Universidade Federal de Uberlândia

Faculdade de Matemática

Número de Condição de Matrizes

Ester Ferreira Dias
Orientando

Santos Alberto Enriquez-Remigio
Orientador

Uberlândia - MG

2020

Ester Ferreira Dias

NÚMERO DE CONDIÇÃO DE MATRIZES

Trabalho de Conclusão de Curso apresentado à Universidade Federal de Uberlândia (UFU), como parte das exigências para a obtenção do título de bacharel em Matemática.

Orientador: Santos Alberto Enriquez-Remigio.

Uberlândia

2020

Ester Ferreira Dias

NÚMERO DE CONDIÇÃO DE MATRIZES

Trabalho de Conclusão de Curso apresentado à Universidade Federal de Uberlândia (UFU), como parte das exigências para a obtenção do título de bacharel em Matemática.

Uberlândia, ____de____________de______

BANCA EXAMINADORA

Prof. Dr. Santos Alberto Enriquez-Remigio
Universidade Federal de Uberlândia (UFU)

Prof. Dr. Jocelino Sato
Universidade Federal de Uberlândia (UFU)

Prof. Dr. Clair do Nascimento
Universidade Federal de Uberlândia (UFU)

Uberlândia

2020

Agradecimentos

Primeiramente, gostaria de agradecer a Deus, pois sem Ele não estaria aqui. Ele me dá a força e a inteligência necessária para passar por uma vida de estudos e conquistar meus objetivos.

Aos meus pais José Elismar e Maria Perpétua e aos meus irmãos Júnior, Jônatas e Emanuel, minha cunhada Thaís e meus sobrinhos Valentina, Paulo, Ana Catarina e Isabela por me darem apoio em continuar meus estudos.

Ao meu namorado Maicon e minha sogra Ivone, por estarem me dando apoio e suporte, principalmente durante este meu último ano de faculdade.

Ao meu orientador, professor Santos, pela paciência de me encaminhar até aqui desde o início do curso de graduação na faculdade.

Aos professores que compõe a banca, Clair e Sato, pela paciência e compreensão que me permitiram melhor apresentação do meu trabalho.

Por fim, aos inúmeros professores que fizeram parte da minha formação durante estes 5 anos de graduação.

"Dê-me, Senhor, agudeza para entender, capacidade para reter, método e faculdade para aprender, sutileza para interpretar, graça e abundância para falar. Dê-me, Senhor, acerto ao começar, direção ao progredir e perfeição ao concluir."

Santo Tomás de Aquino (1225-1274)

Resumo

A solução de um sistema linear de equações $A\vec{x} = \vec{b}$ pode ser fortemente sensível a pequenas perturbações nos dados de entrada da matriz A e/ou do vetor $\vec{b}$. Quando isso acontece, diz-se que o sistema é mal condicionado. Essa sensibilidade está associada ao número de condição da matriz A, $cond(A)$. Quando $cond(A) \approx 1$ o suficiente, o sistema linear é bem condicionado e quando $cond(A) \gg 1$, o sistema é mal condicionado.

No presente trabalho, apresentam-se limitadores superiores (reportados na literatura) para os erros relativos à solução $\vec{x}$, quando: 1) a matriz A é perturbada; 2) o vetor $\vec{b}$ é perturbado; e 3) a matriz A e o vetor $\vec{b}$ são perturbados; assim como todos os conceitos matemáticos necessários, tais como: vetores, matrizes e normas, que levam à dedução matemática desses limitadores. Os respectivos limitadores superiores mostram que o erro relativo em $\vec{x}$, depende do número de condição da matriz A.

Abstract

The solution of a linear system of equations $A\vec{x} = \vec{b}$ can be strongly sensitive to small perturbations in the input data of matrix A and/or vector $\vec{b}$. When this happens, the system is said to be poorly conditioned. This sensitivity is associated with the condition number of matrix A, $cond(A)$. When $cond(A) \approx 1$ is enough, the linear system is well conditioned and when $cond(A) \gg 1$, the system is poorly conditioned.

In the present work, upper limiters (reported in the literature) are presented for the errors related to solution $\vec{x}$, when: 1) matrix A is perturbed; 2) vector $\vec{b}$ is perturbed; and 3) matrix A and vector $\vec{b}$ are perturbed; as well as all the necessary mathematical concepts that lead to the mathematical deduction of these limiters. These constraints indicate that the upper bound for the relative error in $\vec{x}$ depends on the condition number of matrix A.

Sumário

Lista de Figuras

Introdução

No ensino de Cálculo Numérico, dos cursos de graduação, poucas vezes é abordado o assunto do condicionamento de matrizes. Mas condicionamento de matrizes é um tema muito importante, pois, na prática, uma matriz bem condicionada permite que a solução do sistema linear, onde ela aparece, não seja fortemente distante da solução teórica esperada. Mas por que a solução do sistema linear poderia ser distante? Porque na hora do processo de resolução do sistema linear $A\vec{x} = \vec{b}$, resolve-se na verdade um sistema novo $\widehat{A}\vec{\widehat{x}} = \vec{\widehat{b}}$, onde $\widehat{A}$ e $\vec{\widehat{b}}$ são aproximações para A e $\vec{b}$, respectivamente, assim será visto que dependendo das características da matriz A, as soluções $\vec{x}$ e $\vec{\widehat{x}}$ podem ou não ser bem distantes para pequenas diferenças entre os pares $(A, \vec{b})$ e $(\widehat{A}, \vec{\widehat{b}})$.

O estudo ora apresentado considera vetores e matrizes com entradas no conjunto dos números reais. A metodologia usada para o desenvolvimento deste trabalho consistiu em:

1. procurar outros TCC's relacionados com número de condição, para, assim, descobrir estudos que pudessem ajudar a orientar e/ou mudar os objetivos de nossa pesquisa, mas pouco foi encontrado. Contudo, essa busca proporcionou o conhecimento de várias referências de livros em inglês;

2. investigar o conceito de número de condicionamento em livros de cálculo numérico e em materiais disponíveis na internet, publicados na Língua Portuguesa. Os materiais nos quais foram encontradas as informações do número de condição foram: [4, 6, 8] e [9];

3. estudar e/ou revisar conceitos referentes a vetores, matrizes, autovalores, normas, dentre outros assuntos para o entendimento do conceito de condicionamento e de suas propriedades;

4. explorar o conceito de número de condição de uma matriz e suas propriedades, etc.

Os objetivos do presente trabalho são:

1. apresentar uma revisão dos conceitos de vetores e matrizes e suas propriedades, autovalor e autovetor, norma de vetores e matrizes;

2. apresentar o conceito de número de condição de matrizes e exibir limitadores superiores para o erro relativo em $\vec{x}$ da solução exata do sistema linear $A\vec{x} = \vec{b}$, em função do número de condicionamento da matriz e de seus erros relativos em seus dados de entrada A e $\vec{b}$;

3. mostrar exemplos de matrizes bem condicionadas e mal condicionadas e um exemplo de matriz onde vê-se que o número de condição depende da norma escolhida.

Para apresentar o conceito do número de condição de uma matriz e suas propriedades é necessário conhecer suas definições e propriedades relacionadas a: vetores em $\mathbb{R}^n$ e matrizes em $\mathbb{R}^{m\times n}$, aplicações de matrizes e normas. Por conseguinte, este trabalho dividiu-se em quatro capítulos.

No Capítulo 1, apresentam-se dois elementos matemáticos bastante usados na prática: Vetores em $\mathbb{R}^n$ e Matrizes em $\mathbb{R}^{m\times n}$. Também apontam-se conceitos e propriedades relacionados a Autovalores e Autovetores. Quase todas as propriedades apresentadas são demonstradas.

No Capítulo 2, apresentam-se duas aplicações de matrizes, ambas associadas com o número de condição de uma matriz, sendo elas: Transformações Lineares e Sistemas Lineares.

Noções de normas de vetores e de matrizes, assim como propriedades e alguns exemplos, são apresentados no Capítulo 3. E, finalmente, o Capítulo 4 refere-se ao número de condicionamento de matrizes, demonstrando-se limitadores superiores para os erros relativos à solução de $A\vec{x} = \vec{b}$, quando: a matriz A é perturbada, o vetor $\vec{b}$ é perturbado e a matriz A e o vetor $\vec{b}$ são perturbados.

Capítulo 1

Vetores em $\mathbb{R}^n$ e matrizes em $\mathbb{R}^{m\times n}$

Neste capítulo são apresentados dois elementos matemáticos bastante usados na prática: vetor em $\mathbb{R}^n$ e matrizes em $\mathbb{R}^{m\times n}$.

Embora a palavra vetor tenha um significado bem amplo em Álgebra Linear, aqui não será dado esse enfoque; mas, um vetor será entendido como uma lista de números para a qual definem-se duas operações, as quais verificam um conjunto de propriedades.

No caso de matrizes, estas serão focadas em matrizes com elementos ou entradas em $\mathbb{R}$, nas suas operações matriciais e suas respectivas propriedades.

Neste capítulo, será também apresentado o conceito de autovalor e autovetor, assim como as propriedades associadas aos autovalores e autovetores de matrizes de $\mathbb{R}^{m\times n}$.

Para o desenvolvimento deste capítulo, seguimos o apresentado em [1], [2] e [3].

1.1 Vetores do $\mathbb{R}^n$

Definição 1.1.1. *Um* ***vetor*** *do $\mathbb{R}^n$ é uma lista ordenada de n números reais. Dizemos que é uma n-upla de números reais.*

Neste texto, denotaremos vetores com letras minúsculas com uma setinha em cima, como em $\vec{u},\ \vec{v},\ \vec{w}$.

A notação utilizada para representar um vetor é $\vec{u} = (u_1, u_2, \ldots, u_{n-1}, u_n)$ com $u_i \in \mathbb{R}$. Os números u_i's são chamados de **entradas** do vetor $\vec{u}$.

O $\mathbb{R}^n$ é o conjunto de n-uplas ordenadas de números reais.

Exemplo 1.1. *São vetores de $\mathbb{R}^2$: $(-6,-8), (1,2)$; vetores de $\mathbb{R}^4$: $(1,2,3,4)$, $\left(-2, \frac{7}{4}, -1, \frac{2}{3}\right)$ e, vetores de $\mathbb{R}^5$: $(-1,2,4,6,8)$, $\left(1, 2, \frac{7}{4}, -\frac{1}{3}, 3\right)$.*

Observação 1.1.1. *Como um vetor é uma lista ordenada de números, então são distintos entre si os vetores $(-1,2)$ e $(-2,1)$ e também $(1,2,3)$, $(2,3,1)$ e $(3,1,2)$.*

Definição 1.1.2. *Se $\vec{v} = (v_1, v_2, \ldots, v_n)$ e $\vec{w} = (w_1, w_2, \ldots, w_n)$ são vetores de $\mathbb{R}^n$ e se k é um número real qualquer, definimos as seguintes operações:*

$$\vec{v} + \vec{w} = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \quad \textbf{\textit{(Soma de vetores)}} \tag{1.1}$$

$$k\vec{v} = (kv_1, kv_2, \ldots, kv_n) \quad \textbf{\textit{(Multiplicação de um vetor por um número real)}} \tag{1.2}$$

$$-\vec{v} = (-v_1, -v_2, \ldots, -v_n) \quad \textbf{\textit{(Vetor oposto de } \vec{v}\textbf{\textit{)}} \tag{1.3}$$

$$\vec{w} - \vec{v} = \vec{w} + (-\vec{v}) = (w_1 - v_1, w_2 - v_2, \ldots, w_n - v_n) \quad \textbf{\textit{(Diferença de vetores)}} \tag{1.4}$$

Exemplo 1.2. *A soma dos vetores do $\mathbb{R}^4$*

$$\left(1, -1, \frac{1}{4}, -\frac{2}{3}\right) + \left(-2, 2, \frac{3}{4}, \frac{5}{3}\right) = \left(1 - 2, -1 + 2, \frac{1}{4} + \frac{3}{4}, -\frac{2}{3} + \frac{5}{3}\right) = (-1, 1, 1, 1).$$

Exemplo 1.3. *Se $\vec{u} = \left(-1, 3, 1, -2, \frac{3}{2}\right)$, então*

$$2\vec{u} = 2\left(-1, 3, 1, -2, \frac{3}{2}\right) = (-2, 6, 2, -4, 3).$$

Considere $\vec{w} = (-4, 6, 1, -3)$. Então,

$$-\frac{1}{2}\vec{w} = -\frac{1}{2}(-4, 6, 1, -3) = \left(2, -3, -\frac{1}{2}, \frac{3}{2}\right).$$

Teorema 1.1.1. *Se $\vec{u}$, $\vec{v}$ e $\vec{w}$ são vetores de $\mathbb{R}^n$ e se k e m são números reais, então:*

(a) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ ***[Lei da comutatividade da adição]***

(b) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ ***[Lei da associatividade da adição]***

(c) $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$ ***[Elemento neutro da adição]***

(d) $\vec{u} + (-\vec{u}) = \vec{0}$ ***[Elemento oposto]***

(e) $k(\vec{v} + \vec{w}) = k\vec{v} + k\vec{w}$ ***[Lei da distributividade com relação a escalares]***

(f) $(k + m)\vec{v} = k\vec{v} + m\vec{v}$ ***[Lei da distributividade com relação a vetores]***

(g) $k(m\vec{u}) = (km)\vec{u}$ ***[Lei da associatividade da multiplicação]***

(h) $1\vec{u} = \vec{u}$ ***[Elemento neutro da multiplicação]***

Demonstração. As provas dessas propriedades usam as propriedades de mesmo nome relativas aos números reais. Temos,

(a) Sejam $\vec{u} = (u_1, u_2, \ldots, u_n)$ e $\vec{v} = (v_1, v_2, \ldots, v_n)$. Então

$$\begin{aligned}
\vec{u} + \vec{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) \\
&= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \\
&= (v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n) \\
&= (v_1, v_2, \ldots, v_n) + (u_1, u_2, \ldots, u_n) \\
&= \vec{v} + \vec{u}.
\end{aligned}$$

(b) Sejam $\vec{u} = (u_1, u_2, \ldots, u_n)$, $\vec{v} = (v_1, v_2, \ldots, v_n)$ e $\vec{w} = (w_1, w_2, \ldots, w_n)$. Então

$$\begin{aligned}
(\vec{u} + \vec{v}) + \vec{w} &= ((u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n)) + (w_1, w_2, \ldots, w_n) \\
&= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) + (w_1, w_2, \ldots, w_n) \\
&= ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n) \\
&= (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n)) \\
&= (u_1, u_2, \ldots, u_n) + (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \\
&= (u_1, u_2, \ldots, u_n) + ((v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n)) \\
&= \vec{u} + (\vec{v} + \vec{w}).
\end{aligned}$$

(c) Sejam $\vec{u} = (u_1, u_2, \ldots, u_n)$ e $\vec{0} = (0, 0, \ldots, 0)$. Então

$$\begin{aligned}
\vec{u} + \vec{0} &= (u_1, u_2, \ldots, u_n) + (0, 0, \ldots, 0) \\
&= (u_1 + 0, u_2 + 0, \ldots, u_n + 0) \\
&= (u_1, u_2, \ldots, u_n) \\
&= \vec{u}.
\end{aligned}$$

(d) Seja $\vec{u} = (u_1, u_2, \ldots, u_n)$. Então

$$\begin{aligned}
\vec{u} + (-\vec{u}) &= (u_1, u_2, \ldots, u_n) + (-u_1, -u_2, \ldots, -u_n) \\
&= (u_1 - u_1, u_2 - u_2, \ldots, u_n - u_n) \\
&= (0, 0, \ldots, 0) \\
&= \vec{0}.
\end{aligned}$$

(e) Sejam $\vec{v} = (v_1, v_2, \ldots, v_n)$, $\vec{w} = (w_1, w_2, \ldots, w_n)$ e k um escalar. Então

$$\begin{aligned}
k(\vec{v}+\vec{w}) &= k((v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n)) \\
&= k(v_1+w_1, v_2+w_2, \ldots, v_n+w_n) \\
&= (k(v_1+w_1), k(v_2+w_2), \ldots, k(v_n+w_n)) \\
&= (kv_1+kw_1, kv_2+kw_2, \ldots, kv_n+kw_n) \\
&= (kv_1, kv_2, \ldots, kv_n) + (kw_1, kw_2, \ldots, kw_n) \\
&= k(v_1, v_2, \ldots, v_n) + k(w_1, w_2, \ldots, w_n) \\
&= k\vec{v} + k\vec{w}.
\end{aligned}$$

(f) Seja $\vec{v} = (v_1, v_2, \ldots, v_n)$ um vetor de $\mathbb{R}^n$ e, k e m escalares. Então

$$\begin{aligned}
(k+m)\vec{v} &= (k+m)(v_1, v_2, \ldots, v_n) \\
&= ((k+m)v_1, (k+m)v_2, \ldots, (k+m)v_n) \\
&= (kv_1+mv_1, kv_2+mv_2, \ldots, kv_n+mv_n) \\
&= (kv_1, kv_2, \ldots, kv_n) + (mv_1, mv_2, \ldots, mv_n) \\
&= k(v_1, v_2, \ldots, v_n) + m(v_1, v_2, \ldots, v_n) \\
&= k\vec{v} + m\vec{v}.
\end{aligned}$$

(g) Sejam $\vec{u} = (u_1, u_2, \ldots, u_n)$ e $k,\ m$ escalares. Então

$$\begin{aligned}
k(m\vec{u}) &= k(m(u_1, u_2, \ldots, u_n)) \\
&= k(mu_1, mu_2, \ldots, mu_n) \\
&= (k(mu_1), k(mu_2), \ldots, k(mu_n)) \\
&= ((km)u_1, (km)u_2, \ldots, (km)u_n) \\
&= (km)(u_1, u_2, \ldots, u_n) = (km)\vec{u}.
\end{aligned}$$

(h) Seja $\vec{u} = (u_1, u_2, \ldots, u_n)$. Então

$$\begin{aligned}
1\vec{u} &= 1(u_1, u_2, \ldots, u_n) \\
&= (1\cdot u_1, 1\cdot u_2, \ldots, 1\cdot u_n) \\
&= (u_1, u_2, \ldots, u_n) = \vec{u}.
\end{aligned}$$

∎

Definição 1.1.3. *Dizemos que $\vec{v}$ é **múltiplo escalar de** (ou **paralelo a**) $\vec{w}$ se existe um escalar t tal que $\vec{v} = t\vec{w}$.*

Exemplo 1.4. *São paralelos entre si: $(-2, 4, -6, 1)$ e $\left(1, -2, 3, -\frac{1}{2}\right)$ pois*

$$(-2, 4, -6, 1) = -2\left(1, -2, 3, -\frac{1}{2}\right)$$

ou de forma equivalente

$$\left(1, -2, 3, -\frac{1}{2}\right) = -\frac{1}{2}(-2, 4, -6, 1).$$

Exemplo 1.5. *O vetor $\vec{0}$ é paralelo ou múltiplo de qualquer outro vetor, pois $\vec{0} = 0\vec{w}$.*

1.2 Matrizes em $\mathbb{R}^{m\times n}$

Definição 1.2.1. *Uma **matriz** de ordem $m \times n$ é um arranjo de $m \cdot n$ números, disposta em m linhas e n colunas. Toda matriz tem associada um nome que pode ser representada com uma letra maiúscula, por exemplo: A. Para representar os elementos de A escrevemos $A = (a_{ij})$, onde $a_{ij} \in \mathbb{R}$ $(i = 1, \ldots, m$ e $j = 1, \ldots, n)$ ou*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1.5}$$

Embora a definição retrate de uma matriz específica, **matriz real**, vale ressaltar que os elementos da matriz podem ser elementos definidos no conjunto K, onde $K = \mathbb{C}$ (Conjunto dos números complexos), $\mathbb{Q}$ (Conjunto dos números racionais), $\mathbb{Z}$ (Conjunto dos números inteiros), etc.

Uma matriz com somente uma coluna é denominada **matriz coluna**, ou **vetor coluna**, e uma matriz com somente uma linha é denominada **matriz linha**, ou **vetor linha**.

Quando for desejada uma notação mais compacta, a matriz precedente pode ser escrita como

$$[a_{ij}]_{m\times n} \quad \text{ou, simplesmente,} \quad [a_{ij}].$$

A entrada na linha i e na coluna j de uma matriz A também é geralmente denotada pelo símbolo $(A)_{ij}$. Assim, podemos simplificar a notação da matriz (1.5) acima, como

$$(A)_{ij} = a_{ij}.$$

Quando uma matriz contém o mesmo número de linhas e colunas, dizemos que é uma **matriz quadrada**

de ordem $\boldsymbol{n}$, denotada por

$$A = \begin{bmatrix} \boldsymbol{a_{11}} & a_{12} & \cdots & a_{1n} \\ a_{21} & \boldsymbol{a_{22}} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \boldsymbol{a_{nn}} \end{bmatrix}. \tag{1.6}$$

Chamamos de **diagonal principal**, as entradas destacadas na matriz (1.6) de A.

Vejam alguns exemplos de matrizes:

Exemplo 1.6. *Considere as matrizes:*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 5 & 0 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Note que, A é uma matriz 2×3, B é uma matriz coluna 3×1, C é uma matriz linha 1×3 e D é uma matriz quadrada de ordem 3.

Definição 1.2.2. *Duas matrizes são definidas como sendo* ***iguais*** *se tiverem o mesmo tipo e suas entradas correspondentes forem iguais.*

Em outras palavras, sejam A e B duas matrizes de ordem $m \times n$ então, $(A)_{ij} = (B)_{ij}$ se, e somente se,

$$a_{ij} = b_{ij}$$

para quaisquer valores de $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Definição 1.2.3. *Se A for uma matriz quadrada, então o* ***traço de $\boldsymbol{A}$****, denotado por $tr(A)$, é definido pela soma das entradas na diagonal principal de A. O traço de A não é definido se A não for uma matriz quadrada.*

Exemplo 1.7. *Sendo* $A = \begin{bmatrix} \mathbf{1} & 2 & 6 \\ 0 & \mathbf{1} & 5 \\ 2 & 3 & \mathbf{7} \end{bmatrix}$*, então*

$$tr(A) = 1 + 1 + 7 = 9.$$

1.2.1 Operações matriciais

A seguir vamos apresentar algumas operações sobre matrizes que são bastante usadas.

Definição 1.2.4. *Se A e B são matrizes de mesmo tipo $m \times n$, então a* ***soma*** *$A + B$ é a matriz $m \times n$ obtida somando as entradas de B às entradas correspondentes de A, e a* ***diferença*** *$A - B$ é a matriz $m \times n$ obtida subtraindo as entradas de B das entradas correspondentes de A.*

Observe, da Definição 1.2.4, que as operações de soma e subtração não estão definidas para matrizes de tipos distintos.

Em notação matricial, se $A = [a_{ij}]$ e $B = [b_{ij}]$ têm o mesmo tipo, as operações indicadas acima são dadas por

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{e} \quad (A-B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}.$$

Exemplo 1.8. *Sendo* $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ *e* $B = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$*, então*

$$A+B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2+3 & 1+2 & 0+0 \\ 1+0 & 2+1 & 1+0 \end{bmatrix} = \begin{bmatrix} 5 & 3 & 0 \\ 1 & 3 & 1 \end{bmatrix},$$

$$A-B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2-3 & 1-2 & 0-0 \\ 1-0 & 2-1 & 1-0 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Definição 1.2.5. *Se A for uma matriz e c um escalar, então o **produto** cA é a matriz obtida pela multiplicação de cada entrada da matriz A por c. Dizemos que a matriz cA é um **múltiplo escalar** de A.*

Em notação matricial, se $A = [a_{ij}]$, então o elemento da posição ij da matriz cA é dado por

$$(cA)_{ij} = c(A)_{ij} = ca_{ij}.$$

Exemplo 1.9. *Sendo* $A = \begin{bmatrix} 0 & 0 & 2 \\ 6 & 4 & 2 \end{bmatrix}$*, então*

$$\frac{1}{2}A = \frac{1}{2}\begin{bmatrix} 0 & 0 & 2 \\ 6 & 4 & 2 \end{bmatrix} = \begin{bmatrix} \frac{0}{2} & \frac{0}{2} & \frac{2}{2} \\ \frac{6}{2} & \frac{4}{2} & \frac{2}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix}.$$

Definição 1.2.6. *Se A for uma matriz do tipo $m \times r$ e B uma matriz do tipo $r \times n$, então o produto AB é a matriz do tipo $m \times n$ cujas entradas são determinadas como segue. Para obter a entrada na linha i e coluna j de AB, destacamos a linha i de A e a coluna j de B. Multiplicamos as entradas correspondentes da linha e da coluna e então somamos os produtos resultantes.*

Note que, a definição de multiplicação das matrizes A e B, exige que o número de colunas de A seja igual ao número de linhas de B, como descrito abaixo:

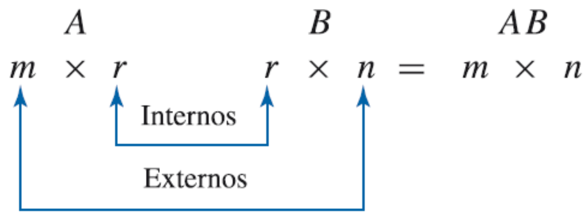


Figura 1.1: Esquema de multiplicação de matrizes [2].

Suponhamos que $A = [a_{ij}]$ é uma matriz $m \times r$ e $B = [b_{ij}]$ é uma matriz $r \times n$. Então, na expressão (1.7) ressalta-se que o elemento $(AB)_{ij}$ desse produto vai combinar a linha i de A com a coluna j de B, como destacado:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a_{i1}} & \boldsymbol{a_{i2}} & \cdots & \boldsymbol{a_{ir}} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{nr} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & \boldsymbol{b_{1j}} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & \boldsymbol{b_{2j}} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & \boldsymbol{b_{rj}} & \cdots & b_{rn} \end{bmatrix}. \tag{1.7}$$

Dessa forma, a entrada $(AB)_{ij}$ é dada por

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} = \sum_{k=1}^{r} a_{ik}b_{kj},$$

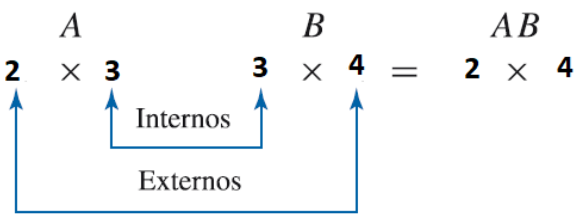
onde o símbolo $\sum$ indica somatório.

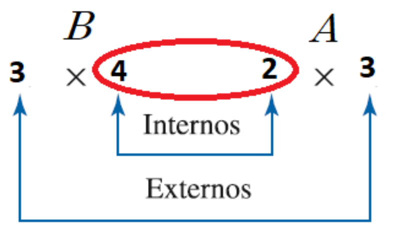
Exemplo 1.10. *Sendo* $A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$ *uma matriz* 2×2 *e* $B = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \end{bmatrix}$ *uma matriz* 2×3*, então devemos encontrar uma matriz* AB *do tipo* 2×3*. Assim,*

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1\cdot 1 + 2\cdot 0 & 1\cdot 2 + 2\cdot 1 & 1\cdot 4 + 2\cdot(-1) \\ 4\cdot 1 + 5\cdot 0 & 4\cdot 2 + 5\cdot 1 & 4\cdot 4 + 5\cdot(-1) \end{bmatrix} \\ &= \begin{bmatrix} 1+0 & 2+2 & 4-2 \\ 4+0 & 8+5 & 16-5 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 4 & 2 \\ 4 & 13 & 11 \end{bmatrix}. \end{aligned}$$

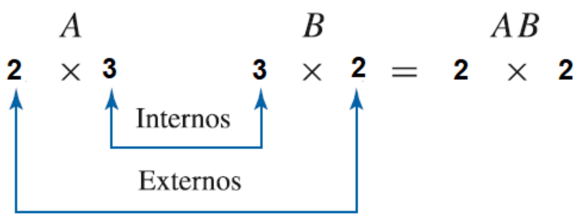
Observação 1.2.1. *Sabe-se que no conjunto dos números reais* $(\mathbb{R})$*, a multiplicação tem a propriedade comutativa, isto é, sempre vale que* $ab = ba$*. Mas, esse fato, não acontece necessariamente na multiplicação de matrizes, ou seja,* AB *nem sempre é igual a* BA*. Vejamos três razões possíveis para que isso não ocorra:*

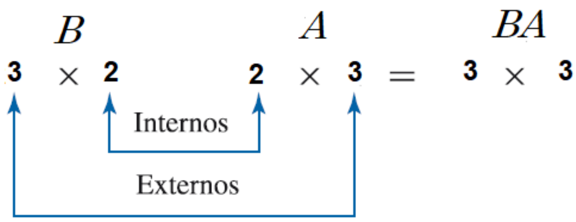
1. AB *pode estar definida e* BA *não (por exemplo, se* A *é uma matriz* 2×3 *e* B *é* 3×4*).*





2. AB *e* BA *podem ambas estar definidas, mas têm tamanhos diferentes (por exemplo, se* A *é uma matriz* 2×3 *e* B *é* 3×2*)*





3. AB *e* BA *podem ambas estar definidas e ter o mesmo tamanho, mas as matrizes podem ser diferentes (conforme ilustrado no exemplo seguinte).*

Exemplo 1.11. *Considere as matrizes*

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}.$$

Multiplicando, obtemos

$$AB = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix},$$

$$BA = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}.$$

Assim, $AB \neq BA$*.*

O exemplo não proíbe a possibilidade de AB e BA serem iguais em *certos* casos, somente que não são iguais em *todos* os casos. Se ocorrer que $AB = BA$, dizemos que as matrizes A e B **comutam**.

Exemplo 1.12. *Considere as matrizes*

$$A = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Multiplicando, obtemos

$$AB = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix},$$

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

Assim, $AB = BA$.

A seguir apresentam-se um conjunto de propriedades das operações matriciais.

Teorema 1.2.1. ***(Propriedades da aritmética matricial)*** *Suponha que as matrizes que aparecem a seguir, sejam tais que as operações indicadas possam ser feitas, e além disso, considere a e b números reais. Então, valem as seguintes propriedades.*

(a) $A + B = B + A$ ***[Lei da comutatividade da adição]***

(b) $A + (B + C) = (A + B) + C$ ***[Lei da associatividade da adição]***

(c) $A(BC) = (AB)C$ ***[Lei da associatividade da multiplicação]***

(d) $A(B + C) = AB + AC$ ***[Lei da distributividade à esquerda]***

(e) $(A + B)C = AC + BC$ ***[Lei da distributividade à direita]***

(f) $A(B - C) = AB - AC$

(g) $(B - C)A = BA - CA$

(h) $a(B + C) = aB + aC$

(i) $a(B - C) = aB - aC$

(j) $(a + b)C = aC + bC$

(k) $(a - b)C = aC - bC$

(l) $a(bC) = (ab)C$

(m) $a(BC) = (aB)C = B(aC)$

Demonstração.

(a) Temos que,

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{e} \quad (B+A)_{ij} = (B)_{ij} + (A)_{ij} = b_{ij} + a_{ij}.$$

Como a_{ij} e b_{ij} são números reais, então

$$a_{ij} + b_{ij} = b_{ij} + a_{ij}.$$

Logo, $A + B = B + A$.

(b) De um lado temos:

$$(A)_{ij} = a_{ij} \quad \text{e} \quad (B+C)_{ij} = (B)_{ij} + (C)_{ij} = b_{ij} + c_{ij}.$$

Assim, $(A + (B + C))_{ij} = a_{ij} + (b_{ij} + c_{ij})$. Por outro lado,

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{e} \quad (C)_{ij} = c_{ij}.$$

Logo, $((A + B) + C)_{ij} = (a_{ij} + b_{ij}) + c_{ij}$. Como a_{ij}, b_{ij} e c_{ij} são números reais, então

$$a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij}.$$

Portanto, $A + (B + C) = (A + B) + C$.

(c) Seja $D = BC$. Assim,

$$(D)_{kj} = d_{kj} = \sum_{l=1}^{q} b_{kl}c_{lj}, \tag{1.8}$$

$$(AD)_{ij} = \sum_{k=1}^{p} a_{ik}d_{kj}. \tag{1.9}$$

Substituindo (1.8) em (1.9):

$$(AD)_{ij} = \sum_{k=1}^{p} a_{ik} \sum_{l=1}^{q} b_{kl}c_{lj} = \sum_{k=1}^{p}\sum_{l=1}^{q} a_{ik}b_{kl}c_{lj}.$$

Seja $Z = AB$, então:

$$((AB)C)_{ij} = (ZC)_{ij} = \sum_{l=1}^{q} z_{il}c_{lj} = \sum_{l=1}^{q}\left(\sum_{k=1}^{p} a_{ik}b_{kl}\right)c_{lj} = \sum_{l=1}^{q}\sum_{k=1}^{p} a_{ik}b_{kl}c_{lj} = (AD)_{ij} = (A(BC))_{ij}.$$

(d) Devemos mostrar que $A(B+C)$ e $AB+AC$ têm o mesmo tipo e que as entradas correspondentes são iguais. Para formar $A(B+C)$, as matrizes B e C devem ter o mesmo tipo, $m \times n$, e então a matriz A deve ter m colunas, de modo que seu tamanho é da forma $r \times m$. Isso faz de $A(B+C)$ uma matriz $r \times n$. Segue que $AB+AC$ também é uma matriz $r \times n$ e, consequentemente, $A(B+C)$ e $AB+AC$ têm o mesmo tamanho.

Suponha que $A = [a_{ik}]$, $B = [b_{kj}]$ e $C = [c_{kj}]$. Queremos mostrar que as entradas correspondentes de $A(B+C)$ e de $AB+AC$ são iguais, ou seja, que

$$[A(B+C)]_{ij} = [AB+AC]_{ij}$$

para todos valores de i e j. Pela definição de soma e produto matriciais, temos

$$\begin{aligned}
[A(B+C)]_{ij} &= a_{i1}(b_{1j}+c_{1j}) + a_{i2}(b_{2j}+c_{2j}) + \cdots + a_{im}(b_{mj}+c_{mj}) \\
&= (a_{i1}b_{1j} + a_{i1}c_{1j} + a_{i2}b_{2j} + a_{i2}c_{2j} + \cdots + a_{im}b_{mj} + a_{im}c_{mj}) \\
&= (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{im}c_{mj}) \\
&= [AB]_{ij} + [AC]_{ij} \\
&= [AB+AC]_{ij}.
\end{aligned}$$

(e) Devemos mostrar que $(A+B)C$ e $AC+BC$ têm o mesmo tipo e que as entradas correspondentes são iguais. Para formar $(A+B)C$, as matrizes A e B devem ter o mesmo tipo, $m \times p$, e então a matriz C deve ter p linhas, de modo que seu tamanho é da forma $p \times n$. Isso faz de $(A+B)C$ uma matriz $m \times n$. Segue que $AC+BC$ também é uma matriz $m \times n$ e, consequentemente, $(A+B)C$ e $AC+BC$ têm o mesmo tamanho.

Suponha que $A = [a_{ik}]$, $B = [b_{ik}]$ e $C = [c_{kj}]$. Queremos mostrar que as entradas correspondentes de $(A+B)C$ e de $AC+BC$ são iguais, ou seja, que

$$[(A+B)C]_{ij} = [AC+BC]_{ij}$$

para todos valores de i e j. Pela definição de soma e produto matriciais, temos

$$\begin{aligned}
[(A+B)C]_{ij} &= (a_{i1}+b_{i1})c_{1j} + (a_{i2}+b_{i2})c_{2j} + \cdots + (a_{ip}+b_{ip})c_{pj} \\
&= (a_{i1}c_{1j} + b_{i1}c_{1j} + a_{i2}c_{2j} + b_{i2}c_{2j} + \cdots + a_{ip}c_{pj} + b_{ip}c_{pj}) \\
&= (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{ip}c_{pj}) + (b_{i1}c_{1j} + b_{i2}c_{2j} + \cdots + b_{ip}c_{pj}) \\
&= [AC]_{ij} + [BC]_{ij} \\
&= [AC+BC]_{ij}.
\end{aligned}$$

(f) Devemos mostrar que $A(B-C)$ e $AB-AC$ têm o mesmo tipo e que as entradas correspondentes são iguais. Para formar $A(B-C)$, as matrizes B e C devem ter o mesmo tipo, $m\times n$, e então a matriz A deve ter m colunas, de modo que seu tamanho é da forma $r\times m$. Isso faz de $A(B-C)$ uma matriz $r\times n$. Segue que $AB-AC$ também é uma matriz $r\times n$ e, consequentemente, $A(B-C)$ e $AB-AC$ têm o mesmo tamanho.

Suponha que $A=[a_{ik}]$, $B=[b_{kj}]$ e $C=[c_{kj}]$. Queremos mostrar que as entradas correspondentes de $A(B-C)$ e de $AB-AC$ são iguais, ou seja, que

$$[A(B-C)]_{ij}=[AB-AC]_{ij}$$

para todos valores de i e j. Pela definição de soma e produto matriciais, temos

$$\begin{aligned}
[A(B-C)]_{ij} &= a_{i1}(b_{1j}-c_{1j})+a_{i2}(b_{2j}-c_{2j})+\cdots+a_{im}(b_{mj}-c_{mj})\\
&= (a_{i1}b_{1j}-a_{i1}c_{1j}+a_{i2}b_{2j}-a_{i2}c_{2j}+\cdots+a_{im}b_{mj}-a_{im}c_{mj})\\
&= (a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{im}b_{mj})-(a_{i1}c_{1j}+a_{i2}c_{2j}+\cdots+a_{im}c_{mj})\\
&= [AB]_{ij}-[AC]_{ij}\\
&= [AB-AC]_{ij}.
\end{aligned}$$

(g) Devemos mostrar que $(B-C)A$ e $BA-CA$ têm o mesmo tipo e que as entradas correspondentes são iguais. Para formar $(B-C)A$, as matrizes B e C devem ter o mesmo tipo, $m\times p$, e então a matriz A deve ter p linhas, de modo que seu tamanho é da forma $p\times n$. Isso faz de $(B-C)A$ uma matriz $m\times n$. Segue que $BA-CA$ também é uma matriz $m\times n$ e, consequentemente, $(B-C)A$ e $BA-CA$ têm o mesmo tamanho.

Suponha que $A=[a_{kj}]$, $B=[b_{ik}]$ e $C=[c_{ik}]$. Queremos mostrar que as entradas correspondentes de $(B-C)A$ e de $BA-CA$ são iguais, ou seja, que

$$[(B-C)A]_{ij}=[BA-CA]_{ij}$$

para todos valores de i e j. Pela definição de soma e produto matriciais, temos

$$\begin{aligned}
[(B-C)A]_{ij} &= (b_{i1}-c_{i1})a_{1j}+(b_{i2}-c_{i2})a_{2j}+\cdots+(b_{ip}-c_{ip})a_{pj}\\
&= (b_{i1}a_{1j}-c_{i1}a_{1j}+b_{i2}a_{2j}-c_{i2}a_{2j}+\cdots+b_{ip}a_{pj}-c_{ip}a_{pj})\\
&= (b_{i1}a_{1j}+b_{i2}a_{2j}+\cdots+b_{ip}a_{pj})-(c_{i1}a_{1j}+c_{i2}a_{2j}+\cdots+c_{ip}a_{pj})\\
&= [BA]_{ij}-[CA]_{ij}\\
&= [BA-CA]_{ij}.
\end{aligned}$$

(h) $(a(B+C))_{ij} = a(b_{ij}+c_{ij}) = ab_{ij}+ac_{ij} = (aB)_{ij}+(aC)_{ij}$.

(i) $(a(B-C))_{ij} = a(b_{ij}-c_{ij}) = ab_{ij}-ac_{ij} = (aB)_{ij}-(aC)_{ij}$.

(j) $((a+b)C)_{ij} = (a+b)c_{ij} = ac_{ij}+bc_{ij} = (aC)_{ij}+(bC)_{ij}$.

(k) $((a-b)C)_{ij} = (a-b)c_{ij} = ac_{ij}-bc_{ij} = (aC)_{ij}-(bC)_{ij}$.

(l) $(a(bC))_{ij} = a(bc_{ij}) = (ab)c_{ij} = ((ab)C)_{ij}$.

(m)
$$\begin{aligned}(a(BC))_{ij} = a\sum_{k=1}^{p} b_{ik}c_{kj} &= \sum_{k=1}^{p}(ab_{ik})c_{kj} = ((aB)C)_{ij}\\ &= \sum_{k=1}^{p} b_{ik}(ac_{kj}) = (B(aC))_{ij}.\end{aligned}$$

∎

Na sequência, três exemplos que ilustram algumas das propriedades demonstradas acima:

Exemplo 1.13. ***(Lei da associatividade da multiplicação)*** *Considere as matrizes*

$$A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 1 & -3 \\ 2 & 7 \end{bmatrix}.$$

Multiplicando BC e AB, respectivamente, temos

$$BC = \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ 7 & 5 \\ 1 & -3 \end{bmatrix},$$

$$AB = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ 9 & 4 \end{bmatrix}.$$

Assim, segue que

$$A(BC) = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 5 & -2 \\ 7 & 5 \\ 1 & -3 \end{bmatrix}}_{BC} = \begin{bmatrix} 9 & -14 \\ 17 & 1 \end{bmatrix},$$

$$(AB)C = \underbrace{\begin{bmatrix} 7 & 1 \\ 9 & 4 \end{bmatrix}}_{AB} \begin{bmatrix} 1 & -3 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 9 & -14 \\ 17 & 1 \end{bmatrix}.$$

Portanto, $A(BC) = (AB)C$.

Exemplo 1.14. ***(Lei da distributividade à esquerda)*** *Considere as matrizes*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & -1 \\ 1 & 4 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 1 & -3 \\ 4 & 1 \end{bmatrix}.$$

Somando as matrizes $B + C$*, e multiplicando* AB *e* AC*, respectivamente,*

$$B + C = \begin{bmatrix} 5 & -1 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 1 & -3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -4 \\ 5 & 5 \end{bmatrix},$$

$$AB = \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 11 & 2 \\ 8 & 11 \\ -5 & 1 \end{bmatrix},$$

$$AC = \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -5 \\ 13 & 0 \\ -1 & 3 \end{bmatrix}.$$

Assim, temos

$$A(B + C) = \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 6 & -4 \\ 5 & 5 \end{bmatrix}}_{B+C} = \begin{bmatrix} 17 & -3 \\ 21 & 11 \\ -6 & 4 \end{bmatrix},$$

$$AB + AC = \underbrace{\begin{bmatrix} 11 & 2 \\ 8 & 11 \\ -5 & 1 \end{bmatrix}}_{AB} + \underbrace{\begin{bmatrix} 6 & -5 \\ 13 & 0 \\ -1 & 3 \end{bmatrix}}_{AC} = \begin{bmatrix} 17 & -3 \\ 21 & 11 \\ -6 & 4 \end{bmatrix}.$$

Portanto, $A(B + C) = AB + AC$.

Exemplo 1.15. ***(Item m)*** *Considere* $a = 3$ *e as matrizes*

$$B = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix}.$$

Assim,

$$a(BC) = 3 \cdot \left(\begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} \right) = 3 \cdot \begin{bmatrix} 7 & 1 \\ 9 & 4 \end{bmatrix} = \begin{bmatrix} 21 & 3 \\ 27 & 12 \end{bmatrix},$$

$$(aB)C = \left(3 \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \right) \cdot \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 9 & -3 & 3 \\ 6 & 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 21 & 3 \\ 27 & 12 \end{bmatrix},$$

$$B(aC) = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \cdot \left(3 \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} \right) = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 9 & 3 \\ 9 & 6 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 21 & 3 \\ 27 & 12 \end{bmatrix}.$$

Portanto, $a(BC) = (aB)C = B(aC)$.

Proposição 1.2.1. *Se A e B são duas matrizes tais que $AB = BA$, então*

(i) $B^kA = AB^k$

(ii) $(AB)^k = A^kB^k$

Demonstração.

(i) Para $k = 1$, temos por hipótese:

$$AB = BA.$$

Suponha que é verdade que $B^nA = AB^n$. Vejamos o que acontece com $B^{n+1}A$.

$$B^{n+1}A = (B^nB)A = B^n(BA) = B^n(AB) = (B^nA)B = (AB^n)B = A(B^nB) = AB^{n+1}.$$

Logo, $B^kA = AB^k$ para todo $k \in \mathbb{N}$.

(ii) Para $k = 1$, $AB = BA$, por hipótese. Suponhamos que $(AB)^n = A^nB^n$. Então,

$$(AB)^{n+1} = (AB)^nAB = (A^nB^n)AB = A^n(B^nA)B = A^n(AB^n)B = (A^nA)(B^nB) = A^{n+1}B^{n+1}.$$

Logo, é verdade que $(AB)^k = A^kB^k$, para todo $k \in \mathbb{N}$.

∎

Em seguida, apresentaremos nesta seção alguns tipos principais de matrizes e suas propriedades para que possamos exemplificar no decorrer dos demais capítulos.

1.2.2 Matriz nula e matriz identidade

Definição 1.2.7. *A **matriz nula**, também conhecida como **matriz zero**, denotada pela letra O, é aquela matriz cujas entradas são todas nulas. Isto é,*

$$(O)_{ij} = 0$$

para quaisquer valores de i e j.

Uma matriz nula $m \times n$ é denotada por $O_{m\times n}$.

Exemplo 1.16. *Alguns exemplos são:*

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \end{bmatrix}.$$

Definição 1.2.8. *A **matriz identidade**, denotada pela letra I, é uma matriz quadrada com entradas 1 na diagonal principal e demais entradas nulas.*

Se for importante enfatizar seu tamanho, escrevemos I_n para a matriz de tamanho $n \times n$. Vejamos alguns exemplos:

Exemplo 1.17. *As matrizes $I_1,\ I_2,\ I_3$ e I_4 são representadas, respectivamente, por:*

$$[1], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Em geral, se A for uma matriz $m \times n$, então vale que,

$$AI_n = A \quad \text{e} \quad I_m A = A.$$

A seguir, um exemplo para ilustrar esta propriedade.

Exemplo 1.18. *Sendo $A = \begin{bmatrix} 2 & -3 & 4 \\ 6 & 2 & 1 \end{bmatrix}$, então*

$$I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 & 4 \\ 6 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 4 \\ 6 & 2 & 1 \end{bmatrix} = A,$$

$$AI_3 = \begin{bmatrix} 2 & -3 & 4 \\ 6 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 4 \\ 6 & 2 & 1 \end{bmatrix} = A.$$

1.2.3 Matriz transposta

Definição 1.2.9. *Se A for uma matriz $m \times n$ qualquer, então a* ***transposta de A****, denotada por A^T, é definida como a matriz $n \times m$ que resulta da troca das linhas com as colunas de A; ou seja, a primeira coluna de A^T é a primeira linha de A, a segunda coluna de A^T é a segunda linha de A, e assim por diante.*

Usando notação matricial, a definição acima pode ser matematicamente escrita como segue:

$$(A^T)_{ij} = (A)_{ji}.$$

Exemplo 1.19. *Sendo* $A = \begin{bmatrix} 3 & 6 & 12 \\ 0 & 3 & -3 \end{bmatrix}$*, então*

$$A^T = \begin{bmatrix} 3 & 0 \\ 6 & 3 \\ 12 & -3 \end{bmatrix}.$$

Agora vejamos os seguintes resultados sobre matriz transposta:

Teorema 1.2.2. *Se os tamanhos das matrizes são tais que as operações indicadas podem ser efetuadas e tomando k um número real, então*

(a) $(A^T)^T = A$

(b) $(A + B)^T = A^T + B^T$

(c) $(A - B)^T = A^T - B^T$

(d) $(kA)^T = kA^T$

(e) $(AB)^T = B^T A^T$

Demonstração.

(a) Os elementos da matriz A de tamanho $m \times n$ são dados por

$$(A)_{ij} = a_{ij},$$

onde $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$. Os elementos da matriz A^T de tamanho $n \times m$ são dados por

$$(A^T)_{ij} = (A)_{ji}.$$

Assim, podemos concluir que os elementos da matriz $(A^T)^T$ serão dados por

$$(A^T)^T_{ij} = ((A)_{ji})^T = (A)_{ij} = a_{ij}.$$

Isto é, $(A^T)^T = A$.

(b) Seja c_{ij}, elemento de $(A+B)^T$, então:

$$\begin{aligned} c_{ij} &= (A+B)_{ji} \\ &= a_{ji} + b_{ji} \\ &= (A^T)_{ij} + (B^T)_{ij} \\ &= (A^T + B^T)_{ij}. \end{aligned}$$

Logo,

$$(A+B)^T = A^T + B^T.$$

(c) Seja c_{ij}, elemento de $(A-B)^T$, então:

$$\begin{aligned} c_{ij} &= (A-B)_{ji} \\ &= a_{ji} - b_{ji} \\ &= (A^T)_{ij} - (B^T)_{ij} \\ &= (A^T - B^T)_{ij}. \end{aligned}$$

Logo,

$$(A-B)^T = A^T - B^T.$$

(d) Seja $k \in \mathbb{R}$ e $A = [a_{ij}]$. Vejamos que,

$$(kA)^T = (ka_{ij})^T = k(a_{ij})^T = ka_{ji} = kA^T \quad \forall i, j.$$

Portanto, $(kA)^T = kA^T$.

(e) Sejam $A = [a_{ij}]_{m\times n}$ e $B = [b_{jk}]_{n\times p}$. Logo, $A^T \in M_{n\times m}$ e $B^T \in M_{p\times n}$. Seja c_{ij} elemento de $(AB)^T$, então:

$$c_{ij} = \sum_{k=1}^{n} (AB)^T_{ij} = \sum_{k=1}^{n} (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Considere d_{ij} elemento de $B^T A^T$, então:

$$d_{ij} = \sum_{k=1}^{n} (B^T)_{ik} (A^T)_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Logo,

$$c_{ij} = d_{ij},$$

para todo $i = 1, 2, \ldots, p$ e $j = 1, 2, \ldots, m$. Então,

$$(AB)^T = B^T A^T.$$

∎

Na sequência, um exemplo que ilustra a propriedade do item (e) do Teorema 1.2.2.

Exemplo 1.20. *Sejam* $A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$ *e* $B = \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix}$. *Então,*

$$AB = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ 9 & 4 \end{bmatrix}.$$

Logo, $(AB)^T = \begin{bmatrix} 7 & 9 \\ 1 & 4 \end{bmatrix}$. *Por outro lado,*

$$B^T A^T = \begin{bmatrix} 3 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 1 & 4 \end{bmatrix}.$$

Portanto, $(AB)^T = B^T A^T$.

1.2.4 Matriz invertível

Definição 1.2.10. *Se A for uma matriz quadrada e se existir uma matriz B de mesmo tamanho tal que $AB = BA = I$, então diremos que A é* ***invertível*** *(ou* ***não singular****) e que B é uma* ***inversa*** *de A. Se não puder ser encontrada tal matriz B, diremos que A é* ***não invertível*** *ou* ***singular****.*

Exemplo 1.21. *Considere as matrizes*

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}.$$

Multiplicando, obtemos

$$AB = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$BA = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Assim, concluímos que a matriz B é a inversa da matriz A e que A é não singular.

Note que a relação $AB = BA = I$ permanece inalterada pela troca de A por B, de modo que se A for invertível e B uma inversa, então também vale que B é invertível e que A é uma inversa de B. Assim, se

$$AB = BA = I$$

dizemos que A e B são inversas uma da outra. Agora observe o resultado a seguir:

Teorema 1.2.3. ***(Unicidade da matriz inversa)*** *Se B e C são ambas inversas da matriz A, então $B = C$.*

Demonstração. Como B é uma inversa de A, temos $BA = I$. Multiplicando ambos os lados à direita por C, dá

$$(BA)C = IC. \tag{1.10}$$

Pela lei da associatividade da multiplicação (Teorema 1.2.1 item (c)), $(BA)C = B(AC)$. Assim, da equação (1.10) resulta

$$B(AC) = C. \tag{1.11}$$

Como C também é uma inversa de A, segue que $AC = I$. Logo, da expressão (1.11),

$$BI = C,$$

o que leva à

$$B = C$$

que é o que queríamos provar. ∎

Como uma consequência desse importante resultado, podemos agora falar da inversa de uma matriz invertível. Se A for invertível, então sua inversa é única e será denotada pelo símbolo A^{-1}. Assim,

$$AA^{-1} = I \quad \text{e} \quad A^{-1}A = I.$$

Teorema 1.2.4. *Se A e B são matrizes invertíveis de mesmo tamanho, então AB é invertível e*

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

Demonstração. Podemos mostrar a invertibilidade e obter a fórmula enunciada ao mesmo tempo mostrando que

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

Vejamos,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I,$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$

Assim, AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$. ∎

Teorema 1.2.5. *Se A for uma matriz invertível e n um inteiro não negativo, então*

(a) A^{-1} *é invertível e* $(A^{-1})^{-1} = A$.

(b) A^n *é invertível e* $(A^n)^{-1} = (A^{-1})^n$.

(c) kA *é invertível com qualquer escalar não nulo* k *e* $(kA)^{-1} = k^{-1}A^{-1}$.

Demonstração.

(a) Por hipótese, A é uma matriz invertível, então existe A^{-1}. A^{-1} é invertível se podemos encontrar uma matriz B tal que

$$A^{-1}B = BA^{-1} = I_n.$$

Tomando $B = A$, segue que A^{-1}, de fato, é invertível e A é a inversa de A^{-1}. Portanto,

$$(A^{-1})^{-1} = A.$$

(b) Podemos mostrar a invertibilidade e obter a fórmula enunciada ao mesmo tempo mostrando que

$$A^n(A^{-1})^n = (A^{-1})^n A^n = I.$$

Como A e A^{-1} comutam entre si, pela Proposição 1.2.1, segue que,

$$A^n(A^{-1})^n = (AA^{-1})^n = I^n = I,$$

$$(A^{-1})^n A^n = (A^{-1}A)^n = I^n = I.$$

Assim, A^n é invertível e $(A^n)^{-1} = (A^{-1})^n$.

(c) Das propriedades (l) e (m) do Teorema 1.2.1 implica que

$$(kA)(k^{-1}A^{-1}) = k^{-1}(kA)A^{-1} = (k^{-1}k)AA^{-1} = (1)I = I,$$

$$(k^{-1}A^{-1})(kA) = (A^{-1}k^{-1})(kA) = A^{-1}(k^{-1}k)A = A^{-1}1A = A^{-1}A = I.$$

Assim, kA é invertível e $(kA)^{-1} = k^{-1}A^{-1}$.

∎

Teorema 1.2.6. *Se A for uma matriz invertível, então A^T também é invertível e*

$$(A^T)^{-1} = (A^{-1})^T.$$

Demonstração. Considerem os produtos $A^T(A^{-1})^T$ e $(A^{-1})^TA^T$. Então temos,

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I \quad \text{e}$$

$$(A^{-1})^TA^T = (AA^{-1})^T = I^T = I.$$

Logo, $(A^T)^{-1} = (A^{-1})^T$.

∎

1.2.5 Matriz diagonal

Definição 1.2.11. *Uma matriz quadrada em que todas as entradas fora da diagonal principal são zero é denominada* ***matriz diagonal****.*

Exemplo 1.22. *Alguns exemplos de matrizes diagonais de ordem 2, 3 e 4:*

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

Uma matriz diagonal arbitrária D de tamanho $n \times n$ pode ser escrita como

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}. \tag{1.12}$$

Uma matriz diagonal é invertível se, e só se, todas as suas entradas na diagonal são não nulas; neste caso a inversa de (1.12) é

$$D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}. \tag{1.13}$$

De fato,

$$DD^{-1} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I,$$

$$D^{-1}D = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I.$$

1.2.6 Matriz triangular

Definição 1.2.12. *Uma matriz quadrada com todas as entradas acima da diagonal principal nulas é denominadas* ***triangular inferior****, e uma matriz quadrada com todas as entradas abaixo da diagonal principal nulas é denominada* ***triangular superior****. Dizemos que uma matriz triangular inferior ou triangular superior é* ***triangular****.*

Exemplo 1.23. ***Matrizes triangulares superiores e inferiores***

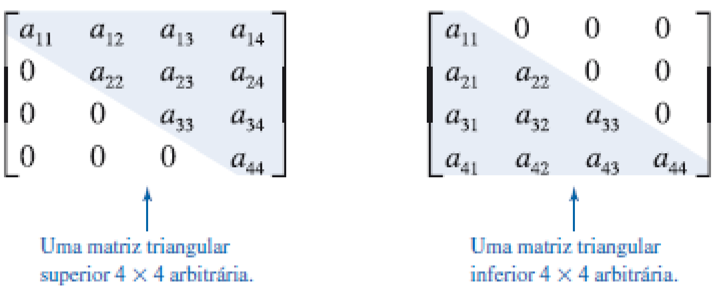


Figura 1.2: Exemplos genéricos de matrizes triangulares de tamanho 4×4 [2].

Observação 1.2.2. *Observe que matrizes diagonais são triangulares inferiores e superiores, pois têm apenas zeros acima e abaixo da diagonal principal.*

1.2.7 Matriz simétrica

Definição 1.2.13. *Uma matriz quadrada A é dita **simétrica** se $A = A^T$.*

Em outras palavras, uma matriz quadrada $A = [a_{ij}]$ é simétrica se, e só se,

$$(A)_{ij} = (A)_{ji}$$

com quaisquer valores de i e j.

Exemplo 1.24. *As seguintes matrizes são simétricas, já que cada uma delas é igual à sua transposta.*

$$\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}.$$

É fácil reconhecer visualmente a simetria de uma matriz: as entradas na diagonal principal não têm restrições, mas as entradas que estão posicionadas simetricamente em relação à diagonal principal devem ser iguais. Toda matriz diagonal, como a terceira matriz do Exemplo 1.24, têm essa propriedade.

O teorema seguinte lista as principais propriedades algébricas das matrizes simétricas.

Teorema 1.2.7. *Sendo A e B matrizes simétricas de mesmo tamanho e k um escalar qualquer, então*

(a) A^T *é simétrica.*

(b) $A + B$ *e* $A - B$ *são simétricas.*

(c) kA *é simétrica.*

Demonstração.

(a) Segue direto da definição, se A é simétrica, então $A = A^T$. Logo, A^T também é simétrica.

(b) Sejam A e B duas matrizes simétricas de mesmo tamanho. Desse modo, $A^T = A$ e $B^T = B$. Assim,

$$(A + B)^T = A^T + B^T = A + B.$$

Analogamente,

$$(A - B)^T = A^T - B^T = A - B.$$

Logo, $A + B$ e $A - B$ são simétricas.

(c) Se A é simétrica, então $A = A^T$. Assim, segue que

$$(kA)^T = k(A^T) = kA.$$

Logo, kA é simétrica.

∎

Não é verdade, em geral, que o produto de matrizes simétricas resulta em uma matriz simétrica. Para ver porquê isso ocorre, segue o seguinte resultado.

Teorema 1.2.8. *O produto de duas matrizes simétricas é uma matriz simétrica se, e só se, as matrizes comutam.*

Demonstração. Sejam A e B matrizes simétricas de mesmo tamanho. Temos

$$(AB)^T = B^T A^T = BA.$$

Assim, $(AB)^T = AB$ se, e só se, $AB = BA$, isto é, se, e só se, A e B comutam. ∎

Em geral, uma matriz simétrica não precisa ser invertível; por exemplo, uma matriz diagonal com um zero na diagonal principal é simétrica, mas não é invertível. Contudo, o próximo teorema mostra que se ocorrer que uma matriz simétrica é invertível, então sua inversa também é simétrica.

Teorema 1.2.9. *Se A for uma matriz simétrica invertível, então A^{-1} é simétrica.*

Demonstração. Suponha que A seja simétrica e invertível. Pelo Teorema 1.2.6 e pelo fato de que $A = A^T$, decorre

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

provando que A^{-1} é simétrica. ∎

Observação 1.2.3. *Numa variedade de aplicações, surgem produtos matriciais da forma AA^T e A^TA. Se A for uma matriz $m \times n$, então A^T é uma matriz $n \times m$, de modo que ambos produtos AA^T e A^TA são matrizes quadradas, a matriz AA^T de tamanho $m \times m$ e a matriz A^TA de tamanho $n \times n$. Esses produtos são sempre simétricos, pois*

$$(AA^T)^T = (A^T)^T A^T = AA^T \quad e \quad (A^TA)^T = A^T(A^T)^T = A^TA.$$

No caso especial em que A é quadrada, temos o seguinte resultado.

Teorema 1.2.10. *Se A for uma matriz invertível, então AA^T e A^TA também serão invertíveis.*

Demonstração. Como A é invertível, pelo Teorema 1.2.6, A^T também é invertível. Assim, segue do Teorema 1.2.4, AA^T e A^TA são invertíveis, por serem produtos de matrizes invertíveis. ∎

1.2.8 Matriz ortogonal

Definição 1.2.14. *Dizemos que uma matriz quadrada A é* ***ortogonal*** *se sua transposta for sua inversa, ou seja, se*

$$A^{-1} = A^T$$

ou, equivalentemente, se

$$AA^T = A^TA = I. \tag{1.14}$$

Exemplo 1.25. *A matriz*

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

é ortogonal, pois

$$A^TA = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Na sequência, alguns resultados das matrizes ortogonais:

Teorema 1.2.11. *Sobre matrizes ortogonais, temos:*

(a) A inversa de uma matriz ortogonal é ortogonal.

(b) Um produto de matrizes ortogonais é ortogonal.

(c) Se A for ortogonal, então $\det(A) = 1$ ou $\det(A) = -1$.

Demonstração.

(a) De fato, se A é ortogonal, então A^T também é. Além disso, segue da definição que $A^{-1} = A^T$, então basta observar que $(A^T)^{-1} = (A^{-1})^{-1} = A$.

(b) Sejam A e B duas matrizes ortogonais. Pelo Teorema 1.2.2 item (e), temos

$$(AB)(AB)^T = (AB)B^TA^T = A(BB^T)A^T = AIA^T = AA^T = I.$$

Fato que demonstra a propriedade.

(c) Para provar a afirmação usamos o Teorema de Binet e lembramos que A^T e A têm o mesmo determinante e que o determinante da identidade é igual a 1. Essas informações se encontram mais aprofundadas no Apêndice 1, sobre Determinantes. Logo,

$$\begin{aligned} \det(AA^T) &= \det(I) \\ \det(A)\cdot\det(A^T) &= \det(I) \\ \det(A)^2 &= 1 \\ \det(A) &= \pm 1. \end{aligned}$$

∎

1.3 Autovalores e autovetores de matrizes

Definição 1.3.1. *Se A for uma matriz $n \times n$, então um vetor não nulo $\vec{x}$ em $\mathbb{R}^n$ é denominado **autovetor** de A se $A\vec{x}$ for um múltiplo escalar de $\vec{x}$, isto é,*

$$A\vec{x} = \lambda\vec{x}$$

*com algum escalar λ. O escalar λ é denominado **autovalor** de A, e dizemos que $\vec{x}$ é um **autovetor associado a λ**.*

Exemplo 1.26. *O vetor $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ é um autovetor de $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ associado ao autovalor $\lambda = 3$, pois*

$$A\vec{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3\vec{x}.$$

Geometricamente, a multiplicação por A expandiu o vetor $\vec{x}$ pelo fator 3.

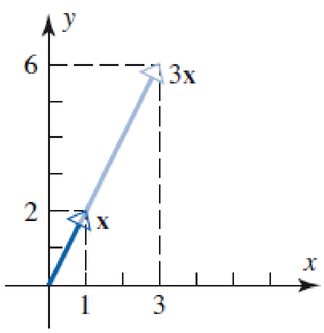


Figura 1.3: Expansão do vetor $\vec{x}$ [2].

Proposição 1.3.1. *Seja λ um autovalor de A, então λ^m é autovalor de A^m.*

Demonstração. De fato, seja λ um autovalor de A, então existe $\vec{x} \neq \vec{0}$, tal que:

$$A\vec{x} = \lambda\vec{x}. \tag{1.15}$$

Multiplicando (1.15) por A à esquerda, temos:

$$\begin{aligned} AA\vec{x} &= A(\lambda\vec{x}) \\ A^2\vec{x} &= \lambda(A\vec{x}) \\ A^2\vec{x} &= \lambda(\lambda\vec{x}) \\ A^2\vec{x} &= \lambda^2\vec{x}. \end{aligned}$$

Por indução matemática, pode-se provar que:

$$A^m\vec{x} = \lambda^m\vec{x}, \tag{1.16}$$

isto é, λ^m é o autovalor de A^m associado ao autovalor $\vec{x}$.

Em geral, seja $P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Define-se $P(A) = a_0I + a_1A + a_2A^2 + \cdots + a_nA^n$. Se $\vec{x}$ é autovetor de A e λ o correspondente autovalor, então:

$$P(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$$

é o autovalor de $P(A)$ associado ao autovetor $\vec{x}$.

De fato,

$$P(A)\vec{x} = \left(\sum_{k=0}^{n} a_kA^k\right)\vec{x}, \tag{1.17}$$

onde $A^0 = I$. Usando (1.16) no lado direito de (1.17), tem-se que:

$$\begin{aligned} P(A)\vec{x} &= \left(\sum_{k=0}^{n} a_k\lambda^k\vec{x}\right) \\ &= \left(\sum_{k=0}^{n} a_k\lambda^k\right)\vec{x} \\ &= P(\lambda)\vec{x}. \end{aligned}$$

∎

Definição 1.3.2. *Uma matriz simétrica A em $\mathbb{R}^{n\times n}$ é denominada **definida positiva**, se $\vec{x}^TA\vec{x} > 0$, para todo $\vec{x} \neq \vec{0}$, $\vec{x}$ em $\mathbb{R}^n$.*

Proposição 1.3.2. *Os autovalores de toda matriz simétrica definida positiva são positivos.*

Demonstração. Seja A uma matriz simétrica definida positiva e $\vec{x}$ um autovetor associado ao autovalor λ, então

$$A\vec{x} = \lambda\vec{x}. \tag{1.18}$$

Multiplicando ambos os lados de (1.18) à esquerda por $\vec{x}^T$, temos

$$\begin{aligned} \vec{x}^T A\vec{x} &= \vec{x}^T(\lambda\vec{x}) \\ &= \lambda\vec{x}^T\vec{x} \\ &= \lambda(x_1^2 + x_2^2 + \cdots + x_n^2) \end{aligned} \tag{1.19}$$

Como A é simétrica definida positiva, então pela Definição 1.3.2, $\vec{x}^T A\vec{x} > 0$ para todo $\vec{x} \neq \vec{0}$. Como $\vec{x}$ é autovetor, então $\vec{x} \neq \vec{0}$. Daí, de (1.19) implica

$$\lambda > 0.$$

∎

Proposição 1.3.3. *Seja $B = A^T A$. Logo, os autovalores de B são números reais não negativos.*

Demonstração. A matriz $B = A^T A$ é simétrica. Seja λ autovalor de A e $\vec{x} \neq \vec{0}$, o respectivo autovetor de B, então

$$B\vec{x} = \lambda\vec{x}.$$

Multiplicando ambos os lados à esquerda por $\vec{x}^T$, temos

$$\vec{x}^T B\vec{x} = \vec{x}^T(\lambda\vec{x}).$$

Como $B = A^T A$, segue que

$$\vec{x}^T(A^T A)\vec{x} = \lambda\vec{x}^T\vec{x}. \tag{1.20}$$

Note que, do lado esquerdo de (1.20), temos

$$\vec{x}^T(A^T A)\vec{x} = (\vec{x}^T A^T)(A\vec{x}) = (A\vec{x})^T(A\vec{x}).$$

Fazendo $A\vec{x} = \vec{y}$ segue de (1.20) que,

$$\vec{y}^T\vec{y} = \lambda\vec{x}^T\vec{x}$$

o que implica em

$$\lambda = \frac{\vec{y}^T\vec{y}}{\vec{x}^T\vec{x}}. \tag{1.21}$$

Como $\vec{y}^T\vec{y} = y_1^2 + y_2^2 + \cdots + y_n^2 \geq 0$ e $\vec{x}^T\vec{x} = x_1^2 + x_2^2 + \cdots + x_n^2 > 0$, segue de (1.21) que

$$\lambda \geq 0.$$

∎

Definição 1.3.3. *O raio espectral da matriz A é denotado por $\rho(A)$ e definido como:*

$$\rho(A) = \max\{|\lambda|,\ \textit{onde } \lambda \textit{ é autovalor de } A\}.$$

Proposição 1.3.4. *Seja A uma matriz não singular, então:*

(a) *Zero não é autovalor de A;*

(b) *λ é autovalor de A se, e somente se, $\dfrac{1}{\lambda}$ é autovalor de A^{-1};*

(c) *$[\rho(A)]^2 = \rho(A^2)$;*

(d) *A^TA e AA^T são matrizes simétricas e possuem os mesmos autovalores;*

(e) *$\rho((A^{-1})^TA^{-1}) = \dfrac{1}{\min\{|\lambda|;\ \lambda \textit{ é autovalor de } A^TA\}}$.*

Demonstração. Vejamos

(a) Suponha que 0 é autovalor de A, então existe $\vec{x} \neq \vec{0}$, tal que:

$$\begin{aligned} A\vec{x} &= 0\vec{x} \\ A\vec{x} &= \vec{0}. \end{aligned} \tag{1.22}$$

Multiplicando (1.22) por A^{-1} à esquerda,

$$\begin{aligned} A^{-1}A\vec{x} &= A^{-1}\vec{0} \\ I\vec{x} &= \vec{0} \\ \vec{x} &= \vec{0}, \end{aligned}$$

que é uma contradição.

(b) Como λ é autovalor de A, então existe $\vec{x} \neq \vec{0}$, tal que:

$$A\vec{x} = \lambda\vec{x}. \tag{1.23}$$

Multiplicando (1.23) à esquerda por A^{-1}, temos:

$$\begin{aligned} A^{-1}A\vec{x} &= A^{-1}(\lambda\vec{x}) \\ I\vec{x} &= \lambda A^{-1}\vec{x} \\ \vec{x} &= \lambda A^{-1}\vec{x}. \end{aligned}$$

Logo,

$$A^{-1}\vec{x} = \frac{1}{\lambda}\vec{x}.$$

Então, $\dfrac{1}{\lambda}$ é autovalor de A^{-1}. A demonstração da recíproca é semelhante. Por outro lado, temos que:

$$\begin{aligned} \rho(A^{-1}) &= \max\{|\lambda|; \lambda \text{ é autovalor de } A^{-1}\} \\ &= \frac{1}{\min\{|\lambda|; \lambda \text{ é autovalor de } A\}}. \end{aligned}$$

(c) Sabe-se que:

- Se λ é autovalor de A, então λ^2 é autovalor de A^2 (Proposição 1.3.1);
- $\rho(A) = \max\{|\lambda|; \lambda$ é autovalor de $A\}$, onde $|\cdot|$ é o módulo do número λ;
- $\rho(A^2) = \max\{|\gamma|; \gamma$ é autovalor de $A^2\}$.

Por outro lado, se γ é autovalor de A^2, então

$$|A^2 - \gamma I| = 0$$

$$|(A - \sqrt{\gamma}I)(A + \sqrt{\gamma}I)| = 0$$

$$|A - \sqrt{\gamma}I|\,|A + \sqrt{\gamma}I| = 0$$

$$|A - \sqrt{\gamma}I| = 0 \quad \text{ou} \quad |A + \sqrt{\gamma}I| = 0$$

$$|A - \sqrt{\gamma}I| = 0 \quad \text{ou} \quad |A - (-\sqrt{\gamma})I| = 0.$$

Logo, $\sqrt{\gamma}$ ou $-\sqrt{\gamma}$ é autovalor de A.

Considere-se os conjuntos:

$$Aut(A) = \{\lambda;\ \lambda \text{ é autovalor de } A\} \quad \text{e}$$

$$Aut(A^2) = \{\gamma;\ \gamma \text{ é autovalor de } A^2\}.$$

Afirmação: Se $\gamma \in Aut(A^2)$, então existe $\lambda \in Aut(A)$ tal que $\gamma = \lambda^2$.

De fato, suponha que existe um $\gamma_0 \in Aut(A^2)$ tal que

$$\gamma_0 \neq \lambda^2, \quad \forall \lambda \in Aut(A)$$

Por hipótese, $\gamma_0 \in Aut(A^2)$, logo $\sqrt{\gamma_0}$ ou $-\sqrt{\gamma_0}$ é autovalor de A. Então, existe $\lambda_0 = \sqrt{\gamma_0}$ ou $\lambda_0 = -\sqrt{\gamma_0}$ tal que

$$\lambda_0^2 = \gamma_0.$$

Contradição. Portanto, a afirmação é verdadeira.

Assim,

$$\begin{aligned}
\rho(A^2) &= \max\{|\gamma|; \gamma \text{ é autovalor de } A^2\} \\
&= \max\{|\lambda^2|; \lambda \text{ é autovalor de } A\} \\
&= \max\{|\lambda|^2; \lambda \text{ é autovalor de } A\} \\
&= (\max\{|\lambda|; \lambda \text{ é autovalor de } A\})^2 \\
&= [\rho(A)]^2.
\end{aligned}$$

(d) Seja λ autovalor de A^TA, então existe $\vec{x} \neq \vec{0}$, tal que

$$A^TA\vec{x} = \lambda\vec{x}. \tag{1.24}$$

Multiplicando A de ambos os lados à esquerda de (1.24), temos:

$$\begin{aligned}
A(A^TA\vec{x}) &= A\lambda\vec{x} \\
AA^T(A\vec{x}) &= \lambda A\vec{x}. \qquad (1.25)
\end{aligned}$$

Fazendo $A\vec{x} = \vec{y}$ em (1.25):

$$AA^T\vec{y} = \lambda\vec{y}.$$

Mas $\vec{y} \neq \vec{0}$, pois A é não singular. Então, λ é autovalor de AA^T. Logo:

$$\rho(A^TA) = \rho(AA^T).$$

(e) Sabe-se que

$$(A^{-1})^T(A^{-1}) = (A^T)^{-1}A^{-1} = (AA^T)^{-1}.$$

Logo, $\rho((A^{-1})^TA^{-1}) = \rho((AA^T)^{-1})$. Usando-se o resultado (b), tem-se que:

$$\begin{aligned}\rho((A^{-1})^TA^{-1}) &= \frac{1}{\min\{|\lambda|;\ \lambda \text{ é autovalor de } AA^T\}}\\ &= \frac{1}{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A^TA\}}.\end{aligned}$$

∎

Definição 1.3.4. *Se A e B forem matrizes quadradas, dizemos que B é **semelhante** a A se existir alguma matriz invertível P tal que $B = P^{-1}AP$.*

Definição 1.3.5. *Uma matriz quadrada A é dita **diagonalizável** se for semelhante a alguma matriz diagonal, ou seja, se existir alguma matriz invertível P tal que $P^{-1}AP$ é diagonal. Nesse caso, dizemos que a matriz P **diagonaliza** A.*

Definição 1.3.6. *Um conjunto $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ de m vetores de $\mathbb{R}^n$ é linearmente independente se a única solução possível da equação*

$$a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_m\vec{u}_m = 0,$$

seja a solução nula. Isto é: $a_1 = a_2 = \cdots = a_m = 0$.

Proposição 1.3.5. *Seja $A = \{u_1, u_2, \cdots, u_n\}$ um conjunto de vetores de $\mathbb{R}^n$. O conjunto de vetores A é linearmente independente se, e somente se,* $\det[u_1, u_2, \cdots, u_n] \neq 0$.

A demonstração desse resultado pode ser vista em [2].

Teorema 1.3.1. *Se A for uma matriz $n \times n$, são equivalentes as afirmações seguintes.*

(a) A é diagonalizável.

(b) A tem n autovetores linearmente independentes.

Demonstração. **(a) ⇒ (b)** Como estamos supondo que A é diagonalizável, existem uma matriz invertível P e uma matriz diagonal D tais que $P^{-1}AP = D$ ou, equivalentemente,

$$AP = PD. \tag{1.26}$$

Denotando os vetores coluna de P por $\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n$ e supondo que as entradas diagonais de D sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$, segue que o lado esquerdo de (1.26) pode ser expresso por

$$AP = A\begin{bmatrix}\vec{p}_1 & \vec{p}_2 & \ldots & \vec{p}_n\end{bmatrix} = \begin{bmatrix}A\vec{p}_1 & A\vec{p}_2 & \ldots & A\vec{p}_n\end{bmatrix}$$

e o lado direito de (1.26) pode ser expresso por

$$PD = \begin{bmatrix} \lambda_1\vec{p}_1 & \lambda_2\vec{p}_2 & \dots & \lambda_n\vec{p}_n \end{bmatrix}.$$

Assim, segue de (1.26) que

$$A\vec{p}_1 = \lambda_1\vec{p}_1, \quad A\vec{p}_2 = \lambda_2\vec{p}_2, \quad \dots, \quad A\vec{p}_n = \lambda_n\vec{p}_n. \tag{1.27}$$

Como P é invertível, então $|P| \neq 0$. Assim, nenhuma coluna de P é combinação linear das outras. Logo, $\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$ é linearmente independente.

Logo, de (1.27), tem-se que os $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n$ são não nulos. Então, A possui n autovetores linearmente independentes.

(b) $\Rightarrow$ **(a)** Suponha que A tenha n autovetores linearmente independentes $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n$ com autovalores associados $\lambda_1, \lambda_2, \dots, \lambda_n$. Escrevendo

$$P = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix}$$

e denotando por D a matriz diagonal de entradas diagonais sucessivas $\lambda_1, \lambda_2, \dots, \lambda_n$, obtemos

$$\begin{aligned} AP &= A\begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} \\ &= \begin{bmatrix} A\vec{p}_1 & A\vec{p}_2 & \dots & A\vec{p}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\vec{p}_1 & \lambda_2\vec{p}_2 & \dots & \lambda_n\vec{p}_n \end{bmatrix} \\ &= PD. \end{aligned}$$

Como os vetores coluna de P são linearmente independentes, segue que P é invertível, de modo que essa última equação pode ser reescrita como $P^{-1}AP = D$, mostrando que A é diagonalizável. ∎

Teorema 1.3.2. *Se $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ forem autovetores de uma matriz A associados a autovalores distintos, então $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ é um conjunto linearmente independente.*

A demonstração desse resultado pode ser vista em [2].

Teorema 1.3.3. *Se uma matriz A de tamanho $n \times n$ tem n autovalores distintos, então A é diagonalizável.*

Demonstração. Se $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ são autovetores associados aos autovalores distintos $\lambda_1, \lambda_2, \dots, \lambda_n$ então, pelo Teorema 1.3.2, $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ são linearmente independentes. Assim, A é diagonalizável pelo Teorema 1.3.1. ∎

Teorema 1.3.4. *Se A for uma matriz simétrica real, então A tem autovalores reais.*

Demonstração. Sejam λ um autovalor de A e $\vec{x}$ um autovetor associado, sendo que λ pode ser complexo e $\vec{x}$ pode estar em $\mathbb{C}^n$. Assim,

$$A\vec{x} = \lambda\vec{x}, \tag{1.28}$$

onde $\vec{x} \neq \vec{0}$. Daí,

$$\begin{aligned} (A\vec{x})^T &= (\lambda\vec{x})^T \\ \vec{x}^T A^T &= \lambda\vec{x}^T. \end{aligned}$$

Como A é simétrica, segue que

$$\vec{x}^T A = \lambda\vec{x}^T. \tag{1.29}$$

Multiplicando ambos os lados à esquerda de (1.28) por $\overline{\vec{x}}^T$, temos

$$\overline{\vec{x}}^T A\vec{x} = \overline{\vec{x}}^T(\lambda\vec{x}) = \lambda\overline{\vec{x}}^T\vec{x} = \lambda\sum_{i=1}^{n}|x_i|^2. \tag{1.30}$$

Por outro lado,

$$\overline{\overline{\vec{x}}^T A\vec{x}} = \vec{x}^T\,\overline{A}\,\overline{\vec{x}}.$$

Como A é uma matriz real, $A = \overline{A}$. Logo,

$$\overline{\overline{\vec{x}}^T A\vec{x}} = \vec{x}^T A\overline{\vec{x}} \overset{(1.29)}{=} \lambda\vec{x}^T\overline{\vec{x}} = \lambda\sum_{i=1}^{n}|x_i|^2. \tag{1.31}$$

Os dois lados de direito de (1.30) e (1.31) são iguais. Logo,

$$\overline{\vec{x}}^T A\vec{x} = \overline{\overline{\vec{x}}^T A\vec{x}}$$

e, $\overline{\vec{x}}^T A\vec{x}$ é um número real. De (1.30), obtém-se:

$$\lambda = \frac{\overline{\vec{x}}^T A\vec{x}}{\displaystyle\sum_{i=1}^{n}|x_i|^2},$$

quociente de dois números reais. Logo, λ é um número real. ∎

Definição 1.3.7. *Sejam A e B matrizes quadradas. Dizemos que A e B são* ***ortogonalmente semelhantes*** *se existir alguma matriz ortogonal P tal que $P^TAP = B$.*

Se A for ortogonalmente semelhante a matriz D, isto é,

$$P^TAP = D,$$

onde D é uma matriz diagonal, então dizemos que A é **ortogonalmente diagonalizável** e que $\boldsymbol{P}$ **diagonaliza $\boldsymbol{A}$ ortogonalmente**.

Teorema 1.3.5. *Se A for uma matriz $n \times n$, então as afirmações são equivalentes.*

(a) A é ortogonalmente diagonalizável.

(b) A tem um conjunto ortonormal de n autovetores.

(c) A é simétrica.

Demonstração. **(a) $\Rightarrow$ (b)** Como A é ortogonalmente diagonalizável, existe alguma matriz ortogonal P tal que $P^{-1}AP$ é diagonal. Como mostramos na prova do Teorema 1.3.1, os n vetores coluna de P são autovetores de A. Como P é ortogonal, esses vetores colunas são ortogonais, de modo que A tem n autovetores ortonormais.

(b) $\Rightarrow$ (a) Suponha que A tenha um conjunto ortonormal $\{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$ de n autovetores. Como mostramos na prova do Teorema 1.3.1, a matriz P que tem esses autovetores como colunas diagonaliza A. Como esses autovetores são ortogonais, P é ortogonal e, assim, diagonaliza A ortogonalmente.

(a) $\Rightarrow$ (c) Na prova de (a) $\Rightarrow$ (b), mostramos que uma matriz quadrada A de ordem n ortogonalmente diagonalizável é ortogonalmente diagonalizada por uma matriz P de tamanho $n \times n$ cujas colunas formam um conjunto ortogonal de autovetores de A. Seja D a matriz diagonal

$$D = P^T AP.$$

Multiplicando ambos os lados da igualdade por P à esquerda e por P^T à direita, temos

$$PDP^T = PP^T APP^T$$

do que segue que

$$A = PDP^T.$$

Assim,

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

o que mostra que A é simétrica.

(c) $\Rightarrow$ (a) Veja [2].

∎

Capítulo 2

Aplicações de matrizes

As matrizes têm muitas aplicações práticas. Aqui serão apresentadas duas dessas aplicações, ambas associadas com o número de condição de uma matriz: transformações lineares de $\mathbb{R}^m$ em $\mathbb{R}^n$ e sistemas de equações lineares.

Embora tenha sido apresentado o conceito de vetor em $\mathbb{R}^n$ (Capítulo 1) como uma lista ordenada de n números, do tipo:

$$\vec{v} = (v_1, v_2, \ldots, v_n),$$

aqui e nos próximos capítulos um vetor deve ser entendido como uma matriz coluna ou vetor coluna, assim o vetor $\vec{v}$ será entendido como:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

As principais referências que utilizamos como base para este capítulo estão em [2] e [3].

2.1 Transformações lineares

Antes de seguirmos com o assunto em questão, apresentaremos algumas definições.

Definição 2.1.1. *Um sistema matemático constituído de um conjunto não vazio K e um par de operações sobre K, respectivamente uma adição $(x, y) \mapsto x + y$ e uma multiplicação $(x, y) \mapsto xy$ (ou $x \cdot y$), é chamado* ***corpo*** *se:*

(a) $a, b, c \in K$, *então* $a + (b + c) = (a + b) + c$ ***[associatividade da adição];***

(b) $a, b \in K$, *então* $a + b = b + a$ ***[comutatividade da adição];***

(c) Existe um elemento $0_K \in K$ *tal que, qualquer que seja* $a \in K$, $a + 0_K = a$ ***[existência de elemento neutro];***

(d) Qualquer que seja $a \in K$, existe um elemento em K, indicado genericamente por $-a$, tal que $a + (-a) = 0_K$ ***[existência de opostos];***

(e) $a, b, c \in K$, então $a(bc) = (ab)c$ ***[associatividade da multiplição];***

(f) $a, b, c \in K$, então $a(b + c) = ab + ac$ e $(a + b)c = ac + bc$ ***[distributividade];***

(g) $a, b \in K$, então $ab = ba$ ***[comutatividade da multiplicação];***

(h) Existe um elemento $1_K \in K$ tal que, qualquer que seja $a \in K$, $a \cdot 1_K = 1_K \cdot a = a$ ***[existência da unidade];***

(i) Para todo $a \in K - \{0_K\}$, existe $b \in K$ tal que $ab = 1$ ***[existência do inverso].***

Exemplo 2.1. *O conjunto dos números racionais $\mathbb{Q}$, o conjunto dos números reais $\mathbb{R}$ e o conjunto dos números complexos $\mathbb{C}$ são exemplos de corpos.*

Definição 2.1.2. *Seja K um conjunto de escalares e seja V um conjunto não vazio qualquer, cujos elementos são denominados vetores. Se em V estiver definida duas operações, a **adição** que associa a cada par de vetores $\vec{u}, \vec{v}$ em V um vetor $\vec{u} + \vec{v}$ em V e a **multiplicação por escalares**, que associa a cada escalar k em K e cada vetor $\vec{u}$ em V um vetor $k\vec{u}$ em V, então V é um **espaço vetorial** sobre K, em relação a essas operações, se as seguintes condições estiverem satisfeitas:*

1. *$\vec{u} + \vec{v} = \vec{v} + \vec{u}$, com qualquer $\vec{u}, \vec{v}$ em V.*

2. *$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$, com qualquer $\vec{u}, \vec{v}, \vec{w}$ em V.*

3. *Existe um único vetor $\vec{0}$ em V, denominado **vetor nulo** de V, ou **vetor zero**, tal que $\vec{0} + \vec{u} = \vec{u} + \vec{0} = \vec{u}$, com qualquer $\vec{u}$ em V.*

4. *Dado qualquer $\vec{u}$ em V, existe um único vetor $-\vec{u}$ em V, denominado **oposto** de $\vec{u}$, tal que $\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = \vec{0}$.*

5. *$k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$, com qualquer k em K e $\vec{u}, \vec{v}$ em V.*

6. *$(k + m)\vec{u} = k\vec{u} + m\vec{u}$, com qualquer k, m em K e $\vec{u}$ em V.*

7. *$k(m\vec{u}) = (km)\vec{u}$, com qualquer k, m em K e $\vec{u}$ em V.*

8. *$1\vec{u} = \vec{u}$, com qualquer $\vec{u}$ em V.*

Observação 2.1.1. *Note que, pela Definição 2.1.1, temos associado a todo espaço vetorial um conjunto de escalares K, sendo K um corpo. Pelo Exemplo 2.1, tal conjunto pode ser considerado o conjunto dos números reais $\mathbb{R}$ ou conjunto dos números complexos $\mathbb{C}$. Porém, no decorrer deste trabalho vamos estabelecer K como sendo o conjunto dos números reais $\mathbb{R}$.*

A seguir apresentam-se alguns exemplos de espaços vetoriais:

Exemplo 2.2. *Seja* $V = \{\vec{0}\}$ *com as operações*

$$\vec{0} + \vec{0} = \vec{0} \quad e \quad k\vec{0} = \vec{0}$$

com escalares k *quaisquer. É fácil verificar que todas as condições da Definição 2.1.1 são satisfeitas. Dizemos que este é o* ***espaço vetorial nulo****.*

Exemplo 2.3. *O conjunto dos números reais* $V = \mathbb{R}$ *munido das operações usuais de adição e multiplicação é um espaço vetorial.*

Exemplo 2.4. ***(*$\mathbb{R}^n$ *é um espaço vetorial)*** *Seja* $V = \mathbb{R}^n$ *e defina as operações de espaço vetorial em* V *como as operações conhecidas de adição e multiplicação por escalar de* n*-uplas, ou seja,*

$$\vec{u} + \vec{v} = (u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$$

$$k\vec{u} = k(u_1, u_2, \dots, u_n) = (ku_1, ku_2, \dots, ku_n).$$

Pelo Teorema 1.1.1, está provado que as condições da Definição 2.1.1 são satisfeitas.

Exemplo 2.5. *Definimos* $M_{m\times n}$*, o conjunto formado pelas matrizes do tipo* $m \times n$*, cujos elementos estão em* $\mathbb{R}$*. Pelo Teorema 1.2.1, está provado que as condições da Definição 2.1.1 são satisfeitas.*

Uma vez que entendemos o conceito de espaços vetoriais, partimos para a transformação linear propriamente dita.

Definição 2.1.3. *Sejam* V *e* W *espaços vetoriais sobre* $\mathbb{R}$*. Uma função* $T : V \to W$ *é denominada* ***transformação linear*** *de* V *em* W *se as duas propriedades seguintes forem válidas com quaisquer vetores* $\vec{u}$ *e* $\vec{v}$ *em* V *e qualquer escalar* k *em* $\mathbb{R}$*.*

(i) $T(k\vec{v}) = kT(\vec{v})$ ***[Homogeneidade]***

(ii) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ ***[Aditividade]***

No caso especial em que $V = W$*, a transformação linear é denominada* ***operador linear*** *do espaço vetorial* V*.*

A homogeneidade e a aditividade de uma transformação linear $T : V \to W$ pode ser usada em combinação para mostrar, de maneira mais geral, que, se $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ forem vetores em V e $k_1, k_2, \dots, k_r$ forem escalares em $\mathbb{R}$, então

$$T(k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_r\vec{v}_r) = k_1T(\vec{v}_1) + k_2T(\vec{v}_2) + \cdots + k_rT(\vec{v}_r). \tag{2.1}$$

O próximo teorema lista duas propriedades básicas de transformação linear.

Teorema 2.1.1. *Se $T : V \to W$ for uma transformação linear, então*

(a) $T(\vec{0}) = \vec{0}$.

(b) $T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$*, para quaisquer que sejam $\vec{u}$ e $\vec{v}$ em V.*

Demonstração.

(a) Seja $\vec{u}$ um vetor qualquer em V. Como $0\vec{u} = \vec{0}$, segue da homogeneidade na Definição 2.1.2 que

$$T(\vec{0}) = T(0\vec{u}) = 0T(\vec{u}) = \vec{0}.$$

(b) Temos que,

$$\begin{aligned} T(\vec{u} - \vec{v}) &= T(\vec{u} + (-1)\vec{v}) \\ &= T(\vec{u}) + (-1)T(\vec{v}) \\ &= T(\vec{u}) - T(\vec{v}). \end{aligned}$$

∎

Observação 2.1.2. *Com as duas partes do Teorema 2.1.1, podemos mostrar que*

$$T(-\vec{v}) = -T(\vec{v})$$

com qualquer $\vec{v}$ em V.

De fato,

$$T(-\vec{v}) = T(\vec{0} - \vec{v}) \stackrel{(b)}{=} T(\vec{0}) - T(\vec{v}) \stackrel{(a)}{=} \vec{0} - T(\vec{v}) = -T(\vec{v}).$$

Vejamos alguns exemplos de transformação linear.

Exemplo 2.6. ***(O operador identidade)*** *Seja V um espaço vetorial qualquer. A aplicação $I : V \to V$ definida por $I(\vec{v}) = \vec{v}$ é denominada* ***operador identidade*** *de V. I é um operador linear. De fato, para $\vec{u}$ e $\vec{v}$ em V e k em $\mathbb{R}$, temos*

(i) $I(k\vec{v}) = kI(\vec{v}) = k\vec{v}$,

(ii) $I(\vec{u} + \vec{v}) = I(\vec{u}) + I(\vec{v}) = \vec{u} + \vec{v}$.

Exemplo 2.7. ***(Transformações de espaços matriciais)*** *Seja $T : M_{m\times n} \to M_{n\times m}$, definida por $T(A) = A^T$. T é uma transformação linear. De fato, segue das partes (b) e (d) do Teorema 1.2.2 que*

(i) $T(kA) = (kA)^T \stackrel{(d)}{=} kA^T = kT(A)$,

(ii) $T(A+B) = (A+B)^T \stackrel{(b)}{=} A^T + B^T = T(A) + T(B)$.

Exemplo 2.8. *Seja* $V = \mathbb{R}^3$ *e* $W = \mathbb{R}^2$ *tal que* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *definida por* $T(x, y, z) = (x - 2z, -x + 3y)$*, então temos,*

$$T(x, y, z) = (x - 2z, -x + 3y) = x(1, -1) + y(0, 3) + z(-2, 0).$$

Assim, podemos reescrever $T(x, y, z) = AX$*, onde* $A = \begin{bmatrix} 1 & 0 & -2 \\ -1 & 3 & 0 \end{bmatrix}$ *é uma matriz denominada matriz de transformação linear e* $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *um vetor em* $\mathbb{R}^3$.

Este exemplo mostra como uma transformação linear pode ser escrita usando matrizes.

Exemplo 2.9. *Seja* A *uma matriz em* $M_{m\times n}$*, então* $T(X) = AX$ *é uma transformação linear de* $\mathbb{R}^m$ *em* $\mathbb{R}^n$*. Fato que pode ser verificado facilmente.*

Proposição 2.1.1. *Seja* $\vec{y} = A\vec{x}$ *uma transformação linear, com* $\vec{x}$, $\vec{y}$ *vetores de* $\mathbb{R}^n$ *e* A *matriz em* $M_{n\times n}$*. Considera-se o segmento de reta* $\overline{x_i x_f}$ *definido por:*

$$\vec{x} = (1 - t)\vec{x}_i + t\vec{x}_f = \vec{x}_i + t(\vec{x}_f - \vec{x}_i),$$

onde $\vec{x}_i$ *é o ponto inicial do segmento,* $\vec{x}_f$ *o ponto final do segmento e* $t \in [0, 1]$*, então* $\vec{y} = A\vec{x}$ *é um segmento de reta que começa em* $A\vec{x}_i$ *e termina em* $A\vec{x}_f$*.*

Demonstração. De fato,

$$\vec{y} = A\vec{x} = A[(1 - t)\vec{x}_i + t\vec{x}_f] = (1 - t)A\vec{x}_i + tA\vec{x}_f,$$

com $t \in [0, 1]$. Note que, quando $t = 0$ temos:

$$\vec{y} = (1 - 0)A\vec{x}_i + 0A\vec{x}_f = 1A\vec{x}_i + 0 = A\vec{x}_i,$$

e quando $t = 1$ temos:

$$\vec{y} = (1 - 1)A\vec{x}_i + 1A\vec{x}_f = 0A\vec{x}_i + A\vec{x}_f = 0 + A\vec{x}_f = A\vec{x}_f.$$

Logo, $\vec{y}$ é um segmento de reta que começa em $A\vec{x}_i$ e termina em $A\vec{x}_f$. ∎

2.2 Sistemas lineares

Considere um sistema de m equações e n incógnitas

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}.$$

Esse sistema pode ser colocado usando a notação de matrizes da seguinte forma:

$$\begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

A matriz $m \times 1$ à esquerda dessa equação pode ser escrita como um produto, resultando

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Denotando essas matrizes por A, X e B, respectivamente, o sistema original de m equações em n incógnitas pode ser substituído pela única equação matricial

$$AX = B. \tag{2.2}$$

A matriz A nesta equação é denominada **matriz de coeficientes** do sistema. A **matriz aumentada** do sistema é obtida pela adjunção de B a A como a última coluna; assim, a matriz aumentada é

$$[A|B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

A barra vertical em $[A|B]$ é só uma maneira conveniente de visualmente separar A de B, não tendo significado matemático.

A equação (2.2) simplifica a expressão de um sistema linear e permite sua análise em função das características da matriz A. Por exemplo, se a matriz A for invertível, então existe A^{-1} e multiplicando (2.2) à esquerda por A^{-1}, tem-se:

$$\begin{aligned} A^{-1}AX &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B. \end{aligned}$$

Isso indica que para calcular a solução, precisa-se determinar a matriz inversa de A.

Capítulo 3

Normas

Neste capítulo, vamos tratar das noções de norma de vetores e matrizes. O termo *norma* é um sinônimo matemático comum para *comprimento* ou *magnitude* de um vetor. Para estudos sobre normas vetoriais, seguimos [2], [3], [4] e também [5] como principais referências. Já para falar sobre normas de matrizes, utilizamos [3], [4], [6] e [7] como principais referências.

3.1 Norma de vetores

Definição 3.1.1. *Seja V um espaço vetorial sobre $\mathbb{R}$. Uma norma em V é uma função real de V em $\mathbb{R}$, denotada por $||\cdot|| : V \to \mathbb{R}$, que satisfaz as seguintes condições, para quaisquer que sejam $\vec{u}, \vec{v}$ em V e para todo escalar k em $\mathbb{R}$, tem-se:*

(i) $||\vec{v}|| \geq 0$ *(positividade);* $||\vec{v}|| = 0$*, se, e somente se,* $\vec{v} = \vec{0}$ *(separação);*

(ii) $||k\vec{v}|| = |k|\ ||\vec{v}||$ *(homogeneidade);*

(iii) $||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ *(desigualdade triangular).*

Para $V = \mathbb{R}^n$, existe uma infinidade de normas vetoriais. Dentre elas temos a norma p, denotada por $||\cdot||_p,\ p \geq 1$, dada por:

$$||\vec{v}||_p := \left(\sum_{i=1}^{n} |v_i|^p \right)^{1/p} = (|v_1|^p + |v_2|^p + \cdots + |v_n|^p)^{1/p}.$$

Em particular, para $p = 1$, temos a Norma 1, também conhecida como, Norma da Soma, definida pela expressão:

$$||\vec{v}||_1 = \sum_{i=1}^{n} |v_i|. \tag{3.1}$$

Para $p = 2$, temos a Norma 2, conhecida também por Norma Euclidiana, definida por:

$$||\vec{v}||_2 = \sqrt{\sum_{i=1}^{n} |v_i|^2}. \tag{3.2}$$

E por fim, para $p = \infty$, temos a Norma ∞ ou Norma do Máximo, dada pela expressão:

$$||\vec{v}||_\infty = \max_{1 \le i \le n} \{|v_i|\}. \tag{3.3}$$

Exemplo 3.1. *Dado o vetor* $\vec{v} = (1, 2, -3, 0)$*, temos:*

$$\begin{aligned} ||\vec{v}||_1 &= |1| + |2| + |-3| + |0| = 1 + 2 + 3 + 0 = 6; \\ ||\vec{v}||_2 &= \sqrt{|1|^2 + |2|^2 + |-3|^2 + |0|^2} = \sqrt{1 + 4 + 9 + 0} = \sqrt{14} \approx 3,74; \\ ||\vec{v}||_\infty &= \max\{|1|, |2|, |-3|, |0|\} = \max\{1, 2, 3, 0\} = 3. \end{aligned}$$

Exemplo 3.2. *Seja* $\mathcal{B} = \{\vec{v} = (x, y) \in \mathbb{R}^2; ||\vec{v}|| \le 1\}$*, onde* $||\cdot||$ *é uma norma vetorial. Dependendo da norma considerada, o conjunto* $\mathcal{B}$ *define uma região determinada de* $\mathbb{R}^2$*, como será mostrada a seguir.*

Para a Norma 1, o conjunto $\mathcal{B}$ *será representado por* $\mathcal{B}_1 = \{\vec{v} = (x, y) \in \mathbb{R}^2; ||\vec{v}||_1 \le 1\}$*. O elemento* $\vec{v}$ *de* $\mathcal{B}_1$*, satisfaz:*

$$\begin{aligned} ||\vec{v}||_1 \le 1 \quad &\Rightarrow \quad |x| + |y| \le 1 \\ &\Rightarrow \quad (\pm x) + (\pm y) \le 1 \\ &\Rightarrow \quad x + y \le 1 \text{ ou } -x + y \le 1 \text{ ou } x - y \le 1 \text{ ou } -x - y \le 1. \end{aligned}$$

Isso significa que os elementos do conjunto $\mathcal{B}_1$ *está compreendido na região cuja fronteira é dada pelo quadrado* $PQRS$*, como se ilustra na Figura 3.1.*

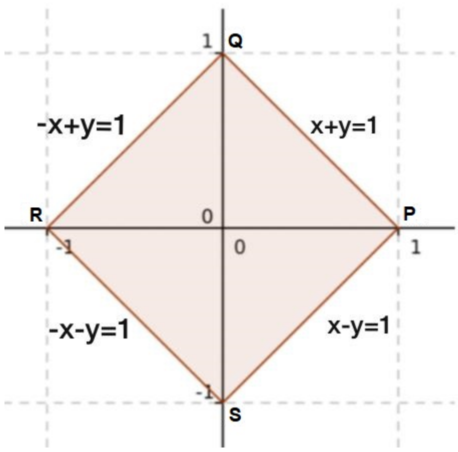


Figura 3.1: $\mathcal{B}_1$ [4].

Para a Norma 2, o conjunto $\mathcal{B}$ *será representada pelo conjunto* $\mathcal{B}_2 = \{\vec{v} = (x, y) \in \mathbb{R}^2; ||\vec{v}||_2 \leq 1\}$. *Seja* $\vec{v} \in \mathcal{B}_2$, *então temos que*

$$||\vec{v}||_2 \leq 1 \Longrightarrow \sqrt{|x|^2 + |y|^2} \leq 1 \Longrightarrow x^2 + y^2 \leq 1.$$

Isso quer dizer que o conjunto $\mathcal{B}_2$ *está compreendido na região cuja fronteira é um círculo de raio 1 (Figura 3.2).*

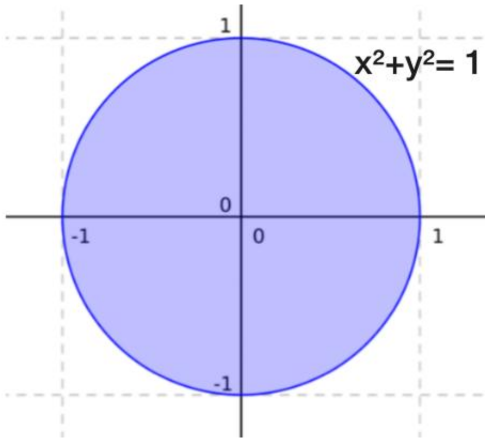


Figura 3.2: $\mathcal{B}_2$ [4].

Já para a Norma ∞, *o conjunto* $\mathcal{B}$ *será representada por* $\mathcal{B}_\infty = \{\vec{v} = (x, y) \in \mathbb{R}^2; ||\vec{v}||_\infty \leq 1\}$. *Os elementos deste conjunto satisfazem a condição*

$$||\vec{v}||_\infty \leq 1 \Rightarrow \max\{|x|, |y|\} \leq 1 \Rightarrow |x| \leq 1 \; ou \; |y| \leq 1 \Rightarrow -1 \leq x \leq 1 \; ou \; -1 \leq y \leq 1.$$

Condição que define uma região cuja fronteira é dada pelo quadrado $PQRS$, *como na Figura 3.3.*

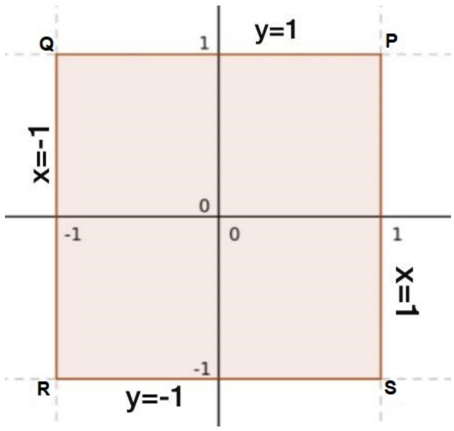


Figura 3.3: $\mathcal{B}_\infty$ [4].

Definição 3.1.2. ***(Normas equivalentes)*** *Duas normas $||\cdot||_p$ e $||\cdot||_q$ são ditas equivalentes se existem constantes $a > 0$ e $b > 0$ tais que:*

$$||\cdot||_p \leq a||\cdot||_q \quad e \quad ||\cdot||_q \leq b||\cdot||_p.$$

Exemplo 3.3. *As normas $||\cdot||_1$, $||\cdot||_2$ e $||\cdot||_\infty$ são equivalentes.*

De fato, seja $\vec{v} = (v_1, v_2, ..., v_n)$ em $\mathbb{R}^n$ e tome $|v_i|$ como a maior das coordenadas do vetor $\vec{v}$. Temos que, por (3.1), (3,2) e (3.3),

$$||\vec{v}||_1 = |v_1| + |v_2| + \cdots + |v_n|,$$

$$||\vec{v}||_2 = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2},$$

$$||\vec{v}||_\infty = \max\{|v_1|, |v_2|, \cdots, |v_n|\}.$$

Segue que

$$\begin{aligned} |v_i|^2 \leq |v_1|^2 + |v_2|^2 + \cdots + |v_n|^2 \quad &\Rightarrow \quad \sqrt{|v_i|^2} \leq \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2} \\ &\Rightarrow \quad |v_i| \leq ||\vec{v}||_2 \\ &\Rightarrow \quad ||\vec{v}||_\infty \leq ||\vec{v}||_2. \end{aligned}$$

A constante de equivalência neste caso é 1. Agora note que,

$$\begin{aligned} (||\vec{v}||_1)^2 &= (|v_1| + |v_2| + \cdots + |v_n|)^2 \\ &= |v_1|^2 + |v_2|^2 + \cdots + |v_n|^2 + \boldsymbol{k} \\ &= (||\vec{v}||_2)^2 + \boldsymbol{k}. \end{aligned}$$

Ou seja, $||\vec{v}||_2 \leq ||\vec{v}||_1$, cuja constante de equivalência também é 1. Por fim, temos

$$||\vec{v}||_1 = |v_1| + |v_2| + \cdots + |v_n| \leq |v_i| + |v_i| + \cdots + |v_i| = n|v_i| = n||\vec{v}||_\infty.$$

Logo, a constante de equivalência aqui é n. Portanto,

$$||\vec{v}||_\infty \leq ||\vec{v}||_2 \leq ||\vec{v}||_1 \leq n||\vec{v}||_\infty.$$

Pelo Exemplo 3.1 podemos observar que dado o vetor $\vec{v} = (1, 2, -3, 0)$, se confirma que

$$||\vec{v}||_\infty \leq ||\vec{v}||_2 \leq ||\vec{v}||_1.$$

Geometricamente, podemos representar os três conjuntos $\mathcal{B}_1$, $\mathcal{B}_2$ e $\mathcal{B}_\infty$ associados as três normas usuais (veja Figura 3.4).

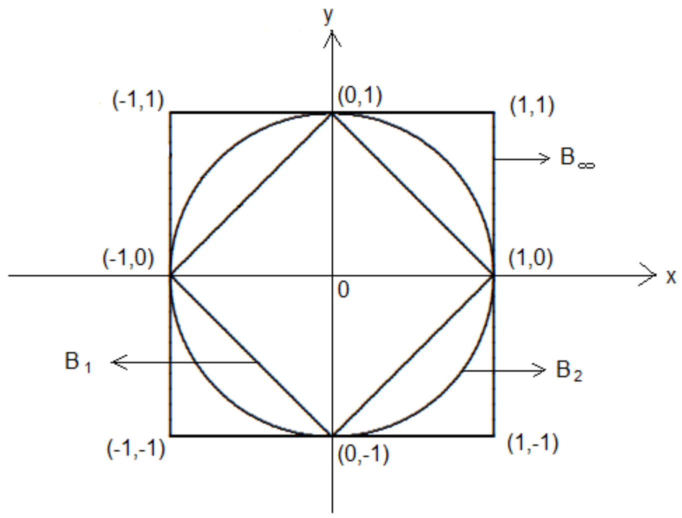


Figura 3.4: Conjuntos $\mathcal{B}_1$, $\mathcal{B}_2$ e $\mathcal{B}_\infty$ [3].

3.2 Norma de Matrizes

Definição 3.2.1. *Seja $M_{n\times n}$ um espaço vetorial sobre $\mathbb{R}$. Uma norma em $M_{n\times n}$ é uma função real de $M_{n\times n}$ em $\mathbb{R}$, denotada como $||\cdot|| : M_{n\times n} \to \mathbb{R}$, que satisfaz as seguintes condições, para quaisquer que sejam A, B em $M_{n\times n}$ e para todo escalar k em $\mathbb{R}$, tem-se:*

(i) $||A|| \geq 0$ *e* $||A|| = 0$*, se, e somente se,* $A = O$*;*

(ii) $||kA|| = |k|\ ||A||$*;*

(iii) $||A + B|| \leq ||A|| + ||B||$.

Segue um tipo particular de norma matricial bastante usada na literatura.

Definição 3.2.2. ***(Normas Matriciais Subordinadas)*** *Dada uma norma de vetor $||\vec{v}||$, pode-se associar uma norma de matriz denominada de Norma Subordinada definida pela seguinte expressão:*

$$||A|| = \sup_{||\vec{x}||=1} ||A\vec{x}||,$$

onde $||A\vec{x}||$ é uma norma vetorial.

O supremo de $\{||A\vec{x}||; ||\vec{x}|| = 1\}$ é assumido sobre todos os vetores n-dimensionais $\vec{x}$ com a norma unitária. Lembrando que $\{\vec{x}; ||\vec{x}|| = 1\}$ é um conjunto fechado e limitado. Logo, pelo Teorema do Valor Extremo[1], a imagem do conjunto $\{||A\vec{x}||; ||\vec{x}|| = 1\}$ possui máximo e mínimo, e assim, o supremo é igual

[1]Este resultado é um teorema de Análise e se encontra mais aprofundado em [10].

ao máximo. Diante disso, a definição acima pode ser reescrita por:

$$||A|| = \max_{||\vec{x}||=1} ||A\vec{x}||. \tag{3.4}$$

Lema 3.2.1. *Uma norma de matriz subordinada satisfaz as seguintes propriedades.*

(i) $||A|| > 0$*, a menos que* $A = O$*, então* $||A|| = 0$*.*

(ii) Para qualquer escalar k *e qualquer* A*,* $||kA|| = |k|\ ||A||$*.*

(iii) Para quaisquer duas matrizes A *e* B*,*

$$||A + B|| \leq ||A|| + ||B||.$$

Demonstração.

(i) Se $A \neq O$*, existe um* $\vec{x} \neq \vec{0}$ *tal que* $A\vec{x} \neq \vec{0}$*. Podemos escalar* $\vec{x}$ *para que* $||\vec{x}|| = 1$ *e ainda tenhamos* $A\vec{x} \neq \vec{0}$*. Assim, pela condição (i) de normas vetoriais,* $||A|| > 0$*. Se* $A = O$*,* $A\vec{x} = \vec{0}$ *para todo* $\vec{x}$ *e, portanto,* $||A|| = 0$*.*

(ii) Na condição (ii) para normas vetoriais, cabe que

$$||kA|| = \max_{||\vec{x}||=1} ||kA\vec{x}|| = \max_{||\vec{x}||=1} |k| \cdot ||A\vec{x}|| = |k| \cdot \max_{||\vec{x}||=1} ||A\vec{x}|| = |k|\ ||A||.$$

(iii) Na condição (iii) para normas vetoriais,

$$\begin{aligned} ||A + B|| &= \max_{||\vec{x}||=1} ||(A + B)\vec{x}|| \\ &= \max_{||\vec{x}||=1} ||A\vec{x} + B\vec{x}|| \\ &\leq \max_{||\vec{x}||=1} [\ ||A\vec{x}|| + ||B\vec{x}||\ ] \\ &\leq \max_{||\vec{x}||=1} ||A\vec{x}|| + \max_{||\vec{x}||=1} ||B\vec{x}|| \\ &= ||A|| + ||B|| \end{aligned}$$

∎

As afirmações (i)*,* (ii) *e* (iii) *indicam que efetivamente a norma subordinada é uma norma. Além disso, satisfaz mais duas propriedades:*

(iv) Para qualquer vetor n*-dimensional* $\vec{x}$ *e qualquer* A*,*

$$||A\vec{x}|| \leq ||A|| \cdot ||\vec{x}||.$$

(v) Para quaisquer duas matrizes A e B,

$$||AB|| \leq ||A|| \cdot ||B||.$$

Tal propriedade é conhecida como Norma Consistente.

Demonstração.

(iv) Para qualquer $\vec{x} \neq \vec{0}$,

$$||A\vec{x}|| = ||\vec{x}|| \cdot \left\| A \left(\frac{1}{||\vec{x}||}\vec{x} \right) \right\| \leq ||\vec{x}|| \cdot ||A||,$$

com

$$\left\| \frac{1}{||\vec{x}||}\vec{x} \right\| = 1$$

e $||A||$ ocorre para o máximo em (3.4). O resultado é trivial para $\vec{x} = \vec{0}$.

(v) $||AB|| = \max\limits_{||\vec{x}||=1} ||AB\vec{x}|| = \max\limits_{||\vec{x}||=1} ||A(B\vec{x})|| \leq \max\limits_{||\vec{x}||=1} [\, ||A|| \cdot ||B\vec{x}||\,]$

por (iv) acima. Portanto

$$||AB|| \leq ||A|| \max_{||\vec{x}||=1} ||B\vec{x}|| = ||A|| \cdot ||B||.$$

∎

Apresentaremos agora, alguns casos clássicos de Normas de Matrizes Subordinadas, que estão associadas a Norma 1 de vetores, Norma ∞ de vetores e Norma 2 de vetores, respectivamente.

Segue pela Definição 3.2.2, que a Norma 1 é dada por

$$||A||_1 = \max_{||\vec{x}||_1=1} ||A\vec{x}||_1,$$

a Norma ∞, pela expressão

$$||A||_\infty = \max_{||\vec{x}||_\infty=1} ||A\vec{x}||_\infty$$

e a Norma 2, por

$$||A||_2 = \max_{||\vec{x}||_2=1} ||A\vec{x}||_2.$$

Seja uma matriz A em $M_{n\times n}$, ao calcular $||A||_\infty$, por exemplo, precisa-se encontrar um vetor $\vec{x}$ tal que $||\vec{x}||_\infty = 1$ e onde $||A\vec{x}||_\infty$ seja o máximo. Eis aí um obstáculo para se obter a $||A||_\infty$ através da definição.

Diante disso, nestes três casos particulares de normas matriciais subordinadas, existem alternativas mais simplificadas de calculá-las. Veja a proposição a seguir.

Proposição 3.2.1. *Seja A uma matriz de tamanho $n \times n$, então temos as seguintes fórmulas equivalentes para calcular as normas subordinadas:*

(i) Norma 1 (ou Norma Coluna)

$$||A||_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|\ ; \tag{3.5}$$

(ii) Norma ∞ (ou Norma Linha)

$$||A||_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|\ ; \tag{3.6}$$

(iii) Norma 2 (ou Norma Euclidiana)

$$||A||_2 = \max\{\sqrt{\lambda};\lambda \textit{ é autovalor de } A^TA\}\ . \tag{3.7}$$

Demonstração. Antes de demonstrarmos estes três resultados, considere $A\vec{x}=\vec{y}$, onde

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix},\ \vec{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\ \text{e}\ \vec{y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

então,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n\\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n\\ \vdots & & \vdots & & \ddots & & \vdots\\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n\end{bmatrix}=\begin{bmatrix}\sum_{j=1}^n a_{1j}x_j\\ \sum_{j=1}^n a_{2j}x_j\\ \vdots\\ \sum_{j=1}^n a_{nj}x_j\end{bmatrix}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}.$$

Logo,

$$y_i=\sum_{j=1}^{n}a_{ij}x_j,\qquad i=1,2,\ldots,n. \tag{3.8}$$

(i) Norma 1 (ou Norma Coluna)

Queremos provar que

$$||A||_1=\max_{||\vec{x}||_1=1}||A\vec{x}||_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|.$$

Temos que,

$$||A\vec{x}||_1=||\vec{y}||_1=\sum_{i=1}^{n}|y_i|.$$

Por (3.8), temos:

$$||A\vec{x}||_1=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right|. \tag{3.9}$$

Por outro lado, aplicando a desigualdade triangular generalizada de números reais

$$|y_i| = \left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \sum_{j=1}^{n} |a_{ij}|\,|x_j|. \tag{3.10}$$

Usando (3.10) em (3.9), temos:

$$||A\vec{x}||_1 = \sum_{i=1}^{n}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|\,|x_j|. \tag{3.11}$$

Mas $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n}\left(\sum_{i=1}^{n}|a_{ij}|\right)|x_j|$. Logo, temos:

$$\begin{aligned}
||A\vec{x}||_1 &\leq \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}||x_j| \\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}|a_{ij}|\right)|x_j| \\
&\leq \sum_{j=1}^{n}\left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)|x_j| \\
&= \left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)\sum_{j=1}^{n}|x_j| \\
&= \left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)||\vec{x}||_1.
\end{aligned}$$

Então,

$$||A\vec{x}||_1 \leq \left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)||\vec{x}||_1. \tag{3.12}$$

Assim, de (3.12), para $||\vec{x}||_1 = 1$, temos que:

$$\begin{aligned}
||A\vec{x}||_1 &\leq \left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)||\vec{x}||_1 \\
&= \left(\max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}|\right)\cdot 1 \\
&= \max_{1\leq k\leq n}\sum_{i=1}^{n}|a_{ik}| \\
&= \max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|.
\end{aligned}$$

E, portanto,

$$||A\vec{x}||_1 \leq \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| \tag{3.13}$$

onde, $\max\limits_{1\leq j\leq n} \sum\limits_{i=1}^{n} |a_{ij}|$ é o limitante para a norma $||A||_1$.

Sabemos que

$$\max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$$

irá acontecer em alguma coluna. Seja p uma das colunas onde acontece o máximo. Logo:

$$\max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{ip}|. \tag{3.14}$$

Agora, considere $\vec{x} = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$, com 1 na posição p. Mas, $||\vec{x}||_1 = 1$. Contudo, dada uma matriz A, multiplicada por este vetor $\vec{x}$, teremos:

$$\vec{y} = A\vec{x} = (a_{1p}, a_{2p}, \ldots, a_{np}).$$

Assim,

$$||A\vec{x}||_1 = \sum_{i=1}^{n} |a_{ip}|.$$

Então, existe $\vec{x}$ tal que $||A||_1$ é igual ao limitante

$$\max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|,$$

segue que

$$||A||_1 = \max_{||\vec{x}||_1=1} ||A\vec{x}||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|.$$

(ii) Norma ∞ (ou Norma Linha)

Queremos provar que

$$||A||_\infty = \max_{||\vec{x}||_\infty=1} ||A\vec{x}||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Temos que,

$$||A\vec{x}||_\infty = ||\vec{y}||_\infty = \max_{1\leq i\leq n} |y_i|.$$

De (3.8), temos:

$$||A\vec{x}||_\infty = \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{ij}x_j \right|. \tag{3.15}$$

Mas,

$$\begin{aligned}\left|\sum_{j=1}^{n} a_{ij}x_j\right| &\leq \sum_{j=1}^{n} |a_{ij}|\ |x_j| \\ &\leq \sum_{j=1}^{n} |a_{ij}|\ ||\vec{x}||_\infty \\ &= \left(\sum_{j=1}^{n} |a_{ij}|\right) ||\vec{x}||_\infty.\end{aligned}$$

Para $||\vec{x}||_\infty = 1$, temos que:

$$\left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \sum_{j=1}^{n} |a_{ij}|.$$

Então,

$$\max_{1\leq i\leq n}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \max_{1\leq i\leq n}\left(\sum_{j=1}^{n} |a_{ij}|\right).$$

Logo, de (3.15)

$$||A||_\infty \leq \max_{1\leq i\leq n}\left(\sum_{j=1}^{n} |a_{ij}|\right). \tag{3.16}$$

Reparem que, sempre existe o seguinte número $\max_{1\leq i\leq n}\left(\sum_{j=1}^{n} |a_{ij}|\right)$. Denominando a linha onde isso acontece de p, temos que $\sum_{j=1}^{n} |a_{pj}|$ representará o máximo.

Seja $\vec{x}$ o vetor cuja componente é dada por $x_j = sinal(a_{pj})$ (função sinal), com $j = 1, 2, \ldots, n$. Assim, se $a_{pj} > 0$, então $x_j = +1$; se $a_{pj} < 0$, então $x_j = -1$; se $a_{pj} = 0$, então $x_j = 0$.

Observe que, o vetor $\vec{x}$ montado tem coordenadas $x_j = 1$ ou -1. Logo, $||\vec{x}||_\infty = 1$. Além disso, temos que $|a_{pj}| = a_{pj}x_j$, pois se $a_{pj} > 0$, então $x_j = 1$ e $a_{pj} \times 1 = a_{pj} > 0$; se $a_{pj} < 0$, então $x_j = -1$ e $a_{pj} \times (-1) = -a_{pj} > 0$; se $a_{pj} = 0$, então $x_j = 0$ e $a_{pj} \times 0 = 0$.

Seja $\vec{y} = \vec{x}$, com $\vec{x}$ sendo o vetor construído acima. Logo, de (3.8), para $i = p$:

$$|y_p| = \left|\sum_{j=1}^{n} a_{pj}x_j\right| = \sum_{j=1}^{n} |a_{pj}|$$

onde $\sum_{j=1}^{n} |a_{pj}|$ é o limitador da norma. Então,

$$||A||_\infty = \max_{||\vec{x}||_\infty=1} ||A\vec{x}||_\infty = \max_{1\leq i\leq n}\left(\sum_{j=1}^{n} |a_{ij}|\right).$$

(iii) Norma 2 (ou Norma Euclidiana)

Queremos provar que

$$||A||_2 = \max_{||\vec{x}||_2=1} ||A\vec{x}||_2 = \max\{\sqrt{\lambda}; \lambda \text{ é o autovalor de } A^T A\} \ .$$

Considerando $A\vec{x} = \vec{y}$, temos:

$$\begin{aligned} (||A\vec{x}||_2)^2 &= (||\vec{y}||_2)^2 \\ &= \vec{y}^T \vec{y} \\ &= (A\vec{x})^T (A\vec{x}) \\ &= \vec{x}^T A^T A\vec{x}. \end{aligned}$$

Considerando $A^T A = B$, temos:

$$(||A\vec{x}||_2)^2 = \vec{x}^T B\vec{x}. \tag{3.17}$$

Neste caso, pela Observação 1.2.3, B é uma matriz simétrica e pelo Teorema 1.3.5, B é ortogonalmente diagonalizável. Logo, existe uma matriz ortogonal P tal que

$$\vec{x}^T B\vec{x} = \vec{x}^T P\Lambda P^T \vec{x},$$

em que Λ é uma matriz diagonal, formada pelos autovalores da matriz B. Fazendo $\vec{z} = P^T\vec{x}$, temos

$$\vec{x}^T B\vec{x} = \vec{z}^T \Lambda\vec{z}. \tag{3.18}$$

Como P é ortogonal, então P^T também é e $||\vec{x}||_2 = ||\vec{z}||_2$. Logo, se $||\vec{x}||_2 = 1$, então $||\vec{z}||_2 = 1$.

De (3.17) e (3.18), tiramos:

$$(||A\vec{x}||_2)^2 = \vec{z}^T \Lambda\vec{z}.$$

É fácil ver que

$$\vec{z}^T \Lambda\vec{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.$$

Portanto,

$$(||A\vec{x}||_2)^2 = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2. \tag{3.19}$$

Baseando-se na Observação 1.2.4, suponha que os λ's estão ordenados tais que:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0.$$

isto é, λ_1 é o maior autovalor da matriz em questão. Assim,

$$\begin{array}{rcl}
\lambda_1 \geq \lambda_2 & \Longrightarrow & \lambda_1 z_2^2 \geq \lambda_2 z_2^2 \\
\lambda_1 \geq \lambda_3 & \Longrightarrow & \lambda_1 z_3^2 \geq \lambda_3 z_3^2 \\
& \vdots & \\
\lambda_1 \geq \lambda_n & \Longrightarrow & \lambda_1 z_n^2 \geq \lambda_n z_n^2
\end{array}$$

então,

$$\begin{array}{rcl}
\lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2 & \leq & \lambda_1 z_1^2 + \lambda_1 z_2^2 + \ldots + \lambda_1 z_n^2 \\
& = & \lambda_1 (z_1^2 + z_2^2 + \ldots + z_n^2).
\end{array}$$

Logo, de (3.19) obtemos:

$$(||A\vec{x}||_2)^2 \leq \lambda_1 (z_1^2 + z_2^2 + \ldots + z_n^2). \tag{3.20}$$

Mas $||\vec{x}||_2 = 1 = ||\vec{z}||_2$, isto é:

$$\begin{array}{rcl}
||\vec{z}||_2 & = & \sqrt{z_1^2 + z_2^2 + \ldots + z_n^2} \\
1 & = & \sqrt{z_1^2 + z_2^2 + \ldots + z_n^2} \\
1^2 & = & \left(\sqrt{z_1^2 + z_2^2 + \ldots + z_n^2}\right)^2 \\
1 & = & z_1^2 + z_2^2 + \ldots + z_n^2.
\end{array}$$

Logo, (3.20) fica:

$$\begin{array}{rcl}
(||A\vec{x}||_2)^2 & \leq & \lambda_1 (z_1^2 + z_2^2 + \ldots + z_n^2) \\
& = & \lambda_1 \cdot 1 \\
& = & \lambda_1.
\end{array}$$

Diante disso, $(||A\vec{x}||_2)^2 \leq \lambda_1$, ou melhor, $||A\vec{x}||_2 \leq \sqrt{\lambda_1}$, onde λ_1 é o maior autovalor de $B = A^T A$.

Escolhemos $\vec{x}$ tal que $\vec{z} = P^T \vec{x} = (1, 0, \ldots, 0)$.

Como existe $\vec{x}$ tal que $||\vec{x}||_2 = 1$ e $(||A\vec{x}||_2)^2 = \lambda_1$. Então:

$$||A||_2 = \max_{||\vec{x}||_2=1} ||A\vec{x}||_2 = \max_{||\vec{x}||_2=1} \{\sqrt{\lambda}; \lambda \text{ é autovalor de } A^TA\}.$$

∎

Proposição 3.2.2. *Seja $A \in \mathbb{M}_{n\times n}$ em $\mathbb{R}$, então*

$$\begin{aligned} ||A||_\infty &= \max\{||A\vec{x}||_\infty;\ ||\vec{x}||_\infty = 1\} \\ &= \max\{||A\vec{x}||_\infty;\ \underbrace{\vec{x} = (x_1, x_2, \ldots, x_n)^T,\ \textit{onde } |x_i| = 1,\ \forall i = 1, 2, \ldots, n}_{\textit{Condição}}\}. \end{aligned}$$

A seguir, apresenta-se a aplicação das três normas matriciais subordinadas.

Exemplo 3.4. *Dado a matriz $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. O cálculo das Normas 1, ∞ e 2 seguindo a Proposição 3.2.1, é dado por:*

De (3.5) e (3.6) segue que,

$$\begin{aligned} ||A||_1 &= \max\{(|1| + |0|), (|2| + |2|)\} = \max\{(1+0), (2+2)\} = \max\{1, 4\} = 4, \\ ||A||_\infty &= \max\{(|1| + |2|), (|0| + |2|)\} = \max\{(1+2), (0+2)\} = \max\{3, 2\} = 3. \end{aligned}$$

Para o cálculo da Norma 2, segundo a expressão (3.7), é necessário que primeiro obtemos a matriz A^TA. Então,

$$A^TA = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix}.$$

Em seguida, deve-se calcular os autovalores da matriz A^TA encontrada. Logo, para o cálculo dos autovalores, usa-se $\det(A^TA - \lambda I) = 0$, isto é,

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 8-\lambda \end{vmatrix} = 0.$$

Assim,

$$\begin{aligned} (1-\lambda)(8-\lambda) - 2\cdot 2 &= 0 \\ 8 - \lambda - 8\lambda + \lambda^2 - 4 &= 0 \\ \lambda^2 - 9\lambda + 4 &= 0. \end{aligned}$$

Disto,

$$\begin{aligned}\Delta &= b^2 - 4ac\\ \Delta &= (-9)^2 - 4\cdot 1\cdot 4\\ \Delta &= 81 - 16\\ \Delta &= 65.\end{aligned}$$

Daí,

$$\begin{aligned}\lambda &= \frac{-b \pm \sqrt{\Delta}}{2a}\\ \lambda &= \frac{-(-9) \pm \sqrt{65}}{2\cdot 1}\\ \lambda &= \frac{9 \pm \sqrt{65}}{2}\\ \lambda_1 &\approx 8,53\\ \lambda_2 &\approx 0,47.\end{aligned}$$

Portanto,

$$||A||_2 \approx \max\{\sqrt{0,47}, \sqrt{8,53}\} \approx \max\{0,68, 2,92\} \approx 2,92.$$

A seguir, apresenta-se o cálculo das Normas 1 e ∞ da mesma matriz A do Exemplo 3.4, porém, seguindo a Definição 3.2.2.

Cálculo de $||A||_1$: Já vimos na seção 3.1 (Norma de vetores), que o conjunto $\mathcal{B}_1$ está compreendido na região cuja fronteira é dada pelo quadrado $PQRS$, como na Figura 3.1. A norma subordinada 1 é dada por

$$||A||_1 = \max_{||\vec{x}||_1=1} ||A\vec{x}||_1.$$

Logo, para calcular $||A||_1$, precisa-se determinar o maior valor da Norma 1 de $\vec{y}$, onde $\vec{y} = A\vec{x}$, com $||\vec{x}||_1 = 1$.

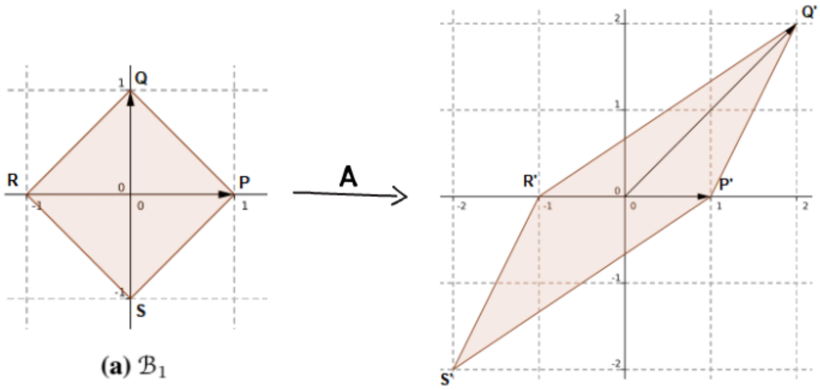
Observação 3.2.1. *Seja $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$, temos:*

$$\begin{aligned}
\vec{y} &= A\vec{x} \\
&= A(x_1\vec{e}_1 + x_2\vec{e}_2) \\
&= x_1 A\vec{e}_1 + x_2 A\vec{e}_2 \\
&= x_1 \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 2x_2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + 2x_2 \\ 2x_2 \end{bmatrix}.
\end{aligned}$$

Considerando $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, então:

$$\begin{cases} y_1 = x_1 + 2x_2 \\ y_2 = 2x_2 \end{cases}.$$

Neste caso, $\vec{y} = (x_1 + 2x_2, 2x_2)$, onde x_1 e x_2 são as coordenadas do vetor $\vec{x}$. A imagem do conjunto $\mathcal{B}_1$, via a transformação $\vec{y} = A\vec{x}$ é ilustrada na Figura 3.5(b).



(a) $\mathcal{B}_1$

(b) $T(A) = \{A\vec{x}; \vec{x} \in \mathcal{B}_1\}$

Figura 3.5: Representação de $\mathcal{B}_1$ e $T(\mathcal{B}_1)$, onde $T(\vec{x}) = A\vec{x}$ [4].

Considere o conjunto

$$C_1 = \{\vec{x} = (x_1, x_2); ||\vec{x}||_1 = 1\} = \{\vec{x} = (x_1, x_2); |x_1| + |x_2| = 1\}.$$

Tal conjunto é o quadrado $PQRS$, isto é, apenas o contorno da região que está compreendida o conjunto $\mathcal{B}_1$.

Podemos observar pela Figura 3.5(a) que os vetores $\overrightarrow{OP}, \overrightarrow{OQ}, \overrightarrow{OR}, \overrightarrow{OS} \in C_1$. Logo, aplicando a transformação temos que para o vetor $\overrightarrow{OP} = (1,0)$, $x_1 = 1$ e $x_2 = 0$. Então,

$$\begin{cases} y_1 = x_1 + 2x_2 = 1 + 2 \cdot 0 = 1 + 0 = 1 \\ y_2 = 2x_2 = 2 \cdot 0 = 0 \end{cases}.$$

Logo, $(y_1, y_2) = (1,0) = \overrightarrow{OP'}$. Para o vetor $\overrightarrow{OQ} = (0,1)$, $x_1 = 0$ e $x_2 = 1$. Assim,

$$\begin{cases} y_1 = x_1 + 2x_2 = 0 + 2 \cdot 1 = 0 + 2 = 2 \\ y_2 = 2x_2 = 2 \cdot 1 = 2 \end{cases}.$$

Logo, $(y_1, y_2) = (2,2) = \overrightarrow{OQ'}$. Para o vetor $\overrightarrow{OR} = (-1,0)$, $x_1 = -1$ e $x_2 = 0$. Então,

$$\begin{cases} y_1 = x_1 + 2x_2 = -1 + 2 \cdot 0 = -1 + 0 = -1 \\ y_2 = 2x_2 = 2 \cdot 0 = 0 \end{cases}.$$

Logo, $(y_1, y_2) = (-1,0) = \overrightarrow{OR'}$. E por fim, para o vetor $\overrightarrow{OS} = (0,-1)$, $x_1 = 0$ e $x_2 = -1$. Assim,

$$\begin{cases} y_1 = x_1 + 2x_2 = 0 + 2(-1) = 0 - 2 = -2 \\ y_2 = 2x_2 = 2(-1) = -2 \end{cases}.$$

Logo, $(y_1, y_2) = (-2,-2) = \overrightarrow{OS'}$. Portanto, obtemos a seguinte relação:

- $\overrightarrow{OP} = (1,0)$ se transforma em $\overrightarrow{OP'} = (1,0)$;
- $\overrightarrow{OQ} = (0,1)$ se transforma em $\overrightarrow{OQ'} = (2,2)$;
- $\overrightarrow{OR} = (-1,0)$ se transforma em $\overrightarrow{OR'} = (-1,0)$;
- $\overrightarrow{OS} = (0,-1)$ se transforma em $\overrightarrow{OS'} = (-2,-2)$.

Assim, pela Proposição 2.1.1, os segmentos $\overline{PQ}$, $\overline{QR}$, $\overline{RS}$ e $\overline{SP}$ são transformados em $\overline{P'Q'}$, $\overline{Q'R'}$, $\overline{R'S'}$ e $\overline{S'P'}$ respectivamente. Portanto, esta relação mostra que o contorno do quadrado $PQRS$ foi transformado no contorno do paralelogramo $P'Q'R'S'$.

Além do mais, pela definição em (3.4),

$$||A||_1 = \max\{||\vec{y}||_1 = ||A\vec{x}||_1; \vec{x} \in C_1\} = \max\{|y_1| + |y_2|; \vec{x} \in C_1\}.$$

Levando em consideração o resultado contido na Proposição 3.2.2, segue que, para $\overrightarrow{OP'} = (1,0)$, temos

$$|y_1| + |y_2| = |1| + |0| = 1 + 0 = 1,$$

para $\overrightarrow{OQ'} = (2,2)$, temos

$$|y_1| + |y_2| = |2| + |2| = 2 + 2 = 4.$$

Para $\overrightarrow{OR'} = (-1,0)$, temos

$$|y_1| + |y_2| = |-1| + |0| = 1 + 0 = 1$$

e, por fim, para $\overrightarrow{OS'} = (-2,-2)$, temos

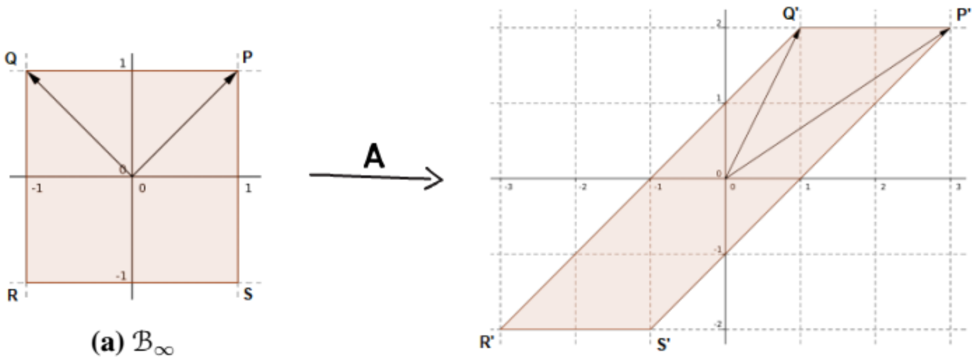
$$|y_1| + |y_2| = |-2| + |-2| = 2 + 2 = 4.$$

Portanto,

$$||A||_1 = \max\{1,4,1,4\} = 4,$$

como calculado no Exemplo 3.4.

Cálculo de $||A||_\infty$: De maneira análoga, a imagem do conjunto $\mathcal{B}_\infty$, via a transformação $\vec{y} = A\vec{x}$ é ilustrada na Figura 3.6(b).



(a) $\mathcal{B}_\infty$ **(b)** $T(A) = \{A\vec{x}; \vec{x} \in \mathcal{B}_\infty\}$

Figura 3.6: Representação de $\mathcal{B}_\infty$ e $T(\mathcal{B}_\infty)$, onde $T(\vec{x}) = A\vec{x}$ [4].

Considere o conjunto

$$C_\infty = \{\vec{x} = (x_1,x_2); ||\vec{x}||_\infty = 1\} = \{\vec{x} = (x_1,x_2); \max\{|x_1|,|x_2|\} = 1\}.$$

Tal conjunto é o quadrado $PQRS$, isto é, apenas o contorno da região que está compreendida o conjunto $\mathcal{B}_\infty$.

Podemos observar pela Figura 3.6(a) que os vetores $\overrightarrow{OP}, \overrightarrow{OQ}, \overrightarrow{OR}, \overrightarrow{OS} \in C_\infty$. Logo, aplicando a transformação temos que para o vetor $\overrightarrow{OP} = (1, 1)$, $x_1 = 1$ e $x_2 = 1$. Então,

$$\begin{cases} y_1 = x_1 + 2x_2 = 1 + 2 \cdot 1 = 1 + 2 = 3 \\ y_2 = 2x_2 = 2 \cdot 1 = 2 \end{cases}.$$

Logo, $(y_1, y_2) = (3, 2) = \overrightarrow{OP'}$. Para o vetor $\overrightarrow{OQ} = (-1, 1)$, $x_1 = -1$ e $x_2 = 1$. Assim,

$$\begin{cases} y_1 = x_1 + 2x_2 = -1 + 2 \cdot 1 = -1 + 2 = 1 \\ y_2 = 2x_2 = 2 \cdot 1 = 2 \end{cases}.$$

Logo, $(y_1, y_2) = (1, 2) = \overrightarrow{OQ'}$. Para o vetor $\overrightarrow{OR} = (-1, -1)$, $x_1 = -1$ e $x_2 = -1$. Então,

$$\begin{cases} y_1 = x_1 + 2x_2 = -1 + 2 \cdot (-1) = -1 - 2 = -3 \\ y_2 = 2x_2 = 2 \cdot (-1) = -2 \end{cases}.$$

Logo, $(y_1, y_2) = (-3, -2) = \overrightarrow{OR'}$. E por fim, para o vetor $\overrightarrow{OS} = (1, -1)$, $x_1 = 1$ e $x_2 = -1$. Assim,

$$\begin{cases} y_1 = x_1 + 2x_2 = 1 + 2(-1) = 1 - 2 = -1 \\ y_2 = 2x_2 = 2(-1) = -2 \end{cases}.$$

Logo, $(y_1, y_2) = (-1, -2) = \overrightarrow{OS'}$. Portanto, obtemos a seguinte relação:

- $\overrightarrow{OP} = (1, 1)$ se transforma em $\overrightarrow{OP'} = (3, 2)$;
- $\overrightarrow{OQ} = (-1, 1)$ se transforma em $\overrightarrow{OQ'} = (1, 2)$;
- $\overrightarrow{OR} = (-1, -1)$ se transforma em $\overrightarrow{OR'} = (-3, -2)$;
- $\overrightarrow{OS} = (1, -1)$ se transforma em $\overrightarrow{OS'} = (-1, -2)$.

Assim, pela Proposição 2.1.1, os segmentos $\overline{PQ}$, $\overline{QR}$, $\overline{RS}$ e $\overline{SP}$ são transformados em $\overline{P'Q'}$, $\overline{Q'R'}$, $\overline{R'S'}$ e $\overline{S'P'}$ respectivamente. Portanto, esta relação mostra que o contorno do quadrado $PQRS$ foi transformado no contorno do paralelogramo $P'Q'R'S'$.

Além do mais, pela definição em (3.4),

$$||A||_\infty = \max\{||\vec{y}||_\infty = ||A\vec{x}||_\infty; \vec{x} \in C_\infty\} = \max\{\max\{|y_1|, |y_2|\}; \vec{x} \in C_\infty\}.$$

Levando em consideração a Proposição 3.2.2, segue que, para $\overrightarrow{OP'} = (3, 2)$, temos

$$\max\{|y_1|, |y_2|\} = \max\{|3|, |2|\} = \max\{3, 2\} = 3,$$

para $\overrightarrow{OQ'} = (1, 2)$, temos

$$\max\{|y_1|, |y_2|\} = \max\{|1|, |2|\} = \max\{1, 2\} = 2.$$

Para $\overrightarrow{OR'} = (-3, -2)$, temos

$$\max\{|y_1|, |y_2|\} = \max\{|-3|, |-2|\} = \max\{3, 2\} = 3$$

e, por fim, para $\overrightarrow{OS'} = (-1, -2)$, temos

$$\max\{|y_1|, |y_2|\} = \max\{|-1|, |-2|\} = \max\{1, 2\} = 2.$$

Portanto,

$$||A||_\infty = \max\{3, 2, 3, 2\} = 3,$$

como obtido no Exemplo 3.4.

A título de curiosidade,

$$||A||_2 = \max\{||A\vec{x}||_2; ||\vec{x}||_2 = 1\} \approx 2,92$$

está representada na Figura 3.7, mas neste trabalho não se encontra a construção.

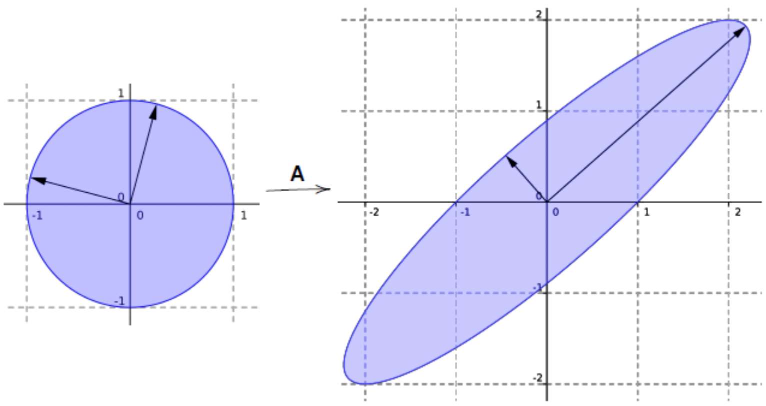


Figura 3.7: Representação de $\mathcal{B}_2$ e $T(\mathcal{B}_2)$, onde $T(\vec{x}) = A\vec{x}$ [4].

3.3 Relações entre as normas de A e A^T

A seguir apresentam-se algumas propriedades de normas, necessárias, para o entendimento das propriedades do número de condição de uma matriz.

Proposição 3.3.1. *Temos que:*

(a) $||I||_p = 1$*;*

(b) Se A é uma matriz ortogonal, então:

(i) $||A\vec{x}||_2 = ||\vec{x}||_2$*;*

(ii) $||A^T\vec{x}||_2 = ||\vec{x}||_2$*;*

(iii) $||A||_2 = 1$*;*

(iv) $||A^T||_2 = ||A^{-1}||_2 = 1$*;*

(c) $||A||_2 = ||A^T||_2$*;*

(d) $||A^{-1}||_2 = \sqrt{\rho((A^{-1})^T A^{-1})} = \dfrac{1}{\sqrt{\min\{|\lambda|;\ \lambda \textit{ é autovalor de } A^TA\}}}$*;*

(e) $||A^TA||_2 = ||A||_2^2$.

(f) $||A||_p \geq |\lambda|$*, para todo autovalor de A.*

Demonstração. Vejamos

(a) $||I||_p = \max\left\{\dfrac{||I\vec{x}||_p}{||\vec{x}||_p}\right\} = 1.$

(b) Temos que:

(i) $||A\vec{x}||_2 = \sqrt{(A\vec{x})^TA\vec{x}} = \sqrt{\vec{x}^TA^TA\vec{x}} = \sqrt{\vec{x}^TI\vec{x}} = \sqrt{\vec{x}^T\vec{x}} = ||\vec{x}||_2.$

(ii) $||A^T\vec{x}||_2 = \sqrt{(A^T\vec{x})^TA^T\vec{x}} = \sqrt{\vec{x}^T(A^T)^TA^T\vec{x}} = \sqrt{\vec{x}^TAA^T\vec{x}} = \sqrt{\vec{x}^TI\vec{x}} = \sqrt{\vec{x}^T\vec{x}} = ||\vec{x}||_2.$

(iii) $||A||_2 = \max\left\{\dfrac{||A\vec{x}||_2}{||\vec{x}||_2}\right\} = 1.$

(iv) $||A^T||_2 = \max\left\{\dfrac{||A^T\vec{x}||_2}{||\vec{x}||_2}\right\} = 1$. Por ser matriz ortogonal, $A^{-1} = A^T$, daí $||A^{-1}||_2 = 1$.

(c) Lembre-se que:

$$||A||_2 = \max\{\sqrt{|\lambda|};\ \lambda \text{ é um autovalor de } A^TA\} = \sqrt{\rho(A^TA)}$$

e norma $||A^T||_2 = \sqrt{\rho(AA^T)}$. Mas $\rho(A^TA) = \rho(AA^T)$. Disso, tem-se que:

$$||A||_2 = ||A^T||_2.$$

(d) Aplicação do resultado *(e)* da Proposição 1.3.2.

(e) Seja λ autovalor de A, então λ^2 é autovalor de A^2. Por outro lado,

$$(A^TA)^T(A^TA) = (A^TA)^2.$$

Mas,

$$\begin{aligned} ||A^TA||_2 &= \sqrt{\rho((A^TA)^T(A^TA))} \\ &= \sqrt{\rho((A^TA)^2)} \\ &= \rho(A^TA) \\ &= ||A||_2^2. \end{aligned}$$

(f) Seja $\vec{x}$ um autovetor de A e λ o correspondente autovalor, então:

$$A\vec{x} = \lambda\vec{x}.$$

Daí,

$$\begin{aligned} ||A\vec{x}||_p &= ||\lambda\vec{x}||_p \\ ||A\vec{x}||_p &= |\lambda| \; ||\vec{x}||_p \\ \frac{||A\vec{x}||_p}{||\vec{x}||_p} &= \frac{|\lambda| \; ||\vec{x}||_p}{||\vec{x}||_p}, \quad \vec{x} \neq \vec{0} \\ \left|\left| A\frac{x}{||\vec{x}||_p} \right|\right|_p &= |\lambda|, \quad \vec{x} \neq \vec{0} \\ ||A\vec{y}||_p &= |\lambda|, \quad ||\vec{y}||_p = 1. \end{aligned}$$

Mas, $||A||_p = \max\limits_{||\vec{y}||_p=1} ||A\vec{y}||_p$. Logo,

$$||A||_p \geq |\lambda|,$$

onde λ é autovalor de A.

∎

Seja $||\cdot||$ uma norma tal que $||I|| = 1$, então segue:

Teorema 3.3.1. *Seja F uma matriz não singular. Se $||F|| < 1$, então $(I - F)$ é não singular e*

$$||(I - F)^{-1}|| \leq (1 - ||F||)^{-1}.$$

A demonstração desse resultado pode ser vista em [9].

Capítulo 4

Condicionamento

Neste capítulo, apresenta-se a definição do número de condição de uma matriz. Número que aparece nos limitadores superiores para os erros relativos da solução de sistemas lineares que sofreram algum tipo de erro nos dados de entrada, os quais podem ser do tipo erro de arredondamento e/ou do tipo erro de falta de exatidão pelo fato de representarem informações de dados experimentais.

O número de condição de uma matriz indica a sensibilidade do sistema linear no cálculo da solução. Sistemas lineares fortemente sensíveis são aqueles em que pequenos erros nos dados de entrada são amplificados no processo de resolução e produzem, portanto, soluções distantes dos procurados. Esses sistemas são denominados de "mal condicionados".

Aqui também serão apresentados alguns exemplos de sistemas lineares bem e mal condicionados. As principais referências que utilizamos como base para este capítulo estão em [3], [8] e [9].

4.1 Introdução

O condicionamento de um problema é um conceito geral. Este conceito preocupa-se com a influência das impurezas nos dados de entrada, nos dados de saída. Sejam x e y os dados originais e ligeiramente perturbados, e sejam $f(x)$ e $f(y)$ as respectivas soluções. Então, temos o seguinte:

Problema bem condicionado. Se y estiver próximo de x, então $f(y)$ estará próximo de $f(x)$.

Problema mal condicionado. Mesmo se y estiver próximo de x, então $f(y)$ pode se afastar de $f(x)$ drasticamente.

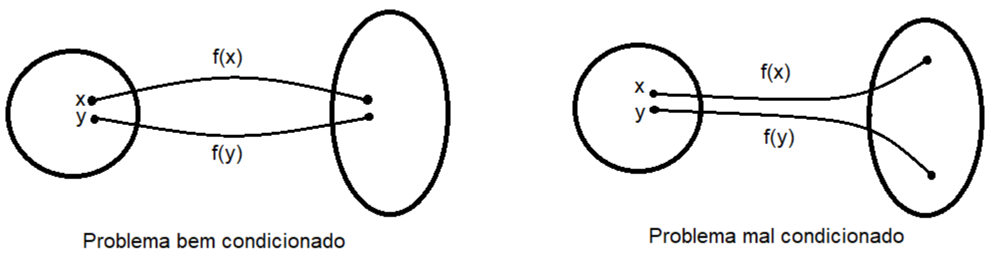


Figura 4.1: Problema bem condicionado (esquerda) e mal condicionado (direita) [9].

No caso de sistemas lineares, considere-se o seguinte exemplo.

Exemplo 4.1. *O cálculo da solução dos seguintes sistemas de duas equações lineares com duas incógnitas:*

$$\begin{bmatrix} 71 & 41 \\ 52 & 30 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \vec{b} = \begin{bmatrix} 100 \\ 70 \end{bmatrix}, \tag{4.1}$$

$$\begin{bmatrix} 71 & 41 \\ 52 & 30 \end{bmatrix} \begin{bmatrix} \widehat{x} \\ \widehat{y} \end{bmatrix} = \widehat{\vec{b}} = \begin{bmatrix} 100,4 \\ 69,3 \end{bmatrix}, \tag{4.2}$$

cujas soluções podem ser obtidas facilmente de forma exata. As soluções são, respectivamente:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -65 \\ 115 \end{bmatrix} \quad e \quad \begin{bmatrix} \widehat{x} \\ \widehat{y} \end{bmatrix} = \begin{bmatrix} -85,35 \\ 150,25 \end{bmatrix}.$$

Observe que o lado direito do segundo sistema de equações (4.2) pode ser visto como uma pequena perturbação do lado direito do primeiro sistema de equações (4.1). Porém, a solução obtida do segundo sistema linear está distante da solução do primeiro sistema. Isto é, pequenas alterações nos dados de entrada produziram soluções distantes. Logo, o sistema é mal condicionado.

A pergunta é: Como saber se o sistema é mal ou bem condicionado? O seguinte resultado mostra um exemplo da necessidade de definir o número de condição de uma matriz. Suponha que tenha dois sistemas lineares

$$A\vec{x} = \vec{b}, \tag{4.3}$$

$$A\widehat{\vec{x}} = \widehat{\vec{b}}. \tag{4.4}$$

onde o lado direito do segundo sistema linear é uma perturbação do lado direito do primeiro, isto é, $\widehat{\vec{b}} = \vec{b} + \vec{r}$, ou ainda, $\vec{r} = \widehat{\vec{b}} - \vec{b}$. Denote-se o erro em $\vec{x}$ por $\vec{e} = \widehat{\vec{x}} - \vec{x}$. Fazendo a equação (4.4) menos a equação (4.3), tem-se a seguinte equação para $\vec{e}$, dada por:

$$\begin{aligned} A\widehat{\vec{x}} - A\vec{x} &= \widehat{\vec{b}} - \vec{b} \\ A(\widehat{\vec{x}} - \vec{x}) &= \widehat{\vec{b}} - \vec{b} \\ A\vec{e} &= \vec{r}. \end{aligned} \tag{4.5}$$

De (4.5), obtém-se:

$$\begin{aligned} A^{-1}A\vec{e} &= A^{-1}\vec{r} \\ \vec{e} &= A^{-1}\vec{r}. \end{aligned} \tag{4.6}$$

De (4.3), (4.5) e (4.6), temos, respectivamente:

$$||\vec{x}|| = ||A^{-1}\vec{b}|| \leq ||A^{-1}||\ ||\vec{b}||, \tag{4.7}$$

$$||\vec{b}|| = ||A\vec{x}|| \leq ||A||\ ||\vec{x}||, \tag{4.8}$$

$$||\vec{r}|| = ||A\vec{e}|| \leq ||A||\ ||\vec{e}||, \tag{4.9}$$

$$||\vec{e}|| = ||A^{-1}\vec{r}|| \leq ||A^{-1}||\ ||\vec{r}||. \tag{4.10}$$

Dividindo-se (4.9) por $||A||\ ||\vec{x}||$ e (4.10) por $||\vec{x}||$, temos, respectivamente:

$$\frac{||\vec{r}||}{||A||\ ||\vec{x}||} \leq \frac{||\vec{e}||}{||\vec{x}||} \quad \text{e} \quad \frac{||\vec{e}||}{||\vec{x}||} \leq \frac{||A^{-1}||\ ||\vec{r}||}{||\vec{x}||}.$$

Usando a relação (4.7) $\left(||\vec{x}|| \leq ||A^{-1}||\ ||\vec{b}||\right)$ e a relação (4.8) $\left(||\vec{b}|| \leq ||A||\ ||\vec{x}|| \Rightarrow \frac{||\vec{b}||}{||A||} \leq ||\vec{x}||\right)$, convenientemente, temos:

$$\frac{||\vec{r}||}{||A||\ ||A^{-1}||\ ||\vec{b}||} \leq \frac{||\vec{e}||}{||\vec{x}||} \quad \text{e} \quad \frac{||\vec{e}||}{||\vec{x}||} \leq \frac{||A^{-1}||\ ||\vec{r}||}{\frac{||\vec{b}||}{||A||}}.$$

Então,

$$\frac{||\vec{r}||}{||A||\ ||A^{-1}||\ ||\vec{b}||} \leq \frac{||\vec{e}||}{||\vec{x}||} \leq \frac{||A||\ ||A^{-1}||\ ||\vec{r}||}{||\vec{b}||} \tag{4.11}$$

ou

$$\frac{1}{||A||\ ||A^{-1}||}\frac{||\vec{r}||}{||\vec{b}||} \leq \frac{||\vec{e}||}{||\vec{x}||} \leq ||A||\ ||A^{-1}||\frac{||\vec{r}||}{||\vec{b}||}. \tag{4.12}$$

A expressão (4.12) define um intervalo para o erro relativo da solução do primeiro sistema linear por causa do uso da solução do segundo sistema linear como aproximação da solução do primeiro. Observe que o maior valor para esse erro relativo é igual a $(||A||\ ||A^{-1}||)\left(\frac{||\vec{r}||}{||\vec{b}||}\right)$. Que indica que o maior erro relativo na solução é proporcional ao número $\frac{||\vec{r}||}{||\vec{b}||}$, onde o fator de proporcionalidade é $||A||\ ||A^{-1}||$. Se esse fator for grande, o erro relativo pode ser grande também.

Se $||A||\ ||A^{-1}|| \approx 1$; então um erro relativo $\frac{||\vec{r}||}{||\vec{b}||}$ pequeno implica um erro relativo para $\vec{x}$ pequeno também. Se $||A||\ ||A^{-1}||$ for grande já não temos um controle do que possa acontecer como o erro relativo para $\vec{x}$.

Definição 4.1.1. *Um problema (com relação a um determinado conjunto de dados) é chamado de problema mal condicionado se uma pequena perturbação relativa nos dados pode causar um grande erro relativo na solução calculada, independentemente do método de solução. Se é chamado de bem condicionado; isto é, um problema é bem condicionado se todas as pequenas perturbações nos dados produzem apenas pequenos erros relativos na solução.*

Definição 4.1.2. *O número de condição de uma matriz quadrada A de ordem n não singular é denotada por $cond(A)$ e definida por:*

$$cond(A) = ||A||\,||A^{-1}||, \tag{4.13}$$

onde $||\cdot||$ é uma norma matricial.

Proposição 4.1.1. *Temos:*

(a) $cond_p(I) = 1$ *para qualquer norma p;*

(b) $cond_p(A) \geq 1$ *para qualquer norma p;*

(c) $cond(kA) = cond(A)$*, onde k é um escalar diferente de zero, para qualquer norma dada;*

(d) $cond_2(A) = 1$ *se, e somente se $A = kB$, onde $k \neq 0$ e B é uma matriz ortogonal;*

(e) $cond_2(A^TA) = (cond_2(A))^2$*;*

(f) $cond_2(A) = cond_2(A^T)$*;*

(g) $cond_1(A) = cond_\infty(A^T)$*;*

(h) $cond_2(A) = \dfrac{\sqrt{\max\{|\lambda|;\ \lambda \text{ é autovalor de } A^TA\}}}{\sqrt{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A^TA\}}}$*;*

(i) Se A é simétrica, então $cond_2(A) = \dfrac{\max\{|\lambda|;\ \lambda \text{ é autovalor de } A\}}{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A\}}$

Demonstração. Vejamos

(a) Temos que:

$$\begin{aligned} cond_p(I) &= ||I||_p \cdot ||I^{-1}||_p \\ &= ||I||_p \cdot ||I||_p \\ &= 1 \cdot 1 \\ &= 1. \end{aligned}$$

Portanto, $cond_p(I) = 1$ para qualquer norma p.

(b) Temos que:

$$\begin{aligned} cond_p(A) &= ||A||_p \cdot ||A^{-1}||_p \\ &\geq ||A \cdot A^{-1}||_p \\ &= ||I||_p \\ &= 1. \end{aligned}$$

Portanto, $cond_p(A) \geq 1$ para qualquer norma p.

(c) Por definição, $cond(A) = ||A||\, ||A^{-1}||$, então

$$
\begin{aligned}
cond(kA) &= ||kA||\, ||(kA)^{-1}|| \\
&= ||kA||\, ||k^{-1}A^{-1}|| \\
&= (|k|\, ||A||)\, (|k^{-1}|\, ||A^{-1}||) \\
&= \underbrace{|k|\, |k^{-1}|}_{=1}\, ||A||\, ||A^{-1}|| \\
&= cond(A).
\end{aligned}
$$

(d) Seja $A = kB$, com $k \neq 0$, onde B é ortogonal $(B^T B = I \Rightarrow B^T = B^{-1})$. Usando propriedades de norma, temos:

$$
\begin{aligned}
||A||_2 &= ||kB||_2 = |k|\, ||B||_2, \\
||A^{-1}||_2 &= ||(kB)^{-1}||_2 = ||k^{-1}B^{-1}||_2 = |k^{-1}|\, ||B^{-1}||_2.
\end{aligned}
$$

Daí,

$$
\begin{aligned}
cond_2(A) &= ||A||_2 \cdot ||A^{-1}||_2 \\
&= |k|\, ||B||_2 \cdot |k^{-1}|\, ||B^{-1}||_2 \\
&= \underbrace{|k|\, |k^{-1}|}_{=1} \cdot ||B||_2\, ||B^{-1}||_2.
\end{aligned}
$$

Pela parte *(b.iii)* e *(b.iv)* da Proposição 3.3.1 segue que,

$$
\begin{aligned}
cond_2(A) &= ||B||_2\, ||B^{-1}||_2 \\
&= 1 \cdot 1 \\
&= 1.
\end{aligned}
$$

(e) Por definição,

$$cond_2(A^T A) = ||A^T A||_2 \cdot ||(A^T A)^{-1}||_2.$$

Pela parte *(e)* da Proposição 3.3.1, $||A^T A||_2 = ||A||_2^2$ e pela parte *(c)* da mesma proposição, $||A||_2 = ||A^T||_2$.

$$||(A^T A)^{-1}||_2 = ||A^{-1}(A^T)^{-1}||_2 = ||((A^{-1})^T)^T (A^{-1})^T||_2 \overset{(e)}{=} ||(A^{-1})^T||_2^2 \overset{(c)}{=} ||A^{-1}||_2^2.$$

Então,

$$\begin{aligned} cond_2(A^TA) &= ||A^TA||_2 \cdot ||(A^TA)^{-1}||_2 \\ &\overset{(e)}{=} ||A||_2^2 \cdot ||A^{-1}||_2^2 \\ &= (||A||_2 \cdot ||A^{-1}||_2)^2 \\ &= (cond(A))^2. \end{aligned}$$

(f) Sabe-se que, pela parte *(c)* da Proposição 3.3.1, $||A||_2 = ||A^T||_2$, daí

$$||A^{-1}||_2 = ||(A^{-1})^T||_2 = ||(A^T)^{-1}||_2,$$

então,

$$\begin{aligned} cond_2(A) &= ||A||_2 \cdot ||A^{-1}||_2 \\ &= ||A^T||_2 \cdot ||(A^T)^{-1}||_2 \\ &= cond_2(A^T). \end{aligned}$$

(g) Novamente, por definição,

$$\begin{aligned} ||A||_1 &= \max\left\{\sum_{s=1}^{n} |a_{is}|;\ s = 1, \ldots, n\right\} = ||A^T||_\infty, \\ ||A^T||_\infty &= \max\left\{\sum_{r=1}^{n} |a_{rj}|;\ r = 1, \ldots, n\right\} = ||A||_1. \end{aligned}$$

Assim, com efeito,

$$\begin{aligned} cond_1(A) &= ||A||_1 \cdot ||A^{-1}||_1 \\ &= ||A^T||_\infty \cdot ||(A^{-1})^T||_\infty \\ &= ||A^T||_\infty \cdot ||(A^T)^{-1}||_\infty \\ &= cond_\infty(A^T). \end{aligned}$$

(h) Observe que,

$$cond_2(A) = ||A||_2 \cdot ||A^{-1}||_2.$$

Pela parte *(d)* da Proposição 3.3.1 temos,

$$\begin{aligned} cond_2(A) &= \sqrt{\rho(A^TA)}\sqrt{\rho((A^{-1})^T(A^{-1}))} \\ &= \frac{\sqrt{\max\{|\lambda|;\ \lambda \text{ é autovalor de } A^TA\}}}{\sqrt{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A^TA\}}}. \end{aligned}$$

(i) Se A é simétrica, então $A^TA = A^2$. Assim,

$$\begin{aligned} cond_2(A) &= \frac{\sqrt{\max\{|\gamma|;\ \gamma \text{ é autovalor de } A^TA\}}}{\sqrt{\min\{|\gamma|;\ \gamma \text{ é autovalor de } A^TA\}}} \\ &= \frac{\max\{\sqrt{\gamma};\ \gamma \text{ é autovalor de } A^2\}}{\min\{\sqrt{\gamma};\ \gamma \text{ é autovalor de } A^2\}} \\ &= \frac{\max\{|\lambda|;\ \lambda \text{ é autovalor de } A\}}{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A\}} \end{aligned}$$

∎

4.2 Teoremas de Condicionamento de Matrizes

Teorema 4.2.1. ***(Teorema da perturbação à direita)***. *Se $\delta\vec{b}$ e $\delta\vec{x}$ são, respectivamente, as perturbações de $\vec{b}$ e $\vec{x}$ no sistema linear $A\vec{x} = \vec{b}$, A é não singular e $\vec{b} \neq \vec{0}$, então*

$$\frac{||\delta\vec{b}||}{||A||\ ||A^{-1}||\ ||\vec{b}||} \leq \frac{||\delta\vec{x}||}{||\vec{x}||} \leq ||A||\ ||A^{-1}||\frac{||\delta\vec{b}||}{||\vec{b}||}$$

ou

$$\frac{1}{cond(A)}\frac{||\delta\vec{b}||}{||\vec{b}||} \leq \frac{||\delta\vec{x}||}{||\vec{x}||} \leq cond(A)\frac{||\delta\vec{b}||}{||\vec{b}||}.$$

Demonstração. Como

$$A\vec{x} = \vec{b} \tag{4.14}$$

e

$$A(\vec{x} + \delta\vec{x}) = \vec{b} + \delta\vec{b},$$

temos

$$A\vec{x} + A\delta\vec{x} = \vec{b} + \delta\vec{b}.$$

De (4.14) tiramos,

$$A\delta\vec{x} = \delta\vec{b}.$$

Multiplicando ambos os lados à esquerda por A^{-1} temos,

$$A^{-1}A\delta\vec{x} = A^{-1}\delta\vec{b}.$$

Ou seja,

$$\delta\vec{x} = A^{-1}\delta\vec{b}.$$

Tomando uma norma matriz-vetor subordinada, obtemos

$$||\delta\vec{x}|| = ||A^{-1}\delta\vec{b}|| \leq ||A^{-1}||\,||\delta\vec{b}||. \tag{4.15}$$

Novamente, considerando a mesma norma em ambos os lados de $A\vec{x} = \vec{b}$, obtemos

$$||\vec{b}|| = ||A\vec{x}|| \leq ||A||\,||\vec{x}||. \tag{4.16}$$

Combinando (4.15) e (4.16), temos

$$||\delta\vec{x}|| \cdot ||\vec{b}|| \leq ||A^{-1}||\,||\delta\vec{b}|| \cdot ||A||\,||\vec{x}||.$$

Dividindo ambos os lados por $||\vec{x}||\,||\vec{b}||$,

$$\frac{||\delta\vec{x}|| \cdot ||\vec{b}||}{||\vec{x}||\,||\vec{b}||} \leq \frac{||A^{-1}||\,||\delta\vec{b}|| \cdot ||A||\,||\vec{x}||}{||\vec{x}||\,||\vec{b}||}.$$

Isto é,

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq ||A||\,||A^{-1}||\frac{||\delta\vec{b}||}{||\vec{b}||}.$$

Por outro lado, $A\delta\vec{x} = \delta\vec{b}$ dá

$$\begin{aligned} ||\delta\vec{b}|| &= ||A\delta\vec{x}|| \\ ||\delta\vec{b}|| &\leq ||A||\,||\delta\vec{x}|| \\ ||\delta\vec{x}|| &\geq \frac{||\delta\vec{b}||}{||A||}. \end{aligned} \tag{4.17}$$

Além disso, de $A\vec{x} = \vec{b}$, temos

$$\begin{aligned} A^{-1}A\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b} \\ ||\vec{x}|| &= ||A^{-1}\vec{b}|| \\ ||\vec{x}|| &\leq ||A^{-1}||\,||\vec{b}|| \\ \frac{1}{||\vec{x}||} &\geq \frac{1}{||A^{-1}||\,||\vec{b}||}. \end{aligned} \tag{4.18}$$

Combinando (4.17) e (4.18), temos

$$\begin{aligned} ||\delta\vec{x}||\,\frac{1}{||\vec{x}||} &\geq \frac{||\delta\vec{b}||}{||A||}\,\frac{1}{||A^{-1}||\,||\vec{b}||} \\ \frac{||\delta\vec{x}||}{||\vec{x}||} &\geq \frac{||\delta\vec{b}||}{||A||\,||A^{-1}||\,||\vec{b}||}. \end{aligned}$$

∎

Teorema 4.2.2. ***(Teorema da perturbação à esquerda)****. Suponha que A seja não singular e $\vec{b} \neq \vec{0}$. Suponha que ΔA e $\delta\vec{x}$ sejam, respectivamente, as perturbações de A e $\vec{x}$ no sistema linear $A\vec{x} = \vec{b}$. Além disso, suponha que ΔA seja tal que $||\Delta A|| < \frac{1}{||A^{-1}||}$. Então*

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq cond(A)\frac{||\Delta A||}{||A||} \Bigg/ \left(1 - cond(A)\frac{||\Delta A||}{||A||}\right).$$

Demonstração. Temos que,

$$(A + \Delta A)(\vec{x} + \delta\vec{x}) = \vec{b}$$

ou

$$(A + \Delta A)\vec{x} + (A + \Delta A)\delta\vec{x} = \vec{b}. \tag{4.19}$$

Como $A\vec{x} = \vec{b}$, temos de (4.19)

$$\begin{aligned} A\vec{x} + \Delta A\vec{x} + A\delta\vec{x} + \Delta A\delta\vec{x} &= \vec{b} \\ \Delta A\vec{x} + A\delta\vec{x} + \Delta A\delta\vec{x} &= \vec{0} \\ A\delta\vec{x} &= -\Delta A\vec{x} - \Delta A\delta\vec{x} \\ A\delta\vec{x} &= -\Delta A(\vec{x} + \delta\vec{x}). \end{aligned}$$

Multiplicando ambos os lados por A^{-1} à esquerda temos,

$$A^{-1} \cdot A\delta\vec{x} = -A^{-1} \cdot \Delta A(\vec{x} + \delta\vec{x}).$$

Isto é,

$$\delta\vec{x} = -A^{-1}\Delta A(\vec{x} + \delta\vec{x}).$$

Passando a norma de ambos os lados, temos

$$\begin{aligned}
||\delta\vec{x}|| &\leq ||A^{-1}||\,||\Delta A|| \cdot (||\vec{x}|| + ||\delta\vec{x}||) \\
&= \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}(||\vec{x}|| + ||\delta\vec{x}||) \\
&= \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\vec{x}|| + \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\delta\vec{x}||,
\end{aligned}$$

ou seja,

$$\begin{gathered}
||\delta\vec{x}|| \leq \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\vec{x}|| + \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\delta\vec{x}|| \\
||\delta\vec{x}|| - \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\delta\vec{x}|| \leq \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}||\vec{x}|| \\
\left(1 - \frac{||A^{-1}||\,||A||\,||\Delta A||}{||A||}\right)||\delta\vec{x}|| \leq \frac{||A||\,||A^{-1}||\,||\Delta A||}{||A||}||\vec{x}||.
\end{gathered}$$

Por hipótese, $||A^{-1}||\,||\Delta A|| < 1$, então a expressão entre parênteses no lado esquerdo é positiva. Podemos, portanto, dividir ambos os lados da desigualdade por esse número sem alterar a desigualdade. Depois disso, se também dividirmos por $||\vec{x}||$, obtemos

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq \frac{||A||\,||A^{-1}||\dfrac{||\Delta A||}{||A||}}{\left(1 - ||A||\,||A^{-1}||\dfrac{||\Delta A||}{||A||}\right)} = cond(A)\frac{||\Delta A||}{||A||} \bigg/ \left(1 - cond(A)\frac{||\Delta A||}{||A||}\right),$$

o que prova o teorema. ∎

Teorema 4.2.3. ***(Teorema de perturbação geral).*** *Suponha que A seja não singular, $\vec{b} \neq \vec{0}$ e $||\Delta A|| < \frac{1}{||A^{-1}||}$. Então*

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq \left(\frac{cond(A)}{1 - cond(A) \cdot \dfrac{||\Delta A||}{||A||}}\right)\left(\frac{||\Delta A||}{||A||} + \frac{||\delta\vec{b}||}{||\vec{b}||}\right).$$

Demonstração. Subtraindo $A\vec{x}$ de

$$(A + \Delta A)(\vec{x} + \delta\vec{x}) = \vec{b} + \delta\vec{b},$$

temos

$$\begin{aligned} (A + \Delta A)(\vec{x} + \delta\vec{x}) - A\vec{x} &= \vec{b} + \delta\vec{b} - A\vec{x} \\ (A + \Delta A)(\vec{x} + \delta\vec{x}) - A\vec{x} &= \delta\vec{b}, \end{aligned}$$

pois $A\vec{x} = \vec{b}$. Assim,

$$\begin{aligned} (A + \Delta A)\vec{x} + (A + \Delta A)\delta\vec{x} - A\vec{x} &= \delta\vec{b} \\ A\vec{x} + \Delta A\vec{x} + (A + \Delta A)\delta\vec{x} - A\vec{x} &= \delta\vec{b} \\ \Delta A\vec{x} + (A + \Delta A)\delta\vec{x} &= \delta\vec{b}. \end{aligned}$$

Subtraindo $\Delta A\vec{x}$ em ambos os lados,

$$\begin{aligned} \Delta A\vec{x} + (A + \Delta A)\delta\vec{x} - \Delta A\vec{x} &= \delta\vec{b} - \Delta A\vec{x} \\ (A + \Delta A)\delta\vec{x} &= \delta\vec{b} - \Delta A\vec{x}. \end{aligned} \tag{4.20}$$

Note que

$$A(I - A^{-1}(-\Delta A)) = AI - AA^{-1}(-\Delta A) = A + \Delta A$$

então, podemos reescrever (4.20) como,

$$A(I - A^{-1}(-\Delta A))\delta\vec{x} = \delta\vec{b} - \Delta A\vec{x}. \tag{4.21}$$

Seja $A^{-1}(-\Delta A) = F$. Então,

$$||F|| = ||A^{-1}(-\Delta A)|| \leq ||A^{-1}||\,||\Delta A|| < 1 \quad \text{(por hipótese).}$$

Como $||F|| < 1$, então $I - F$ é invertível (veja Teorema 3.3.1), e então de (4.21) temos

$$A(I - F)\delta\vec{x} = \delta\vec{b} - \Delta A\vec{x}.$$

Multiplicando ambos os lados à esquerda por $(A(I - F))^{-1}$ temos,

$$\begin{aligned} (A(I - F))^{-1} \cdot A(I - F)\delta\vec{x} &= (A(I - F))^{-1} \cdot (\delta\vec{b} - \Delta A\vec{x}) \\ \delta\vec{x} &= (I - F)^{-1}A^{-1}(\delta\vec{b} - \Delta A\vec{x}). \end{aligned}$$

Novamente, usando o Teorema 3.3.1, podemos escrever

$$||(I-F)^{-1}|| \leq \frac{1}{1-||F||}.$$

Assim,

$$\begin{aligned} ||\delta\vec{x}|| &= ||(I-F)^{-1}A^{-1}(\delta\vec{b}-\Delta A\vec{x})|| \\ &\leq ||(I-F)^{-1}||\;||A^{-1}||\;(||\delta\vec{b}||+||\Delta A||\;||\vec{x}||) \\ &\leq \frac{1}{1-||F||}\;||A^{-1}||\;(||\delta\vec{b}||+||\Delta A||\;||\vec{x}||), \end{aligned}$$

ou seja,

$$||\delta\vec{x}|| \leq \frac{||A^{-1}||}{1-||F||}(||\delta\vec{b}||+||\Delta A||\;||\vec{x}||).$$

Dividindo ambos os lados por $||\vec{x}||$, temos

$$\begin{aligned} \frac{||\delta\vec{x}||}{||\vec{x}||} &\leq \frac{||A^{-1}||}{1-||F||}\left(\frac{||\delta\vec{b}||+||\Delta A||\;||\vec{x}||}{||\vec{x}||}\right) \\ &= \frac{||A^{-1}||}{1-||F||}\left(\frac{||\delta\vec{b}||}{||\vec{x}||}+||\Delta A||\right). \end{aligned}$$

Note que,

$$||\vec{b}|| \leq ||A||\;||\vec{x}|| \quad \Longrightarrow \quad \frac{1}{||\vec{x}||} \leq \frac{||A||}{||\vec{b}||}.$$

Isto é,

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq \frac{||A^{-1}||}{1-||F||}\left(\frac{||\delta\vec{b}||\;||A||}{||\vec{b}||}+||\Delta A||\right).$$

Colocando $||A||$ em evidência, temos

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \leq \frac{||A^{-1}||\;||A||}{1-||F||}\left(\frac{||\delta\vec{b}||}{||\vec{b}||}+\frac{||\Delta A||}{||A||}\right). \tag{4.22}$$

Observe que

$$||F|| = ||A^{-1}(-\Delta A)|| \leq ||A^{-1}||\;||\Delta A|| \implies -||F|| \geq -||A^{-1}||\;||\Delta A||.$$

Somando 1 em ambos os lados, temos

$$1-||F|| \geq 1-||A^{-1}||\;||\Delta A||.$$

Como por hipótese, $||A^{-1}||\;||\Delta A|| < 1$, segue que $1-||F|| > 0$ e, consequentemente, $||F|| < 1$.

De novo

$$||F|| = ||A^{-1}(-\Delta A)|| \le ||A^{-1}||\;||\Delta A|| = \frac{||A^{-1}||\;||A||}{||A||}||\Delta A||. \tag{4.23}$$

Como $||F|| < 1$, podemos escrever de (4.22) e (4.23)

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \le \left(\frac{||A^{-1}||\;||A||}{1 - \dfrac{||A^{-1}||\;||A||}{||A||}||\Delta A||}\right)\left(\frac{||\delta\vec{b}||}{||\vec{b}||} + \frac{||\Delta A||}{||A||}\right).$$

Isto é,

$$\frac{||\delta\vec{x}||}{||\vec{x}||} \le \left(\frac{cond(A)}{1 - \dfrac{cond(A)}{||A||}||\Delta A||}\right)\left(\frac{||\delta\vec{b}||}{||\vec{b}||} + \frac{||\Delta A||}{||A||}\right).$$

∎

A seguir, alguns exemplos que matrizes bem condicionadas.

Exemplo 4.2. *A matriz identidade é um exemplo de matriz bem condicionada, pois pelo item (a) da Proposição 4.1.2,* $cond_p(I) = 1$ *para qualquer norma* p*.*

Exemplo 4.3. *Uma matriz ortogonal* A *é um exemplo de matriz bem condicionada, pois pelo item (d) da Proposição 4.1.2,* $cond_2(A) = 1$ *para qualquer norma.*

Exemplo 4.4. *Seja* $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$*, matriz simétrica. Pode-se verificar que* $A^{-1} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$*. De fato,*

$$\begin{aligned} AA^{-1} &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \\ A^{-1}A &= \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I. \end{aligned}$$

Pelo item (i) da Proposição 4.1.2, se A *é simétrica, então*

$$cond_2(A) = \frac{\max\{|\lambda|;\ \lambda \text{ é autovalor de } A\}}{\min\{|\lambda|;\ \lambda \text{ é autovalor de } A\}}.$$

Para o cálculo dos autovalores, usa-se $\det(A - \lambda I) = 0$*, isto é,*

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = 0.$$

Assim,

$$\begin{aligned}(1-\lambda)(3-\lambda)-2\cdot 2 &= 0\\ 3-\lambda-3\lambda+\lambda^2-4 &= 0\\ \lambda^2-4\lambda-1 &= 0.\end{aligned}$$

Disto,

$$\begin{aligned}\Delta &= b^2-4ac\\ \Delta &= (-4)^2-4\cdot 1\cdot(-1)\\ \Delta &= 16+4\\ \Delta &= 20.\end{aligned}$$

Daí,

$$\begin{aligned}\lambda &= \frac{-b\pm\sqrt{\Delta}}{2a}\\ \lambda &= \frac{-(-4)\pm\sqrt{20}}{2\cdot 1}\\ \lambda &= \frac{4\pm\sqrt{20}}{2}\\ \lambda_1 &\approx 4,24\\ \lambda_2 &\approx -0,24.\end{aligned}$$

Portanto,

$$\begin{aligned}cond_2(A) &= \frac{\max\{|\lambda|;\ \lambda \textit{ é autovalor de } A\}}{\min\{|\lambda|;\ \lambda \textit{ é autovalor de } A\}}\\ &\approx \frac{|4,24|}{|-0,24|}\\ &\approx \frac{4,24}{0,24}\\ &\approx 17,67.\end{aligned}$$

Agora, seguem alguns exemplos de matrizes mal condicionadas.

Exemplo 4.5. *Considere o sistema de equações lineares:*

$$\begin{cases} 7x_1+10x_2=1\\ 5x_1+7x_2=0,7\end{cases}$$

cuja solução é $x_1 = 0$ *e* $x_2 = 0,1$. *Então,*

$$A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}, \quad onde \quad A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix},$$

$cond_1(A) = cond_\infty(A) = 289$ *e* $cond_2(A) \approx 223$.

Com esses números de condicionamento, conclui-se que o sistema pode ser mal condicionado, isto é, o sistema pode ser sensível a pequenas perturbações introduzidas no vetor constante $\vec{b}$. *De fato, considere o sistema perturbado*

$$\begin{cases} 7\overline{x}_1 + 10\overline{x}_2 = 1,01 \\ 5\overline{x}_1 + 7\overline{x}_2 = 0,69 \end{cases}$$

cuja solução é $\overline{x}_1 = -0,17$ *e* $\overline{x}_2 = 0,22$.

Vejamos os cálculos:

(1) Cálculo de A^{-1}:

$$A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \Longrightarrow A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}.$$

De fato, seja

$$\left[\begin{array}{cc|cc} 7 & 10 & 1 & 0 \\ 5 & 7 & 0 & 1 \end{array}\right].$$

Aplicando as seguintes operações elementares $L_1 \to \frac{1}{7}L_1$ e $L_2 \to \frac{1}{7}L_2$, tem-se:

$$\left[\begin{array}{cc|cc} 1 & \frac{10}{7} & \frac{1}{7} & 0 \\ \frac{5}{7} & 1 & 0 & \frac{1}{7} \end{array}\right].$$

Fazendo $L_2 \to \left(-\frac{5}{7}\right) L_1 + L_2$, obtém-se:

$$\left[\begin{array}{cc|cc} 1 & \frac{10}{7} & \frac{1}{7} & 0 \\ 0 & -\frac{1}{49} & -\frac{5}{49} & \frac{1}{7} \end{array}\right].$$

Aplicando-se $L_2 \to (-49)L_2$:

$$\left[\begin{array}{cc|cc} 1 & \frac{10}{7} & \frac{1}{7} & 0 \\ 0 & 1 & 5 & -7 \end{array}\right].$$

Por fim, fazendo $L_1 \to \left(-\dfrac{10}{7}\right) L_2 + L_1$:

$$\left[\begin{array}{cc|cc} 1 & 0 & -7 & 10 \\ 0 & 1 & 5 & -7 \end{array}\right].$$

Logo:

$$A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}.$$

(2) Cálculo da perturbação relativa em $\vec{b} = \begin{bmatrix} 1 \\ 0,7 \end{bmatrix}$.

Seja $\vec{\bar{b}} = \begin{bmatrix} 1,01 \\ 0,69 \end{bmatrix}$, então $\delta\vec{b} = \vec{\bar{b}} - \vec{b} = \begin{bmatrix} 1,01 - 1 \\ 0,69 - 0,7 \end{bmatrix} = \begin{bmatrix} 0,01 \\ -0,01 \end{bmatrix}$. Assim,

$$\begin{aligned} ||\delta\vec{b}||_2 &= \sqrt{(0,01)^2 + (-0,01)^2} \\ &= \sqrt{0,0001 + 0,0001} \\ &= \sqrt{0,0002} \\ &\simeq 0,014142135. \end{aligned}$$

Além disso,

$$\begin{aligned} ||\vec{b}||_2 &= \sqrt{1^2 + (0,7)^2} \\ &= \sqrt{1 + 0,49} \\ &= \sqrt{1,49} \\ &\approx 1,220655562. \end{aligned}$$

Logo, a perturbação relativa em $\vec{b}$ é:

$$\frac{||\delta\vec{b}||_2}{||\vec{b}||_2} = \frac{0,014142135}{1,220655562} \approx 0,011585688.$$

(3) Cálculo da perturbação relativa em $\vec{x} = \begin{bmatrix} 0 \\ 0,1 \end{bmatrix}$.

Temos $\vec{\tilde{x}} = \begin{bmatrix} -0,17 \\ 0,22 \end{bmatrix}$. Logo, $\delta\vec{x} = \vec{\tilde{x}} - \vec{x} = \begin{bmatrix} -0,17 - 0 \\ 0,22 - 0,1 \end{bmatrix} = \begin{bmatrix} -0,17 \\ 0,12 \end{bmatrix}$. Daí:

$$\begin{aligned} ||\delta\vec{x}||_2 &= \sqrt{(-0,17)^2 + (0,12)^2} \\ &= \sqrt{0,0289 + 0,0144} \\ &= \sqrt{0,0433} \\ &\approx 0,20808652. \end{aligned}$$

A Norma 2 de $\vec{x}$ é:

$$\begin{aligned} ||\vec{x}||_2 &= \sqrt{0^2 + (0,1)^2} \\ &= \sqrt{0 + 0,01} \\ &= \sqrt{0,01} \\ &= 0,1. \end{aligned}$$

Então, a pertubação relativa em $\vec{x}$ é:

$$\frac{||\delta\vec{x}||_2}{||\vec{x}||_2} = \frac{0,20808652}{0,1} \approx 2,080865205.$$

(4) Cálculo de $cond_2(A) = ||A||_2 \cdot ||A^{-1}||_2$.

Pela expressão (3.7) da Proposição 3.2.1, tem-se que:

$$||A||_2 = \max\{\sqrt{\lambda}; \lambda \text{ é o maior autovalor de } A^T A\}.$$

Mas $A^T A$ é

$$B = A^T A = \begin{bmatrix} 7 & 5 \\ 10 & 7 \end{bmatrix} \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} = \begin{bmatrix} 49+25 & 70+35 \\ 70+35 & 100+49 \end{bmatrix} = \begin{bmatrix} 74 & 105 \\ 105 & 149 \end{bmatrix}.$$

Para o cálculo dos autovalores, usa-se $\det(B - \lambda I) = 0$, isto é,

$$\begin{vmatrix} 74-\lambda & 105 \\ 105 & 149-\lambda \end{vmatrix} = 0.$$

Assim,

$$\begin{aligned}
(74-\lambda)(149-\lambda)-11025 &= 0\\
11026-74\lambda-149\lambda+\lambda^2-11025 &= 0\\
\lambda^2-223\lambda+1 &= 0.
\end{aligned}$$

Disto,

$$\begin{aligned}
\Delta &= b^2-4ac\\
\Delta &= (-223)^2-4\cdot 1\cdot 1\\
\Delta &= 49729-4\\
\Delta &= 49725.
\end{aligned}$$

Daí,

$$\begin{aligned}
\lambda &= \frac{-b\pm\sqrt{\Delta}}{2a}\\
\lambda &= \frac{-(-223)\pm\sqrt{49725}}{2\cdot 1}\\
\lambda &= \frac{223\pm\sqrt{49725}}{2}\\
\lambda_1 &\approx 222,9955156\\
\lambda_2 &\approx 0,004484395.
\end{aligned}$$

Portanto,

$$||A||_2 \approx \sqrt{222,9955156} \approx 14,93303437.$$

Novamente, pela expressão (3.7) da Proposição 3.2.1, tem-se:

$$||A^{-1}||_2 = \max\{\sqrt{\lambda};\lambda \text{ é o maior autovalor de } (A^{-1})^T A^{-1}\}.$$

Denominando-se $C = (A^{-1})^T A^{-1}$, o cálculo dos autovalores de C é feito de forma semelhante. Assim,

$$C=(A^{-1})^T A^{-1} = \begin{bmatrix} -7 & 5\\ 10 & -7\end{bmatrix}\begin{bmatrix} -7 & 10\\ 5 & -7\end{bmatrix} = \begin{bmatrix} 49+25 & -70-35\\ -70-35 & 100+49\end{bmatrix} = \begin{bmatrix} 74 & -105\\ -105 & 149\end{bmatrix}.$$

De $\det(C - \lambda I) = 0$, tem-se:

$$\begin{vmatrix} 74-\lambda & -105 \\ -105 & 149-\lambda \end{vmatrix} = 0.$$

Assim,

$$\begin{aligned} (74-\lambda)(149-\lambda) - 11025 &= 0 \\ 11026 - 74\lambda - 149\lambda + \lambda^2 - 11025 &= 0 \\ \lambda^2 - 223\lambda + 1 &= 0. \end{aligned}$$

Como já vimos acima,

$$\begin{aligned} \lambda_1 &\approx 222,9955156 \\ \lambda_2 &\approx 0,004484395. \end{aligned}$$

Portanto,

$$||A^{-1}||_2 \approx \sqrt{222,9955156} \approx 14,93303437.$$

Assim,

$$cond_2(A) = ||A||_2 \cdot ||A^{-1}||_2 \approx (14,93303437)^2 \approx 223.$$

(5) Cálculo de $cond_1(A)$ e $cond_\infty(A)$.

Considere

$$A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \quad \text{e} \quad A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}.$$

Pela Proposição 3.2.1 parte (i), segue que:

$$||A||_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|.$$

Assim,

$$\begin{aligned} ||A||_1 &= \max\{(|7| + |5|), (|10| + |7|)\} = \max\{(7+5), (10+7)\} = \max\{12, 17\} = 17, \\ ||A^{-1}||_1 &= \max\{(|-7| + |5|), (|10| + |-7|)\} = \max\{(7+5), (10+7)\} = \max\{12, 17\} = 17. \end{aligned}$$

Logo,

$$cond_1(A) = ||A||_1 \cdot ||A^{-1}||_1 = 17 \cdot 17 = 289.$$

E, novamente, pela Proposição 3.2.1 parte (ii), tem-se:

$$||A||_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Então,

$$\begin{aligned} ||A||_\infty &= \max\{(|7| + |10|), (|5| + |7|)\} = \max\{(7+10), (5+7)\} = \max\{17, 12\} = 17, \\ ||A^{-1}||_\infty &= \max\{(|-7| + |10|), (|5| + |-7|)\} = \max\{(7+10), (5+7)\} = \max\{17, 12\} = 17. \end{aligned}$$

Portanto,

$$cond_\infty(A) = ||A||_\infty \cdot ||A^{-1}||_\infty = 17 \cdot 17 = 289.$$

A variação na solução pode ser considerada grande quando comparada com a variação do vetor $\vec{b}$. Logo, esse sistema é mal condicionado.

Exemplo 4.6. ***(Aplicação do Teorema 4.2.1)*** *Seja* $A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix}$. *Pode-se verificar que* $A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}$. *De fato,*

$$\begin{aligned} AA^{-1} &= \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \\ A^{-1}A &= \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I. \end{aligned}$$

Logo, $||A||_\infty = 17$ *e* $||A^{-1}||_\infty = 17$. *Então,* $cond_\infty(A) = ||A||_\infty ||A^{-1}||_\infty = 17 \cdot 17 = 289$. *Considere o sistema linear:*

$$A\vec{x} = \vec{b} = \begin{bmatrix} 17 \\ 12 \end{bmatrix}, \tag{4.24}$$

cuja solução exata é $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. *Note que,*

$$\begin{aligned} ||\vec{b}||_\infty &= \max\{|17|, |12|\} = \max\{17, 12\} = 17, \\ ||\vec{x}||_\infty &= \max\{|1|, |1|\} = \max\{1, 1\} = 1. \end{aligned}$$

Além disso, considere o sistema linear

$$A\vec{\hat{x}} = \vec{\hat{b}} = \begin{bmatrix} 17,01 \\ 11,98 \end{bmatrix}, \tag{4.25}$$

cuja solução espera-se que esteja perto de $\vec{x}$, pois o vetor do lado direito de (4.25) é o lado direito de (4.24) perturbado em

$$\delta\vec{b} = \widehat{\vec{b}} - \vec{b} = \begin{bmatrix} 17,01 \\ 11,98 \end{bmatrix} - \begin{bmatrix} 17 \\ 12 \end{bmatrix} = \begin{bmatrix} 0,01 \\ -0,02 \end{bmatrix}.$$

Assim,

$$||\delta\vec{b}||_\infty = \max\{|0,01|, |-0,02|\} = \max\{0,01, 0,02\} = 0,02.$$

Resolvendo-se (4.25) de forma exata, pois a inversa é conhecida de forma exata, tem-se que

$$\widehat{\vec{x}} = \begin{bmatrix} 0,73 \\ 1,19 \end{bmatrix}.$$

De fato, como $A\widehat{x} = \widehat{b}$, então

$$\begin{aligned} \widehat{x} &= A^{-1}\widehat{b} \\ &= \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 17,01 \\ 11,98 \end{bmatrix} \\ &= \begin{bmatrix} -119,07 + 119,8 \\ 85,05 - 83,86 \end{bmatrix} \\ &= \begin{bmatrix} 0,73 \\ 1,19 \end{bmatrix}. \end{aligned}$$

Logo,

$$\delta\vec{x} = \widehat{\vec{x}} - \vec{x} = \begin{bmatrix} 0,73 \\ 1,19 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0,27 \\ 0,19 \end{bmatrix}.$$

Daí,

$$||\delta\vec{x}||_\infty = \max\{|-0,27|, |0,19|\} = \max\{0,27, 0,19\} = 0,27.$$

Então, o erro relativo em $\vec{x}$ é

$$\frac{||\delta\vec{x}||_\infty}{||\vec{x}||_\infty} = \frac{0,27}{1} = 0,27.$$

O erro relativo em $\vec{b}$ é

$$\frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty} = \frac{0,02}{17} = 0,00176.$$

E, além do mais,

$$||A||_\infty ||A^{-1}||_\infty \frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty} = 0,339864.$$

Pode-se observar que:

$$\frac{||\delta\vec{x}||_\infty}{||\vec{x}||_\infty} < ||A||_\infty ||A^{-1}||_\infty \frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty}. \tag{4.26}$$

Exemplo 4.7. *Considere* $A = \begin{bmatrix} 1 & 0,99 \\ 0,99 & 0,98 \end{bmatrix}$. *Sabe-se que:*

$$A^{-1} = -\frac{1}{10^{-4}} \begin{bmatrix} 0,98 & -0,99 \\ -0,99 & 1 \end{bmatrix},$$

ou ainda,

$$\begin{aligned} A^{-1} &= -10000 \cdot \begin{bmatrix} 0,98 & -0,99 \\ -0,99 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 10000 \cdot 0,98 & 10000 \cdot (-0,99) \\ 10000 \cdot (-0,99) & 10000 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} 9800 & -9900 \\ -9900 & 10000 \end{bmatrix}. \end{aligned}$$

Considere $A\vec{x} = \vec{b} = \begin{bmatrix} 1,99 \\ 1,97 \end{bmatrix}$, *cuja solução exata é*

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Já o sistema linear $A\hat{\vec{x}} = \hat{\vec{b}} = \begin{bmatrix} 1,989903 \\ 1,970106 \end{bmatrix}$, *tem por solução:*

$$\hat{\vec{x}} = \begin{bmatrix} 3 \\ -1,0203 \end{bmatrix}.$$

A perturbação em $\vec{b}$ *é*

$$\delta\vec{b} = \hat{\vec{b}} - \vec{b} = \begin{bmatrix} 1,989903 \\ 1,970106 \end{bmatrix} - \begin{bmatrix} 1,99 \\ 1,97 \end{bmatrix} = \begin{bmatrix} -0,000097 \\ 0,000106 \end{bmatrix}$$

e gerou uma perturbação

$$\delta\vec{x} = \vec{\hat{x}} - \vec{x} = \begin{bmatrix} 3 \\ -1,0203 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2,0203 \end{bmatrix}.$$

Observe que,

$$\begin{aligned} ||A||_\infty &= \max\{|1| + |0,99|, |0,99| + |0,98|\} = \max\{1,99, 1,97\} = 1,99 \qquad e \\ ||A^{-1}||_\infty &= \max\{|9800| + |-9900|, |-9900| + |10000|\} = \max\{19700, 19900\} = 19900. \end{aligned}$$

Note também que,

$$\begin{aligned} ||\vec{b}||_\infty &= \max\{|1,99|, |1,97|\} = \max\{1,99, 1,97\} = 1,99 \\ ||\vec{x}||_\infty &= \max\{|1|, |1|\} = \max\{1, 1\} = 1 \\ ||\delta\vec{b}||_\infty &= \max\{|-0,000097|, |0,000106|\} = \max\{0,000097, 0,000106\} = 0,000106 \\ ||\delta\vec{x}||_\infty &= \max\{|2|, |-2,0203|\} = \max\{2, 2,0203\} = 2,0203. \end{aligned}$$

Então, o erro relativo em $\vec{x}$ é

$$\frac{||\delta\vec{x}||_\infty}{||\vec{x}||_\infty} = \frac{2,0203}{1} = 2,0203.$$

O erro relativo em $\vec{b}$ é

$$\frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty} = \frac{0,000106}{1,99} = 0,000053266.$$

E, além do mais,

$$||A||_\infty ||A^{-1}||_\infty \frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty} = (1,99)(19900)(0,000053266) \approx 2,1094.$$

Daí,

$$\frac{||\delta\vec{x}||_\infty}{||\vec{x}||_\infty} < ||A||_\infty ||A^{-1}||_\infty \frac{||\delta\vec{b}||_\infty}{||\vec{b}||_\infty}. \tag{4.27}$$

Assim, pode-se observar que a relação (4.26) continua sendo satisfeita.

O seguinte exemplo mostra como o número de condição depende da norma escolhida.

Exemplo 4.8. *Considere-se a seguinte matriz triangular inferior $A \in \mathbb{R}^{n\times n}$ definida por:*

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

Pode-se verificar que sua inversa é:

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

As normas $||A||_\infty$ e $||A^{-1}||_\infty$ são, respectivamente, 2 e 2. Já às normas $||A||_1$ e $||A^{-1}||_1$ são, respectivamente, n e n. Assim, os números de condição nas normas ∞ e 1, são:

$$\begin{aligned} cond_\infty(A) &= ||A||_\infty \cdot ||A^{-1}||_\infty = 2 \cdot 2 = 4, \\ cond_1(A) &= ||A||_1 \cdot ||A^{-1}||_1 = n \cdot n = n^2. \end{aligned}$$

Observe-se que para $n \gg 1$, $cond_1(A) \gg cond_\infty(A)$. Também observe que o determinante de A é 1. Para calcular o condicionamento da matriz A na norma 2, precisa-se determinar os autovalores da matriz A^TA. A matriz simétrica A^TA é dada por:

$$A^TA = \begin{bmatrix} n & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & & & & \\ 1 & & 1 & & & \\ \vdots & & & \ddots & & \\ 1 & & & & 1 & \\ 1 & 0 & & \cdots & 0 & 1 \end{bmatrix}. \tag{4.28}$$

Pode-se provar que os autovalores de A^TA são, respectivamente:

$$1, \quad \lambda_1 = \frac{n+1+\sqrt{(n+1)^2-4}}{2} \quad e \quad \lambda_2 = \frac{n+1-\sqrt{(n+1)^2-4}}{2}.$$

Sendo o maior e menor autovalor em módulo, respectivamente, λ_1 e 1. Logo,

$$cond_2(A) = \frac{\lambda_1}{1} = \lambda_1 = \frac{n+1+\sqrt{(n+1)^2-4}}{2} = \frac{(n+1)}{2}\left[1+\sqrt{1-\frac{4}{(n+1)^2}}\right].$$

Na Figura que segue apresenta-se as funções $cond(A)$ para a matriz A acima em função do tamanho da matriz.

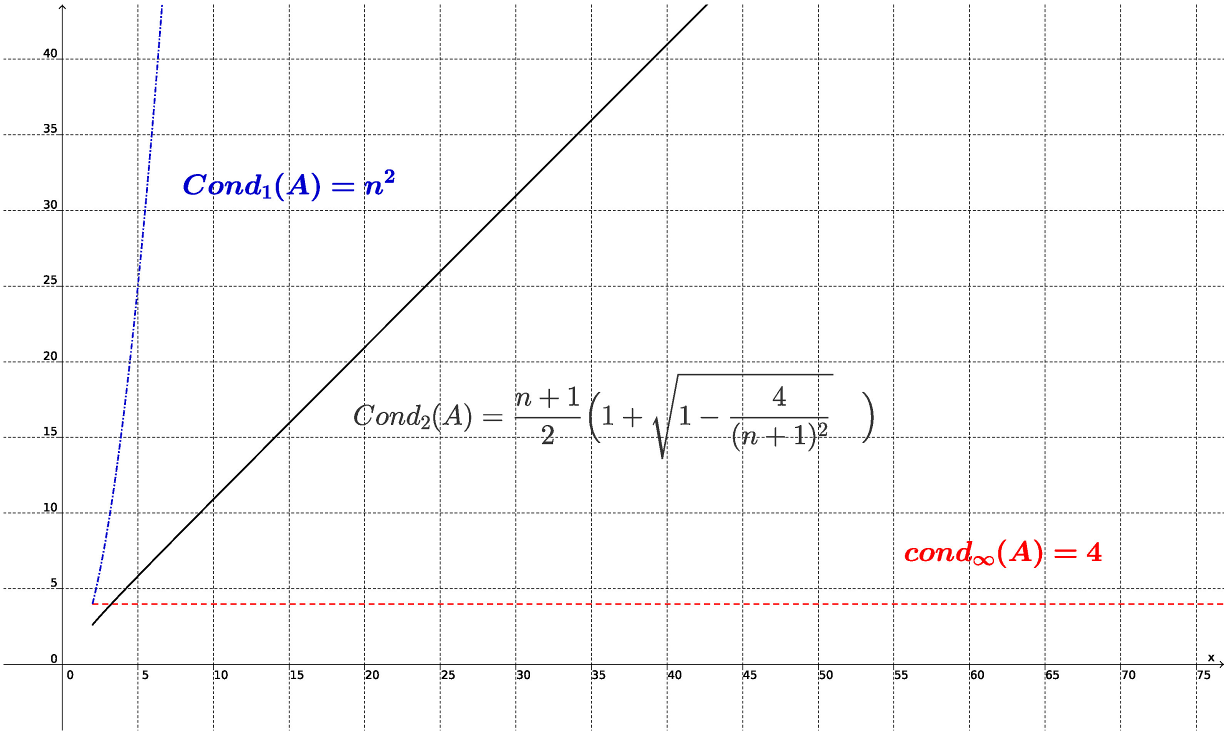


Figura 4.2: Número de condição da matriz A em função do tamanho da matriz.

Conclusão

Nesse trabalho foram apresentados limitadores superiores para os erros relativos à solução de sistemas lineares do tipo $A\vec{x} = \vec{b}$, quando: a matriz A é perturbada; o vetor $\vec{b}$ é perturbado; e a matriz A e o vetor $\vec{b}$ são perturbados. Além disso, evidenciaram-se todos os conceitos matemáticos necessários, tais como vetores, matrizes e normas que levam à dedução matemática desses limitadores. Desta forma, os objetivos propostos no presente trabalho foram atingidos.

Este trabalho permitiu revisar e aprender conceitos matemáticos de Álgebra Linear, assim como a demonstração matemática de muitas das propriedades apresentadas. Também, proporcionou o contato com metodologias de pesquisa que permitiram cumprir os objetivos propostos. Fruto desta pesquisa, o autor apresentou dois trabalhos em eventos científicos: Norma Matricial Subordinada e seu cálculo usando as imagens da Transformação Linear associada [11]; e Cálculo de Normas Subordinadas de Matrizes [12]. Além da experiência adquirida na realização do estudo, conquistou-se a bagagem matemática fundamental para a realização de investigações futuras relacionadas a condicionamento de matrizes.

Para trabalhos futuros, sugere-se:

1. Determinar os números de condição de matrizes associadas a Problemas de Valor de Contorno.

2. Implementar métodos numéricos para cálculo aproximado do número de condição de matrizes.

Apêndice

.1 Determinantes

Este apêndice é produto do que foi apresentado nos livros [2] e [7].

Definição .1.1. *O **determinante** de uma matriz A de tamanho $n \times n$ é definido recursivamente da seguinte maneira:*

(i) Se $A = [a_{11}]$, que é uma matriz quadrada de tamanho 1×1, o determinante de A, denotado por $\det(A)$, é a_{11}.

*(ii) Se A é uma matriz quadrada de tamanho $n \times n$, $n \geq 2$, definimos A_{ij} **o menor** de A como determinante da matriz de tamanho $(n-1) \times (n-1)$ obtida pela exclusão da i-ésima linha e da j-ésima coluna de A. Usando o chamado desenvolvimento de Lagrange, definimos*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{ij}, \quad \text{para qualquer } j, \quad 1 \leq j \leq n, \tag{29}$$

ou

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A_{ij}, \quad \text{para qualquer } i, \quad 1 \leq i \leq n, \tag{30}$$

Observe que obtemos o mesmo resultado, independentemente de usarmos (29) ou (30) e independentemente do valor de j em (29) ou de i em (30).

Se

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

representamos o determinante de A por:

$$\det\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{ou} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{31}$$

Para uma matriz quadrada de ordem 2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

temos,

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

e, para uma matriz quadrada de ordem 3,

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

Seguem alguns resultados envolvendo Determinantes.

Teorema .1.1. *A matriz*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

é invertível se, e só se, $\det(A) \neq 0$*, caso em que a inversa é dada pela fórmula*

$$A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Teorema .1.2. *Se A for uma matriz triangular $n \times n$ (triangular superior, inferior ou diagonal), então* $\det(A)$ *é o produto das entradas na diagonal principal da matriz, ou seja,* $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Exemplo .9. ***(Determinante de uma matriz triangular inferior)*** *As contas a seguir mostram que o determinante de uma matriz triangular inferior 4×4 é o produto de suas entradas diagonais. Cada parte da* *conta usa uma expansão em cofatores ao longo da primeira linha.*

$$\begin{aligned}
\begin{vmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \\
&= a_{11}a_{22} \begin{vmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{vmatrix} \\
&= a_{11}a_{22}a_{33}|a_{22}| \\
&= a_{11}a_{22}a_{33}a_{22}.
\end{aligned}$$

O seguinte lema é uma consequência direta do Teorema .1.1.

Lema .1.1. *Seja I a matriz identidade, então* $\det(I) = 1$.

Teorema .1.3. *Uma matriz quadrada A é invertível se, e só se,* $\det(A) \neq 0$.

Teorema .1.4. ***(Teorema de Binet)*** *Se A e B são matrizes quadradas de mesmo tamanho, então*

$$\det(AB) = \det(A)\det(B).$$

As demonstrações destes teoremas se encontram em [2].

Referências Bibliográficas

[1] CABRAL, Marco A. P. e GOLDFELD, Paulo, *Curso de Álgebra Linear*, ISBN, Rio de Janeiro, 2012.

[2] ANTON, Howard e RORRES, Chris, *Álgebra Linear com Aplicações*, bookman, Porto Alegre, 2012.

[3] SPERANDIO, Décio, MENDES, João Teixeira, e SILVA, Luiz Henry Monken, *Cálculo Numérico*, Pearson Prentice Hall, São Paulo, 2006.

[4] VALLE, Marcos Eduardo, *Aula 5 – Normas de vetores e matrizes. Condicionamento de uma Matriz*, UNICAMP, São Paulo.

[5] BERTOLOTO, Fábio José, *Notas de aula / Análise 3 – Noções topológicas no R^n*. UFU, Uberlândia, 2021.

[6] FRANCO, Neide Maria Bertoldi, *Cálculo Númerico*, Instituto de Ciências Matemáticas e de Computação, São Paulo.

[7] PHILLIPS, G. M. M. e TAYLOR, Peter J. *Theory and Applications of Numerical Analysis*, Second edition, September 1996.

[8] SILVA, Marcelo Frankye Azevedo da, *TCC Pré-condicionamento do Método dos Gradientes Conjugados para Sistemas Lineares*, UFAP, Amapá, 2012.

[9] DATTA, Biswa Nath,*Numerical Linear Algebra and Applications*, Second Edition, Philadelphia, 2010.

[10] VALLE, Marcos Eduardo, *Aula 10 – Teorema do Valor Extremo e o Método dos Multiplicadores de Lagrange*, UNICAMP, São Paulo.

[11] DIAS, Ester Ferreira e ENRIQUEZ-REMIGIO, Santos Alberto, *Norma Matricial Subordinada e seu cálculo usando as imagens da Transformação Linear associada*, Anais da Semat e Semest, Uberlândia, 2020.

[12] DIAS, Ester Ferreira e ENRIQUEZ-REMIGIO, Santos Alberto, *Cálculo de Normas Subordinadas de Matrizes*, Anais da Semat e Semest, Uberlândia, 2020.