema 20. *Sejam V, W e Z espaços vetoriais sobre o corpo $\mathbb{K}$. Sejam $T: V \to W$ e $U: W \to Z$ transformações lineares. Então a função composta UT, definida por $(UT)(\mathbf{v}) = U(T(\mathbf{v}))$ é uma transformação linear de V em Z.*

Demonstração:

$$
\begin{aligned}
(UT)(\alpha\mathbf{v}+\mathbf{u}) &= U[T(\alpha\mathbf{v}+\mathbf{u}] \\
&= U(\alpha T(\mathbf{v})+T(\mathbf{u})) \\
&= \alpha[U(T(\mathbf{v}))]+U(T(\mathbf{u})) \\
&= \alpha(UT)(\mathbf{v})+(UT)(\mathbf{u})
\end{aligned}
$$

■

No que segue, estaremos interessados nas transformações lineares de um espaço vetorial nele mesmo.

Definição 30. *Se V é um espaço vetorial sobre o corpo $\mathbb{K}$, um operador linear sobre V é uma transformação linear de V em V.*

Aplicando o Teorema (20) no caso em que $V = W = Z$, de modo que U e T sejam operadores lineares sobre V, vemos que a composição UT é ainda um operador linear sobre V. Assim, o espaço $L(V,V)$ possui uma "multiplicação", definida sobre si por meio de composição. Neste caso, o operador UT também está definido e devemos notar que em geral $UT \neq TU$, isto é, $UT - TU \neq 0$. Particularmente, se T é um operador linear sobre V, podemos compor T com T. Usaremos a notação $T^2 = TT$ e em geral, $T^n = \underbrace{T\ldots T}_{n \text{ vezes}}$ para $n \in \mathbb{N}$. Definimos $T^0 = I$ se $T \neq 0$.

Com base na observação feita acima, vamos então defnir uma importante classe de operadoradores lineares.

Definição 31. *Um operador linear $T \in L(V,V)$ é chamado de nilpotente se existir um inteiro $m > 0$ tal que $T^m = 0$. O índice de nilpotência de um tal operador será o menor índice com esta propriedade.*

Mesmo não sendo comutativa, a "multiplicação"que temos sobre $L(V,V)$ está bastante relacionada com as operações de espaço vetorial de $L(V,V)$.

Lema 2. *Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$, U, T_1 e T_2 operadores lineares sobre V e α um elemento de $\mathbb{K}$. Então:*

i) $IU = UI = U$;

ii) $U(T_1+T_2) = UT_1+UT_2; (T_1+T_2)U = T_1U+T_2U$;

iii) $\alpha(UT_1) = (\alpha U)T_1 = U(\alpha T_1)$.

Demonstração:

i) Esta propriedade da função idêntica é óbvia.

ii) Podemos escrever

$$\begin{aligned} U(T_1+T_2)(\mathbf{v}) &= U[(T_1+T_2)(\mathbf{v})] \\ &= U[T_1(\mathbf{v})+T_2(\mathbf{v})] \\ &= U[T_1(\mathbf{v})]+U[T_2(\mathbf{v})] \\ &= (UT_1)(\mathbf{v})+(UT_2)(\mathbf{v}). \end{aligned}$$

Portanto $U(T_1+T_2) = UT_1+UT_2$. Além disso

$$\begin{aligned} [(T_1+T_2)U](\mathbf{v}) &= (T_1+T_2)(U(\mathbf{v})) \\ &= T_1(U(\mathbf{v}))+T_2(U(\mathbf{v})) \\ &= (T_1U)(\mathbf{v})+(T_2U)(\mathbf{v}) \end{aligned} \tag{3.4}$$

Assim sendo, temos $(T1+T_2)U = T_1U+T_2U$.

iii) Pode ser provado de forma similar ao item anterior.

∎

Teorema 21. *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{K}$ e seja T uma transformação linear de V em W. Se T é injetora e sobrejetora, então a função inversa T^{-1} é uma transformação linear de W em V.*

Demonstração: Lembrando que T ser injetora significa que $T(\mathbf{v}) \neq T(\mathbf{u})$ sempre que $\mathbf{v} \neq \mathbf{u}$ e que T ser sobrejetora significa que a imagem de T é todo o espaço W. Quando T é injetora e sobrejetora, existe uma função inversa T^{-1}, determinada de modo único, que leva W sobre V tal que $T^{-1}T$ é a função idêntica de V e TT^{-1} é a função idêntica de W. O que estamos demonstrando aqui é que, se uma função linear T é inversível, então a inversa T^{-1} também é linear.

Sejam $\mathbf{u}_1$ e $\mathbf{u}_2$ vetores em W e seja α um escalar. Queremos mostrar que

$$T^{-1}(\alpha\mathbf{u}_1+\mathbf{u}_2) = \alpha T^{-1}(\mathbf{u}_1)+T^{-1}(\mathbf{u}_2).$$

Seja $v_i = T^{-1}(\mathbf{u})_i$, $i = 1,2$, isto é, seja $\mathbf{v}_i$ o único vetor em V tal que $T(\mathbf{v}_i) = \mathbf{u}_i$. Como T é linear

$$\begin{aligned} T(\alpha\mathbf{v}_1+\mathbf{v}_2) &= \alpha T(\mathbf{v}_1)+T(\mathbf{v}_2) \\ &= \alpha\mathbf{u}_1+\mathbf{u}_2. \end{aligned}$$

Assim, $\alpha\mathbf{v}_1+\mathbf{v}_2$ é o único vetor em V que é levado por T em $\alpha\mathbf{u}_1+\mathbf{u}_2$, portanto

$$\begin{aligned} T^{-1}(\alpha\mathbf{u}_1+\mathbf{u}_2) &= \alpha\mathbf{v}_1+\mathbf{v}_2 \\ &= \alpha(T^{-1}(\mathbf{u}_1))+T^{-1}(\mathbf{u}_2) \end{aligned}$$

e T^{-1} é linear. ■

Se $T: V \to W$ é uma transformação linear, diremos que T é não singular se o núcleo de T consistir apenas do vetor nulo, o que equivale a dizer que T é injetora, pois quando T é linear, dados dois vetores $\mathbf{u},\mathbf{v} \in v$, $T(\mathbf{u}) = T(\mathbf{v})$ se e somente se $T(\mathbf{u}-\mathbf{v}) = \mathbf{0}$. Nesta linguagem, T é inversível se e somente se T é não singular e a imagem de T é todo o espaço W.

Teorema 22. *Considere a transformação linear $T: V \to W$. Então T é não singular se e somente se T leva todo subconjunto linearmente independente de V sobre um conjunto linearmente independente de W.*

Demonstração: Suponhamos primeiro que T seja não singular. Seja S um subconjunto linearmente independente de V. Se $\mathbf{v}_1,\ldots,\mathbf{v}_k$ são vetores em S, então os vetores $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_k)$ são linearmente independentes, pois se

$$\alpha_1 T(\mathbf{v}_1)+\ldots+\alpha_k T(\mathbf{v}_k) = \mathbf{0}$$

então

$$T(\alpha_1\mathbf{v}_1+\ldots+\alpha_k\mathbf{v}_k) = \mathbf{0}$$

e como T é não singular,

$$\alpha_1\mathbf{v}_1+\ldots+\alpha_k\mathbf{v}_k = \mathbf{0}$$

de que segue que cada $\alpha_i = 0$ pois S é um conjunto independente. Este argumento mostra que a imagem de S por meio de T é independente.

Suponhamos que T leve subconjuntos independentes sobre subconjuntos independentes. Seja $\mathbf{v}$ um vetor não nulo em V. Então o conjunto S constituído apenas pelo vetor $\mathbf{v}$ é independente. A imagem de S é o conjunto constituído apenas pelo vetor $T(\mathbf{v})$ e este conjunto é independente. Portanto $T(\mathbf{v}) \neq \mathbf{0}$, pois o conjunto constituído apenas pelo vetor nulo é dependente. Isto mostra que o núcleo de T é o subespaço nulo, isto é, T é não singular. ■

Teorema 23. *Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$, tais que $dim(V) = dim(W)$. Se T é uma transformação linear de V em W, as seguintes afirmações são equivalentes:*

i) *T é inversível;*

ii) *T é não singular;*

iii) *A imagem de T é W; item[iv)] Se $\{\mathbf{v}_1, ..., \mathbf{v}\alpha_n\}$ é uma base arbitrária de V, então $\{T(\mathbf{v}_1)$, ..., $T(\mathbf{v}_n)\}$ é uma base de W;*

v) *Existe pelo menos uma base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V tal que $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ deja uma base de W.*

Demonstração:

$(i) \to (ii)$ Se T é inversível, T é não singular.

$(ii) \to (iii)$ Suponhamos que T seja não singular. Seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base de V. Pelo Teorema (22), $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ é um conjunto linearmente independente de vetores em W e como a dimensão de W também é n, este conjunto de vetores é uma base de W. Agora seja $\mathbf{u}$ um vetor arbitrário em W. Existem escalares $\alpha_1, \ldots, \alpha_n$ tais que

$$\begin{aligned} \mathbf{u} &= \alpha_1 T(\mathbf{v}_1) + \ldots + \alpha_n T(\mathbf{v}_n) \\ &= T(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n) \end{aligned}$$

o que mostra que u está na imagem de T.

$(iii) \to (iv)$ Suponhamos agora que T seja sobrejetora. Se $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base arbitrária de V, os vetores $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ geram a imagem de T, que é todo o espaço W, por hipótese. Como a dimensão de W é n, estes n vetores precisam ser linearmente independentes, isto é, precisam formar uma base de W.

$(iv) \to (v)$ Não requer nenhum comentário.

$(v) \to (i)$ Suponhamos que exista alguma base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V tal que $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ seja uma base de W. Como os $T(\mathbf{v}_i)$ geram W, é evidente que a imagem de T coincide com W. Se $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$ está no núcleo de T, então

$$T(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n) = \mathbf{0}$$

ou

$$\alpha_1 T(\mathbf{v}_1) + \ldots + \alpha_n T(\mathbf{v}_n) = \mathbf{0}$$

e como os $T(\mathbf{v}_i)$ são independentes, cada $\alpha_i = 0$ e assim $\mathbf{v} = \mathbf{0}$. Mostramos que a imagem de T é W e que T é não singular, logo T é inversível.

■

Se V é espaço vetorial de dimensão finita, o Teorema (23) nos diz o que segue a respeito de operadores lineares sobre V. Se T é um operador linear sobre V que é injetor (não singular), então a imagem de T é, necessariamente, todo V, logo tem que ser inversível. Se a imagem de T é todo V, então T tem que ser não singular, logo inversível. Estas duas afirmações possuem uma demonstração mais simples que a que fizemos e que usa o Teorema (17). Seja r o posto de T e k a nulidade de T. Pelo Teorema (17), $r + k = n$. A afirmação que T é não singular significa que $k = 0$, ao passo que a afirmação de que a imagem de T é V significa que $r = n$. Como $r + k = n$ estas afirmações sobre T são obviamente equivalentes.

Lema 3. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ e sejam U e T operadores lineares inversíveis sobre V.Então UT é inversível e $(UT)^{-1} = T^{-1}U^{-1}$*

Demonstração: Verificar simplesmente que

$$(UT)(T^{-1}U^{-1}) = (T^{-1}U^{-1})(UT) = I$$

■

3.2 Isomorfismos

Nesta seção iremos estudar as transformações lineares bijetoras (injetoras e sobrejetoras simultaneamente) e nessas condições identificar os espaços vetoriais envolvidos.

Definição 32. *Se V e W são espaços vetoriais sobre o corpo $\mathbb{K}$, uma transformação linear bijetora $T : V \to W$ é denominada um isomorfismo de V em W. Se existir tal isomorfismo, diremos que V é isomorfo a W.*

É fácil perceber que V é trivialmente isomorfo a V, pois o operador idêntico é um isomorfismo de V em V. Além disso, se V é isomorfo a W por meio de um isomorfismo T, então W é isomorfo a V, uma vez que T^{-1} é um isomorfismo de W em V - por esse motivo, ao invés dizermos que V é isomorfo a W ou vice-versa, costumamos dizer que V e W são isomorfos, simplismente. Também é simples verificar que se V é isomorfo a W e W é isomorfo a Z, então V

é isomorfo a Z. Sendo assim, podemos dizer que o isomorfismo é uma relação de equivalência sobre a classe dos espaços vetoriais.

Teorema 24. *Se V é um espaço vetorial sobre o corpo $\mathbb{K}$ tal que $dim(V) = n$, então V é isomorfo a $\mathbb{K}^n$*

Demonstração: Seja V um espaço n-dimensional sobre o corpo $\mathbb{K}$ e seja $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ordenada de V. Definamos uma função T de V em $\mathbb{K}^n$, como segue: Se $\mathbf{v}$ está em V, seja $T(\mathbf{v})$ a n-upla $(\alpha_1, \ldots, \alpha_n)$ das coordenadas de $\mathbf{v}$ em relação à base ordenada $\mathfrak{B}$, isto é, a n-upla tal que

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n.$$

Em nossa discussão de coordenadas no Capítulo 2, verificamos que esta T é linear, injetora e leva V sobre $\mathbb{K}^n$. ■

Para muitos objetivos frequentemente, consideram-se espaços vetoriais isomorfos como sendo o mesmo, apesar de que os vetores e as operações nestes espaços possam ser bem diferentes.

Suponhamos que T seja um isomorfismo de V em W. SE S é um conjunto de V, o Teorema (22) nos diz que S é linearmente independente se , e somente se, o cojunto $T(S)$ em W é independente. Portanto, ao decidirmos se S é independente, não importa se considerarmos S ou $T(S)$. A partir disso vê-se que um isomorfismo "conserva a dimensão", isto é, todo subespaço de V de dimensão finita tem a mesma dimensão que sua imagem por meio de T. Para ilustrar essa ideia, suponha que $\mathbf{A}$ seja uma matriz de ordem $m \times n$ sobre o corpo $\mathbb{K}$. O espaço solução da matriz $\mathbf{A}$ pode ser dado por via de duas definições. A primeira definição do espaço solução é o conjunto das n-uplas $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{K}^n$ que satisfazem cada uma das equações do sistema $\mathbf{AX} = \mathbf{0}$. A segunda definição é dada pelo conjunto das $n \times 1$ matrizes coluna $\mathbf{X}$ tais que $\mathbf{AX} = \mathbf{0}$. O primeiro espaço solução é portanto um subespaço de $\mathbb{K}^n$ e o segundo é um subespaço do espaço de todas as $n \times 1$ matrizes sobre $\mathbb{K}$. Agora existe um isomorfismo evidente entre $\mathbb{K}^n$ e as $n \times 1$ matrizes colunas sobre $\mathbb{K}$, a saber,

$$(\mathbf{x}_1, \ldots, \mathbf{x}_n) \to \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$$

Por meio deste isomorfismo, o primeiro espaço solução de $\mathbf{A}$ é levado sobre o segundo espaço solução. Estes espaços têm a mesma dimensão, portanto se quisermos demonstrar um teorema sobre a dimensão do espaço solução, não importará qual espaço resolvamos discutir.

3.3 Matrizes de Transformações Lineares

Seja V um espaço vetorial n-dimensional sobre o corpo $\mathbb{K}$ e seja W um espaço vetorial m-dimensional sobre $\mathbb{K}$. Sejam $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ordenada de V e $\mathfrak{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ uma base ordenada de W. Se T é uma transformação linear arbitrária de V em W, então T é determinada por seu efeito sobre os vetores $\mathbf{v}_j$. Cada um dos n vetores $T(\mathbf{v}_j)$ pode ser expresso de modo único como uma combinação linear

$$T(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{u}_i \tag{3.5}$$

dos $\mathbf{u}_i$, sendo os escalares $a_{ij}, \ldots, a_{mj}$ as coordenadas de $T(\mathbf{v}_j)$ em relação à base ordenada $\mathfrak{B}'$. Consequentemente a transformação T é determinada pelos mn escalares a_{ij} por meio das fórmulas (3.5). A $m \times n$ matriz $\mathbf{A}$ definida por $\mathbf{A} = (a_{ij})$ é denominada a matriz de T em relação ao par de bases ordenadas $\mathfrak{B}$ e $\mathfrak{B}'$.

Se $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ é um vetor em V, então

$$\begin{aligned} T(\mathbf{v}) &= T\left(\sum_{j=1}^{n} \alpha_j \mathbf{v}_j\right) \\ &= \sum_{j=1}^{n} \alpha_j T(\mathbf{v}_j) \\ &= \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij}\mathbf{u}_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\alpha_j \mathbf{u}_i\right) \end{aligned}$$

Se $\mathbf{X}$ é a matriz das coordenadas de $\mathbf{v}$ em relação à base ordenada $\mathfrak{B}$, o cálculo acima mostra que $\mathbf{AX}$ é a matriz das coordenadas do vetor $T(\mathbf{v})$ em relação à base ordenada $\mathfrak{B}'$, uma vez que o escalar

$$\sum_{j=1}^{n} a_{ij}\alpha_j$$

é o elemento da i-ésima linha da matriz coluna $\mathbf{AX}$. Observemos também que se $\mathbf{A}$ é uma $m \times n$ matriz arbitrária sobre o corpo $\mathbb{K}$, então

$$T\left(\sum_{j=1}^{n} \alpha_j \mathbf{v}_j\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\alpha_j\right)\mathbf{u}_i \tag{3.6}$$

define uma transformação linear T de V em W, cuja matriz é $\mathbf{A}$, em relação à $\mathfrak{B}$, $\mathfrak{B}'$. Para formalizar este conceito, segue o teorema:

Teorema 25. *Seja V um espaço vetorial n-dimensional sobre o corpo $\mathbb{K}$ e W um espaço vetorial* *m-dimensional sobre $\mathbb{K}$. Seja $\mathfrak{B}$ uma base ordenada de V e $\mathfrak{B}'$ uma base ordenada de W. Para cada transformação linear T de V em W, existe uma $m \times n$ matriz A sobre o corpo $\mathbb{K}$, a matriz de T em relação à $\mathfrak{B}$, $\mathfrak{B}'$, tal que*

$$[T(\mathbf{v})]_{\mathfrak{B}'} = \mathbf{A}[\mathbf{v}]_{\mathfrak{B}}$$

para todo vetor $\mathbf{v} \in V$. Além disso, $T \rightarrow A$ é uma correspondência bijetora entre o conjunto das transformações lineares de V em W e o conjunto das $m \times n$ matrizes sobre o corpo $\mathbb{K}$

Suponhamos agora que T e U sejam transformações lineares de V em W e que a matriz de T em relação a $\mathfrak{B}$, $\mathfrak{B}'$ seja $\mathbf{B}$. Sendo $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$, temos que

$$\begin{aligned}(T+U)(\mathbf{v}) &= T(\mathbf{v}) + U(\mathbf{v}) \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\alpha_j\right)\mathbf{u}_i + \sum_{i=1}^{m}\left(\sum_{j=1}^{n} b_{ij}\alpha_j\right)\mathbf{u}_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} (a_{ij}+b_{ij})\alpha_j\right)\mathbf{u}_i.\end{aligned}$$

Assim, a matriz de $T+U$ em relação a $\mathfrak{B}$, $\mathfrak{B}'$ é a soma das matrizes $\mathbf{A}$ e $\mathbf{B}$. Analogamente, pode-se verificar com facilidade que se α é um escalar arbitrário, então a matriz de (αT) é $\alpha\mathbf{A}$. Estas duas observações implicam no fato que a correspondência entre transformações lineares e matrizes definida por $\mathfrak{B}$ e $\mathfrak{B}'$ é linear.

Agora, suponhamos que a matriz de T em relação a $\mathfrak{B}$, $\mathfrak{B}'$ seja $\mathbf{A}$ e que a matriz de U em relação ao mesmo par seja $\mathbf{B}$. Se $\mathbf{v}$ é um vetor arbitrário em V,

$$\begin{aligned}[T(\mathbf{v})]_{\mathfrak{B}'} &= \mathbf{A}[\mathbf{v}]_{\mathfrak{B}} \\ [U(\mathbf{v})]_{\mathfrak{B}} &= {}'\mathbf{B}[\mathbf{v}]_{\mathfrak{B}}\end{aligned}$$

e como sabemos que

$$(\alpha\mathbf{A} + \mathbf{B})\mathbf{X} = \alpha\mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{X}$$

temos

$$[\alpha T(\mathbf{v}) + U(\mathbf{v})]_{\mathfrak{B}'} = (\alpha\mathbf{A} + \mathbf{B})[\mathbf{v}]_{\mathfrak{B}}$$

Portanto, deve-se ter que $\alpha\mathbf{A} + \mathbf{B}$ é a matriz de $\alpha T + U$ em relação ao par S, S'.

Teorema 26. *Seja V um espaço vetorial n-dimensional sobre o corpo $\mathbb{K}$ e seja W um espaço vetorial m-dimensional sobre $\mathbb{K}$. Para cada par de bases ordenadas $\mathfrak{B}$, $\mathfrak{B}'$ de V e W, respectivamente, a função que associa a uma transformação linear T sua matriz em relação a $\mathfrak{B}$, $\mathfrak{B}'$ é um isomorfismo entre o espaço $L(V,W)$ e o espaço das $m \times n$ matrizes sobre o corpo $\mathbb{K}$.*

Demonstração: Observamos acima que a função em questão é linear e, como está enunciado no Teorema (25), esta função é injetora e leva $L(V,W)$ sobre o conjunto das $m\times n$ matrizes. ■

Estaremos mais interessados na representação por matrizes de transfomações lineares de um espaço em si mesmo, isto é, operadores lineares sobre um espaço V. Neste caso, é mais conveniente usar a mesma base ordenada em casa caso, isto é, tomar $\mathfrak{B}=\mathfrak{B}'$. A matriz representante será então denominada simplesmente a matriz de T em relação à base ordenada $\mathfrak{B}$.

Relembrando, se T é um operador linear sobre o espaço vetorial V de dimensão finita e $\mathfrak{B}=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ é uma base ordenada de V, a matriz de T em relação a $\mathfrak{B}$ é a $n\times n$ matriz $\mathbf{A}$ cujos elementos a_{ij} são definidos pelas equações

$$T(\mathbf{v}_j)=\sum_{i=1}^{n}a_{ij}\mathbf{v}_i,\quad j=1,\ldots,n. \tag{3.7}$$

Deve-se ter sempre em mente que esta matriz que representa T depende da base ordenada $\mathfrak{B}$ e que existe uma matriz que representa T em relação a cada base ordenada de V.

Denotaremos a matriz do operador linear T em relação à base ordenada $\mathfrak{B}$ poe $[T]_{\mathfrak{B}}$. A maneira como esta matriz e a base ordenada descrevem T é que, para cada $v\in V$,

$$[T(\mathbf{v})]_{\mathfrak{B}}=[T]_{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}}.$$

Exemplo 9. *Seja $\mathbb{K}$ um corpo e seja T o operaor sobre $\mathbb{K}^2$ definido por*

$$T(\mathbf{x}_1,\mathbf{x}_2)=(\mathbf{x}_1,0)$$

É fácil ver que T é um operador linear sobre $\mathbb{K}^2$. Seja $\mathfrak{B}$ a base ordenada canônica de $\mathbb{K}^2$, $\mathfrak{B}=\{\mathbf{e}_1,\mathbf{e}_2\}$. Então

$$T(\mathbf{e}_1)=T(1,0)=(1,0)=1\mathbf{e}_1+0\mathbf{e}_2$$

$$T(\mathbf{e}_2)=T(0,1)=(0,0)=0\mathbf{e}_1+0\mathbf{e}_2$$

de modo que a matriz de T em relação à base ordenada $\mathfrak{B}$ é

$$[T]_{\mathfrak{B}} = \begin{pmatrix} 1 & 0\\ 0 & 0\end{pmatrix}$$

Sejam V, W e Z espaços vetoriais sobre o corpo $\mathbb{K}$, de dimensões n, m e p, resectivamente. Seja T uma transformação linear de V em W e U uma transformação linear de W em Z. Suponhamos que existam bases ordenadas

$$\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\},\ \mathfrak{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\},\ \mathfrak{B}'' = \{\mathbf{t}_1, \ldots, \mathbf{t}_n\}$$

para os espaços V, W e Z, respectivamente. Seja $\mathbf{A}$ a matriz de T em relação ao par $\mathfrak{B}$ e $\mathfrak{B}'$ e $\mathbf{B}$ a matriz de U em relação ao par $\mathfrak{B}'$ e $\mathfrak{B}''$. É fácil ver então que a matriz $\mathbf{C}$ da transformação UT em relação ao par $\mathfrak{B}$, $\mathfrak{B}''$ é o produto de B por A, pois se v é um vetor arbitrário em V

$$\begin{aligned} [T(\mathbf{v})]_{\mathfrak{B}'} &= \mathbf{A}[\mathbf{v}]_{\mathfrak{B}} \\ [U(T(\mathbf{v}))]_{\mathfrak{B}''} &= \mathbf{B}[T(\mathbf{v})]_{\mathfrak{B}'} \\ [(UT)(\mathbf{v})]_{\mathfrak{B}''} &= \mathbf{BA}[\mathbf{v}]_{\mathfrak{B}}. \end{aligned}$$

Logo, pela definição e unicidade da matriz representante, temos, necessariamente, $\mathbf{C} = \mathbf{BA}$. Isto também pode ser visto efetuando os cálculos

$$\begin{aligned} (UT)(\mathbf{v}_j) &= U(T(\mathbf{v}_j)) \\ &= U\left(\sum_{k=1}^{m} a_{kj}\mathbf{u}_k\right) \\ &= \sum_{k=1}^{m} a_{kj}(U\mathbf{u}_k)) \\ &= \sum_{k=1}^{m} a_{kj}\sum_{i=1}^{p} b_{ik}\mathbf{t}_i \\ &= \sum_{i=1}^{p}\left(\sum_{k=1}^{m} b_{ik}a_{kj}\right)\mathbf{t}_i \end{aligned}$$

de modo que temos

$$c_{ij} = \sum_{k=1}^{m} b_{ik}c_{kj}. \tag{3.8}$$

Formalizando este conceito, segue o teorema.

Teorema 27. *Sejam V, W e Z espaços vetoriais de dimensão finita sobre o corpo $\mathbb{K}$; seja T uma transformação linear de V em W e U uma transformação linear de W em Z. SE $\mathfrak{B}$, $\mathfrak{B}'$ e $\mathfrak{B}''$ são bases ordnadas dos espaços V, W e Z, respectivamente, se $\mathbf{A}$ é a matriz de T em relação ao par $\mathfrak{B}$ e $\mathfrak{B}'$ e $\mathbf{B}$ é a matriz de U em relação ao par $\mathfrak{B}'$ e $\mathfrak{B}''$, então a matriz da composta UT em relação ao par $\mathfrak{B}$, $\mathfrak{B}''$ é a matriz produto $\mathbf{C} = \mathbf{BA}$.*

É importante notar que se T e U são operadores lineares sobre um espaço V e se estamos usando apenas uma base ordenada $\mathfrak{B}$, então o Teorema (25) toma a forma simples $[UT]_{\mathfrak{B}} = [U]_{\mathfrak{B}}[T]_{bs}$. Assim, nesse caso, a correspondência que $\mathfrak{B}$ determina entre operadores lineares e matrizes é não somente um isomorfismo de espaço vetorial, mas conserva também produtos.

Uma consequência simples disso é que o operador linear T é inversível se, e somente se, $[T]_{\mathfrak{B}}$ é uma matriz inversível. De fato, o operador idêntico I é representado pela matriz identidade em relação a qualquer base ordenada, portanto,

$$UT = TU = I$$

é equivalente a

$$[U]_{\mathfrak{B}}[T]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[U]_{\mathfrak{B}} = \mathbf{I}.$$

Evidentemente, quando T é inversível

$$[T^{-1}]_{\mathfrak{B}} = [T]_{\mathfrak{B}}^{-1}$$

Agora consideremos a seguinte questão: seja T um operador linear sobre o espaço de dimensão finita V e sejam

$$\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\},\ \mathfrak{B}' = \{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$$

duas bases ordenadas de V. Sabemos que existe uma única $n \times n$ matriz inversível $\mathbf{P}$ tal que

$$[\mathbf{v}]_{\mathfrak{B}} = \mathbf{P}[\mathbf{v}]_{\mathfrak{B}'} \tag{3.9}$$

para todo vetor $\mathbf{v} \in V$. Por definição

$$[T(\mathbf{v})]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} \tag{3.10}$$

Aplicando (3.9) ao vetor $T(\mathbf{v})$ temos

$$[T(\mathbf{v})]_{\mathfrak{B}} = \mathbf{P}[T(\mathbf{v})]_{\mathfrak{B}'} \tag{3.11}$$

Combinando (3.9), (3.10) e (3.11), obtemos

$$[T]_{\mathfrak{B}}\mathbf{P}[\mathbf{v}]_{\mathfrak{B}'} = \mathbf{P}[T(\mathbf{v})]_{\mathfrak{B}'}$$

ou

$$\mathbf{P}^{-1}[T]_{\mathfrak{B}}\mathbf{P}[\mathbf{v}]_{\mathfrak{B}'} = [T(\mathbf{v})]_{\mathfrak{B}'}$$

e então é necessário que

$$[T]_{\mathfrak{B}'} = \mathbf{P}^{-1}[T]_{\mathfrak{B}}\mathbf{P}. \tag{3.12}$$

Antes de enunciarmos formalmente este resultado, observemos um fato. Existe um único operador linear U que leva $\mathfrak{B}$ sobre $\mathfrak{B}'$ definido por

$$U(\mathbf{v}_j) = \mathbf{v}'_j,\ j = 1, \ldots, n.$$

Este operador U é inversível uma vez que leva uma base de V sobre uma base de V. A matriz $\mathbf{P}$ (dada como acima) é exatamente a matriz do operador U em relação à base ordenada $\mathfrak{B}$. De fato, $\mathbf{P}$ é definida por

$$\mathbf{v}'_j = \sum_{i=1}^{n} \mathbf{P}_{ij}\mathbf{v}_i$$

e como $U(\mathbf{v}_j) = \mathbf{v}'_j$, esta equação pode ser escrita como

$$U(\mathbf{v}_j) = \sum_{i=1}^{n} \mathbf{P}_{ij}\mathbf{v}_i.$$

Portanto $\mathbf{P} = [U]_{\mathfrak{B}}$ por definição.

Teorema 28. *Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e sejam*

$$\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\},\ \mathfrak{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$$

bases ordenadas de V. Suponhamos que T seja um operador linear sobre V. Se $\mathbf{P}$ é a $n \times n$ matriz que exprime as coordenadas de cada vetor de V em relação a $\mathfrak{B}$ em têrmos de suas coordenadas em relação a $\mathfrak{B}'$, então

$$[T]_{\mathfrak{B}'} = \mathbf{P}^{-1}[T]_{\mathfrak{B}} + \mathbf{P}.$$

Alternativamente, se U é o operador inversível sobre V definido por $U(\mathbf{v}_j) = \mathbf{v}'_j$, $j = 1, \ldots, n$, então

$$[T]_{\mathfrak{B}'} = [U]_{\mathfrak{B}}^{-1}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}}$$

4 Normas, Produto Interno e Ortogonalidade

4.1 Normas

Definição 33. *Dado um espaço vetorial V sobre o corpo $\mathbb{K}$ dos números reais ou complexos, uma função $\|\cdot\| : V \to \mathbb{R}^+$ é chamada de norma se, para quaisquer $\mathbf{u}, \mathbf{v} \in V$ e todo $\alpha \in \mathbb{K}$:*

i) $\|\mathbf{u}\| = 0 \Rightarrow \mathbf{u} = \mathbf{0}$. *Se esta condição não for atendida, a função será no máximo uma Semi-norma.*

ii) $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$.

iii) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. *(Desigualdade Triangular)*

Se o espaço vetorial V tem uma norma, ele passa a ser chamado de espaço normado.

Há algumas normas usuais para vetores quando trabalha-se sobre o corpo $\mathbb{K}$ (sendo $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Segue abaixo uma lista das normas mais comuns.

Consideremos V um $\mathbb{K}$-espaço vetorial normado de dimensão finita n e $\mathbf{v} \in V$. Sendo assim, são normas vetoriais sobre V:

- **Norma Euclidiana**: $\|\mathbf{v}\| \triangleq \sqrt{\bar{\mathbf{v}}^T \mathbf{v}} = \left(\sum_{i=1}^{n} |\mathbf{v}_i|^2 \right)^{\frac{1}{2}}$

- **Norma-p**: $\|\mathbf{v}\|_p \triangleq \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}, \;\; 1 \leq p < \infty$

- **Norma-∞**: $\|\mathbf{v}\|_\infty \triangleq \max\limits_{1 \leq i \leq n} |v_i|$

No MATLAB, os comandos que computam tais normas são dados por

- **Norma Euclidiana**: `norm(v) = norm(v,2) = sum(abs(v).^2)^(1/2)`
- **Norma-p**: `norm(v,p) = sum(abs(v).^p)^(1/p)`
- **Norma-∞**: `norm(v,inf) = max(abs(V))`

O conceito de norma vetorial pode ser estendido para matrizes, no sentido de podemos mensurar matrizes.

Definição 34. *Uma norma matricial uma é uma função $\|\cdot\| : \mathbb{K}^{m\times n} \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), sendo $m,n < \infty$, que satisfaz as seguintes propriedades:*

i) $\|\mathbf{A}\| \geq 0$ *e* $\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = 0,\ \forall\, \mathbf{A} \in \mathbb{K}^{m\times n}$

ii) $\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|,\ \forall\, \alpha \in \mathbb{K},\ \forall\, \mathbf{A} \in \mathbb{K}^{m\times n}$

iii) $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|,\ \forall\, \mathbf{A},\mathbf{B} \in \mathbb{K}^{m\times n}$

A norma matricial de uso mais frequente na matemática, e também nesta pesquisa, é a norma de Frobenius, definida por

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}$$

sendo $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$.

Definição 35. *Seja $\|\cdot\|$ uma norma vetorial em $\mathbb{K}^n$. A norma da matriz $\mathbf{A} \in \mathbb{K}^{m\times n}$ induzida pela norma vetorial é dada por*

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| = \max_{\|\mathbf{x}\|\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$$

sendo $\mathbf{x}$ pertencente a $\mathbb{K}^{n\times 1}$.

Observação 7. *Se a norma matricial for induzida por uma norma-p vetorial, denotaremos a norma de uma matriz $\mathbf{A}$ por $\|\mathbf{A}\|_p$.*

Proposição 5. *Sejam $\mathbf{A} \in \mathbb{K}^{m\times p}$, $\mathbf{B} \in \mathbb{K}^{p\times n}$ e $\mathbf{x} \in \mathbb{K}^{p\times 1}$. Então temos que*

i) $\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|\ \|\mathbf{x}\|$

ii) $\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\ \|\mathbf{B}\|$

Demonstração:

i) Se $||x|| = 0$, o caso é trivial. Suponha $||x|| \neq \mathbf{0}$. Então, da Definição (35), temos que

$$||\mathbf{A}|| \geq \frac{||\mathbf{Ax}||}{||\mathbf{x}||} \Leftrightarrow ||\mathbf{Ax}|| \leq ||\mathbf{A}||\ ||\mathbf{x}||$$

ii) Para mostrarmos que esta propriedade vale, basta ver que do item (i), para qualquer vetor não nulo $\mathbf{x}$,

$$\frac{||\mathbf{ABx}||}{||\mathbf{x}||} \leq ||\mathbf{A}|| \frac{||\mathbf{Bx}||}{||\mathbf{x}||}$$

e, portanto,

$$||\mathbf{AB}|| = \max_{||\mathbf{x}|| \neq 0} \frac{||\mathbf{ABx}||}{||\mathbf{x}||} \leq ||\mathbf{A}|| \max_{||\mathbf{x}|| \neq 0} \frac{||\mathbf{Bx}||}{||\mathbf{x}||} = ||\mathbf{A}||\ ||\mathbf{B}||$$

■

Consideremos o espaço vetorial $\mathbb{K}^{m \times n}$ normado e $\mathbf{v} \in V$. Sendo assim, segue abaixo alguns exemplos de normas matriciais sobre $\mathbb{K}^{m \times n}$:

- **Norma da soma**: $||\mathbf{A}||_1 = \max\limits_{j=1,\ldots,n} \left\{ \sum\limits_{i=1}^{m} |a_{ij}| \right\}$

- **Norma do máximo**: $||\mathbf{A}||_\infty = \max\limits_{i=1,\ldots,m} \left\{ \sum\limits_{j=1}^{n} |a_{ij}| \right\}$

No MATLAB, os comandos que computam tais normas são dados por

- **Norma da soma**: `norm(A,1)`

- **Norma do máximo**: `norm(A,inf)`

4.2 Produto Interno

Um produto interno é uma função $V \times V \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), que associa a cada par de vetores $\mathbf{u}, \mathbf{v} \in V$ um escalar (real ou complexo) $\langle \mathbf{u}, \mathbf{v} \rangle$, chamado o produto interno de $\mathbf{u}$ por $\mathbf{v}$ de modo que as seguintes propriedades sejam satisfeitas, para quaisquer $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ e $\alpha \in \mathbb{K}$:

1. $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$,

2. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$

3. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$;

4. $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ se $\mathbf{u} \neq \mathbf{0}$

Observação 8.

i) *Segue facilmente das propriedades acima que*

- $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = \mathbf{0}, \ \forall \mathbf{v} \in V$
- $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Longleftrightarrow \mathbf{v} = \mathbf{0}$

ii) *É facil ver que vale*

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle, \ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$$

De fato, utilizando as propriedades $(1.)$ *e* $(3.)$*, teremos para* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *que*

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \overline{\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle} = \overline{\langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle} = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} + \overline{\langle \mathbf{w}, \mathbf{u} \rangle} = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$$

iii) *Utilizando as propriedades* $(2.)$ *e* $(3.)$*, chega-se a :*

$$\langle \mathbf{u}, \alpha \mathbf{v} \rangle = \overline{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle, \ \forall \alpha \in \mathbb{K}, \ \forall \mathbf{u}, \mathbf{v}, \in V$$

De fato, se $\alpha \in \mathbb{K}$ *e* $\mathbf{u}, \mathbf{v} \in V$*, temos que*

$$\langle \mathbf{u}, \alpha \mathbf{v} \rangle = \overline{\langle \alpha \mathbf{v}, \mathbf{u} \rangle} = \overline{\alpha \langle \mathbf{v}, \mathbf{u} \rangle} = \overline{\alpha}\, \overline{\langle \mathbf{v}, \mathbf{u} \rangle} = \overline{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle$$

iv) *No caso de* $\mathbb{K} = \mathbb{R}$*, a prorpiedade* $(3.)$ *implica a igualdade* $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ *para* $\mathbf{u}, \mathbf{v} \in V$*, pois, neste caso, teremos que* $\overline{\langle \mathbf{v}, \mathbf{u} \rangle} = \langle \mathbf{u}, \mathbf{v} \rangle \ \forall \mathbf{u}, \mathbf{v} \in V$*. Essa simetria se perde no caso complexo. De fato, se* V *é um espaço vetorial complexo, temos por* $(4.)$ *que* $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *e* $\langle i\mathbf{v}, i\mathbf{v} \rangle > 0$*. Se tivéssemos aqui* $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ *para todos* $\mathbf{u}, \mathbf{v} \in V$*, então usando* $(2.)$*, chegaríamos a*

$$\langle i\mathbf{v}, i\mathbf{v} \rangle = i \langle \mathbf{v}, i\mathbf{v} \rangle = i \langle i\mathbf{v}, \mathbf{v} \rangle = i^2 \langle \mathbf{v}, \mathbf{v} \rangle = -\langle \mathbf{v}, \mathbf{v} \rangle > 0,$$

uma contradição.

Definição 36. *Em um espaço vetorial* V *munido um um produto interno* $\langle \ , \ \rangle$*, para cada* $\mathbf{u} \in V$*, o número não negativo* $\|\mathbf{u}\| = \sqrt{\langle u, u \rangle}$ *é denominado a norma do vetor* $\mathbf{u}$ *induzida pelo produto interno* $\langle \ , \ \rangle$*.*

Quando $\|\mathbf{u}\| = 1$ diz-se que $\mathbf{u} \in V$ é um vetor unitário. Todo vetor $\mathbf{u} \neq \mathbf{0}$ se escreve como $\mathbf{u} = \|\mathbf{u}\| \cdot \mathbf{u}'$, em que $\mathbf{u}'$ é um vetor unitário. Para isto, basta escrever $\mathbf{u}' = \|\mathbf{u}\|^{-1} \cdot \mathbf{u}$.

Exemplo 10. *No espaço euclidiano* $\mathbb{R}^n$*, o produto interno canônico de* $\mathbf{u} = (\alpha_1, \cdots, \alpha_n)$ *com* $\mathbf{v} = (\beta_1, \cdots, \beta_n)$ *é definido por* $\langle \mathbf{u}, \mathbf{v} \rangle = \alpha_1 \beta_1 + \cdots + \alpha_n \beta_n$*.*

Todo espaço vetorial de dimensão finita pode ser munido de um produto interno. Assim, quando nos referimos a um espaço munido de um produto interno, não é atribuída a esse espaço uma propriedade especial, mas sim estamos dizendo que entre os possíveis produtos internos que nele podem ser introduzidos, um em particular foi escolhido e fixado.

4.3 Ortogonalidade

Definição 37. *Seja V um espaço vetorial com produto interno. Dois vetores $\mathbf{u},\mathbf{v} \in V$ são ditos ortogonais quando $\langle \mathbf{u},\mathbf{v}\rangle = 0$. Denota-se então por $\mathbf{u} \perp \mathbf{v}$.*

Em particular, $\mathbf{0}$ é ortogonal a qualquer vetor de V pois $\langle \mathbf{0},\mathbf{v}\rangle = \langle \mathbf{0}+\mathbf{0},\mathbf{v}\rangle = \langle \mathbf{0},\mathbf{v}\rangle + \langle \mathbf{0},\mathbf{v}\rangle$ e então segue que $\langle \mathbf{0},\mathbf{v}\rangle = \langle \mathbf{v},\mathbf{0}\rangle = \mathbf{0}$ para todo $\mathbf{v} \in V$.

Definição 38. *Um conjunto $X \subset V$ diz-se ortogonal quando dois vetores distintos quaisquer em X são ortogonais. Se, além disso, todos os vetores de X são unitários então X chama-se um conjunto ortonormal.*

Proposição 6. *Seja V um espaço vetorial com produto interno e seja A um subconjunto ortogonal de V formado por vetores não nulos. Se $\mathbf{v}$ pertence ao conjunto gerado pelos vetores $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, com $\mathbf{v}_i \in A$, então*

$$\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v},\mathbf{v}_i\rangle}{\|\mathbf{v}_i\|^2}\mathbf{v}_i.$$

Demonstração: Seja $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$, com $\alpha_i \in \mathbb{K}$, $i = 1,\ldots,n$. Então como $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ é um conjunto ortogonal, temos, para $j = 1,\ldots,n$, que

$$\langle \mathbf{v},\mathbf{v}_j\rangle = \langle \sum_{i=1}^{n} \alpha_i \mathbf{v}_i,\mathbf{v}_j\rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{v}_i,\mathbf{v}_j\rangle = \alpha_j \langle \mathbf{v}_j,\mathbf{v}_j\rangle$$

Portanto,

$$\alpha_j = \frac{\langle \mathbf{v},\mathbf{v}_j\rangle}{\|\mathbf{v}_i\|^2},\ \forall j = 1,\ldots,n, \text{ e } \mathbf{v} = \textstyle\sum_{i=1}^{n} \frac{\langle \mathbf{v},\mathbf{v}_i\rangle}{\|\mathbf{v}_i\|^2}\mathbf{v}_i.$$

■

Proposição 7. *Num espaço vetorial V com produto interno, todo conjunto ortogonal X de vetores não nulos é linearmente independente.*

Demonstração: Sejam $\mathbf{v}_1,\ldots,\mathbf{v}_n \in X$. Temos $\langle \mathbf{v}_i,\mathbf{v}_j\rangle = 0$ se $i \neq j$. Se $\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = 0$ é uma combinação linear nula desses vetores então, para cada $i = 1,2,\ldots,n$, tomando o produto interno de ambos os membros desta igualdade por $\mathbf{v}_i$, temos

$$\alpha_1\langle \mathbf{v}_1, \mathbf{v}_i\rangle + \ldots + \alpha_n\langle \mathbf{v}_n, \mathbf{v}_i\rangle = 0,$$

logo, $\alpha_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle = \alpha_i\|\mathbf{v}_i\|^2 = 0$ pois todos os produtos internos $\langle \mathbf{v}_i, \mathbf{v}_j\rangle$, com $i \neq j$, são nulos em virtude da ortogonalidade de X. Além disso, como os vetores pertencentes ao conjunto X são todos não nulos, resulta de $\alpha_i\|\mathbf{v}_i\|^2 = 0$ que $\alpha_i = 0$. Assim, os coeficientes da combinação linear $\sum \alpha_i \mathbf{v}_i = 0$ são todos iguais a zero e os vetores do conjunto X são, portanto, linearmente independentes. ∎

Corolário 5. *Seja V um espaço vetorial sobre $\mathbb{K}$ com produto interno e seja $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ortonormal de V. Então para $\mathbf{v} \in V$, temos*

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v}_i\rangle \mathbf{v}_i.$$

Definição 39. *Seja $\mathbf{u} \in V$ um vetor unitário e $\mathbf{v} \in V$ um vetor qualquer. Então o vetor $\langle \mathbf{u}, \mathbf{v}\rangle\mathbf{u}$ chama-se a projeção ortogonal de $\mathbf{v}$ sobre o subespaço gerado por $\mathbf{u}$.*

Esta definição é justificada pelo fato de que $\mathbf{w} = \mathbf{v} - \langle \mathbf{u}, \mathbf{v}\rangle\mathbf{u}$ implica que $\mathbf{v} = \langle \mathbf{u}, \mathbf{v}\rangle\mathbf{u} + \mathbf{w}$, em que $\mathbf{w}$ é perpendicular a $\mathbf{u}$. Com efeito, tomando o produto interno de $\mathbf{u}$ por ambos os membros da igualdade $\mathbf{w} = \mathbf{v} - \langle \mathbf{u}, \mathbf{v}\rangle\mathbf{u}$, tem-se

$$\langle \mathbf{u}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle\langle \mathbf{u}, \mathbf{u}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle = 0$$

pois $\langle \mathbf{u}, \mathbf{u}\rangle = 1$.

Quando se tem apenas $\mathbf{u} \neq \mathbf{0}$, o subespaço gerado por $\mathbf{u}$ é o mesmo gerado pelo vetor unitário $\mathbf{u}' = \mathbf{u}/|\mathbf{u}|$. A projeção ortogonal de $\mathbf{v}$ sobre este subespaço é, portanto, igual a $\langle \mathbf{u}', \mathbf{v}\rangle\mathbf{u}'$, ou seja, $(\langle \mathbf{u}, \mathbf{v}\rangle/\langle \mathbf{u}, \mathbf{u}\rangle\mathbf{u}$. Denotamos por

$$pr_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$$

a projeção ortogonal do vetor $\mathbf{v}$ sobre o subespaço gerado pelo vetor não nulo $\mathbf{u}$.

Teorema 29. ***(Desigualdade de Schwarz)*** *Seja V um espaço vetorial com produto interno. Então*

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|, \ \forall \mathbf{u}, \mathbf{v} \in V.$$

A igualdade vale se, e somente se $\{\mathbf{u}, \mathbf{v}\}$ for linearmente dependente.

Demonstração: Para $\alpha, \beta \in \mathbb{K}$ e $\mathbf{u}, \mathbf{v} \in V$, teremos:

$$\begin{aligned}
\langle \alpha\mathbf{u} - \beta\mathbf{v}, \alpha\mathbf{u} - \beta\mathbf{v} \rangle &= \langle \alpha\mathbf{u}, \alpha\mathbf{u} \rangle - \langle \beta\mathbf{v}, \alpha\mathbf{u} \rangle - \langle \alpha\mathbf{u}, \beta\mathbf{v} \rangle + \langle \beta\mathbf{v}, \beta\mathbf{v} \rangle \\
&= \alpha\overline{\alpha}\langle \mathbf{u}, \mathbf{u} \rangle - \beta\overline{\alpha}\langle \mathbf{v}, \mathbf{u} \rangle - \alpha\overline{\beta}\langle \mathbf{u}, \mathbf{v} \rangle + \beta\overline{\beta}\langle \mathbf{v}, \mathbf{v} \rangle \\
&= |\alpha|^2\|\mathbf{u}\|^2 - 2Re(\alpha\overline{\beta}\langle \mathbf{u}, \mathbf{v} \rangle) + |\beta|^2\|\mathbf{v}\|^2
\end{aligned}$$

(na última igualdade, usamos $\beta\overline{\alpha}\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\alpha\overline{\beta}\langle \mathbf{u}, \mathbf{v} \rangle}$). Se tomarmos $\alpha = \|\mathbf{v}\|^2$ e $\beta = \langle \mathbf{u}, \mathbf{v} \rangle$, obteremos

$$\alpha\overline{\beta}\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{v}\|^2\overline{\beta}\beta = \|\mathbf{v}\|^2|\beta|^2 \in \mathbb{R}.$$

Daí segue que

$$\begin{aligned}
0 &\leq \langle \alpha\mathbf{u} - \beta\mathbf{v}, \alpha\mathbf{u} - \beta\mathbf{v} \rangle \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^4 - 2\|\mathbf{v}\|^2|\langle \mathbf{u}, \mathbf{v} \rangle|^2 + |\langle \mathbf{u}, \mathbf{v} \rangle|^2\|\mathbf{v}\|^2 \\
&= \|\mathbf{v}\|^2(\|\mathbf{u}\|^2\|\mathbf{v}\|^2 - |\langle \mathbf{u}, \mathbf{v} \rangle|^2).
\end{aligned}$$

Portanto,

$$\|\mathbf{u}\|^2\|\mathbf{v}\|^2 - |\langle \mathbf{u}, \mathbf{v} \rangle|^2 \geq 0$$

ou, equivalentemente

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

■

Suponha agora que $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$. Usando o cálculo acima teremos $\langle \alpha\mathbf{u} - \beta\mathbf{v}, \alpha\mathbf{u} - \beta\mathbf{v} \rangle = 0$ quando $\alpha = \|\mathbf{v}\|^2$ e $\beta = \langle \mathbf{u}, \mathbf{v} \rangle$. Se $\mathbf{v} = \mathbf{0}$, então $\{\mathbf{u}, \mathbf{v}\}$ é, com certeza, linearmente dependente. Suponha então que $\mathbf{v} \neq \mathbf{0}$. Logo $\alpha \neq 0$ e daí $\alpha\mathbf{u} = \beta\mathbf{v}$, o que implica que $\mathbf{u} = \frac{\beta}{\alpha}\mathbf{v}$. Portanto, o conjunto $\{\mathbf{u}, \mathbf{v}\}$ é linearmente dependente.

Supondo agora que $\{\mathbf{u}, \mathbf{v}\}$ é linearmente dependente, isto é, supondo que exista $\lambda \in \mathbb{K}$ tal que $\mathbf{u} = \lambda\mathbf{v}$, uma conta simples nos leva à igualdade $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$.

Corolário 6. ***(Desigualdade Triangular)*** *Seja V um espaço vetorial com produto interno. Então*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \ \forall \mathbf{u}, \mathbf{v} \in V$$

Demonstração: Vimos acima que

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2Re(\langle \mathbf{u}, \mathbf{v} \rangle) + \|\mathbf{v}\|^2.$$

Como $Re(z) \leq |z|$ para todo $z \in \mathbb{C}$, teremos então que

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \; \|\mathbf{v}\| + \|\mathbf{v}\|^2 \end{aligned}$$

Portanto

$$\|\mathbf{u}+\mathbf{v}\|^2 \leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$$

do qual segue o resultado, pois $\|\mathbf{u}+\mathbf{v}\| \geq 0$ e $\|\mathbf{u}\| + \|\mathbf{v}\| \geq 0$. ∎

Mostraremos agora que existem bases ortonormais em todo espaço vetorial de dimensão finita provido de um produto interno.

O Processo de Ortogonalização de Gram-Schmidt

Seja V um espaço vetorial sobre $\mathbb{K}$ com produto interno $\langle \ , \ \rangle$. Considere $A = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset V$ um conjunto linearmente independente. Vamos construir outro conjunto $A' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\} \subset V$ que seja ortogonal e tal que os subespaços gerados por A e por A' sejam os mesmos. Esta construção é feita indutivamente como segue

- $\mathbf{w}_1 = \mathbf{v}_1$
- $\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1$.

  Observe que $\mathbf{w}_2 \neq \mathbf{0}$, já que $\{\mathbf{v}_1, \mathbf{v}_2\}$ é linearmente independente, além do fato que $\mathbf{w}_2 \perp \mathbf{w}_1$. De fato,

$$\begin{aligned} \langle \mathbf{w}_2, \mathbf{w}_1\rangle &= \langle \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1, \mathbf{w}_1\rangle \\ &= \langle \mathbf{v}_2, \mathbf{w}_1\rangle - \frac{\langle \mathbf{v}_2, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2} \cdot \langle \mathbf{w}_1, \mathbf{w}_1\rangle \\ &= 0. \end{aligned}$$

- Definidos $\mathbf{w}_1, \ldots, \mathbf{w}_n$, $1 < k < n$, podemos definir $\mathbf{w}_{k+1}$ como sendo:

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{v}_{k+1} - \frac{\langle \mathbf{v}_{k+1}, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \cdots - \frac{\langle \mathbf{v}_{k+1}, \mathbf{w}_k\rangle}{\|\mathbf{w}_k\|^2}\mathbf{w}_k \\ &= \mathbf{v}_{k+1} - \sum_{j=1}^{k} \frac{\langle \mathbf{v}_{k+1}, \mathbf{w}_j\rangle}{\|\mathbf{w}_j\|^2}\mathbf{w}_j \end{aligned}$$

Pelos resultados vistos nas proposições (6) e (7), fica fácil ver que o conjunto $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ é ortogonal e , em particular, lineamente independente. Observe também que, para cada $i = 1, \ldots, n$ $\mathbf{w}_i$ pertence ao conjunto W gerado pelos vetores $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $W = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. Como

$dim(W) = n$, segue que $A' = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ é uma base de W, o que mostra a igualdade dos subespaços gerados por A e por A'.

Corolário 7. *Todo espaço vetorial de dimensão finita com produto interno admite uma base ortogonal.*

O próximo resultado explicita uma condição suficiente para que um espaço vetorial tenha uma base ortonormal.

Teorema 30. *Todo espaço vetorial de dimensão finita $n \geq 1$ com produto interno possui uma base ortonormal.*

Demonstração: Para provarmos este resultado, vamos usar o processo de ortogonalização de Gram-Schmidt descrito anteriormente.

Seja V um espaço vetorial de dimensão $n \geq 1$ e seja $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ uma base de V. Então, pelo corolário (7), existe um conjunto ortogonal $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ que gera V. Como todo conjunto ortogonal é linearmente independente, segue que $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ é uma base ortogonal de V. Por fim, $\left\{ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \dots, \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|} \right\}$ é uma base ortonormal, como desejávamos. ■

Teorema 31. *Se $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ é uma base ortonormal de um espaço com produto interno V e $\mathbf{u}$ é um vetor qualquer de V, então*

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \dots, + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n$$

Demonstração: Como $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ é uma base, o vetor $\mathbf{u}$ pode ser escrito na forma

$$\mathbf{u} = k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \dots + k_n \mathbf{v}_n$$

Vamos mostrar que $k_i = \langle \mathbf{u}, \mathbf{v}_i \rangle$ para $i = 1, \dots, n$. Para cada vetor $\mathbf{v}_i$ de β nós temos:

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v}_i \rangle &= \langle k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \dots + k_n \mathbf{v}_n, \mathbf{v}_i \rangle \qquad (4.1) \\ &= k_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + k_2 \langle \mathbf{v}_2, \mathbf{v}_i \rangle + \dots + k_n \langle \mathbf{v}_n, \mathbf{v}_i \rangle \end{aligned}$$

Pelo fato que $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ é um conjunto ortonormal, temos que

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & se\ i = j \\ 0 & se\ i \neq j \end{cases}$$

Portanto a equação (4.2) pode ser simplificada a

$$\langle \mathbf{u}, \mathbf{v}_i \rangle = k_i,\ i = 1, \dots, n$$

■

Definição 40. *Seja W um subespaço do espaço vetorial V. Um vetor $\mathbf{v} \in V$ é dito ortogonal a W se ele for ortogonal a todos os vetores em W. O conjunto de todos os vetores em V que são ortogonais a W é chamado complemento ortogonal de W em V e é denotado por $W^{\perp}$.*

Observação 9. *Seja W um subconjunto de um espaço vetorial V com produto interno.*

i) *O conjunto $W^{\perp}$ é um subespaço vetorial de V, mesmo que W não tenha estrutura de espaço vetorial. De fato,*

- $\mathbf{0} \in W^{\perp}$, *pois* $\langle \mathbf{0}, \mathbf{v} \rangle = 0,\ \forall\, \mathbf{v} \in V$.
- *Se* $\mathbf{v}_1, \mathbf{v}_2 \in W^{\perp}$ *então* $\langle \mathbf{v}_1, \mathbf{u} \rangle = \langle \mathbf{v}_2, \mathbf{u} \rangle = 0,\ \forall\, \mathbf{u} \in W$.

  Portanto, temos $\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{u} \rangle = \langle \mathbf{v}_1, \mathbf{u} \rangle + \langle \mathbf{v}_2, \mathbf{u} \rangle = 0,\ \forall\, \mathbf{u} \in W$ *e então* $\mathbf{v}_1 + \mathbf{v}_2 \in W^{\perp}$
- *Analogamente, se* $\alpha \in \mathbb{K}$ *e* $\mathbf{v} \in W^{\perp}$, *então* $\alpha \mathbf{v} \in W^{\perp}$.

ii) *Se* $W = \{\mathbf{0}\}$, *então* $W^{\perp} = V$.

iii) *Se W contiver uma base de V, então* $W^{\perp} = \{\mathbf{0}\}$.

iv) *É claro que* $W^{\perp} = \{\mathbf{v} \in V; \langle \mathbf{u}, \mathbf{v} \rangle = 0,\ \forall\, \mathbf{u} \in W\}$.

Proposição 8. *Seja V um espaço vetorial sobre $\mathbb{K}$ munido de um produto interno. Sejam $W \subseteq V$ um subespaço e $\mathfrak{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ um conjunto gerador para W. Então $\mathbf{v} \in W^{\perp}$ se, e somente se $\langle \mathbf{v}, \mathbf{w}_i \rangle = 0$ para cada $i = 1, \ldots, k$.*

Demonstração: Seja $\mathbf{w} \in W$. Então existem $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, tais que $\mathbf{w} = \alpha_1 \mathbf{w}_1 + \ldots + \mathbf{w}_k \alpha_k$. Portanto,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{k} \overline{\alpha_i} \langle \mathbf{v}, \mathbf{w}_i \rangle.$$

Se assumirmos que $\langle \mathbf{v}, \mathbf{w}_i \rangle = 0$ para cada $1 \leq i \leq k$, segue que $\langle \mathbf{v}, \mathbf{w} \rangle = 0$, ou seja, $\mathbf{v} \in W^{\perp}$. Reciprocamente, se $\mathbf{v} \in W^{\perp}$, então $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ para cada $\mathbf{w} \in W$. Em particular, $\langle \mathbf{v}, \mathbf{w}_i \rangle = 0$ para cada $1 \leq i \leq k$. ■

Proposição 9. *Seja V um espaço vetorial de dimensão $n \geq 1$ e com produto interno e seja $W \subset V$ um subespaço próprio de V. Então $V = W \oplus W^{\perp}$.*

Demonstração: Precisamos mostrar que

i) $V = W + W^{\perp}$, isto é, cada vetor $\mathbf{v} \in V$ se escreve como uma soma $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$ com $\mathbf{w}_1 \in W$ e $\mathbf{w}_2 \in W^{\perp}$.

ii) $W \bigcap W^{\perp} = \{\mathbf{0}\}$

Se $W = 0$, não há nada a provar. Assuma que $W \neq 0$ e seja $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ uma base ortogonal de W. Considere uma base $\mathfrak{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$ de V, também ortogonal e contendo $\mathfrak{B}$. Segue de (6) que se $\mathbf{v} \in V$, então

$$\mathbf{v} = \sum_{i=1}^{n} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i = \sum_{i=1}^{m} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i + \sum_{i=m+1}^{n} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i$$

Obviamente, $\sum_{i=1}^{m} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i \in W$. Por outro lado, para cada $j = 1, \ldots, m$

$$\left\langle \mathbf{v}_j, \sum_{i=m+1}^{n} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i \right\rangle = \sum_{i=m+1}^{n} \frac{\overline{\langle \mathbf{v}, \mathbf{v}_i \rangle}}{\|\mathbf{v}_i\|^2} \langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$$

Assim, segue da proposição (8) que $\sum_{i=m+1}^{n} \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i \in W^{\perp}$. Logo, $V = W + W^{\perp}$ e isto prova (i).

Para provarmos (ii), seja $\mathbf{w} \in W \bigcap W^{\perp}$. Como $\mathbf{w} \in W^{\perp}$, temos $\langle \mathbf{w}, \mathbf{v} \rangle = 0$ para todo $\mathbf{v} \in W$. Em particular, $\langle \mathbf{w}, \mathbf{w} \rangle = 0$ e, portanto, $\mathbf{w} = 0$, como queríamos. ∎

Corolário 8. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ de dimensão finita com produto interno e seja W um subespaço de V. Então*

$$dim(V) = dim(W) + dim(W^{\perp}$$

4.4 Fatoração QR

Vamos mostrar agora um tipo de fatoração de matrizes, denominada fatoração **QR**. Esse tipo de fatoração é muito utilizado em softwares computacionais para resolver sistemas lineares, encontrar aproximações de mínimos quadrados e encontrar os autovalores de uma matriz. Esta últiplica aplicação da fatoraçao QR, mostraremos mais adiante nesta pesquisa.

Teorema 32. *Se* $\mathbf{A}$ *é uma matriz $m \times n$ com colunas linearmente independentes, então* $\mathbf{A}$ *pode ser fatorada como* $\mathbf{A} = QR$, *na qual* $\mathbf{Q}$ *é uma matriz $m \times n$ cujas colunas formam uma base ortonormal para o espaço coluna de* $\mathbf{A}$ *e* $\mathbf{R}$ *é uma matriz triangular superior inversível $n \times n$.*

Demonstração: Vamos representar por $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ os vetores coluna linearmente independentes de $\mathbf{A}$, que formam uma base para o espaço coluna de $\mathbf{A}$. Usando o processo de Gram-Schmidt, podemos obter uma base ortonormal $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ para o espaço coluna de $\mathbf{A}$.

Vamos relembrar como obtivemos essa base ortonormal. Primeiramente construímos uma base ortogonal $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ da seguinte maneira:

$$\mathbf{v}_1 = \mathbf{u}_1$$

e, para $i = 2,3,\ldots,n$, temos

$$\mathbf{v}_i = \mathbf{u}_i - \frac{\langle \mathbf{u}_i, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{u}_i, \mathbf{v}_{i-1} \rangle}{\|\mathbf{v}_{i-1}\|^2}\mathbf{v}_{i-1} \tag{4.2}$$

Finalmente, $\mathbf{w}_i = \frac{1}{\|\mathbf{v}_i\|}\mathbf{v}_i$ para $i = 1,2,\ldots,n$. Então, cada um dos vetores $\mathbf{u}_i$ pode ser escrito como uma combinação linear dos vetores $\mathbf{w}_1,\ldots,\mathbf{w}_n$:

$$\begin{aligned}
\mathbf{u}_1 &= r_{11}w_1 + r_{21}w_2 + \ldots + r_{n1}w_n \\
\mathbf{u}_2 &= r_{12}w_1 + r_{22}w_2 + \ldots + r_{n2}w_n \\
&\vdots \\
\mathbf{u}_n &= r_{1n}w_1 + r_{2n}w_2 + \ldots + r_{nn}w_n
\end{aligned} \tag{4.3}$$

Do teorema (31), temos

$$r_j i = \langle \mathbf{u}_i, \mathbf{w}_j \rangle$$

Além disso, da equação (4.2), vemos que $\mathbf{u}_i$ pertence ao espaço gerado por $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ que é o mesmo espaço gerado por $\{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$. Como $\mathbf{w}_j$ é ortogonal a $[\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_i]$ para $j > i$, $\mathbf{w}_j$ é ortogonal a $\mathbf{u}_i$. Portanto $r_{ij} = 0$ para $j > i$. Seja $\mathbf{Q}$ a matriz cujas colunas são $\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_n$. Seja

$$\mathbf{r}_j = \begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{pmatrix}$$

Então, as equações (4.3) podem ser escritas em forma matricial como

$$\mathbf{A} = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_n] = [\mathbf{Qr}_1\ \mathbf{Qr}_2\ \cdots\ \mathbf{Qr}_n] = \mathbf{QR}$$

em que $\mathbf{R}$ é a matriz cujas colunas são $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$. Logo,

$$\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \cdots & r_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

Vamos mostrar agora que $\mathbf{R}$ é inversível. Seja $\mathbf{x}$ uma solução do sistema linear $\mathbf{Rx} = \mathbf{0}$. Multiplicando essa equação a esquerda por $\mathbf{Q}$, obtemos

$$\mathbf{Q}(\mathbf{Rx}) = (\mathbf{QR})\mathbf{x} = \mathbf{Ax} = \mathbf{Q0} = \mathbf{0}$$

O sitema homogêneo $\mathbf{Ax} = \mathbf{0}$ pode ser escrito como

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \ldots + x_n\mathbf{u}_n = 0$$

em que $x_1, x_2, \ldots, x_n$ são as componentes do vetor $\mathbf{x}$. Como as colunas de $\mathbf{A}$ são linearmente independentes, temos

$$x_1 = x_2 = \ldots = x_n = 0,$$

de modo que $\mathbf{x}$ tem que ser o vetor nulo. Logo, $\mathbf{R}$ é inversível, o que finaliza a demonstração. ■

5 *Formas Canônicas*

Neste capítulo, T representará um operador linear sobre um $\mathbb{K}$-espaço vetorial n-dimensional V.

5.1 Autovalores e Autovetores

Nesta seção, consideraremos apenas matrizes em $\mathbb{C}^{nxn}$ e operadores sobre $\mathbb{C}$. Sabemos que uma função $T : V \to V$ definida por $T(\mathbf{v}) = A\mathbf{v}$ para $\mathbf{v} \in V$ é uma transformação linear. Dessa forma, poderíamos pensar na seguinte questão:

Dada uma transformação linear T de um espaço vetorial V nele mesmo, $T : V \to V$, quais vetores são levados em um múltiplo deles mesmos por esta transformação?

Definição 41. *Seja $T : V \to V$ um operador linear sobre o espaço vetorial V sobre o corpo $\mathbb{C}$. Então $\lambda \in \mathbb{C}$ é dito um autovalor de T se existe $\mathbf{v} \neq \mathbf{0}$ tal que $T(\mathbf{v}) = \lambda\mathbf{v}$. Todo vetor não nulo $\mathbf{v}$ satisfazendo a equação acima é chamado de um autovetor de T associado ao autovalor λ. Autovalores são também chamados de valores próprios ou valores característicos. Nesses casos, os autovetores são denominados vetores próprios ou vetores característicos, respectivamente.*

Lembremos que dada uma matriz quadrada $\mathbf{A}$ com entradas complexas podemos definir um operador $T : \mathbb{C}^n \to \mathbb{C}^n$ via

$$T\mathbf{v} = \mathbf{A}\mathbf{v}$$

onde $\mathbf{v}$ é um vetor em $\mathbb{C}^n$.

Podemos pois, repetindo a definição acima para operadores, definir autovalores e autovetores para matrizes. Além disso, como as matrizes que representam um mesmo operador em bases diferentes são semelhantes segue que os autovalores de um operador coincidem com os autovalores de qualquer matriz que o represente. Por outro lado usando a representação matricial de um operador fazemos o seguinte:

Se λ autovalor de $\mathbf{A}$ então existe $\mathbf{v} \neq 0$ tal que $\mathbf{Av} = \lambda \mathbf{v}$, isto é

$$\mathbf{Av} = \lambda \mathbf{Iv} \Leftrightarrow \mathbf{Av} - \lambda \mathbf{Iv} = 0 \Leftrightarrow [\mathbf{A} - \lambda \mathbf{I}]\, \mathbf{v} = 0,$$

e portanto concluímos que λ é autovalor de $\mathbf{A}$ se e somente se o operador $\mathbf{A} - \lambda \mathbf{I}$ não é inversível. Em particular

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

Segue que os autovalores de $\mathbf{A}$ são as raízes complexas do polinômio

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

que se denomina de polinômio característico de $\mathbf{A}$. É fácil ver que matrizes semelhantes tem o mesmo polinômio característico e assim podemos definir o polinômio característico de um operador T como o polinômio característico de qualquer matriz que o represente.

Definição 42. *Se $T : V \rightarrow V$ é um operador linear sobre o espaço vetorial V, denotamos por $V_(\lambda)$ o subespaço de V gerado por todos os autovetores associados a λ. Da mesma forma, sendo $\mathbf{A}$ uma matriz $n \times n$, denotamos o subespaço gerado pelos autovetores de $\mathbf{A}$ associados a λ, por V_λ.*

Assim:

$$V_\lambda = \{\mathbf{v} \in V \,|\, T(\mathbf{v}) = \lambda \mathbf{v}\} = N(T - \lambda I).$$

$$V_\lambda = \{\mathbf{v} \in \mathbb{C}^n \,|\, \mathbf{Av} = \lambda \mathbf{v}\} = N(\mathbf{A} - \lambda \mathbf{I}).$$

Conforme resulta da definição clássica de determinante, $p(\lambda) = det(A - \lambda I)$ é um polinômio de grau n em λ, cujo termo de maior grau é $(-1)^n \lambda^n$ e, pelo Teorema Fundamental da Álgebra, possuirá n raízes complexas. Segue que um operador sobre um espeço vetorial de dimensão n tem no máximo n autovalores distintos. Quanto aos autovetores temos o seguinte resultado:

Teorema 33. *Seja $T : V \rightarrow V$ um operador linear e $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ os autovetores correspondentes aos autovalores $\{\lambda_1, \cdots, \lambda_n\}$ de T. Se $\lambda_i \neq \lambda_j$ para $i \neq j$, então o conjunto $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ é linearmente independente.*

Demonstração: Dado o operador linear $T : V \rightarrow V$, sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vetores não nulos em V tais que $T(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1, \ldots, T(\mathbf{v}_n) = \lambda_n \mathbf{v}_n$, em que os números reais $\lambda_1, \ldots, \lambda_n$ são dois a dois diferentes. Provaremos, por indução, que $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ é um conjunto linearmente independente.

A afirmação é óbvia quando $n = 1$. Supondo-a verdadeira para $n-1$ vetores, inferiremos daí sua validez para n. Dada a combinação linear nula

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = 0 \tag{5.1}$$

aplicaremos o operador T a ambos os membros da igualdade, levando em conta o fato que $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$. Resulta então que

$$\lambda_1\alpha_1\mathbf{v}_1 + \ldots + \lambda_n\alpha_n\mathbf{v}_n = 0 \tag{5.2}$$

Multiplicando a igualdade (5.1) por λ_n e subtraindo de (5.2) vem:

$$(\lambda_1 - \lambda_n)\alpha_1\mathbf{v}_1 + \ldots + (\lambda_{n-1} - \lambda_n)\alpha_{n-1}\mathbf{v}_{n-1} = 0$$

Pela hipótese de indução, os vetores $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ são independentes. Logo

$$(\lambda_1 - \lambda_n)\alpha_1 = \ldots = (\lambda_{n-1} - \lambda_n)\alpha_{n-1} = 0.$$

Como os autovalores são todos diferentes, as $n-1$ diferenças nos parênteses acima são diferentes de zero, logo $\alpha_1 = \ldots = \alpha_{n-1} = 0$. Isto reduz a igualdade (5.1) a $\alpha_n\mathbf{v}_n = 0$. Como $\mathbf{v}_n \neq \mathbf{0}$, segue-se que $\alpha_n = 0$. Assim, a igualdade (5.1) só pode ocorrer quando todo os coeficientes α_i forem nulos, o que prova o teorema ■

Corolário 9. *Seja V um espaço vetorial de dimensão n. Se um operador linear $T : V \to V$ possui n autovalores diferentes, então existe uma base $\{\mathbf{v}_1, \cdots, v_n\} \subset V$ em relacão a qual a matriz de T é diagonal.*

Com efeito, se $T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1, \cdots, T(\mathbf{v}_n) = \lambda_n\mathbf{v}_n$ com cada $\mathbf{v}_i \neq \mathbf{0}$ e os λ_i dois a dois distintos, então $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ é, por consequência do teorema 33, uma base de V. A matriz de T nesta base é dada por

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Vamos agora estudar um pouco mais sobre os polinômios que se anulam em um operador linear T, cujas raízes nos trazem importantes informações sobre o comportamento de T.

Definição 43. *O polinômio minimal (ou polinônio mínimo) de um operador linear T em $L(V,V)$ é o polinômio mônico $m_T(x)$ de menor grau, tal que $m_T(T)(\mathbf{v}) = 0$, $\forall\ \mathbf{v} \in V$*

Já definimos dois polinômios relacionados a uma transformação linear T, a saber, o polinômio característico $p_T(x)$ e o polinômio minimal $m_T(x)$. Vamso olhas melhor as relações entre eles. iremos mostrar inicialmente que $p_T(x)$ se anula em T e, portanto, $m_T(x)$ é um divisor de $p_T(x)$. Também iremos mostrar que eles possuem as mesmas raízes. Comecemos inicialmente enunciando o chamado Teorema de Cayley-Hamilton.

Teorema 34. ***(Cayley-Hamilton)*** *Seja T um operador linear, $T \in L(V,V)$ e $p_T(x)$ seu polinômio característico. Então $p_T(T) = 0$.*

Demonstração: A prova deste teorema pode ser encontrada em [8]. ■

O que faremos agora é relacionar as raízes do polinômio característico e minimal.

Proposição 10. *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão $n \geq 1$ e $T \in L(V,V)$. Então, os polinômios característico e minimal de T têm as mesmas raízes a menos de multiplicidade.*

Demonstração: Sejam $p_T(x)$ e $m_T(x)$ os polinômios característico e minimal de T, respectivamente, e seja $\lambda \in \mathbb{K}$. Precisamos mostrar que $p_T(\lambda) = 0$ se e somente se $m_T(\lambda) = 0$. Suponha inicialmente que $p_T(\lambda) = 0$, isto é, que λ seja um autovalor de T. Então existe $\mathbf{0} \neq \mathbf{v} \in V$ tal que $T(\mathbf{v}) = \lambda\mathbf{v}$.

Se $T : V \to V$ for um operador linear então já vimos que, para cada $i \geq 0$, $T^i = \underbrace{T \ldots T}_{i \text{ vezes}}$ também será um operador em $L(V,V)$. Por outro lado, observe que, para cada $i \geq 1$, temos $T^i(\mathbf{v}) = \lambda^i\mathbf{v}$. Agora, se escrevermos $m_t(x) = \sum_{i=0}^{m} a_i x^i$, teremos então que

$$0 = m_T(T)(\mathbf{v}) = \left(\sum_{i=0}^{m} a_i T^i\right)(\mathbf{v}) = \left(\sum_{i=0}^{m} a_i \lambda^i\right)\mathbf{v}$$

e, portanto, $m_T(\lambda) = \sum_{i=0}^{m} a_i \lambda^i = 0$ pois $\mathbf{v} \neq \mathbf{0}$. logo, λ é uma raiz de $m_T(x)$. Suponhamos agora que $m_T(\lambda) = 0$. Então, $m_T(x) = (x-\lambda)q(x)$. Pela condição de minimalidade no grau do polinômio m_T, segue que $q(T) \neq 0$ e, portanto, existe $\mathbf{u} \in V$ tal que $q(T)(\mathbf{u}) \neq 0$. Se denotamos $\mathbf{v} = q(T)(\mathbf{u})$, teremos então

$$0 = m_T(T)(\mathbf{u}) = (T - \lambda I)(q(T)(\mathbf{u})) = (T - \lambda I)(\mathbf{v})$$

e, portanto $\mathbf{v}$ é um autovetor de T associado ao autovalor λ. Logo, $p_T(\lambda) = 0$ e o resultado está demonstrado. ■

5.2 Operadores Diagonalizáveis

Definição 44. *Seja T um operador linear sobre o espaço V de dimensão finita. Dizemos que T é diagonalizável se existe uma base de V formada por vetores característicos de T.*

A razão para o nome deveria ser evidente; de fato, se existe uma base ordenada $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V na qual cada $\mathbf{v}_i$ é um vetor característico de T, então a matriz T em relação à base ordenada $\mathfrak{B}$ é diagonal. Se $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$, então

$$[T]_{\mathfrak{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Definição 45. *Uma matriz* $\mathbf{A}$ *é diagonalizável se existe uma matriz inversível* $\mathbf{P}$ *tal que*

$$\mathbf{AP} = \mathbf{PD}$$

em que $\mathbf{D}$ *é uma matriz diagonal.*

É fácil concluir que uma matriz $\mathbf{A}$ é diagonalizável se e só se cada coluna j de $\mathbf{P}$ é autovetor de $\mathbf{A}$ associado ao autovalor d_{ii}. Ou seja, $\mathbf{A}$ é diagonalizável se e só se existe uma base de $\mathbb{C}^n$ formada por autovetores de $\mathbf{A}$. Daí, podemos formalizar a seguinte definição:

Teorema 35. *Seja T um operador linear sobre o espaço V de dimensão finita e sejam $\lambda_1, \cdots, \lambda_k$ os autovalores de T, dois a dois distintos. Então*

i) *Se $\mathbf{v}_1 + \cdots + \mathbf{v}_k = 0$ com $\mathbf{v}_i \in V_{\lambda_i}$, $i = 1, \cdots, k$, então $\mathbf{v}_i = 0$ para cada i.*

ii) *Para cada $i = 1, \cdots, k$, seja β_i um conjunto linearmente independente contido em V_{λ_i}. Então $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ é linearmente independente.*

Demonstração:

i) Mostremos este item pelo princípio da indução em $k \geq 1$.

Para $k = 1$, o caso é trivial. Seja agora $k > 1$ e suponha que nossa hipótese de indução vale pata todo $j < k$. Vamos mostrar que esse resultado vale também para $j = k$. Seja

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k = 0 \tag{5.3}$$

com $\mathbf{v}_i \in V_{\lambda_i}$, $\forall i = 1, \cdots, k$. Calculando T em (5.3), teremos que

$$0 = T(0) = T(\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k) = T(\mathbf{v}_1) + T(\mathbf{v}_2) + \cdots + T(\mathbf{v}_k)$$

e portanto,

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_k \mathbf{v}_k = 0 \tag{5.4}$$

Multiplicando a equação (5.3) por λ_1 e subtraindo (5.4), temos

$$(\lambda_1 - \lambda_2)\mathbf{v}_2 + \cdots + (\lambda_1 - \lambda_k)\mathbf{v}_k = 0$$

Usando-se a hipótese de indução, teremos que $(\lambda_1 - \lambda_i)\mathbf{v}_i = 0$ para cada $i = 2, \cdots, k$. Como $\lambda_1 \neq \lambda_i$ se $i \neq 1$, temos que $\mathbf{v}_i = 0$, para todo $i = 2, \cdots, k$. Substituindo em (5.3), teremos que também $\mathbf{v}_1 = 0$, como queríamos.

ii) Para cada $i = 1, \cdots, k$, considere o conjunto $\beta_i = \{\mathbf{v}_{i1}, \cdots, \mathbf{v}_{in_i}\}$. Vamos mostrar que

$$\{\mathbf{v}_{11}, \cdots, \mathbf{v}_{1n_1}, \cdots, \mathbf{v}_{k1}, \cdots, \mathbf{v}_{kn_k}\}$$

é linearmente independente. Sendo assim, assuma que

$$\alpha_{11}\mathbf{v}_{11} + \cdots + \alpha_{1n_1}\mathbf{v}_{1n_1} + \cdots + \alpha_{k1}\mathbf{v}_{k1} + \cdots + \alpha_{kn_k}\mathbf{v}_{kn_k} = 0$$

onde $\alpha_{ij} \in \mathbb{C}$, para todos i e j. Como $\sum_{j=1}^{n_i} \alpha_{ij}\mathbf{v}_{ij} \in V_{\lambda_i}$ para cada i, segue do item (i) que $\sum_{j=1}^{n_i} \alpha_{ij}\mathbf{v}_{ij} = 0$, $\forall i = 1, \cdots, k$. Como β_i é um conjunto linearmente independente, teremos que $\alpha_{ij} = 0$ para todo $i = 1, \cdots, k$ e todo $j = 1, \cdots, n_i$ e, portanto, $\beta_1 \cup \cdots \cup \beta_k$ é linearmente independente.

■

Corolário 10. *Seja $T : V \to V$ um operador linear, onde V é um espaço vetorial de dimensão finita. Se $\lambda_i, \cdots, \lambda_k$ são todos os autovalores de T, então T é diagonalizável se e somente se $dim(V) = \sum_{i=1}^{k} dim(V(\lambda_i))$.*

Teorema 36. *Se o operador linear $T : V \to V$ possuir todos os autovalores distintos, então T é diagonalizável.*

Demonstração: Supondo que $dim(V) = n$, sejam $\mathbf{w}_1, \cdots, \mathbf{w}_n$ os autovetores de T correspondentes aos seus autovalores, dados por $\lambda_1, \cdots, \lambda_n$. Seja W a matriz cujas colunas são os autovetores $\mathbf{w}_1, \cdots, \mathbf{w}_n$. Com um rápido cálculo podemos ver que, sendo $(a_{ij})_{nxn}$ a matriz que representa T em uma certa base, teremos:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \\ \\ \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ou seja,

$$\begin{pmatrix} \\ T(\mathbf{w}_1) & T(\mathbf{w}_2) & \cdots & T(\mathbf{w}_n) \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \lambda_1\mathbf{w}_1 & \lambda_2\mathbf{w}_2 & \cdots & \lambda_n\mathbf{w}_n \\ \\ \end{pmatrix}$$

Sabemos que a matriz W é inversível, pois como os autovalores λ_i são distintos para todo i, os autovetores $\mathbf{w}_i$ são linearmente independentes. Logo,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \\ \\ \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_n \\ \\ \end{pmatrix}^{-1}$$

Portanto, o operador T é diagonalizável.

∎

Se um operador T possui polinômio característico com raízes repetidas, então T pode ou não ser diagonalizável. Neste caso, pode-se escrever o polinômio característico de T como um produto de n fatores da seguinte forma

$$(\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_r)^{k_r},$$

no qual λ_i, $i = 1, \cdots, r$ são os autovalores distintos de T e $k_1, \cdots, k_r$ são naturais cuja soma é n.

Definição 46. *Seja λ_0 um autovalor de um operador $T : V \to V$. A multiplicidade geométrica de λ_0 é a dimensão do subespaço vetorial $V_{\lambda_0} = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda_0\mathbf{v}\}$. A multiplicade algébrica de λ_0 é sua multiplicidade como raiz do polinômio característico de T, isto é, o maior inteiro m tal que $p_T(\lambda) = (\lambda_0 - \lambda)^m \cdot q(\lambda)$, sendo $q(\lambda)$ ainda um polinômio.*

Proposição 11. *Seja λ um autovalor do operador $T : V \to V$, em que V é um espaço vetorial de dimensão finita. Então λ possui multiplicidade geométrica menor ou igual à sua multiplicidade algébrica.*

Demonstração: Seja $W = V_\lambda$ e assuma que $dim(W) = s$. Sejam $\mathfrak{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$ uma base de W e $\mathfrak{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_s, \mathbf{w}_{s+1}, \ldots, \mathbf{w}_n\}$ uma base de V contendo $\mathfrak{B}'$. Como $T(\mathbf{w}_i) = \lambda\mathbf{w}_i$

para $i = 1, \ldots, s$, podemos escrever $[T]_{\mathfrak{B}}$ na forma de blocos

$$[T]_{\mathfrak{B}} = \begin{pmatrix} \lambda & 0 & \cdots & 0 & \\ 0 & \lambda & \cdots & 0 & \mathbf{A}_1 \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & \lambda & \\ & & & & \\ 0 & 0 & \cdots & 0 & \\ \vdots & \vdots & & \vdots & \mathbf{A}_2 \\ 0 & 0 & \cdots & 0 & \end{pmatrix}$$

em que $\mathbf{A}_1 \in \mathbb{K}^{s \times n-s}$ e $\mathbf{A}_2 \in \mathbb{K}^{(n-s)\times(n-s)}$. Um cálculo simples nos dá

$$p_T(x) = det(xI - [T]_{\mathfrak{B}}) = (x-\lambda)^s \cdot det(x\mathbf{I} - \mathbf{A}_2).$$

Por definição, temos que $ma(\lambda)$ é o maior índice j tal que $(x-\lambda)^j$ divide $p_T(x)$. Portanto, $mg(\lambda) = s \leq ma(\lambda)$, como queríamos. ■

5.3 Subespaços T-invariantes

Dado um operador linear $T : V \to V$, por vezes é conveniente considerarmos a sua restrição a algum subespaço dado W de V. No entanto, nem sempre a imagem desta restrição está contida no próprio subespaço. Isto nos leva à seguinte definição:

Definição 47. *Seja T um operador linear sobre o espaço vetorial V e seja $W \subseteq V$ um subespaço de V. Dizemos que W é um subespaço T-invariante de V se $T(\mathbf{w}) \in W$ para todo $\mathbf{w} \in W$.*

Observação 10. *Se T é um operador linear sobre V, então*

i) *os subespaços $N(T)$ e $Im(T)$ são T-invariantes;*

ii) *se λ for um autovalor de T, então $Aut_T(\lambda)$ é um subespaço T-invariante de V. De fato, se $\mathbf{v} \in V_\lambda$, então $T(\mathbf{v}) = \lambda\mathbf{v} \in V_\lambda$;*

iii) *se W é um subespaço T-invariante, então a restrição de T a W é um operador linear em $L(W,W)$.*

Exemplo 11. *Seja $T : \mathbb{C}^3 \to \mathbb{C}^3$ definida por $T(x,y,z) = (0,x,y)$. Se $W = [e_1, e_2]$, então $T(W) = [e_2, e_3]$ e assim W não é um subespaço T-invariante de $\mathbb{C}^3$. Se $W' = [e_2, e_3]$, teremos então que $T(W') = [e_3] \subseteq W'$ e segue assim que W' é um subespaço T-invariante de V.*

Seja T um operador linear sobre o $\mathbb{K}$-espaço vetorial V de dimensão $n \geq 1$ e $W \subset V$ um subespaço T-invariante de V de dimensão m com $1 \leq m < n$. Considere $\mathfrak{B}'$ uma base de W e estenda-a a uma base $\mathfrak{B}$ de V. Como observado acima, a restrição de T a W, isto é, a função $T' : W \to W$ dada por $T'(\mathbf{w}) = T(\mathbf{w})$, $\forall \mathbf{w} \in W$, é um operador linear. Daí, a matriz $[T]_{\mathfrak{B}}$ é escrita da seguinte maneira

$$[T]_{\mathfrak{B}} \quad = \quad \begin{pmatrix} [T']_{\mathfrak{B}'} & \mathbf{A} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$$

onde $\mathbf{0}$ indica a matriz nula em $\mathbb{M}_{(n-m)\times m}(\mathbb{K})$, $\mathbf{A} \in \mathbb{M}_{m\times(n-m)}(\mathbb{K})$ e $\mathbf{B} \in \mathbb{M}_{(n-m)\times(n-m)}(\mathbb{K})$. Muitas vezes, pode-se escrever o espaço vetorial V como a soma direta de dois (ou mais) subespaços T-invariantes e, como neste caso a restrição de T a cada um destes subespaços é um operador linear, podemos descrever a matriz de T usando os blocos das matrizes destas restrições. Sendo mais específico, seja $V = W_1 \oplus \ldots \oplus W_r$ e suponha que cada subespaço W_i seja invariante. Sejam $\mathfrak{B}_1, \ldots, \mathfrak{B}_r$ bases de $W_1, \ldots, W_r$, respectivamente. Como a soma $W_1 + \ldots + W_r$ é direta, segue que $\mathfrak{B} = \mathfrak{B}_1 \bigcup \ldots \bigcup \mathfrak{B}_r$ é uma base de V. Não é difícil ver então que a matriz $[T]_{\mathfrak{B}}$ tem a seguinte forma:

$$[T]_{\mathfrak{B}} \quad = \quad \begin{pmatrix} [T]_{\mathfrak{B}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & [T]_{\mathfrak{B}} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & [T]_{\mathfrak{B}} \end{pmatrix}$$

em que cada T_i indica a restrição de T ao subespaço W_i e os números $\mathbf{0}$ indicam as matrizes nulas correspondentes.

Neste caso, a descrição de $[T]_{\mathfrak{B}}$ será reduzida à descrição das matrizes $[T_1]_{\mathfrak{B}_1}, \ldots, [T_r]_{\mathfrak{B}_r}$. Com isso, também escrevemos $T = T_1 \oplus \ldots \oplus T_r$ e dizemos que o operador T é a soma direta dos operadores $T_1, \ldots, T_r$.

5.4 Decomposição de Schur

Teorema 37. ***(Teorema de Schur)*** *Se* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *então existe uma matriz unitária* $\mathbf{U}$ *e existe uma matriz triangular superior* $\mathbf{S}$ *tais que* $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^H$.

Demonstração: Esste teorema é obviamente válido para $n = 1$. Vamos supor, por indução, que o teorema é válido para todo k, $1 \leq k < n$. Seja $\mathbf{A}$ uma matriz complexa de ordem n. Então $\mathbf{A}$ tem um autovetor $\mathbf{v}$, $\|\mathbf{v}\|_2 = 1$, associado a um autovalor λ. Seja $\mathbf{U}_0$ uma matriz unitária tal que a primeira coluna seja o vetor $\mathbf{v}$. Então $\mathbf{S}_0 = \mathbf{U}_0^H \mathbf{A} \mathbf{U}_0$ é uma matriz tal que $\mathbf{S}_0(2:n,1) = 0$. Pela hipótese de indução aplicada à matriz $\mathbf{A}_1 = \mathbf{S}_0(2:n,2:n)$, existe uma matriz unitária $\mathbf{M}$ de ordem $n-1$ e existe uma matriz triangular $\mathbf{T}_1$ tais que $\mathbf{T}_1 = \mathbf{M}^H \mathbf{A}_1 \mathbf{M}$. Seja

$$\mathbf{U}_1 = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{M} \end{pmatrix}.$$

Note que $\mathbf{U}_1$ é unitária. Seja $\mathbf{U} = \mathbf{U}_0 \mathbf{U}_1$. Então

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \begin{pmatrix} \lambda & w^{\mathbf{S}} \\ 0 & \mathbf{S}_1 \end{pmatrix},$$

para algum vetor $w \in \mathbb{C}^{n-1}$ ■

Uma consequência do Teorema de Schur é a seguinte: os autovalores dependem continuamente das entradas da matriz. Para provar esse fato, considere uma sequência $\{\mathbf{A}_k\}$ de matrizes que convergem para $\mathbf{A}$ quando $k \to \infty$. Seja

$$\mathbf{A}_k = \mathbf{Q}_k \mathbf{S}_k \mathbf{Q}_k^H,$$

em que $\mathbf{Q}_k$ é unitária e $\mathbf{S}_k$, triangular superior (forma de Schur de $\mathbf{A}_k$. O conjunto das matrizes unitárias é um conjunto compacto de $\mathbb{C}^{n\times n}$.Portanto, existe uma subsequência $\{\mathbf{Q}_{k_i}\}$ de $\{\mathbf{Q}_k\}$ tal que $\mathbf{Q}_{k_i}$ converge a uma matriz unitária $\mathbf{Q}$. Por conseguinte,

$$\lim_{k\to\infty} \mathbf{Q}_k^H \mathbf{A}_k \mathbf{Q}_k = \lim_{k_i\to\infty} \mathbf{Q}_{k_i}^H \mathbf{A}_{k_i} \mathbf{Q}_{k_i} = \mathbf{Q}^H \mathbf{A} \mathbf{Q}$$

Como para todo k $\mathbf{Q}_k^H \mathbf{A}_k \mathbf{Q}_k$ é triangular superior, $\mathbf{Q}^H \mathbf{A} Q$ só pode ser triangular superior. Conclusão: os autovalores de $\mathbf{A}_k$, que são as entradas diagonais de $\mathbf{S}_k$, convergem para os autovalores de $\mathbf{A}$, que estão todos na diagonal de $\mathbf{Q}^H \mathbf{A} \mathbf{Q}$.

5.5 Forma de Jordan

Nesta seção vamos obter a forma canônica de Jordan de um operador. No sentido de entendermos com mais profundidade a ação de um operador linear, procuraremos decompô-lo em operadores de um tipo particular bem simples. Seja T um operador linear sobre um espaço vetorial V, complexo de dimensão $n \geq 1$. Seja

$$p_T(t) = \prod_{i=1}^{k} (t - \lambda_i)^{n_i}$$

a fatoração do polinômio característico de T onde $\lambda_1,\ldots,\lambda_k$ são os autovalores distintos de T, e o inteiro $n_i \geq 1$ é a multiplicidade de λ_i com $i = 1..k$. Denotamos o autoespaço de T associado ao autovalor λ_i por

$$V_{\lambda_i}(T) = N(T - \lambda_i I).$$

Definição 48 (Autoespaços Generalizados)**.** *Definimos o* autoespaço generalizado *de T associado ao autovalor λ_i como o subespaço*

$$V_{[\lambda_i]}(T) = N((T - \lambda_i)^{n_i}) \subset V.$$

Para simplificar a notação vamos denotar o autoespaço generalizado associado a λ_i por V_i. É fácil ver que este subespaço é um subespaço invariante por T. O primeiro terorema no caminho de nossa forma canônica do operador se chama de teorema da decomposição primária que enunciamos agora e cuja prova colocamos no apêndice.

Teorema 38 (Decomposição Primária)**.** *Seja T um operador sobre V, espaço complexo.Então V é a soma direta dos subespaços generalizados de T. A dimensão de cada autoespaço generalizado coincide com a multiplicidade algébrica de correspondente autovalor.*

$$V = E(T,\lambda_1) \oplus \cdots \oplus E(T,\lambda_k).$$

Como consequência deste teorema vamos provar que T pode ser expresso como uma soma direta

$$T = T_1 \oplus \cdots \oplus T_k \qquad \text{com } T_i \in L(V_i).$$

Cada T_i pode ser expresso como uma soma

$$T_i = S_i + N_i; \qquad S_i, N_i \in L(V_i) \quad i = 1..k,$$

onde S_i é um operador diagonalizável e N_i é um operador nilpotente. Além disso $S_i N_i = N_i S_i$ para $i = 1..n$.

Suponhamos a princípio que T tenha apenas um autovalor λ de multiplicidade $n = \dim V$. Segue do teorema que $V = V_\lambda$. Seja então

$$N = T - \lambda I, \quad S = \lambda I.$$

É imediato que $T = S + N$, $SN = NS$, S é diagonal em qualquer base e N é nilpotente pois

$$N(N^n) = V_\lambda = V.$$

Para o caso geral fazemos

$$T_i = T|V_{\lambda_i}.$$

Então $T = T_1 \oplus \cdots \oplus T_k$. Como cada T_i tem apenas um autovalor λ_i, podemos pelo raciocínio anterior, escrever

$$T_i = S_i + N_i \qquad S_i, N_i \in L(V_i),$$

onde $S_i = \lambda_i I$ em V_i, e $N_i = T_i - S_i$ operador nilpotente de ordem n_i para $i = 1..k$. Segue que

$$T = S + N,$$

onde

$$S = S_1 \oplus \cdots \oplus S_k$$

$$N = N_1 \oplus \cdots \oplus N_k.$$

é imediato que $SN = NS$ e que N é nilpotente e S diagonalizável. Provamos assim o seguinte resultado.

Teorema 39. *Seja $T : V \to V$ um operador linear sobre o corpo dos complexos. Então*

$$T = S + N,$$

onde $SN = NS$, S é um operador diagonalizável e N é um operador nilpotente.

Definição 49. *Um bloco de Jordan elementar de ordem m associado a λ é uma matriz quadrada $m \times m$ da forma*

$$J_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

O teorema que provaremos no apêndice afirma que:

Teorema 40. *Seja o operador linear $T : V \to V$ onde V é um espaço complexo. Então se*

$$p_T(t) = \prod_{i=1}^{k} (t - \lambda_i)^{n_i}$$

é o polinômio característico de T, existe uma base de V na qual T pode se representado por *uma matriz de blocos de Jordan da forma*

$$A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{bmatrix}$$

onde cada J_i para $i = 1..k$ é um bloco de Jordan de ordem $n_i \times n_i$ associado ao autovalor λ_i.

Observe que para provarmos o teorema e sabendo que $T = S + N$, S diagonalizável e N nilpotente com $SN = NS$, basta obtermos uma forma de Jordan para os operadores nilpotentes, isto é uma base na qual S se representa por uma matriz diagonal e N por uma matriz de blocos em que cada bloco é da da forma:

$$N = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

o que deixaremos para o apêndice.

6 Adjuntos

6.1 Funcionais Lineares e Adjuntos

Definição 50. *Se V é um espaço vetorial sobre o corpo $\mathbb{K}$, um funcional linear sobre V é uma transformação linear de V em $\mathbb{K}$.*

Sejam V um espaço vetorial sobre o corpo $\mathbb{K}$ com produto interno e $\mathbf{w} \in V$. Podemos, a partir de $\mathbf{w}$, definir um funcional linear em V^H da seguinte maneira:

$$\begin{array}{rcl} f_{\mathbf{w}} : V & \to & \mathbb{K} \\ \mathbf{u} & \mapsto & f_{\mathbf{w}}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{v} \rangle. \end{array}$$

A linearidade de $f_{\mathbf{w}}$ é garantida pelas propriedades (4.2) e (4.2) da definição de produto interno. Podemos perguntar se vale a recíproca da observação acima, isto é, de dado um funcional linear $f \in V^H$, existe um vetor $\mathbf{w} \in V$ tal que $f = f_{\mathbf{w}}$. A seguir apresentaremos um resultado que responde afirmativamente a esta pergunta.

Sejam V um $\mathbb{K}$ espaço vetorial de dimensão finita $n \geq 1$ e $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ uma base ortonormal de V. Considere o elemento $\mathbf{w} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \in V$ e calculemos $f_{\mathbf{w}}$ como definida acima em um vetor $\mathbf{v}_j$ da base de $\mathfrak{B}$:

$$f_{\mathbf{w}}(\mathbf{v}_j) = \langle \mathbf{v}_j, \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \rangle = \sum_{i=1}^{n} \overline{\alpha_i} \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \overline{\alpha_i}.$$

Logo $w = \sum_{i=1}^{n} \overline{f_{\mathbf{w}}(\mathbf{v}_j)} \mathbf{v}_j$. Esta conta será útil na demonstração do próximo resultado.

Proposição 12. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita $n \geq 1$. Se $f \in V^H$, então existe um único $\mathbf{w} \in V$ tal que $f(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w} \rangle$ para todo $\mathbf{u} \in V$.*

Demonstração: Segue do Teorema (30) que existe uma base ortonormal $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V. Considere o elemento $\mathbf{w} = \sum_{j=1}^{n} \overline{f(\mathbf{v}_j)} \mathbf{v}_j$. Vamos mostrar que $f = f_{\mathbf{w}}$, isto é, que $f(\mathbf{u}) = f_{\mathbf{w}}(\mathbf{u}) =$

$\langle \mathbf{u}, \mathbf{w} \rangle$, para todo $\mathbf{u} \in V$. De fato, para $k = 1, \ldots, n$, temos

$$f_{\mathbf{w}}(\mathbf{v}_k) = \langle \mathbf{v}_k, \mathbf{w} \rangle = \left\langle \mathbf{v}_k, \sum_{j=1}^{n} \overline{f(\mathbf{v}_j)} \mathbf{v}_j \right\rangle = \sum_{j=1}^{n} \overline{\overline{f(\mathbf{v}_j)}} \langle \mathbf{v}_k, \mathbf{v}_j \rangle = f(\mathbf{v}_k)$$

e, portanto, como f e $f_{\mathbf{w}}$ coincidem nos elementos de uma base, segue que $f = f_{\mathbf{w}}$, como queríamos.

Para mostrar a unicidade, considere $\mathbf{w}_1, \mathbf{w}_2 \in V$ tais que $f = f_{\mathbf{w}1} = f_{\mathbf{w}2}$, isto é, tais que $f(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w}_1 \rangle$ e $f(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w}_2 \rangle$ para todo $\mathbf{u} \in V$. Então $\langle \mathbf{u}, \mathbf{w}_1 \rangle = \langle \mathbf{u}, \mathbf{w}_2 \rangle$ para todo $\mathbf{u} \in V$. Assim, $\langle \mathbf{u}, \mathbf{w}_1 - \mathbf{w}_2 \rangle = 0$, $\forall\, \mathbf{u} \in V$. Segue da definição de produto interno que $\mathbf{w}_1 - \mathbf{w}_2 = 0$ e a unicidade está provada. ■

Observação 11. *Quando a dimensão de V for infinita, a proposição acima não é verdadeira de um modo geral. Convém, no entanto, observar que, no contexto da teoria de Análise Funcional, teremos que a proposição acima é verdadeira para todos os funcionais lineares contínuos sobre os denominados espaços de Hilbert de dimensão infinita. Tal resultado é conhecido como Teorema de Riesz e sua demonstração pode ser encontrada em qualquer texto básico de Análise Funcional.*

Teorema 41. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita. Se $T \in L(V,V)$, então existe um único operador $T^H \in L(V,V)$ tal que $\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^H(\mathbf{v}) \rangle$ para todos $\mathbf{u}, \mathbf{v} \in V$.*

Demonstração: Seja $\mathbf{v} \in V$. Queremos definir $T^H(\mathbf{v}) \in V$. para tanto, vamos considerar o seguinte funcional linear

$$\begin{array}{rcl} f_{\mathbf{w}} : V & \to & \mathbb{K} \\ \mathbf{u} & \mapsto & f(\mathbf{u}) = \langle T(\mathbf{u}), \mathbf{v} \rangle. \end{array}$$

Observe que f é linear, pois se $\mathbf{u}_1, \mathbf{u}_2 \in V$ e $\alpha \in \mathbb{K}$, então

$$\begin{array}{rcl} f(\mathbf{u}_1 + \alpha \mathbf{u}_2) & = & \langle T(\mathbf{u}_1 + \alpha \mathbf{u}_2), \mathbf{v} \rangle \\ & = & \langle T(\mathbf{u}_1) + \alpha T(\mathbf{u}_2), \mathbf{v} \rangle \\ & = & \langle T(\mathbf{u}_1), \mathbf{v} \rangle + \alpha \langle T(\mathbf{u}_2), \mathbf{v} \rangle \\ & = & f(\mathbf{u}_1) + \alpha f(\mathbf{u}_2) \end{array}$$

Pela proposição (12) sabemos que existe um único $\mathbf{w} \in V$ tal que $f(\mathbf{u}) = \langle \mathbf{u}, \mathbf{w} \rangle$, para todo $\mathbf{u} \in V$, isto é, tal que $\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle$, para todo $\mathbf{u} \in V$. Como $\mathbf{w}$ é determinado de modo único por

$\mathbf{v}$, definimos $T^H(\mathbf{v}) = \mathbf{w}$. Por construção, teremos então

$$\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^H(\mathbf{v}) \rangle, \quad \forall\, \mathbf{u}, \mathbf{v} \in V$$

Para mostrarmos que T^H definida acima é linear, considere vetores $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2 \in V$ e um escalar $\alpha \in \mathbb{K}$. Então

$$\begin{aligned}
\langle \mathbf{u}, T^H(\mathbf{v}_1 + \alpha \mathbf{v}_2) \rangle &= \langle T(\mathbf{u}), \mathbf{v}_1 + \alpha \mathbf{v}_2 \rangle \\
&= \langle T(\mathbf{u}), \mathbf{v}_1 \rangle + \overline{\alpha} \langle T(\mathbf{u}), \mathbf{v}_2 \rangle \\
&= \langle \mathbf{u}, T^H(\mathbf{v}_1) \rangle + \overline{\alpha} \langle \mathbf{u}, T^H * (\mathbf{v}_2) \rangle \\
&= \langle \mathbf{u}, T^H(\mathbf{v}_1) + \alpha T^H(\mathbf{v}_2) \rangle
\end{aligned}$$

Portanto, $\langle \mathbf{u}, T^H(\mathbf{v}_1 + \alpha \mathbf{v}_2 - T^H(\mathbf{v}_1) + \alpha T^H(\mathbf{v}_2) \rangle = 0$, para todo $\mathbf{v} \in V$, o que implica que $T^H(\mathbf{v}_1 + \alpha \mathbf{v}_2) = T^H * (\mathbf{v}_1) + \alpha T^H(\mathbf{v}_2)$. Portanto, T^H é linear. A unicidade decorre facilmente da construção feita. ∎

Definição 51. *Seja $T \in L(V,V)$, em que V é um espaço vetorial sobre o corpo $\mathbb{K}$ com produto interno. Dizemos que T possui um adjunto se existir um operador linear $T^H \in L(V,V)$ tal que $\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^H(\mathbf{v}) \rangle$, para todos $\mathbf{u}, \mathbf{v} \in V$. Diremos, neste caso, que T^H é o adjunto de T.*

Observação 12. **i)** *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita com base ortonormal $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $T \in L(V,V)$. Observe que combinando as construções feitas em (12) e (41), podemos dar a seguinte fórmula explícita para T^H:*

$$T^H(\mathbf{v}) = \sum_{j=1}^{n} \overline{\langle T(\mathbf{v}_j), \mathbf{v} \rangle} \mathbf{v}_j, \quad \forall\, \mathbf{v} \in V.$$

ii) *O Teorema (41) garante que se V é de dimensão finita, então todo operador $T \in L(V,V)$ possui um adjunto. Isto, porém, não é verdade de modo geral quando dim$(V) = \infty$. Neste caso, é possível mostrar que todo operador linear contínuo entre os chamados espaços de Hilbert admite adjunto. Tais conceitos, bem como a demonstração da afirmação feita, podem ser encontrados em textos básicos de Análise Funcional.*

Proposição 13. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ com produto interno. Sejam $T, S \in L(V,V)$ operadores lineares que admitem adjuntos T^H e S^H, respectivamente e $\alpha \in \mathbb{K}$. Então*

i) *$T + S$ admite adjunto e $(T+S)^H = T^H + S^H$.*

ii) *αT admite adjunto e $(\alpha T)^H = \overline{\alpha} T^H$.*

iii) *TS admite adjunto e $(TS)^H = S^H T^H$.*

iv) *T^H admite adjunto e $(T^H)^H = T$.*

Demonstração:

i) Se $\mathbf{u}, \mathbf{v} \in V$, teremos que

$$\begin{aligned} \langle (T+S)(\mathbf{u}), \mathbf{v} \rangle &= \langle T(\mathbf{u}) + S(\mathbf{u}), \mathbf{v} \rangle \\ &= \langle T(\mathbf{u}), \mathbf{v} \rangle + \langle S(\mathbf{u}), \mathbf{v} \rangle \\ &= \langle \mathbf{u}, T^H(\mathbf{v}) \rangle + \langle \mathbf{u}, S^H(\mathbf{v}) \rangle \\ &= \langle \mathbf{u}, (T^H + S^H)(\mathbf{v}) \rangle \end{aligned}$$

Como a igualdade dada acima vale para todos os vetores $\mathbf{u}, \mathbf{v} \in V$, segue que $T+S$ admite adjunto e $(T+S)^H = T^H + S^H$.

ii) Para $\mathbf{u}, \mathbf{v} \in V$, teremos que

$$\begin{aligned} \langle (\alpha T)(\mathbf{u}), \mathbf{v} \rangle &= \alpha \langle T(\mathbf{u}), \mathbf{v} \rangle \\ &= \alpha \langle \mathbf{u}, T^H(\mathbf{v}) \rangle \\ &= \langle \mathbf{u}, \overline{\alpha} T^H(\mathbf{v}) \rangle \\ &= \langle \mathbf{u}, (\overline{\alpha} T^H)(\mathbf{v}) \rangle \end{aligned}$$

e, portanto, αT admite adjunto e $(\alpha T)^H = \overline{\alpha T^H}$.

iii) Para $\mathbf{u}, \mathbf{v} \in V$, temos que

$$\begin{aligned} \langle (TS), \mathbf{v} \rangle &= \langle T(S(\mathbf{u})), \mathbf{v} \rangle \\ &= \langle S(\mathbf{u}), T^H(\mathbf{v}) \rangle \\ &= \langle \mathbf{u}, S^H(T^H(\mathbf{v})) \rangle \\ &= \langle \mathbf{u}, S^H T^H(\mathbf{v}) \rangle \end{aligned}$$

Logo, TS possui adjunto e $(TS)^H = S^H T^H$.

iv) Para $\mathbf{u}, \mathbf{v} \in V$, temos que

$$\begin{aligned} \langle T^H(\mathbf{u}), \mathbf{v} \rangle &= \overline{\langle \mathbf{v}, T^H(\mathbf{u}) \rangle} \\ &= \overline{\langle T(\mathbf{v}), \mathbf{u} \rangle} \\ &= \langle \mathbf{u}, T(\mathbf{v}) \rangle \end{aligned}$$

e, portanto, T^H possui adjunto e $(T^H)^H = T$.

■

Segue em particular da proposição acima que o conjunto dos operadores que admitem adjuntos é um subespaço de $L(V,V)$. É claro que o operador nulo admite adjunto que é o próprio operador.

Quando o espaço vetorial V tiver dimensão finita, vimos acima que qualquer operador $T \in L(V,V)$ admite adjunto T^H. Para a descrição deste adjunto, é muitas vezes conveniente utilizar as matrizes de T e T^H com relação a uma base ortonormal fixada e ver como elas estão relacionadas.

Proposição 14. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita. Sejam $\mathfrak{B} = \{v_1, \ldots, v_n\}$ uma base ortonormal de V e $T \in L(V,V)$. Se $[T]_{\mathfrak{B}} = (a_{ij})$, então $a_{ij} = \langle T(v_j), v_i\rangle$, $\forall\, i, j = 1, \ldots, n$.*

Demonstração: Segue da definição de $[T]_{\mathfrak{B}}$ que

$$T(\mathbf{v}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{v}_i, \quad \forall\; j = 1, \ldots, n \tag{6.1}$$

Por outro lado, como $\mathfrak{B}$ é uma base ortonormal, então segue do Corolário (5) que para todo $\mathbf{v} \in V$, $\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v}_i\rangle \mathbf{v}_i$. Em particular, temos que

$$T(\mathbf{v}_j) = \sum_{i=1}^{n} \langle T(\mathbf{v}_j), \mathbf{v}_i\rangle \mathbf{v}_i, \quad \forall\; j = 1, \ldots, n. \tag{6.2}$$

Comparando-se as equações (6.1) e (6.2) (que nos dão ambas as coordenadas de $T(\mathbf{v}_j)$ em termos da base $\mathfrak{B}$), concluímos então que $a_{ij} = \langle T(\mathbf{v}_j), \mathbf{v}_i\rangle$ para todo $i, j = 1, \ldots, n$, como queríamos. ■

Teorema 42. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno de dimensão finita, e seja $T \in L(V,V)$. Em relação a qualquer base ortonormal de V, a matriz de T^H é igual à transposta conjugada da matriz de T.*

Demonstração: Seja $\mathfrak{B} = \{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ uma base ortonormal do espaço V e considere $[T]_{\mathfrak{B}} = (a_{ij})$ e $[T^H]_{\mathfrak{B}} = (c_{ij})$ as matrizes dos operadores T e T^H, respectivamente, com relação à base $\mathfrak{B}$. Segue da Proposição (14) que $a_{ij} = \langle T(\mathbf{v}_j), \mathbf{v}_i\rangle$ e $c_{ij} = \langle , \mathbf{v}_i T^H(\mathbf{v}_j)\rangle$ para todo $i, j = 1, \ldots, n$. Usando-se a definição de T^H e as propriedades do produto interno, segue que

$$c_{ij} = \langle T^H(\mathbf{v}_j), \mathbf{v}_i\rangle = \overline{\langle \mathbf{v}_i, T^H(\mathbf{v}_j)\rangle} = \overline{\langle T(\mathbf{v}_i), \mathbf{v}_j\rangle} = \overline{a_{ji}}.$$

Portanto, $\overline{[T]_{\mathfrak{B}}}^T = [T^H]_{\mathfrak{B}}$, como queríamos mostrar. ■

6.2 Operadores Auto-Adjuntos

Uma importante classe de operadores lineares é formada pelos operadores que coincidem com os respectivos adjuntos. Estudar tais operadores é o principal objetivo desta seção.

Definição 52. *Seja $T \in L(V,V)$, em que V é um $\mathbb{K}$-espaço vetorial com produto interno. Dizemos que T é auto-adjunto se T admite adjunto T^H e $T^H = T$. No caso em que $\mathbb{K} = \mathbb{C}$ usamos também o termo hermitiano e no caso em que $\mathbb{K} = \mathbb{R}$, usamos também o termo simétrico.*

Proposição 15. *Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita e $T \in L(V,V)$. As seguintes afirmações são equivalentes:*

i) *T é auto-adjunto*

ii) $\overline{[T]_{bs}}^T = [T]_{\mathfrak{B}}$ *para toda base ortonormal $\mathfrak{B}$ de V*

iii) *Existe uma base ortonormal $\mathfrak{B}$ de V tal que* $\overline{[T]_{bs}}^T = [T]_{\mathfrak{B}}$

Demonstração: Seja $\mathfrak{B}$ uma base ortonormal de V. Vimos no Teorema (42) que $[T^H]_{\mathfrak{B}} = \overline{[T]_{bs}}^T$. Assumindo T auto-adjunto, segue que $\overline{[T]_{bs}}^T = [T]_{\mathfrak{B}}$ o que prova a implicação $(i) \Rightarrow (ii)$. Por outro lado, se assumirmos que $\overline{[T]_{bs}}^T = [T]_{\mathfrak{B}}$, então $[T^H]_{\mathfrak{B}} = [T]_{bs}$ e, portanto, T é auto adjunto, o que prova $(iii) \Rightarrow (i)$. A implicação $(ii) \Rightarrow (iii)$ é trivial. ∎

Corolário 11. *Sejam V um espaço vetorial de dimensão finita e $\mathfrak{B}$ uma base ortonormal de V. Se $T \in L(V,V)$ for um operador linear auto-adjunto e se $[T]_{\mathfrak{B}} = (a_{ij})$, então $a_{ij} = \overline{a_{ji}}$, $\forall\, i,j$. Em particular, os elementos da diagonal de $[T]_{\mathfrak{B}}$ são números reais.*

Lema 4. *Sejam V um $\mathbb{C}$-espaço vetorial com produto interno e $T \in L(V,V)$. As seguintes afirmações são equivalentes:*

i) $T = 0$

ii) $\langle T(\mathbf{u}), \mathbf{u} \rangle = 0, \quad \forall\, \mathbf{u} \in V$

iii) $\langle T(\mathbf{u}), \mathbf{v} \rangle = 0, \quad \forall\, \mathbf{u}, \mathbf{v} \in V$

Demonstração:

$(i) \Rightarrow (ii)$ Trivial.

$(ii) \Rightarrow (iii)$ Sejam $\mathbf{u},\mathbf{v} \in V$ e $\alpha,\beta \in \mathbb{C}$ e considere $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v} \in V$. Então

$$\begin{aligned}
0 &= \langle T(\mathbf{w}),\mathbf{w}\rangle \\
&= \langle T(\alpha\mathbf{u}+\beta\mathbf{v}),\alpha\mathbf{u}+\beta\mathbf{v}\rangle \\
&= |\alpha|^2\langle T(\mathbf{u}),\mathbf{u}\rangle + \overline{\alpha}\beta\langle T(\mathbf{v}),\mathbf{u}\rangle + \alpha\overline{\beta}\langle T(\mathbf{u}),\mathbf{v}\rangle + |\beta|^2\langle T(\mathbf{v}),\mathbf{v}\rangle \\
&= \overline{\alpha}\beta\langle T(\mathbf{v}),\mathbf{u}\rangle + \alpha\overline{\beta}\langle T(\mathbf{u}),\mathbf{v}\rangle
\end{aligned}$$

Como a igualdade acima vale para todos os valores $\alpha,\beta \in \mathbb{C}$, escolhendo-se $\alpha = \beta = 1$, teremos que $\langle T(\mathbf{v}),\mathbf{u}\rangle + \langle T(\mathbf{u}),\mathbf{v}\rangle = 0$. Por outro lado, escolhendo-se os valores $\alpha = i$ e $\beta = 1$, teremos $-i\langle T(\mathbf{v}),\mathbf{u}\rangle + i\langle T(\mathbf{u}),\mathbf{v}\rangle = 0$. Resolvendo-se o sistema

$$\begin{cases} \langle T(\mathbf{v},\mathbf{u}\rangle + \langle T(\mathbf{u},\mathbf{v}\rangle = 0 \\ -i\langle T(\mathbf{v},\mathbf{u}\rangle + i\langle T(\mathbf{u},\mathbf{v}\rangle = 0 \end{cases}$$

Segue que $T(\mathbf{u},\mathbf{v}\rangle = 0$ e o resultado está provado.

$(iii) \Rightarrow (i)$ Como $(\forall\, \mathbf{u},\mathbf{v} \in V)$ $\langle T(\mathbf{u}),\mathbf{v}\rangle = 0$, podemos escolher $\mathbf{v} = T(\mathbf{u})$. Assim, $(\forall\, \mathbf{u},\mathbf{v} \in V)$ $\langle T(\mathbf{u}),T(\mathbf{u})\rangle = 0$.

■

Observação 13. *A equivalência acima não é verdadeira se considerarmos espaços vetoriais se considerarmos espaços vetoriais sobre $\mathbb{R}$. Na verdade, a equivalência das condições (i) e (iii) continua valendo, assim como a implicação $(iii) \Rightarrow (ii)$. O que não é verdade é a implicação $(ii) \Rightarrow (iii)$.*

Proposição 16. *Sejam V um $\mathbb{C}$-espaço vetorial com produto interno e $T \in L(V,V)$. Então T é um operador hermitiano se e somente se $\langle T(\mathbf{v}),\mathbf{v}\rangle \in \mathbb{R}$, $\forall\, \mathbf{v} \in V$.*

Demonstração:

$(\Rightarrow)$ Se T é hermitiano, então $T = T^H$. Portanto, para cada $\mathbf{v} \in V$, temos que

$$\langle T(\mathbf{v}),\mathbf{v}\rangle = \langle \mathbf{v},T^H(\mathbf{v})\rangle = \langle \mathbf{v},T(\mathbf{v})\rangle = \overline{\langle T(\mathbf{v}),\mathbf{v}\rangle}$$

e então $\langle T(\mathbf{v}),\mathbf{v}\rangle \in \mathbb{R}$.

$(\Leftarrow)$ Suponha que $\langle T(\mathbf{v}),\mathbf{v}\rangle \in \mathbb{R}$, $\forall\, \mathbf{v} \in V$. Então

$$\langle T(\mathbf{v}),\mathbf{v}\rangle = \overline{\langle T(\mathbf{v}),\mathbf{v}\rangle} = \overline{\langle \mathbf{v},T^H(\mathbf{v})\rangle} = \langle T^H(\mathbf{v}),\mathbf{v}\rangle.$$

Assim, $\langle T(\mathbf{v}) - T^H(\mathbf{v}), \mathbf{v}\rangle = 0, \forall\, \mathbf{v} \in V$. Segue do Lema (4) que $T(\mathbf{v}) = T^H(\mathbf{v}),\ \forall\, \mathbf{v} \in V$, como queríamos.

■

É claro que a Proposição (16) não é válida em geral se V for um espaço vetorial sobre o corpo $\mathbb{R}$.

6.3 Operadores Unitários

Definição 53. *Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $T \in L(V,V)$. Dizemos que T é unitário se for um isomorfismo de espaços com produto interno.*

Observação 14. *Sejam $T_1, T_2 \in L(V,V)$, em que V é um $\mathbb{K}$-espaço vetorial com produto interno. Então, se T_1 e T_2 forem operadores unitários, T_1T_2 também o será. Além disso, se T_1 é unitário, então T_1^{-1} também é unitário.*

Proposição 17. *Seja $T \in L(V,V)$, em que V é um $\mathbb{K}$-espaço vetorial com produto interno. Então T é unitário se, e somente se, o adjunto T^H existir e $TT^H = T^HT = I$.*

Demonstração: Suponha, em primeiro lugar, que T seja unitário. Então T é invertível e, como preserva produto interno, teremos que

$$\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle T(\mathbf{u}), (TT^{-1})(\mathbf{v})\rangle = \langle \mathbf{u}, T^{-1}(\mathbf{v})\rangle$$

para todo $\mathbf{u}, \mathbf{v} \in V$. Logo, T^{-1} é o adjunto de T. É claro que $TT^H = T^HT = I$. Reciprocamente, suponha que T^H exista e que $TT^H = T^HT = I$. Então T é inversível e $T^{-1} = T^H$. Falta provar que T preserva produtos internos. De fato,

$$\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, (T^HT)(\mathbf{v})\rangle = \langle \mathbf{u}, I \cdot \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$$

para todos $\mathbf{u}, \mathbf{v} \in V$ e o resultado está provado. ■

Definição 54. *Seja $\mathbf{A} \in \mathbb{K}^{n\times n}$. Dizemos que $\mathbf{A}$ é unitária se $\mathbf{A}\overline{\mathbf{A}}^T\mathbf{A} = \mathbf{I}$. Quando $\mathbb{K} = \mathbb{R}$, também dizemos que $\mathbf{A}$ é ortogonal.*

6.4 Operadores Normais

Vamos discutir nesta seção a existência de bases ortonormais formadas por autovetores de um dado operador linear T.

Vamos assumir que V é um $\mathbb{K}$-espaço vetorial com produto interno de dimensão finita $n \geq 1$ e que $T \in L(V,V)$. Suponha que exista uma base ortonormal $\mathfrak{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ cujos elementos são autovetores de T, isto é, que $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ para certos λ_i's em $\mathbb{K}$, $i = 1, \ldots, n$. Assim, $[T]_{\mathfrak{B}}$ tem os elementos $\lambda_1, \ldots, \lambda_n$ na diagonal principal e zero nas outras posições. Como a base $\mathfrak{B}$ é ortonormal, então $[T^H]_{\mathfrak{B}}$ tem os elementos $\overline{\lambda_1}, \ldots, \overline{\lambda_n}$ na diagonal principal e zero nas outras posições. Se $\mathbb{K} = \mathbb{R}$, então $\lambda_i = \overline{\lambda_i}$ para cada $i = 1, \ldots, n$ e portanto, $[T]_{\mathfrak{B}} = [T^H]_{\mathfrak{B}}$, isto é, T é auto-adjunto. Se $\mathbb{K} = \mathbb{C}$, então T não é necessariamente auto-adjunto, mas vale a relação $[T]_{\mathfrak{B}} \cdot [T^H]_{\mathfrak{B}} = [T^H]_{\mathfrak{B}} \cdot [T]_{\mathfrak{B}}$, ou melhor, T comuta com T^H: $TT^H = T^H T$. Iremos mostrar que a recíproca do resultado acima também vale, isto é, se T é tal que $TT^H = T^H T$, então existe uma base ortonormal de V cujos elementos são autovetores de T.

Definição 55. *Sejam V um espaço vetorial com produto interno e $T \in L(V,V)$. Dizemos que T é normal se existir T^H e $TT^H = T^H T$.*

Observação 15. **i)** *Todo operador auto-adjunto é normal.*

i) *Todo múltiplo escalar de um operador normal é normal. De fato, se T é normal e $\alpha \in \mathbb{K}$,*

$$\begin{aligned}
(\alpha T)^H (\alpha T) &= (\overline{\alpha} T^H)(\alpha T) \\
&= \overline{\alpha}\alpha (T^H T) \\
&= \alpha\overline{\alpha} (T T^H) \\
&= (\alpha T)(\overline{\alpha} T^H) \\
&= (\alpha T)(\alpha T)^H
\end{aligned}$$

iii) *A soma de operadores normais não é necessariamente normal. A verificação pode ser feita facilmente, porém não a faremos aqui.*

Proposição 18. *Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $T \in L(V,V)$ um operador normal. Então,*

i) $\|T(\mathbf{v})\| = \|T^H(\mathbf{v})\|, \quad \forall\, \mathbf{v} \in V.$

ii) *Se $T(\mathbf{v}) = \lambda \mathbf{v}$ para $\lambda \in \mathbb{K}$ e $\mathbf{v} \in V$, então $T^H(\mathbf{v}) = \overline{\lambda}\mathbf{v}$.*

iii) *Se $T(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1$ e $T(\mathbf{v}_2) = \lambda_2 \mathbf{v}_2$, para $\mathbf{v}_1, \mathbf{v}_2 \in V$ e $\lambda_1, \lambda_2 \in \mathbb{K}$, com $\lambda_1 \neq \lambda_2$, então $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$*

Demonstração:

i) Seja $\mathbf{v} \in V$, então

$$\langle T(\mathbf{v}), T(\mathbf{v})\rangle = \langle \mathbf{v}, T^H(T(\mathbf{v}))\rangle = \langle \mathbf{v}, T(T^H(\mathbf{v}))\rangle = \overline{\langle T(T^H(\mathbf{v})), \mathbf{v}\rangle}$$

Como $\langle T(\mathbf{v}), T(\mathbf{v})\rangle$ é um número real, segue que

$$\overline{\langle T(T^H(\mathbf{v})), \mathbf{v}\rangle} = \langle T(T^H(\mathbf{v})), \mathbf{v}\rangle = \langle T^H(\mathbf{v}), T^H(\mathbf{v})\rangle$$

. Portanto, $\|T(\mathbf{v})\| = \|T^H(\mathbf{v})\|$, como queríamos.

ii) Se $T(\mathbf{v}) = \lambda\mathbf{v}$, então $(T - \lambda I)(\mathbf{v}) = 0$. Logo, $\|(T - \lambda I)(\mathbf{v})\| = 0$. Usando o item (i), concluímos que $\|(T - \lambda i)^H(\mathbf{v})\| = 0$. Então $(T - \lambda I)^H(\mathbf{v}) = 0$ e, portanto, $T^H(\mathbf{v}) = \overline{\lambda}\mathbf{v}$.

iii) Observe que

$$\langle T(\mathbf{v}_1), \mathbf{v}_2\rangle = \langle \mathbf{v}_1, T^H(\mathbf{v}_2)\rangle = \langle \mathbf{v}_1, \overline{\lambda_2}\mathbf{v}_2\rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle.$$

Por outro lado,

$$\langle T(\mathbf{v}_1), \mathbf{v}_2\rangle = \langle \lambda_1\mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle.$$

Daí, $\lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle$ e, portanto, $(\lambda_1 - \lambda_2)\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$. Como $\lambda_1 \neq \lambda_2$, segue que $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$.

■

Teorema 43. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita com produto interno. Se $T \in L(V,V)$ é auto-adjunto, então T possui um autovetor.*

Demonstração: Observe inicialmente que se $\mathbb{K} = \mathbb{C}$, então $p_T(x)$ tem raízes e elas são os autovalores de T, como queríamos. Vamos assumir então que $\mathbb{K} = \mathbb{R}$. Suponha $dim(V) = n$ e seja $T \in L(V,V)$ um operador auto-adjunto. Sejam $\mathfrak{B}$ uma base ortonormal de V e $\mathbf{A} = [T]_{\mathfrak{B}}$. Como $T = T^H$, temos, pelo Teorema (42) que $\mathbf{A} = \overline{\mathbf{A}}^T$. Considere $W = \mathbb{C}^{n\times 1}$ com produto interno $\langle \mathbf{X}, \mathbf{Y}\rangle = \overline{\mathbf{Y}}^T\mathbf{X}$ e $S : W \to W$ o operador linear dado por $S(\mathbf{X}) = \mathbf{A}\mathbf{X}$. Sabemos que $S^H(\mathbf{X}) = \overline{\mathbf{A}}^T\mathbf{X} = \mathbf{A}\mathbf{X}$ e, portanto, S é auto-adjunto. Por outro lado, não é difícil ver que $p_T(x) = p_S(x)$. Seja λ uma raiz de p_S. Como W é um espaço vetorial sobre $\mathbb{C}$, segue que λ é um autovalor de S.

Afirmação: λ é um valor real. De fato, se $\mathbf{v} \neq \mathbf{0}$ for um autovetor asociado ao autovalor λ, então $\langle S(\mathbf{v}), \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \lambda\langle \mathbf{v}, \mathbf{v}\rangle$ e, por outro lado,

$$\langle S(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, S^H(\mathbf{v})\rangle = \langle \mathbf{v}, S(\mathbf{v})\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle.$$

Daí, $\lambda\langle \mathbf{v}, \mathbf{v}\rangle = \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle$ e então $(\lambda - \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle = 0$. Como $\langle \mathbf{v}, \mathbf{v}\rangle \neq 0$, segue que $\lambda - \overline{\lambda} = 0$ e, consequentemente, $\lambda \in \mathbb{R}$. Observe que então λ é uma raiz real de $p_T(x)$ e, portanto, λ é um autovalor de T como queríamos. ■

Vamos usar o resultado acima para mostrar que se T é um operador auto-adjunto em $L(V,V)$, então V tem uma base ortonormal cujos elementos são autovetores de T. Para tanto, precisamos do seguinte lema:

Lema 5. *Sejam V um $\mathbb{K}$ espaço vetorial com produto interno e de dimensão finita e $T \in L(V,V)$. Se W é um subespaço T-invariante de V, então $W^\perp$ é T^H-invariante.*

Demonstração: Temos que mostrar que $T^H(\mathbf{w}) \in W^\perp$, para cada $\mathbf{w} \in W^\perp$, isto é, que $\langle \mathbf{v}, T^H(\mathbf{w})\rangle = 0$, $\forall\ \mathbf{v} \in W$. Sejam $\mathbf{v} \in W$ e $\mathbf{w} \in W^\perp$. Como W é T-invariante, então $T(\mathbf{v}) \in W$ e, portanto, $\langle T(\mathbf{v}), \mathbf{w}\rangle = 0$. O resultado agora segue do fato de

$$\langle \mathbf{v}, T^H(\mathbf{w})\rangle = \langle T(\mathbf{v}), \mathbf{w}\rangle = 0$$

■

Proposição 19. *Seja V um $\mathbb{K}$-espaço vetorial com produto interno e de dimensão finita. Se $T \in L(V,V)$ é auto-adjunto, então existe uma base ortonormal de V cujos vetores são autovetores de T.*

Demonstração: Vamos supor que $dim(V) = n \geq 1$. Pel Teorema (43), T possui um autovetor $\mathbf{v}_1$. Se $dim(V) = 1$, então $\left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}\right\}$ é uma base, como queríamos. Vamos agora supor que $n > 1$ e que o resultado vale para todo espaço vetorial de dimensão $n-1$. Seja $W = [\mathbf{v}_1]$, em que $\mathbf{v}_1$ é o autovetor acima. É fácil ver que W é invariante por T. Pelo Lema (5), $W^\perp$ é T^H-invariante. Como $T^H = T$ segue que $W^\perp$ é T-invariante. Agora, como $W^\perp$ é um espaço de dimensão $n-1$, segue da hipótese de indução que $W^\perp$ possui uma base ortonormal $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ formada por autovetores. Logo, $\mathfrak{B} = \left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \mathbf{v}_2, \ldots, \mathbf{v}_n\right\}$ é um conjunto ortonormal com $dim(V)$ elementos e, portanto, uma base de V. Por construção, todos os elementos de $\mathfrak{B}$ são autovetores e o resultado está provado. ■

Corolário 12. *Seja $\mathbf{A} \in \mathbb{R}^{n\times n}$ uma matriz simétrica. Então existe uma matriz inversível $\mathbf{M} \in \mathbb{R}^{n\times n}$ tal que $\mathbf{M}^T\mathbf{A}\mathbf{M}$ é diagonal.*

Teorema 44. *Sejam V um $\mathbb{C}$-espaço vetorial com produto interno e de dimensão finita e $T \in L(V,V)$. Então T será um operador normal se, e somente se existir uma base ortonormal de V cujos vetores sejam autovetores de T.*

Demonstração: Seja $\mathbf{v}_1 \in V$ um autovetor de T ($\mathbf{v}_1$ existe pois V é um espaço vetorial complexo). Sem perda de generalidade podemos supor que $\|\mathbf{v}_1\| = 1$. Considere $W = [\mathbf{v}_1]$.

Assim, W é T-invariante. Da Proposição (18) segue que $\mathbf{v}_1$ é autovetor de T^H e portanto W é T^H-invariante. Pelo Lema (5) concluímos então que $W^\perp$ é invariante por $(T^H)^H = T$. A restrição de T a $W^\perp$ é um operador normal. Usando o mesmo argumento de indução usado na Proposição (19), mostra-se que existe uma base ortonormal de autovetores. A recíproca foi mostrada no início desta seção. ∎

7 Métodos Iterativos para a Computação de Autovalores

Exemplo 12. *Considere a famosa sequência de Fibonacci*

$$x_0 = 0,\ \ x_1 = 1,\ \ \forall k > 1,\ \ x_k = x_{k-1} + x_{k-2}.$$

Em termos matriciais, para $k \geq 1$,

$$\begin{aligned} \begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ x_k \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^k \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \end{aligned}$$

ou seja, $\mathbf{u}_k = \mathbf{A}^k \mathbf{u}_0$. Os autovalores de $\mathbf{A}$ são $\lambda_1 = \frac{1+\sqrt{5}}{2}$ e $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Note que $(1, \lambda_1)$ e $(1, \lambda_2)$ são dois autovetores de A associados respectivamente a λ_1 e a λ_2. Assim,

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

em que

$$\mathbf{S} = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \quad e \quad \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Logo, $\mathbf{u}_k = \mathbf{S}\mathbf{D}^k\mathbf{S}^{-1}\mathbf{u}_0$ e, como

$$\begin{aligned} \mathbf{S}^{-1}\mathbf{u}_0 &= \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\lambda_1 - \lambda_2} \\ \frac{1}{\lambda_2 - \lambda_1} \end{pmatrix} \\ \mathbf{u}_k &= \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1^k - \lambda_2^k \\ \lambda_1^{k+1} - \lambda_2^{k+1} \end{pmatrix}. \end{aligned}$$

Assim,

$$x_k = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right].$$

Agora, para todo k, $\frac{|\lambda_2|^k}{\sqrt{5}} < 0.5$. *Concluímos então que* x_k *é o inteiro mais próximo de* $\frac{\lambda_1^k}{\sqrt{5}}$

Vemos no exemplo acima que, se uma matriz é diagonalizável, as potências da matriz são facilmente computadas, uma vez que conheçamos sua decomposição espectral $\mathbf{A} = \mathbf{PDP}^{-1}$. Porém para calcularmos os autovalores de uma matriz de ordem maior do que 4, mesmo conhecendo os coeficientes de seu polinômio característico, é necessário a utilização de métodos iterativos, pois não há fórmulas algébricas para encontrar raízes de polinômios de grau ≥ 5, conforme a Teoria de Galois.

Os métodos para a computação de autovalores são baseados em álgebra matricial e neste capítulo vamos mostrar dois importantes métodos: o método da Potência e o método de Iteração de Subespaço. Vamos aqui trabalhar apenas com matrizes diagonalizáveis, fato que não restringe nossa pesquisa visto que o conjunto das matrizes diagonalizáveis sobre $\mathbb{C}^{m\times n}$ é denso no conjunto de todas as matrizes sobre $\mathbb{C}^{m\times n}$.

7.1 Método de Potência

Teorema 45. *Seja* $\mathbf{A} \in \mathbb{C}^{nxn}$. *Se* λ *é o único autovalor de* $\mathbf{A}$ *satisfazendo* $|\lambda| = \rho(\mathbf{A})$ *e sendo* λ *autovalor simples, então a sequência de matrizes* $(\lambda^{-1}\mathbf{A})^k$ *converge para a matriz* $\mathbf{uv}^T$ *quando* $k \to \infty$, *em que* $\mathbf{u}$ *e* $\mathbf{v}$ *são os autovetores normalizados de* $\mathbf{A}$ *e* $\mathbf{A}^T$, *respectivamente, associados a* λ, *de modo que* $\mathbf{v}^T\mathbf{u} = 1$.

Demonstração: Como já mostramos neste trabalho, uma matriz $\mathbf{A} \in \mathbb{C}^{nxn}$ é semelhante a uma matriz na forma normal de Jordan cujo bloco associado a λ é (pelo fato que λ é autovalor simples) de ordem 1; além disso, a razão de $\frac{\lambda}{\lambda_i}$, para $i = 1,...,\lambda_{n-1}$ é menor que 1 sob as hipóteses do teorema. Isso significa que existe uma matriz não singular $\mathbf{P}$ tal que:

$$\mathbf{PAP}^{-1} = \mathbf{A}_J = \begin{pmatrix} \lambda & 0 \\ 0 & Z \end{pmatrix}, \rho(\mathbf{Z}) < |\lambda|$$

Portanto,

$$\mathbf{P}(\lambda^{-1}\mathbf{A})^k\mathbf{P}^{-1} = (\lambda^{-1}\mathbf{A}_J)^k = \begin{pmatrix} 1 & 0 \\ 0 & (\lambda^{-1}Z)^k \end{pmatrix},$$

E então

$$\lim_{k\to\infty}(\lambda^{-1}\mathbf{A_J})^k = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{e}_1\mathbf{e}_1^T,$$

em que $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$. De modo geral, teremos

$$\lim_{k\to\infty} (\lambda^{-1}\mathbf{A})^k = (\mathbf{P}^{-1}\mathbf{e}_1)(\mathbf{e}_1^T\mathbf{P}) \tag{7.1}$$

Mas $\mathbf{Au} = \lambda\mathbf{u}$, $\mathbf{u} \neq \mathbf{0}$ e, além disso, $\mathbf{u}$ é determinado de maneira única exceto por um fator não nulo. Ademais, $\mathbf{A}_J\mathbf{e}_1 = \lambda\mathbf{e}_1$, a partir do qual,

$$\mathbf{A}(\mathbf{P}^{-1}\mathbf{e}_1) = \mathbf{P}^{-1}(\mathbf{PAP}^{-1})e_1 = \mathbf{P}^{-1}\mathbf{A}_J\mathbf{e}_1 = \mathbf{P}^{-1}(\lambda\mathbf{e}_1) = \lambda(\mathbf{P}^{-1}\mathbf{e}_1).$$

Por isso,

$$\mathbf{u} = \alpha\mathbf{P}^{-1}\mathbf{e}_1, \qquad \alpha \neq 0. \tag{7.2}$$

Similarmente temos que $\mathbf{A}^T\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, sendo $\mathbf{v}$ também determinado de maneira única exceto por um fator não nulo. Agora, $A_J^T e_1 = \lambda e_1$ e consequentemente

$$\mathbf{A}^T(\mathbf{P}^T\mathbf{e}_1) = \mathbf{P}^T((\mathbf{P}^T)^{-1}\mathbf{A}^T\mathbf{P}^T)\mathbf{e}_1 = \mathbf{P}^T\mathbf{A}_J^T\mathbf{e}_1 = \mathbf{P}^T(\lambda\mathbf{e}_1) = \lambda(\mathbf{P}^T\mathbf{e}_1).$$

Portanto,

$$\mathbf{v} = \beta\mathbf{P}^T\mathbf{e}_1, \qquad \beta \neq 0. \tag{7.3}$$

Por hipótese, $\mathbf{v}^T\mathbf{u} = 1$. Substituindo 7.2 em 7.1, temos que

$$\beta(\mathbf{e}_1^T\mathbf{P})\alpha(\mathbf{P}^{-1}\mathbf{e}_1) = 1,$$

e então

$$\alpha\beta = 1$$

desde que $\mathbf{e}_1^T\mathbf{e}_1 = 1$.

Usando 7.2 e 7.3, descobrimos por 7.1 que

$$\lim_{k\to\infty} (\lambda^{-1}\mathbf{A})^k = (\alpha^{-1})\mathbf{u}(\beta^{-1})\mathbf{v}^T = \mathbf{uv}^T,$$

como queríamos provar. ∎

Teorema 46. *Seja* $\mathbf{A}$ *uma matriz quadrada que possui um único auvalor satisfazendo* $|\lambda| = \rho(\mathbf{A})$*; suponha que* λ *é simples.*

Seja $\mathbf{v}$ *um autovetor de* $\mathbf{A}^T$ *associado a* λ*. Se* $\mathbf{z}$ *é um vetor arbitrário tal que* $\mathbf{v}^T\mathbf{z} \neq \mathbf{0}$*, então*

$$\lim_{k\to\infty} (\lambda^{-1}\mathbf{A})^k\mathbf{z} = \mathbf{y} \neq \mathbf{0}$$

em que $\mathbf{y}$ *é um autovetor de* $\mathbf{A}$ *associado ao autovalor* λ*.*

Demonstração: Pelo Teorema 45,

$$\lim_{k\to\infty}(\lambda^{-1}\mathbf{A})^k\mathbf{z} = (\mathbf{u}\mathbf{v}^T)\mathbf{z} = \mathbf{u}(\mathbf{v}^T\mathbf{z}),$$

em que $\mathbf{u}$ é um autovetor de $\mathbf{A}$ associado ao autovalor λ satisfazendo $\mathbf{v}^T\mathbf{u} = 1$. Como $\mathbf{v}^T\mathbf{z} \neq \mathbf{0}$, o vetor $\mathbf{y} = (\mathbf{v}^T\mathbf{z})\mathbf{u}$ é também um autovetor de $\mathbf{A}$ associado a λ, como queríamos demonstrar. ■

Observação 1: O teorema 46 resulta que para um valor suficientemente grande de k os vetores $A^{k+1}z$ e A^kz são aproximadamente proporcionalis,

$$A^{k+1}z \approx \mu A^k z;$$

o fator de proporcionalidade μ é aproximadamente igual a λ, e o vetor $\mathbf{A}^{k+1}z$ é aproximadamente igual ao autovalor de $\mathbf{A}$ associado a λ. Este processo pode ser utilizado para determinar aproximadamente λ e seu autovetor correspondente.

Observação 2: A precisão da aproximação de λ ou do autovetor u obtidos depois de um número finito de iterações depende da taxa $\xi = |\lambda_2|/|\lambda|$, em que $|\lambda_2|$ é o módulo máximo dos autovalores remanescentes. Se ξ é muito próximo de 1, a convergência é lenta, se ξ é muito próximo de 0, a convergência é rápida. O número necessário de iterações, k, é a grosso modo, tal que ξ^k tenha a mesma ordem de magnitude que a necessária precisão relativa do resultado.

7.2 Iterações de Subespaço

Ainda não lidamos com a possibilidade de autovalores com módulos muito próximos. Suponhamos então que uma certa matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ possua autovalores satisfazendo $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Se a razão entre os dois autovalores de maior módulo é próxima a um, não podemos esperar que o método de potência apresentado na seção anterior convirja rapidamente. Neste caso, ao invés de olharmos para o autovetor associado a λ_1, olharemos para o subespaço invariante associado a λ_1 e λ_2 (e talvez a mais autovalores, se estiverem todos muito próximos a λ_1). Essa é a ideia por trás das Iterações de Subespaço.

Veremos agora dois importantes métodos de Iteração de Subespaço, o Método de Iteração Ortogonal e o Método LOPSI [10]. O método da Iteração Ortogonal foi originalmente introduzido por Bauer [2], que chamou o método Treppeniteration (iteração escada). É um método simples para calcular os autovalores de maiores módulos de uma matriz real não hermitiana e é uma generalização do método de potência. Vários métodos que empregam técnicas de iteração subespaço têm sido desenvolvidos. Um deles é denominado LOPSI, o qual veremos com mais detalhes posteriormente nesta pesquisa.

7.2.1 Iteração Ortogonal

Na iteração ortogonal, ao invés de olharmos para $\mathbf{A}^K\mathbf{x}_0$, com um vetor inicial $\mathbf{x}_0$, nós olhamos para $\mathbf{V}_k = \mathbf{A}^k\mathbf{V}_0$, em que $\mathbf{V}_0$ é um subespaço inicial arbitrário. Se $\mathbf{V}_0$ é um espaço p-dimensional, então sob algumas hipóteses o espaço $\mathbf{V}_k$ convergirá assintoticamente para o subespaço invariante p-dimensional de $\mathbf{A}$ associado aos p autovalores de $\mathbf{A}$ de maior módulo. A análise é, basicamente, a mesma análise feita para o método de potência. Para iniciar os cálculos, primeiramente é necessário encontrar bases para os subespaços $\mathbf{V}_k$. Vamos definir essas bases pela recorrência

$$\mathbf{Q}_{k+1}\mathbf{R}_{k+1} = \mathbf{A}\mathbf{Q}_k \tag{7.4}$$

em que $\mathbf{Q}_0$ é a matriz com p colunas ortonormais e $\mathbf{Q}_{k+1}\mathbf{R}_{k+1}$ representa uma fatoração **QR** econômica. Essa recorrência leva o nome de Iteração Ortogonal, pelo fato que as colunas de $\mathbf{Q}_{k+1}$ são uma base ortonormal para o espaço coluna de $\mathbf{A}\mathbf{Q}_k$, e o espaço gerado por $\mathbf{Q}_k$ é o mesmo espaço gerado por $\mathbf{A}^k\mathbf{Q}_0$. Note que, se $p = 1$, então teríamos apenas o método de potência. Supondo que exista um intervalo entre $|\lambda_p|$ e $|\lambda_{p+1}|$, a iteração ortogonal muito provavelmente convergirá para uma base ortonormal do subespaço invariante gerado pelos primeiros p autovetores de $\mathbf{A}$. Mas é interessante não olhar somente para o comportamento do subepaço, mas também para os espaços gerados de cada autovetor. Por exemplo, observemos que a primeira coluna $\mathbf{q}_{k,1}$ de $\mathbf{Q}_k$ satisfaz a recorrência

$$\mathbf{q}_{k+1,1}\mathbf{r}_{k+1,1} = \mathbf{A}\mathbf{q}_{k+1},$$

o que significa que os vetores $\mathbf{q}_{k+1}$ evoluem de acordo com o método de potência. Assim, ao longo do tempo, esperamos que a primeira coluna de $\mathbf{Q}_k$ convirja para o autovetor dominante. Similarmente, esperamos que as duas primeiras colunas de $\mathbf{Q}_k$ convirjam para uma base do espaço invariante bidimensional associado aos dois primeiros autovetores; as três primeiras colunas convergirão para uma base do espaço invariante tridimensional associado aos três primeiros autovetores. Isto sugere que somos capazes de obter uma lista completa de subespaços invariantes encaixados considerando uma matriz quadrada inicial $\mathbf{Q}_0$.

A fim de analisar o comportamento dessa iteração, suponha que

$$\mathbf{Q}^H\mathbf{A}\mathbf{Q} = \mathbf{T} = diag(\lambda_i) + \mathbf{N} \quad |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \tag{7.5}$$

seja a decomposição de Schur de uma matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$. Definimos como o desvio da normalidade de $\mathbf{A}$ como

$$\|\mathbf{N}\|_F^2 = \|\mathbf{A}\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2.$$

Assuma que $1 \leq p < n$ e considere as partições das matrizes $\mathbf{Q}$, $\mathbf{T}$ e $\mathbf{N}$, como segue:

$$\mathbf{Q} = \begin{array}{c} \begin{array}{cc} p & n-p \end{array} \\ \begin{pmatrix} \mathbf{Q}_\alpha & \mathbf{Q}_\beta \end{pmatrix} \end{array}$$

$$\mathbf{T} = \begin{array}{cc} \begin{array}{cc} p & n-p \end{array} & \\ \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{0} & \mathbf{T}_{22} \end{pmatrix} & \begin{array}{c} p \\ n-p \end{array} \end{array}$$

$$\mathbf{N} = \begin{array}{cc} \begin{array}{cc} p & n-p \end{array} & \\ \begin{pmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{0} & \mathbf{N}_{22} \end{pmatrix} & \begin{array}{c} p \\ n-p \end{array} \end{array}$$

em que os valores p e $n-p$ acima de cada bloco representam a quantidade de colunas do referente bloco e os valores p e $n-p$ ao lado de cada bloco representam a quantidade de linhas do refente bloco.

Se $|\lambda_p| > |\lambda_{p+1}|$, então o subespaço gerado pelos p autovalores associados com os p autovalores dominantes de $\mathbf{A}$, $D_p(\mathbf{A}) = C(\mathbf{Q}_\alpha)$, é dito um subespaço invariante dominante (pois é gerado pelos autovetores associados aos autovalores dominantes de $\mathbf{A}$). Este é o único subespaço invariante associado aos autovalores $\lambda_1, \ldots, \lambda_p$.

Observação 16. *Lembrando que $C(\mathbf{A})$ é a notação que representa o espaço coluna de uma matriz $\mathbf{A}$.*

Vamos definir agora a noção de distância entre subespaços, que será útil para provarmos um resultado importante posteriormente.

Definição 56. *Suponha que S_1 e S_2 sejam subespaços de $\mathbb{C}^n$ e que $dim(S_1) = dim(S_2)$. Definimos a distância entre dois espaços por*

$$dist(S_1, S_2) = \|P_1 - P_2\|_2$$

em que P_1 é a projeção ortogonal em S_i.

O teorema a seguir mostra que com algumas hipóteses razoáveis, os subespaços $C(\mathbf{Q}_k)$ gerados por (7.4) convergem para $D_p(\mathbf{A})$ com uma taxa proporcional a $|\dfrac{\lambda_{p+1}}{\lambda_p}|^k$.

Teorema 47. *Seja decomposição de Schur da matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$, $n \geq 2$, dada como em (7.5) e considere as representações das matrizes $\mathbf{Q}$, $\mathbf{T}$ e $\mathbf{N}$, como dadas anteriormente. Assuma que $|\lambda_p| > |\lambda_{p+1}|$ e que um dado $\theta \geq 0$ satisfaz*

$$(1+\theta)|\lambda_p| > \|N\|_F.$$

Se $\mathbf{Q}_0 \in \mathbb{C}^{n\times p}$ possui colunas ortonormais e

$$d = dist(D_r(\mathbf{A}^H), C(\mathbf{Q}_0)) < 1,$$

então as matrizes $\mathbf{Q}_k$ geradas por (7.4), satisfazem

$$dist(D_r(\mathbf{A}), C(\mathbf{Q}_k)) \leq \frac{(1+\theta)^{n-2}}{\sqrt{1-d^2}}\left(1+\frac{\|\mathbf{T}_{12}\|_F}{sep(\mathbf{T}_{11},\mathbf{T}_{22})}\right)\left(\frac{|\lambda_{r+1}|+\|\mathbf{N}\|_F/(1+\theta)}{|\lambda_r|-\|\mathbf{N}\|_F/(1+\theta)}\right)^k$$

no qual $sep(\mathbf{T}_{11},\mathbf{T}_{22}) = \min\limits_{\mathbf{X}\neq 0} \dfrac{\|\mathbf{T}_{11}\mathbf{X}-\mathbf{X}\mathbf{T}_{22}\|_F}{\|\mathbf{X}\|_F}$, $\mathbf{X} \in \mathbb{C}^{p\times p}$.

Demonstração: Não faremos aqui a demonstração deste teorema, porém o leitor pode encontrá-la em [5]. ■

A condição $d < 1$ no Teorema (47) garante que a matriz inicial $\mathbf{Q}_0$ deve satisfazer:

$$D < 1 \leftrightarrow D_r(\mathbf{A}^H)^{\perp} \bigcap C(\mathbf{Q}_0) = \{\mathbf{0}\}.$$

O teorema essencialmente diz que se esta condição é satisfeita e se θ é escolhido grande o suficiente, então,

$$dist(D_r(\mathbf{A}, C(\mathbf{Q}_k)) \leq c|\frac{\lambda_{p+1}}{\lambda_p}|^k$$

em que c depende de $sep(\mathbf{T}_{11},\mathbf{T}_{22})$ e no desvio da normalidade de $\mathbf{A}$. É claro que a convergência será lenta se a distância entre $|\lambda_p|$ e $|\lambda_{p+1}|$ não for grande o suficiente.

Serão expostos agora alguns exemplos que consistem em resultados de experimentos nos quais utilizamos um método de iteração de subespaço. Para gerar tais exemplos, utilizamos algumas matrizes da galeria do MATLAB [6]. Tais matrizes foram escolhidas pelo fato de terem características especiais em sua forma que tornam seus autovalores sensíveis, isto é, pequenas mudanças na matriz podem induzir grandes mudanças em seus autovalores.

O método de iteração ortogonal tem início com a escolha de um subespaço arbitrário Z - cuja dimensão deve ser menor ou igual a da matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ a ser iterada. Em nossos testes, optamos por utilizar o subespaço gerado pelas p colunas de uma matriz arbitrária $n \times p$, que no MATLAB é gerada através do comando $\mathbf{Z} = rand(n, p)$, para ser o subspaço inicial de cada exemplo. Conforme supracitado, ao escolhermos um subespaço de dimensão p para iniciar o método, esperamos que ele convirja para um subespaço invariante associado aos p autovalores de $\mathbf{A}$ de maior módulo. O teste de convergência será a comparação do erro com a tolerância exigida, no qual o erro é o módulo da diferença entre os autovalores obtidos pelo método e pelos autovalores exatos de $\mathbf{A}$, calculados pela função *eig* do MATLAB.

Em cada exemplo, a figura representa o plano complexo no qual plotamos os valores resultantes das iterações, representados pelo símbolo “o” e também os valores exatos dos autovalores de $\mathbf{A}$ - calculados a partir da função `eig(A)` do MATLAB - representados pelo símbolo “+”.

As matrizes utilizadas nos quatro exemplos podem ser classificadas como segue.

1. Matriz Frank - matriz com autovalores mal-condicionados.

O comando $\mathbf{A} = gallery('frank', n, k)$ retorna a matriz Frank de ordem n. É uma matriz Hessenberg superior, isto é, todos os elementos abaixo da primeira subdiagonal são zero. Independente de sua ordem, seu determinante é sempre igual a 1. Se $k = 0$, a matriz é da forma padrão, se $k = 1$, os elementos se refletem sobre o anti-diagonal $(1, q) - (q, 1)$. Os autovalores de $\mathbf{A}$ podem ser obtidos em termos de zeros dos polinômios de Hermite. Eles são positivos e ocorrem em pares de inversos, assim se n é ímpar, 1 é um autovalor.

Um algoritmo que cria uma matriz Frank no MATLAB pode ser dado por:

```
if nargin == 1, k = 0; end
p = n:-1:1;
F = triu( p( ones(n,1), :) - diag( ones(n-1,1), -1), -1 );
if k ~= 0
   F = F(p,p)';
end
```

2. Matris Lesp - matriz tridiagonal com autovalores reais, sensíveis. O comando $\mathbf{B} = ('lesp', n)$ retorna uma matriz $n \times n$ cujos autovalores são reais e bem distribuídos no intervalo $[-2 \cdot n - 3.5, -4, 5]$.

A sensibilidade dos autovalores aumenta exponencialmente à medida que os autovalores crescem negativamente. A matriz $\mathbf{B}$ é semelhante à matriz simétrica tridiagonal com as mesmas entradas diagonais e fora da diagonal com entradas 1, através de uma transformação de semelhança com $\mathbf{D} = diag(1!, 2!, ..., n!)$.

Um algoritmo que calcula uma matriz Lesp no MATLAB pode ser dado por:

```
x = 2:n;
T = full(tridiag( ones(size(x))./x, -(2*[x n+1]+1), x));
```

3. Matriz Grcar - uma matriz de Toeplitz com autovalores sensíveis.

O comando $\mathbf{C} = grcar(n,k)$ retorna uma matriz $n \times n$ com -1 na primeira subdiagonal, 1 na diagonal, e 1 nas k superdiagonais. O padrão é $k = 3$. Os valores próprios desta matriz formam um padrão interessante no plano complexo, que veremos posteriormente no exemplo.

Um algoritmo que calcula uma matriz Grcar no MATLAB pode ser dado por:

```
if nargin == 1, k = 3; end
G = tril(triu(ones(n)), k) - diag(ones(n-1,1), -1);
```

4. Matriz Hanowa - matriz cujos autovalores ficam em uma linha vertical no plano complexo.

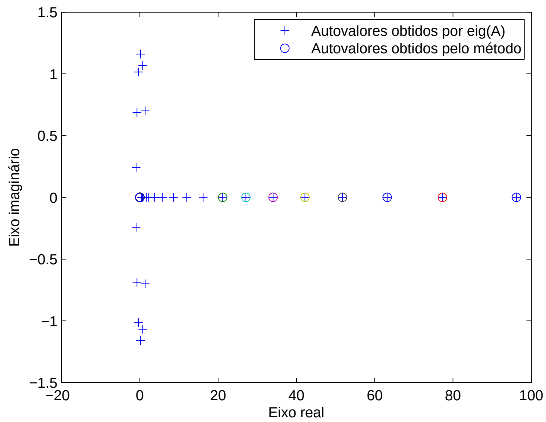
O comando $\mathbf{D} = ('hanowa', n, d)$ retorna um bloco $n \times n$ da forma:

```
[d*eye(m) -diag(1:m); diag(1:m) d*eye(m)]
```

O argumento n é um número inteiro par tal que $n = 2 \cdot m$. Matriz $\mathbf{D}$ tem autovalores complexos da forma $d \pm k \cdot i$, para $1 \leq k \leq m$. O valor padrão de d é -1.

Exemplo 13. $\mathbf{A} = gallery('frank', 30, 1)$

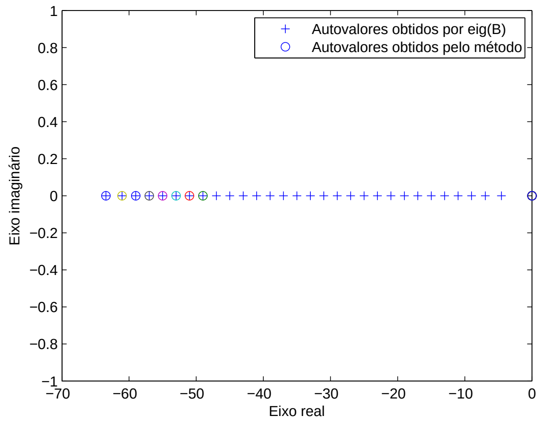
O resultado que obtivemos utilizando o método de iteração ortogonal, após 89 iterações, pode ser visto no gráfico que segue.



Podemos notar que os 8 autovalores que obtivemos pelo método de iteração ortognal coincidem com os 8 autovalores de maior módulo da matriz $\mathbf{A}$

Exemplo 14. $\mathbf{B} = gallery('lesp', 30)$

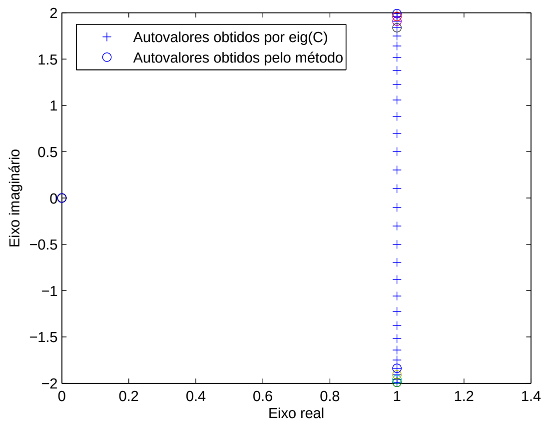
O resultado que obtivemos utilizando o método de iteração ortogonal, após 591 iterações, pode ser visto no gráfico que segue.



Apesar de ter a convergência um pouco mais lenta do que no exemplo anterior, podemos perceber que o método mais um vez chegou ao resultado esperado, convergindo para os 8 autovalores de maior módulo da matriz **B***.*

Exemplo 15. $\mathbf{C} = gallery('grcar'(n,k))$

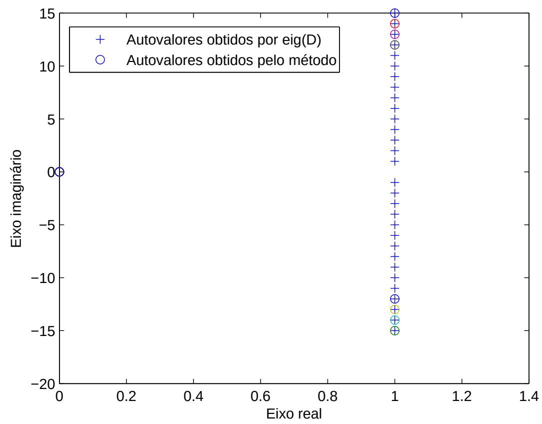
O resultado que obtivemos utilizando o método de iteração ortogonal, após 298 iterações, pode ser visto no gráfico que segue.



Com um número menor de iterações do que no exemplo anterior, conseguimos chegar ao resultado esperado mais uma vez.

Exemplo 16. $\mathbf{D} = gallery('hanowa', n, d)$

O resultado que obtivemos utilizando o método de iteração ortogonal, após 143 iterações, pode ser visto no gráfico que segue.



Mais uma vez, percebemos que o método convergiu para os 8 autovalores de maior médulo da matriz dada.

7.2.2 O método LOPSI

O código LOPSI desenvolvido por Stewart e Jennings [10], foi baseado no trabalho realizado em 1970 por Clint e Jennings. Ele utiliza iteração de subespaço combinada com uma projeção oblíqua assimétrica, lopsided em inglês, para calcular os p autovalores de maiores módulos e os autovetores correspondentes, referentes à uma matriz dada. O algoritmo utilizado pelo LOPSI, que é descrito em Stewart e Jennings [10], tem a seguinte estrutura geral

1. Escolhemos um subespaço de dimensão p representado por uma matriz $\mathbf{V}$ $n \times p$, como na iteração de subespaço.

2. Calcule $\mathbf{Z} = \mathbf{AV}$.

3. Projeção oblíqua: resolva o sistema linear

$$\mathbf{X} = (\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\mathbf{Z}.$$

4. Encontre, usando a função $eig(X)$ do MATLAB, os autovalores de $\mathbf{X}$. Em seguida, ordene os autovalores e os correspondentes autovetores em ordem decrescente de valores absolutos. Defina como *vet* a matriz cujas colunas são esses autovetores.

5. Defina agora $\mathbf{Z} = \mathbf{Zvet}$.

6. Para controlar as entradas da matriz $\mathbf{Z}$, fazemos a correção $\mathbf{V} = \frac{\mathbf{Z}}{\|\mathbf{Z}\|_F}$

7. Se o erro não for menor que a tolerância exigida, retorne à etapa (2).

O seguinte argumento mostra que os métodos de iteração de subespaço e LOPSI são equivalentes.

Seja $\mathbf{A} \in \mathbb{C}^{n \times n}$. Consideremos uma iteração de LOPSI:

$$\begin{aligned} \mathbf{Z} &= \mathbf{AV} \quad (7.6)\\ \mathbf{X} &= (\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H\mathbf{Z} \end{aligned}$$

Se $\mathbf{V} = \mathbf{QR}$ é a decomposição QR de $\mathbf{V}$, então temos que $\mathbf{Z} = \mathbf{AQR}$ e

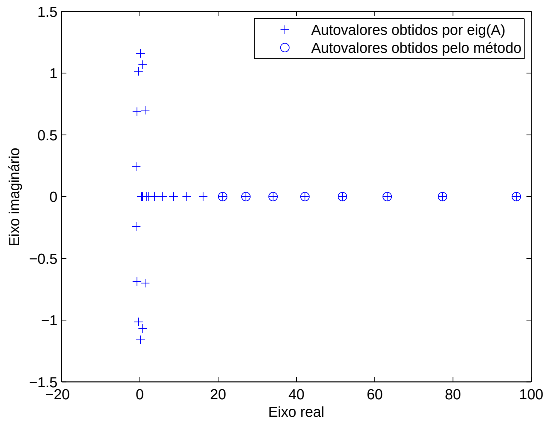
$$\begin{aligned} \mathbf{X} &= (\mathbf{R}^H\mathbf{Q}^H\mathbf{QR})^{-1}\mathbf{R}^H\mathbf{Q}^H\mathbf{AQR} \quad (7.7)\\ &= \mathbf{R}^{-1}(\mathbf{R}^H)^{-1}\mathbf{R}^H\mathbf{Q}^H\mathbf{AQR} \quad (7.8)\\ &= \mathbf{R}^{-1}(\mathbf{Q}H\mathbf{AQ})\mathbf{R}. \quad (7.9) \end{aligned}$$

Segue que a matriz $\mathbf{X}$ gerada pelo LOPSI e a matriz $\mathbf{Q}^T\mathbf{AQ}$ são equivalentes, e portanto têm os mesmos autovalores. Assim, em cada etapa dos métodos obtemos os mesmos autovalores.

Vamos expor agora resultados de experimentos nos quais utilizamos o Método LOPSI. As matrizes escolhidas para gerar tais exemplos são exatamente as mesmas utilizadas nos testes do Método de Iteração Ortogonal, citados anteriormente.

Exemplo 17. $\mathbf{A} = gallery('frank', 30, 1)$

O resultado que obtivemos utilizando o Método LOPSI, pode ser visto no gráfico que segue.



Podemos perceber na figura que o objetivo foi alcançado, porém foram efetuadas 1500 iterações. Aumentando a tolerância de 10^{-10} *para* 10^{-4}*, conseguimos chegar ao objetivo com apenas 48 iterações.*

Exemplo 18. $\mathbf{B} = gallery('lesp', 30)$

Os resultados que obtivemos utilizando o Método LOPSI, pode ser visto no gráfico que segue.

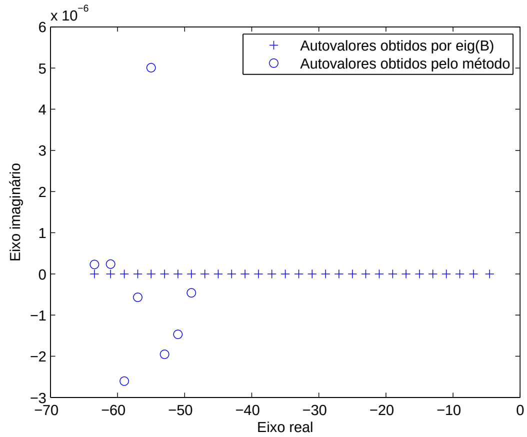


Figura 1:

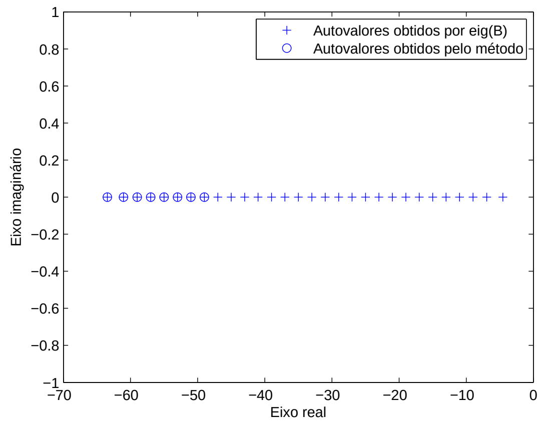
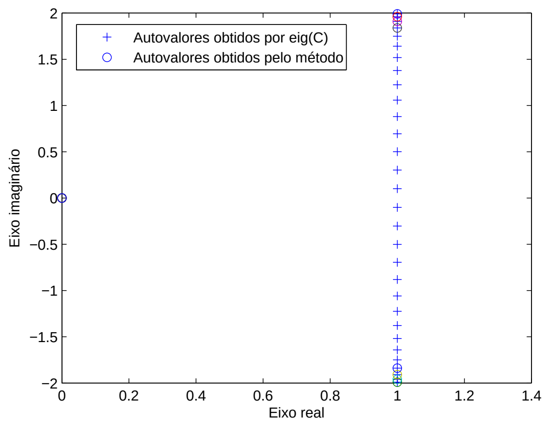


Figura 2:

Na Figura 1, o número de iterações realizadas chegou a 1500 e os autovalores obtidos pelo método, não coincidiram exatamente com os 8 autovalores de maior módulo da matriz, mais especificamente, coincidiram apenas na parte real e não na parte imaginária. Na Figura 2 fizemos um teste, aumentando a tolerância para 10^{-4} *e iniciando o método não mais com um matriz arbitrária do MATLAB, e sim com uma matriz formada pelas 8 primeiras colunas da própria matriz* **A**. *Aassim obtivemos nosso resultado esperado, após 286 iterações.*

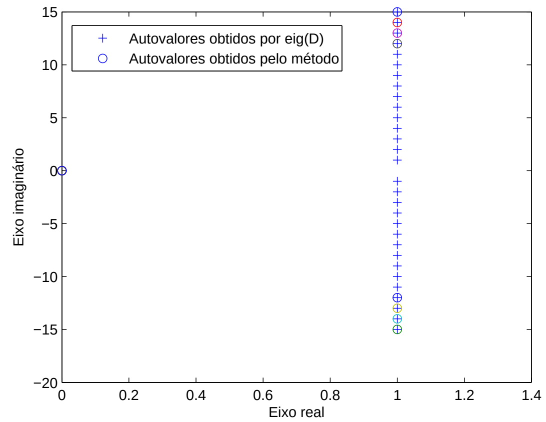
Exemplo 19. $\mathbf{C} = gallery('grcar'(n,k))$

O resultado que obtivemos utilizando o Método LOPSI, após 411 iterações, pode ser visto no gráfico que segue.



Exemplo 20. $\mathbf{D} = gallery('hanowa', n, d)$

O resultado que obtivemos utilizando o método de iteração ortogonal, após 131 iterações, pode ser visto no gráfico que segue.



Apêndice

O objetivo deste apêndice é provar o teorema abaixo.

Teorema 48. *Seja N um operador nilpotente em um espaço vetorial real ou complexo, V. Então V possui uma base $\mathfrak{B}$ tal que a matriz $\mathbf{A} = [N]_{\mathfrak{B}}$ é da forma*

$$\mathbf{A} = diag\{\mathbf{A}_1, \ldots \mathbf{A}_r\},$$

em que $\mathbf{A}_i$ é um bloco nilpotente elementar, e a dimensão de $\mathbf{A}_k$ é uma função de k não crescente. As matrizes $\mathbf{A}_1, \ldots, \mathbf{A}_r$ são unicamente depterminadas pelo operador N.

Nesta seção, V representará um espaço vetorial real ou complexo.

Um subespaço $W \subset V$ é dito um subespaço cíclico de um operador T em V se $T(W) \subset W$ e se existir um vetor $\mathbf{v} \in W$ tal que W é gerado por $T^n(\mathbf{v})$, $n \in \mathbb{N}$. Neste caso, chamamos $\mathbf{v}$ de um vetor cíclico de W.

Qualquer vetor $\mathbf{v}$ gera um subespaço cíclico pelas iterações de $\mathbf{v}$ por T, isto é, $\mathbf{v}$, $T(\mathbf{v})$, $T^2(\mathbf{v})$, ..., gera um subespaço que evidentemente é cíclico. Denotamos este subespaço por $Z(\mathbf{v})$ ou $Z(\mathbf{v}, T)$.

Suponha que $N : V \to V$ é um operador nilpotente. Para cada $\mathbf{v} \in V$ existe um menor inteiro positivo n, denotado por $nil(\mathbf{v})$ ou $nil(\mathbf{v}, N)$, tal que $N^n(\mathbf{v}) = 0$. Se $\mathbf{v} \neq 0$, então $N + (\mathbf{v}) \neq 0$ para $0 \leq k < nil(\mathbf{v})$.

Lema 6. *Seja $nil(\mathbf{v}, N) = n$. Então os vetores $N^k(\mathbf{v})$, $0 \leq k \leq n-1$ formam uma base para $Z(\mathbf{v}, N)$.*

Demonstração: É claro que os vetores $N^k(\mathbf{v})$, $0 \leq k \leq n-1$ geram $Z(\mathbf{v})$. Se eles fossem linearmente dependentes, então existiria uma relação

$$\sum_{k=0}^{n-1} a_k N^k(\mathbf{v}) = 0$$

na qual nem todos os a_k fossem nulos. Considere j como sendo o menor índice, $0 \leq j \leq n-1$, tal que $a_j = 0$. Então

$$\begin{aligned} 0 &= N^{n-j-1}\left(\sum_{k=j}^{n-1} a_k N^k(\mathbf{v})\right) \\ &= \sum_{k=j}^{n-1} a_k N^{n+k-j-1}(\mathbf{v}) \\ &= a_j N^{n-1}(\mathbf{v}) + \sum_{k=j+1}^{n-1} a_k N^{n+k-j-1}(\mathbf{v}) \\ &= a_j N^{n-1}(\mathbf{v}) \end{aligned}$$

considerando que $n+k-j-1 \geq n$ se $k \geq j+1$. Então $a_j N^{n-1}(\mathbf{v}) = 0$ e asim $N^{n-1}(\mathbf{v}) = 0$ pois $a_j \neq 0$. Mas isto contradiz o fato que $n = nil(\mathbf{v}, N)$. Logo, os vetores $N^k(\mathbf{v})$, $0 \leq k \leq n-1$ formam uma base para $Z(\mathbf{v}, N)$. ∎

Este resultado prova que com relação à base $\{\mathbf{v}, N(\mathbf{v}), \ldots, N^{n-1}(\mathbf{v})\}$, com $n = nil(\mathbf{v})$, o operador $N|Z(\mathbf{v})$ é representado pela matriz

$$\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$$

com 1's na subdiagonal abaixo da diagonal principal e zeros nas outras entradas. Daqui é de onde provêm os 1's abaixo da diagonal na forma canônica de Jordan.

Um argumento similar para provar o Lema (6) mostra que se $\sum_{k=0}^{q} a_k N^k(\mathbf{v})$, então $a_k = 0$ para $k < nil(\mathbf{v}, N)$.

É conveniente adotarmos a notação $p(T)$ para denotar o operador $\sum_{k=0}^{q} a_k T^k$ se p for o polinômio

$$p(t) = a_q t^q + \ldots + a_1 t + a_0.$$

Então a afirmação provado acima pode ser enunciado sa seguinte forma:

Lema 7. *Seja $n = nil(\mathbf{v}, N)$. Se $p(t)$ é um polinômio tal que $p(N)\mathbf{v} = 0$ então t^n divide $p(t)$, isto é, existe um polinômio $p_1(t)$ tal que $p(t) = t^n p_1(t)$.*

vamos agora provar a existência da forma canônica para um operador nilpotente N. Considerando a matriz discutida acima para $N|_{Z(x)}$, a prova da forma cnônica se reduz a provar:

Proposição 20. *Seja $N : V \to V$ um operador nilpotente. Então V é uma soma direta de* *subespaços cíclicos.*

Demonstração: Vamos fazer esta prova por indução sobre $dim(V)$. O caso $dim(V) = 0$ é trivial. Se $dim(V) > 0$, então $N(V) < dim(V)$, desde que N possua um núcleo não trivial. Portanto existem vetores não nulos $\mathbf{y}_1, \ldots, \mathbf{y}_r \in N(V)$ tais que

$$N(V) = Z(\mathbf{y}_1) \oplus \ldots \oplus Z\mathbf{y}_r.$$

Seja $x_i \in V$ um vetor não nulo tal que

$$N(x_i) = y_i, \quad j = 1, \ldots, r.$$

Nós provamos que os subespaços $Z(x_1), \ldots, Z(x_r)$ são independentes.

Observe que $nil(x_i) \geq 2$, desde que

$$N(x_i) = \mathbf{y}_i \neq 0.$$

Se os subespaços $Z(x_i)$ não são independentes, existem vetores $\mathbf{u}_j \in Z(x_i)$, nem todos nulos, tais que $\sum_{j=1}^{r} \mathbf{u}_j = 0$. Portanto, $\sum_j N(\mathbf{u}_j) = 0$. Como $N(\mathbf{u}_j \in N(Z(x_i)) = Z\mathbf{y}_i$ e $Z(\mathbf{y}_i)$ são independentes por hipótese, segue que $\mathbf{u}_j$ pertence ao núclo de N, para $j = 1, \ldots, r$. Agora, cada $\mathbf{u}_j$ tem a forma

$$\sum_{k=0}^{n_j-1} a_{jk} N^k(x_j), \quad n_j = nil(x_j).$$

Consequentemente $\mathbf{u}_j = p_j(N)x_j$ sendo p o polinômio $p_j(t) = \sum_{k=0}^{n_j-1} a_{jk}t^k$. Portanto $N(\mathbf{u}_j = p_j(N)\mathbf{y}_j = 0$. Pelo Lema(7), $p_j(t)$ é divisível por t^m se $m \leq nil(\mathbf{y}_j)$. Como $1 \leq nil(\mathbf{y}_j)$, podemos escrever

$$p_j(t) = s_j(t)t$$

para algum polinômio $s_j(t)$.

Mas agora, substituindo N por t, teremos

$$\begin{aligned} \mathbf{u}_j &= s_j(N)N(x_i) \\ &= s_j(N)\mathbf{y}_i \in Z(\mathbf{y}_i). \end{aligned}$$

Portanto $\mathbf{u}_j = \mathbf{0}$ se os $Z(\mathbf{y}_j)$ são independentes.

Agora vamos mostrar que

$$V = Z(x_1) \oplus \ldots \oplus Z(x_r) \oplus L \tag{7.1}$$

com L pertencente ao núcleo de N. Seja K o núcleo de N e seja L um subespaço de K tal que

$$K = (K \bigcap N(V)) \oplus L.$$

Então L é independente com os $Z(x_i)$. Para ver isso, tome $\mathbf{v} \in (Z(x_1) \oplus \ldots \oplus Z(x_r)) \bigcap L$. Então $\mathbf{v} \in (Z(x_1) \oplus \ldots \oplus Z(x_r)) \bigcap K$, e por um argmento similar ao citado anteriormente, isso implica que $\mathbf{v} \in N(V)$. Mas $N(V) \bigcap L = \mathbf{0}$, logo, $\mathbf{v} = \mathbf{0}$. é claro que cada subespaço cíclico em K, e portanto em L, é unidimensional. Portanto $L = Z(\mathbf{w}_1) \oplus \ldots \oplus Z(\mathbf{w}_s$, em que $[\mathbf{w}_1, \ldots, \mathbf{w}_s]$ é uma base de L. Finalmente,

$$V = Z(x_1) \oplus \ldots \oplus Z(x_r) \oplus Z(\mathbf{w}_1) \oplus \ldots \oplus Z(\mathbf{w}_s.$$

Esta proposição implica o teorema, exceto pela unicidade das matrizes $\mathbf{A}_1, \ldots, \mathbf{A}_r$. A unicidade é equivalente à afirmação que o operador N determina o tamanho dos blocos $\mathbf{A}_i$ (ou as dimensões dos subespaços cíclicos). Isto pode ser feito por indução em $dim(V)$.

Considere a restrição de N à sua imagem $N(V) = F$:

$$N|_F : F \to F.$$

Seja V a soma direta de subespaços cíclicos, isto é, $V = Z_1 \oplus \ldots \oplus Z_r \oplus W_1$, em que W_1 pertence ao núcleo de N e cada $Z_k = [x_k, Ax_k, A^2x_k, \ldots]$, $k = 1 : r$, com $dim(Z_k) > 1$. Então $N(V)$ é a soma direta

$$N(Z_1) \oplus \ldots \oplus N(Z_r),$$

em que $N(Z_k)$ é cíclico e $dim(N(Z_k)) = dim(Z_k) - 1$. Como $dim(N(F)) < dim(V)$, os números $\{dim(Z_k) - 1\}$ são determinados por $N|_F$. Portanto, são determinados por N. Segue que os números $\{dim(Z_k)\}$ são também determinados por N.

Isto conclui a prova do teorema.

$\blacksquare$

8 Considerações Finais

Analisados os resultados dos experimentos numéricos, pudemos perceber que os métodos de iteração de subespaço funcionam bem com certas matrizes com autovalores sensíveis. Apesar de que o método de Iteração Ortogonal e o Método LOPSI são equivalentes como provamos, temos que o primeiro é computacionalmente mais caro do que o segundo, visto que a Iteração Ortogonal utiliza a fatoração **QR** que é mais cara do que a resolução do sistema linear utilizado no Método LOPSI. Em contrapartida, como o método LOPSI resolve um sistema linear, podemos ter problemas com matrizes muito próximas de se tornarem singulares.

Referências

[1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe e H. van der Vorst, Templates for the solution of Algebraic eigenvalue Problems: A Practical Guide, Philadelphia: SIAM, 2000.

[2] F. L. Bauer. Das Verfahren der Treppeniteration und verwandte Verfahren zur Losung algebraischer Eigenwertprobleme. Zeitschrift fiir angewandte Mathematik und Physik, 8:214-235, 1957.

[3] L. H. Bezerra, Matrizes Especiais em Matemática Numérica, 28^o Colóquio Brasileiro de Matemática, Rio de Janeiro: IMPA, 2011.

[4] M. Fiedler,Special matrices and their applications in numerical mathematics, Publishers of Technical Literature, Prague, Czecho- slovakia, 1986.

[5] G. H. Golub e C. F. Van Loan, Matrix Computations, 3. ed., Baltimore: Johns Hopkins University Press, 1996.

[6] N. J. Higham. The Matrix Computation Toolbox. http://www.ma.man.ac.uk/ higham/mctoolbox.

[7] M. W. Hirsch and S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra, San Diego: Academic Press,

[8] K. Hofmman, R. Kunze, Álgebra Linear, Livros Técnicos e Científicos Editora, Rio de Janeiro, 1976.

[9] G. W. Stewart, Matrix Algorithms Volume II: Eigensystems, University of Maryland, College Park, Maryland, SIAM, Philadelphia, 2001.

[10] W. J. Stewart and A. Jennings. Algorithm 570: LOPSI: A Simultaneous Iteration Method for Real Matrices. ACM Trans. Math. Software, 1981.

[11] G. Strang, Linear Algebra and its Aplications, Academic Press, 1992.