

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL -
PROFMAT**

WILLIAN BURGARDT DE SOUZA

**MÉTODO DOS MÍNIMOS QUADRADOS APLICADO A UM PROBLEMA DE
GEOPOSICIONAMENTO**

CURITIBA

2018

WILLIAN BURGARDT DE SOUZA

**MÉTODO DOS MÍNIMOS QUADRADOS APLICADO A UM PROBLEMA DE
GEOPOSICIONAMENTO**

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional da Universidade Tecnológica Federal do Paraná em Curitiba - PROFMAT-UTCT como requisito parcial para obtenção do grau de Mestre.

Orientadora: Denise de Siqueira

Coorientador: Rodolfo Gotardo Begiato

CURITIBA

2018

Dados Internacionais de Catalogação na Publicação

S729m Souza, Willian Burgardt de
2018 Método dos mínimos quadrados aplicado a um problema
de geoposicionamento / Willian Burgardt de Souza.--
2018.
51 f.: il.; 30 cm.

Disponível também via World Wide Web.
Texto em português com resumo em inglês.
Dissertação (Mestrado) - Universidade Tecnológica
Federal do Paraná. Programa de Mestrado Profissional em
Matemática em Rede Nacional, Curitiba, 2018.
Bibliografia: f. 51.

1. Sistemas lineares. 2. Mínimos quadrados. 3. Álgebra
linear. 4. Sistema de posicionamento global. 5. Solução
de problemas. 6. Matemática - Estudo e ensino. 7.
Matemática - Dissertações. I. Siqueira, Denise de, orient.
II. Begiato, Rodolfo Gotardi, coorient. III. Universidade
Tecnológica Federal do Paraná. Programa de Mestrado
Profissional em Matemática em Rede Nacional. IV. Título.

CDD: Ed. 23 - 510

Biblioteca Central do Câmpus Curitiba - UTFPR
Bibliotecária: Luiza Aquemi Matsumoto CRB-9/794

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 48

A Dissertação de Mestrado intitulada **Método dos mínimos quadrados aplicado a um problema de geoposicionamento**, defendida em sessão pública pelo(a) candidato(a) **Willian Burgardt de Souza**, no dia 08 de fevereiro de 2018, foi julgada para a obtenção do título de Mestre, área de concentração Matemática, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Matemática em Rede Nacional.

BANCA EXAMINADORA:

Prof(a). Dr(a). Denise Siqueira - Presidente - UTFPR

Prof(a). Dr(a). Rodrigo Garcia Eustáquio - UTFPR

Prof(a). Dr(a). Mael Sachine - UFPR

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 08 de fevereiro de 2018.

Carimbo e Assinatura do(a) Coordenador(a) do Programa

AGRADECIMENTOS

Aos professores do programa, agradeço por todo incentivo e conselhos, de ordem profissional ou pessoal. Em especial, a minha orientadora professora Denise de Siqueira, por todo o auxílio e compreensão, assim como ao meu orientador Rodolfo Gotardo Begiato, que não poderiam ter sido melhores aliados durante a elaboração deste trabalho.

A minha família agradeço todo tipo de apoio que oportunizou estes estudos desde o início. Em especial, ao meu pai e ao meu sogro por todo suporte. A minha esposa Ariel Marczaki, que sempre me apoiou e auxiliou nos momentos mais difíceis.

Aos membros da banca examinadora pelas sugestões e contribuições, que auxiliaram neste trabalho.

À CAPES pelo apoio financeiro, através de bolsa de estudos, à minha pesquisa.

RESUMO

SOUZA, Willian Burgardt. **Método dos mínimos quadrados aplicado a um problema de geoposicionamento**. 47 f. Dissertação - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

O objetivo deste trabalho é apresentar o método dos mínimos quadrados para resolver sistemas lineares sobredeterminados, ou seja, sistemas da forma $A\mathbf{x} = \mathbf{b}$, em que $A_{m \times n}$, com $m > n$. Neste sentido, veremos como a resolução destes sistemas estão relacionados com encontrar a projeção ortogonal \mathbf{b} sobre o subespaço gerado pelas colunas de A . Este tipo de sistema é usado ainda para modelar um problema de geoposicionamento, cujo objetivo é determinar a posição de um receptor que recebe o sinal de vários satélites.

Palavras-chave: Sistemas Lineares. Mínimos Quadrados. Geoposicionamento.

ABSTRACT

SOUZA, Willian Burgardt. **Least square method applied to a geo-positioning problem.** 47 pg. Dissertation - Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Universidade Tecnológica Federal do Paraná. Curitiba, 2018

The main goal of this work is to present the least squares method to solve overdetermined linear systems, that is, systems of the form $Ax = \mathbf{b}$, where $A_{m \times n}$, with $m > n$. In this sense, we showed that the resolution of these systems is related to the orthogonal projection problem of \mathbf{b} on the subspace generated by the columns of A . This type of system is used to model a problem of geo-positioning, whose objective is to determine the position of a receiver that receives the signal from several satellites.

Keywords: Linear System. Least Square Method. Geo-positioning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Imagem $\mathbf{b} = A\mathbf{x}$ para dois vetores do \mathbb{R}^2	21
Figura 2 – Ação de A sobre os vetores definidos em um quadrado unitário	22
Figura 3 – Representação de $Im(A)$ e $N(A)$	23
Figura 4 – Representação do $N(A)$	24
Figura 5 – Ilustrações de $B[0, 1]$ utilizando diferentes normas	27
Figura 6 – Ilustração da norma 2 de matrizes	29
Figura 7 – Ilustração da norma 1 de matrizes	29
Figura 8 – Representação dos conjuntos S e S^\perp	32
Figura 9 – Projeção do vetor \mathbf{u} sobre o vetor \mathbf{v}	33
Figura 10 – Projeção de \mathbf{u} sobre \mathbf{v}	34
Figura 11 – Projeção do vetor \mathbf{u} sobre S	35
Figura 12 – Representação da superfície esférica de centro (a, b, c)	44
Figura 13 – Representação da interseção de duas superfícies esféricas de centros (a_1, b_1, c_1) e (a_2, b_2, c_2)	45
Figura 14 – Representação da interseção de três superfícies esféricas de centros (a_1, b_1, c_1) , (a_2, b_2, c_2) , e (a_3, b_3, c_3)	45
Figura 15 – Localização do receptor com 4 satélites	47
Figura 16 – Representação da posição real do receptor	49
Figura 17 – Representação dos sinais “multi-path”	50

SUMÁRIO

	INTRODUÇÃO	15
1	FUNDAMENTAÇÃO TEÓRICA	17
1.1	Espaços e subespaços vetoriais	17
1.2	Base e dimensão	24
1.3	Norma de Vetores	26
1.4	Ortogonalidade	29
2	SISTEMAS SOBREDETERMINADOS	33
2.1	Projeção Ortogonal	33
2.1.1	Projeção de um vetor sobre outro vetor	33
2.1.2	Projeção de um vetor sobre um subespaço	34
2.2	Mínimos quadrados	39
3	APLICAÇÃO	43
3.1	Sistema de Geoposicionamento por Satélite	43
3.2	Determinando a distância entre o satélite e o receptor	43
4	CONCLUSÃO	51
	REFERÊNCIAS	53

INTRODUÇÃO

O uso de sistemas lineares é empregado na resolução de diversos tipos de problemas dentro da Matemática. No caso de sistemas lineares sobredeterminados, ou seja, sistemas lineares com mais equações do que incógnitas, em geral, ou o sistema possui infinitas soluções ou não possui nenhuma solução. Este trabalho estudará o segundo caso, em que o sistema não possui solução e neste caso o método dos mínimos quadrados é empregado para encontrar uma "solução aproximada".

O método dos mínimos quadrados consiste em determinar uma aproximação para o sistema de forma a tornar o resíduo o menor possível.

Ao longo deste trabalho, serão explorados todos os conceitos necessários para estabelecer uma sistemática de resolução deste tipo de problema.

Como aplicação dos resultados teóricos estabelecidos é apresentado um problema de geoposicionamento, ou seja, o objetivo é determinar a posição de um receptor em um sistema de geoposicionamento global. Como exemplo deste tipo de sistema, o mais conhecido é o GPS (*Global Positioning System*), traduzido do inglês como Sistema de Posicionamento Global.

Para isso este trabalho está assim dividido:

No Capítulo 1 são apresentados os conceitos básicos da álgebra linear, fundamentais para o desenvolvimento do trabalho, explorando as ideias de espaços e subespaços vetoriais, base e dimensão de um subespaço vetorial, norma de vetores e por fim ortogonalidade.

O Capítulo 2 inicia com as principais propriedades de projeção ortogonal e posteriormente, é apresentado o método dos mínimos quadrados, objeto principal de estudo deste trabalho.

Por fim, no Capítulo 3 é apresentado um sistema de geoposicionamento por satélite (GPS) explicando o seu funcionamento de forma breve e desenvolvendo um modelo matemático para determinar a posição de um receptor, com base nas informações obtidas pelos satélites.

1 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é apresentar uma revisão dos conceitos da álgebra linear que serão abordados neste trabalho. Sempre que necessário será apresentado um exemplo para ilustrar algum conceito. Todos os resultados apresentados aqui foram consultados em (GOLUB; LOAN, 1996), (LIMA, 2009), (MEYER, 2000), (STRANG, 2009) e (TREFETHEN; III, 1997).

1.1 ESPAÇOS E SUBESPAÇOS VETORIAIS

A noção de espaço vetorial será de grande importância para o estudo apresentado aqui. Esta seção apresentará resultados sobre espaços, subespaços, espaço linha, espaço coluna, núcleo e imagem, além de apresentar exemplos para ilustrar esses assuntos. Ainda será dada uma atenção especial ao espaço das matrizes.

Definição 1.1. *Um espaço vetorial V é um conjunto munido das operações de adição e multiplicação por escalar, em que dados $\mathbf{u}, \mathbf{v} \in V$ e $\alpha, \beta \in \mathbb{R}$ tem-se que $\alpha\mathbf{u} + \mathbf{v} \in V$ e satisfaz as seguintes propriedades:*

1. *comutatividade: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;*
2. *associatividade: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ e $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;*
3. *elemento neutro da adição: existe um vetor $\mathbf{0} \in V$, chamado de vetor nulo, tal que $\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}$ para todo $\mathbf{v} \in V$;*
4. *inverso aditivo: para cada vetor $\mathbf{v} \in V$ existe um vetor $-\mathbf{v} \in V$ chamado o inverso aditivo, ou o simétrico de \mathbf{v} tal que $-\mathbf{v} + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$;*
5. *distributividade: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ e $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$;*
6. *elemento neutro da multiplicação: $1 \cdot \mathbf{v} = \mathbf{v}$.*

Exemplo 1.2. *O conjunto $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$ munido das operações usuais de adição e multiplicação por escalar é um espaço vetorial.*

O elemento neutro da adição em \mathbb{R}^n é o vetor $\mathbf{0} = (0, 0, \dots, 0)$ e o elemento neutro da multiplicação é o escalar 1.

Exemplo 1.3. *O conjunto $M_{m \times n}$ de todas as matrizes $m \times n$ munidos da operação de adição ($[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$) e da operação de multiplicação por escalar ($\alpha[a_{ij}] = [\alpha a_{ij}]$) com $i = 1, \dots, m$ e $j = 1, \dots, n$, é um espaço vetorial. Neste caso, o elemento neutro da adição em $M_{m \times n}$ é a matriz $m \times n$ em que todos os seus elementos são nulos, o elemento neutro da multiplicação é o escalar 1 e o inverso aditivo da matriz $[a_{ij}]$ é a matriz $[-a_{ij}]$.*

A partir deste momento sempre que for citado o conjunto $M_{m \times n}$ trata-se do conjunto das matrizes $m \times n$ descrito conforme o Exemplo 1.3.

Muitas vezes necessitamos trabalhar com subconjuntos de um espaço vetorial. Estes subconjuntos, munidos de algumas propriedades dão origem ao que chamamos de subespaços vetoriais.

Definição 1.4. *Seja V um espaço vetorial e W um subconjunto não vazio de V . Dizemos que W é um subespaço vetorial de V se, dados $\mathbf{u}, \mathbf{v} \in W$ e $\alpha \in \mathbb{R}$ as seguintes propriedades forem satisfeitas:*

1. $\mathbf{0} \in W$,
2. $\mathbf{u} + \alpha\mathbf{v} \in W$.

Note que o subconjunto W satisfaz todas as propriedades de espaço vetorial, sendo ele próprio um espaço vetorial.

Exemplo 1.5. *Seja $S = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$ um subconjunto de \mathbb{R}^n com $A \in M_{m \times n}$. S assim definido é um subespaço vetorial de \mathbb{R}^n . De fato, observe inicialmente que $\mathbf{0} \in S$, pois $A\mathbf{0} = \mathbf{0}$. Além disso, dados $\mathbf{x}_1, \mathbf{x}_2 \in S$ e $\alpha \in \mathbb{R}$, tem-se que*

$$\begin{aligned} A(\mathbf{x}_1 + \alpha\mathbf{x}_2) &= A(\mathbf{x}_1) + A(\alpha\mathbf{x}_2) \\ &= A(\mathbf{x}_1) + \alpha A(\mathbf{x}_2) \\ &= \mathbf{0} + \alpha\mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

Portanto S é um subespaço vetorial de \mathbb{R}^n .

Exemplo 1.6. *Seja $S = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{b}, \mathbf{b} \neq \mathbf{0}\}$ um subconjunto de \mathbb{R}^n com $A \in M_{m \times n}$. S assim definido não é um subespaço vetorial de \mathbb{R}^n , pois dados $\mathbf{x}_1, \mathbf{x}_2 \in S$, tem-se que*

$$\begin{aligned} A(\mathbf{x}_1 + \mathbf{x}_2) &= A\mathbf{x}_1 + A\mathbf{x}_2 \\ &= 2\mathbf{b} \notin S. \end{aligned}$$

Exemplo 1.7. *Seja $V = M_{2 \times 3}$ e $S = \{A \in V \mid a_{j,3} = 1, j = 1, 2\}$. S não é um subespaço vetorial de V pois $\mathbf{0}_{2 \times 3} \notin S$.*

Uma das características de um espaço vetorial é que ele pode ser caracterizado por somente uma parte de seus vetores. Nos casos em que trabalharemos ele pode ser caracterizado por um conjunto finito de vetores. Veremos a seguir como fazer essa caracterização.

Definição 1.8. *O vetor \mathbf{u} é dito ser combinação linear de $\mathbf{u}_1, \dots, \mathbf{u}_n$ se existem escalares $\alpha_1, \dots, \alpha_n$ tais que*

$$\mathbf{u} = \alpha_1\mathbf{u}_1 + \dots + \alpha_n\mathbf{u}_n = \sum_{i=1}^n \alpha_i\mathbf{u}_i.$$

Definição 1.9. *Sejam V um espaço vetorial e $X \subset V$ um conjunto não vazio. O conjunto de todas as combinações lineares de X é chamado espaço gerado por X e denotado por $[X]$. Ou seja*

$$[X] = \left\{ \mathbf{u} \in V \mid \mathbf{u} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n = \sum_{i=1}^n \alpha_i \mathbf{u}_i, \text{ onde } \mathbf{u}_1, \dots, \mathbf{u}_n \in X \text{ e } \alpha_1, \dots, \alpha_n \in \mathbb{R} \right\}.$$

Neste caso, dizemos que X é o conjunto gerador de $[X]$.

Exemplo 1.10. *Seja $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$, $\mathbf{u} = (a, b, c) \in \mathbb{R}^3$ e $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. Os vetores $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ geram o \mathbb{R}^3 . De fato, a equação vetorial*

$$\alpha_1(1, 0, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 0, 1) = (a, b, c) \quad (1.1)$$

ou ainda

$$(\alpha_1, \alpha_2, \alpha_3) = (a, b, c) \quad (1.2)$$

tem solução $\alpha_1 = a$, $\alpha_2 = b$, $\alpha_3 = c$. Portanto $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ geram o \mathbb{R}^3 .

Proposição 1.11. *Sejam V um espaço vetorial e $X \subset V$ um conjunto não vazio. O espaço gerado por X é um subespaço vetorial.*

Demonstração. 1. $\mathbf{0} \in [X]$. De fato, dado $\mathbf{v} \in X$ temos que $\mathbf{0} = 0\mathbf{v}$.

2. $\mathbf{u} + \alpha\mathbf{v} \in [X]$. De fato, dados $\mathbf{u}, \mathbf{v} \in [X]$ temos que existem $\mathbf{u}_1, \dots, \mathbf{u}_n \in X$ e $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{R}$ tais que $\mathbf{u} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n$ e $\mathbf{v} = \beta_1 \mathbf{u}_1 + \cdots + \beta_n \mathbf{u}_n$. Assim, $\mathbf{u} + \alpha\mathbf{v} = (\alpha_1 + \alpha\beta_1)\mathbf{u}_1 + \cdots + (\alpha_n + \alpha\beta_n)\mathbf{u}_n \in [X]$

□

Afim de explorar um pouco mais os espaços das matrizes considere o conjunto $M_{m \times n}$ e defina

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \quad (1.3)$$

os vetores formados pelas colunas de $A \in M_{m \times n}$. Logo a matriz A pode ser reescrita como

$$A = \left(\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \right). \quad (1.4)$$

Tome agora o conjunto de todas as combinações lineares das colunas de A , a este conjunto denotamos de $C(A)$,

$$C(A) = \left\{ \sum_{i=1}^n \alpha_i \mathbf{a}_i \mid \alpha_i \in \mathbb{R}, \mathbf{a}_i \in \mathbb{R}^m \right\}. \quad (1.5)$$

$C(A)$ assim definido é um subespaço vetorial de \mathbb{R}^m . Este subespaço é chamado de **espaço coluna de A** .

Do mesmo modo definindo

$$\bar{\mathbf{a}}_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix}, \dots, \bar{\mathbf{a}}_m = \begin{pmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (1.6)$$

os vetores formados pelas linhas de $A \in M_{m \times n}$, tem-se que A pode ser reescrita como

$$A = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \bar{\mathbf{a}}_2 \\ \vdots \\ \bar{\mathbf{a}}_m \end{pmatrix}. \quad (1.7)$$

De modo análogo a (1.5) define-se **espaço linha** da matriz A por

$$L(A) = \left\{ \sum_{i=1}^m \alpha_i \bar{\mathbf{a}}_i \mid \alpha_i \in \mathbb{R}, \bar{\mathbf{a}}_i \in \mathbb{R}^n \right\}. \quad (1.8)$$

Observe que $L(A)$ é um subespaço de \mathbb{R}^n .

Exemplo 1.12. Dada a matriz

$$A = \begin{pmatrix} 1 & 3 & 3 \\ 2 & 4 & 6 \end{pmatrix} \quad (1.9)$$

o espaço coluna de A é formado por todos os vetores da forma

$$\mathbf{v} = \alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \alpha_3 \begin{pmatrix} 3 \\ 6 \end{pmatrix}. \quad (1.10)$$

O espaço linha de A é formado por todos os vetores da forma

$$\mathbf{v} = \alpha_1 \begin{pmatrix} 1 & 3 & 3 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 & 4 & 6 \end{pmatrix}, \quad (1.11)$$

com $\alpha_1, \alpha_2 \in \mathbb{R}$.

Outros dois subespaços de grande importância para o nosso estudo são o núcleo e a imagem de uma matriz.

Definição 1.13. Dada a matriz $A \in M_{m \times n}$ define-se

$$Im(A) = \{ \mathbf{b} \in \mathbb{R}^m \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \in \mathbb{R}^n \}. \quad (1.12)$$

Note que se $\mathbf{x} = (\alpha_1, \dots, \alpha_n)$ com $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ e $A \in M_{m \times n}$ então \mathbf{Ax} é o conjunto de todas as combinações lineares das colunas de A , ou seja, $Im(A) = C(A)$. Assim sendo,

$Im(A)$ é gerada pelas colunas de A ,

$$\mathbf{b} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \quad (1.13)$$

$$= \begin{pmatrix} \mathbf{a}_1 \end{pmatrix} \alpha_1 + \begin{pmatrix} \mathbf{a}_2 \end{pmatrix} \alpha_2 + \cdots + \begin{pmatrix} \mathbf{a}_n \end{pmatrix} \alpha_n \in C(A). \quad (1.14)$$

Definição 1.14. Dada a matriz $A_{m \times n}$ define-se

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}. \quad (1.15)$$

Os próximos exemplos ilustram estes subespaços.

Exemplo 1.15. Considere $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. A $Im(A) = \mathbb{R}^2$. De fato, $Im(A)$ é gerada por $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e, dado $(a, b) \in \mathbb{R}^2$

$$\begin{pmatrix} a \\ b \end{pmatrix} = (-a) \begin{pmatrix} -1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A Figura 1 ilustra a imagem de dois vetores após a aplicação da matriz A .

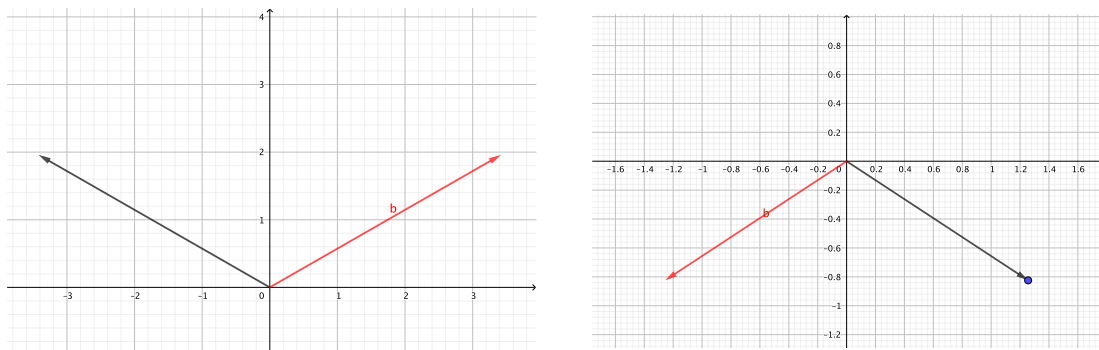


Figura 1 – Imagem $\mathbf{b} = A\mathbf{x}$ para dois vetores do \mathbb{R}^2

O $N(A)$ é o conjunto formado por todos os vetores $\mathbf{x} = (x_1, x_2)^T$ tais que

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.16)$$

Desta forma $N(A) = \{(0, 0)^T\}$.

Exemplo 1.16. Considere $A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$. A $\text{Im}(A) = \mathbb{R}^2$. De fato, $\text{Im}(A)$ é gerada por $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e, dado $(a, b) \in \mathbb{R}^2$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{(3a - b)}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{(b - a)}{2} \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

A Figura 2 ilustra a ação da matriz A sobre os vetores definidos em um quadrado unitário.

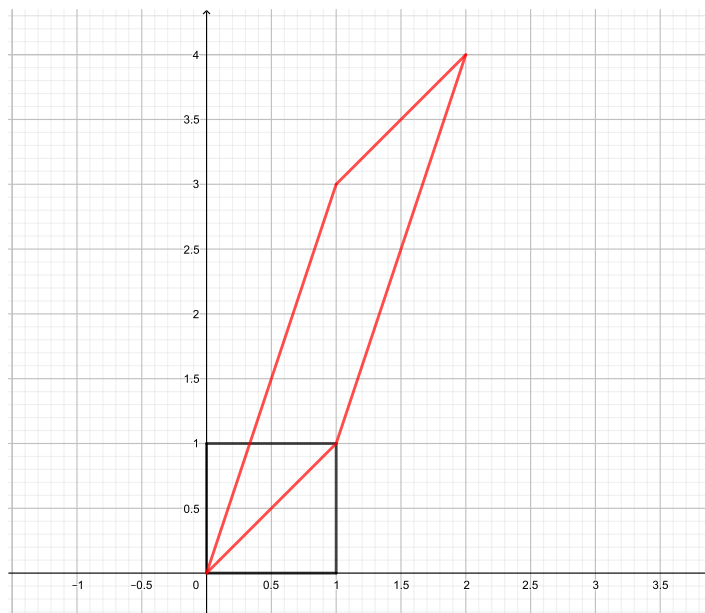


Figura 2 – Ação de A sobre os vetores definidos em um quadrado unitário

O $N(A)$ é o conjunto formado por todos os vetores $\mathbf{x} = (x_1, x_2)^T$ tais que

$$\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.17)$$

Desta forma $N(A) = \{(0, 0)^T\}$.

Exemplo 1.17. Considere $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. $\text{Im}(A)$ não gera \mathbb{R}^2 , por exemplo $(1, 1)^T$ não é gerado pois

$$C(A) = \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 4 \end{pmatrix} = (\alpha + 2\beta) \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Logo $C(A)$ são todos os múltiplos de $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ conforme ilustrado na Figura 3.

Além disso, $N(A)$ é o conjunto formado por todos os vetores $\mathbf{x} = (x_1, x_2)^T$ tais que

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.18)$$

Desta forma $N(A) = \{x_2(-2, 1)^T \mid x_2 \in \mathbb{R}\}$. Assim $N(A)$ é gerado por $(-2, 1)^T$ conforme ilustra a Figura 3.

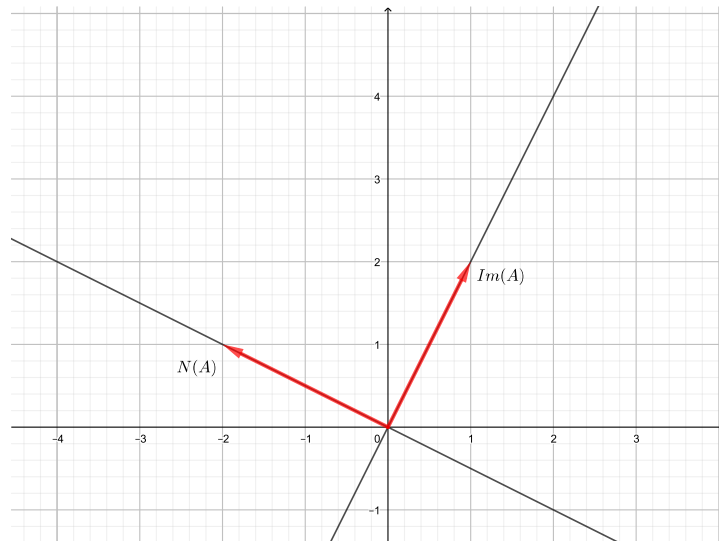


Figura 3 – Representação de $Im(A)$ e $N(A)$

Exemplo 1.18. Considere $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 6 \end{pmatrix}$. A $Im(A) = \mathbb{R}^2$. De fato, dado $a, b, c \in \mathbb{R}$,

$$C(A) = a \begin{pmatrix} 1 \\ 2 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c \begin{pmatrix} 2 \\ 6 \end{pmatrix} = (a + 2c) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (b + 2c) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Logo $Im(A)$ é gerado por $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

O $N(A)$ é o conjunto formado por todos os vetores $\mathbf{x} = (x_1, x_2, x_3)^T$ tais que

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.19)$$

Desta forma $N(A) = \{(-2x_3, -2x_3, x_3)^T : x_3 \in \mathbb{R}\}$ ou seja, $N(A)$ são todos os vetores que estão sobre a reta gerada por $\mathbf{v} = (-2, -2, 1)^T$ conforme Figura 4.

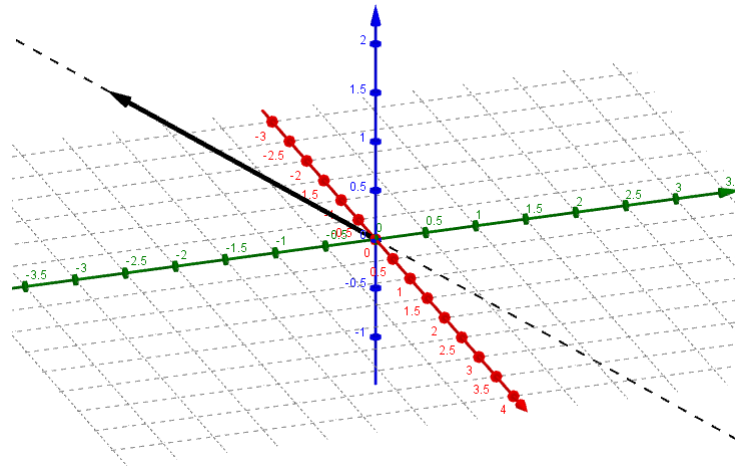


Figura 4 – Representação do $N(A)$

1.2 BASE E DIMENSÃO

Os espaços vetoriais de dimensão finita possuem uma estrutura algébrica que pode ser evidenciada pelas ideias de base e dimensão. Determinada uma base para um espaço vetorial, seus elementos são meras combinações lineares dos vetores que compõem a base. Nesta seção serão apresentados as definições de base e dimensão mas para isso vejamos os conceitos de dependência e independência linear.

Definição 1.19. *Seja $X = \{v_1, v_2, \dots, v_n\}$ um subconjunto de um espaço vetorial V . Dizemos que X é **linearmente independente** (abreviadamente **L.I.**) se $\alpha_1 v_1 + \dots + \alpha_n v_n = \mathbf{0}$ implicar que $\alpha_i = 0$ para todo $i = 1, \dots, n$. Se um conjunto não for L.I. dizemos que ele é **linearmente dependente** (abreviadamente **L.D.**).*

Exemplo 1.20. *Seja $S \subset \mathbb{R}^3$ definido por $S = \{(1, 2, 3)^T, (4, 5, 6)^T, (7, 8, 9)^T\}$. S assim definido é um conjunto L.D. De fato, $(7, 8, 9)^T + (1, 2, 3)^T + (-2)(4, 5, 6)^T = (0, 0, 0)^T$.*

Exemplo 1.21. *Seja $S \subset \mathbb{R}^3$ definido por $S = \{(1, 2, 3)^T, (4, 5, 6)^T, (1, 8, 9)^T\}$. S assim definido é um conjunto L.I.. Verificar que este conjunto é L.I. é equivalente a resolver o sistema*

$$\begin{pmatrix} 1 & 4 & 1 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (1.20)$$

Resolvendo o sistema, percebe-se que temos solução única $\alpha_1 = \alpha_2 = \alpha_3 = 0$, o que indica que o conjunto é L.I..

Observe que se considerarmos uma matriz A , cujas colunas são formadas pelos vetores \mathbf{a}_i , dizer que as colunas de A são L.I. é equivalente a dizer que

$$\alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n = \mathbf{0}$$

tem solução única $\alpha_i = 0$ para todo $i = 1, \dots, n$. Ou seja, o sistema

$$A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \mathbf{0}$$

deve ter solução única $\alpha_i = 0$ para todo $i = 1, \dots, n$, o que nos indica que o núcleo da matriz A deve conter somente o vetor nulo.

Definição 1.22. Uma **base** de um espaço vetorial V é um conjunto $X \subset V$ linearmente independente que gera V .

Proposição 1.23. Se o espaço vetorial V admite uma base com n elementos, qualquer outra base de V conterá também n elementos.

Demonstração. Ver (LIMA, 2009), pag. 30, Corolário 2. □

Dessa maneira, se X é uma base de V então o número de elementos de X é a chamada **dimensão** de V , denotada por $\dim(V)$.

Seja $A \in M_{m \times n}$, denominamos de **posto coluna** de A a dimensão do espaço coluna de A , ou seja, $\dim(C(A))$. Observe portanto que $\dim(C(A)) \leq n$. Do mesmo modo, denominamos de **posto linha** de A a dimensão do espaço linha da matriz A , ou seja, $\dim(L(A))$ e neste caso $\dim(L(A)) \leq m$.

Teorema 1.24. Dada uma matriz $A \in M_{m \times n}$ tem-se que o posto linha é igual ao posto coluna.

Demonstração. Seja p o posto coluna de $A \in M_{m \times n}$, logo existem vetores $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p\}$ que formam uma base para $C(A)$. Além disso denotamos cada \mathbf{w}_k por,

$$\mathbf{w}_k = \begin{pmatrix} w_{1k} \\ w_{2k} \\ \vdots \\ w_{mk} \end{pmatrix}.$$

Sendo assim para cada $\mathbf{a}_j \in C(A)$ com $j = 1, \dots, n$ tem-se

$$\mathbf{a}_j = \alpha_{j1}\mathbf{w}_1 + \alpha_{j2}\mathbf{w}_2 + \dots + \alpha_{jp}\mathbf{w}_p. \quad (1.21)$$

Tomando a i -ésima coordenada de cada elemento de (1.21) temos que

$$\begin{aligned} a_{i1} &= \alpha_{11}w_{i1} + \alpha_{12}w_{i2} + \dots + \alpha_{1p}w_{ip} \\ a_{i2} &= \alpha_{21}w_{i1} + \alpha_{22}w_{i2} + \dots + \alpha_{2p}w_{ip} \\ &\vdots = \vdots \\ a_{in} &= \alpha_{n1}w_{i1} + \alpha_{n2}w_{i2} + \dots + \alpha_{np}w_{ip} \end{aligned}$$

Observe com isso que, para cada elemento da linha de A , $\bar{\mathbf{a}}_i = (a_{i1} \ a_{i2} \cdots a_{in})$ pode ser escrito como,

$$\bar{\mathbf{a}}_i = w_{i1} \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{pmatrix} + w_{i2} \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{n2} \end{pmatrix} + \cdots + w_{ip} \begin{pmatrix} \alpha_{1p} \\ \alpha_{2p} \\ \vdots \\ \alpha_{np} \end{pmatrix}, \quad (1.22)$$

ou seja, as linhas de A são geradas por vetores da forma $\alpha_k = (\alpha_{11}, \alpha_{21}, \dots, \alpha_{nk})^T$, com $k = 1, \dots, p$. Assim $\dim(L(A)) \leq p$ e portanto $p = \dim(C(A)) \geq \dim(L(A))$.

Repetindo este mesmo processo na matriz A^T , como as colunas de A^T são as linhas de A , é possível mostrar que $\dim(C(A^T)) = \dim(L(A)) \geq \dim(L(A^T)) = \dim(C(A))$ e portanto o resultado segue.

□

Por fim definimos o posto de uma matriz denotado por $\rho(A)$, como sendo $\dim(C(A))$ (ou $\dim(L(A))$). Observe portanto que $\rho(A) \leq \min\{m, n\}$.

Observe que tanto no Exemplo 1.15 quanto no Exemplo 1.18 vale $\rho(A) = \dim(\text{Im}(A)) = 2$, visto que as matrizes A desses exemplos possuem duas colunas L.I.

1.3 NORMA DE VETORES

Os vetores desempenham um papel importante na matemática e na física. Através deles podemos representar velocidade, aceleração ou as forças que agem sobre um objeto por exemplo. Nesta seção será apresentada a definição e os principais tipos de norma de vetor. Além disso será definida norma de matrizes e sempre que possível serão apresentadas ilustrações para facilitar o entendimento.

Definição 1.25. Dados $\mathbf{u} = (u_1, \dots, u_n)^T$, $\mathbf{v} = (v_1, \dots, v_n)^T \in \mathbb{R}^n$, o produto interno usual do \mathbb{R}^n é a função $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 \cdot v_1 + \dots + u_n \cdot v_n. \quad (1.23)$$

Definição 1.26. Dado $\mathbf{x} \in \mathbb{R}^n$, a norma usual de \mathbb{R}^n é a função $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Note que a norma satisfaz as seguintes propriedades:

1. $\|\mathbf{x}\| \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$ e $\|\mathbf{x}\| = 0$ se e somente se $\mathbf{x} = 0$,
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$3. \|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|, \alpha \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n.$$

É possível definir ainda outras normas em \mathbb{R}^n , as mais usuais são as chamadas *normas p*, em que dado $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$,

$$\|\mathbf{x}\|_p = (|x_1|^p + \dots + |x_n|^p)^{\frac{1}{p}}, \quad p \geq 1. \quad (1.24)$$

Vejamos alguns casos:

1. $p = 1$, $\|\mathbf{x}\|_1 = |x_1| + |x_2| + \dots + |x_n|$
2. $p = 2$, $\|\mathbf{x}\|_2^2 = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2$
3. $p = 3$, $\|\mathbf{x}\|_3^3 = |x_1|^3 + |x_2|^3 + \dots + |x_n|^3$
4. $p = \infty$, $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$

Observação 1.27. Quando $p=2$ temos a norma usual.

Afim de ilustrar geometricamente algumas das p -normas, considere em \mathbb{R}^2 o seguinte conjunto:

$$B[0, 1] = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_p \leq 1\}.$$

A Figura 5, ilustra este conjunto para diferentes escolhas de normas.

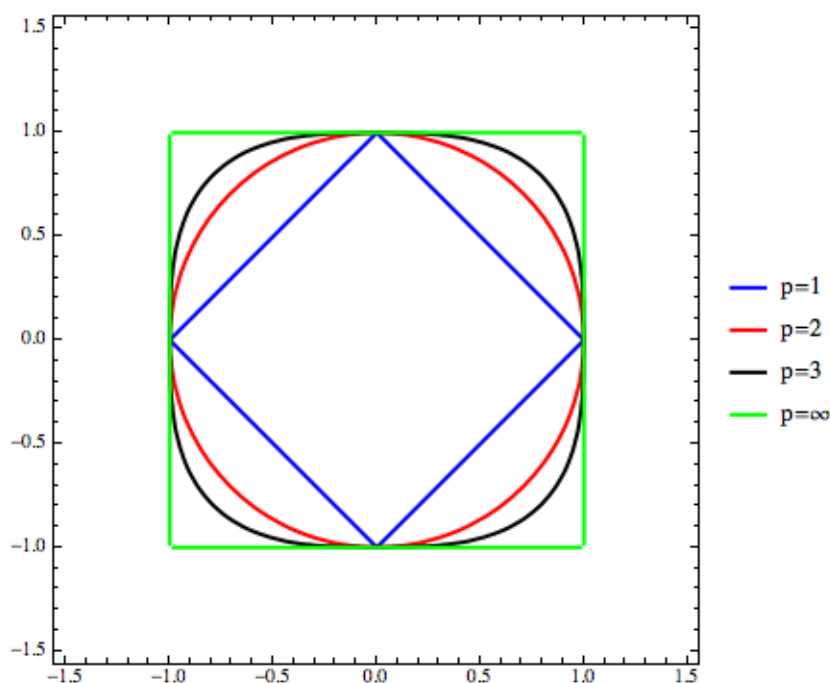


Figura 5 – Ilustrações de $B[0, 1]$ utilizando diferentes normas

Como existem diferentes tipos de normas é importante nos questionarmos sobre qual delas devemos escolher para solucionar um determinado problema. Para isso, o conceito de equivalência de norma é importante.

Definição 1.28. Dadas $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$ normas, dizemos que $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$ são equivalentes se existem constantes positivas, c_1 e c_2 tais que:

$$c_1 \|\mathbf{x}\|_\alpha \leq \|\mathbf{x}\|_\beta \leq c_2 \|\mathbf{x}\|_\alpha. \quad (1.25)$$

Exemplo 1.29. Seja $\mathbf{v} = (1, -2, 3)^T$ um vetor de \mathbb{R}^3 . Assim temos que:

$$\|\mathbf{v}\|_1 = |1| + |-2| + |3| = 6 \quad (1.26)$$

$$\|\mathbf{v}\|_2 = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}. \quad (1.27)$$

Observe que, tomando $c_1 = 0$ e $c_2 = 1$ tem-se que

$$0 \cdot 6 \leq \sqrt{14} \leq 1 \cdot 6. \quad (1.28)$$

O que ilustra a equivalência das normas $\|\cdot\|_1$ e $\|\cdot\|_2$ em \mathbb{R}^3 .

O exemplo anterior ilustra um resultado mais geral no qual afirma que em \mathbb{R}^n todas as normas são equivalentes, logo a escolha da norma pode ser tomada da forma mais conveniente, de acordo com o problema abordado.

Para o espaço das matrizes $M_{m \times n}$ é possível definir as seguintes normas:

i) **Norma de Frobenius**

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (1.29)$$

ii) **Norma p**

$$\|A\|_p = \sup_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|_p \quad (1.30)$$

Afim de ilustrar o comportamento da norma do espaço de matriz considere o seguinte exemplo.

Exemplo 1.30. Seja $A_{2 \times 2}$ dada por

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad (1.31)$$

$\mathbf{u} = (x_1, x_2)^T \in \mathbb{R}^2$. Logo $\|A\|_2 = \sup_{\|\mathbf{u}\|=1} \|A\mathbf{u}\|_2$. Observe que se $\|\mathbf{u}\|_2 = 1$ então o vetor \mathbf{u} está sobre a circunferência centrada na origem de raio 1, conforme ilustrado na Figura 6 (a). Ao calcular $A\mathbf{u}$ o que se espera é a ação de A sobre cada vetor nesta circunferência. O que pode ser visto na Figura 6 (b). Logo, neste caso $\|A\|_2$ representa o maior valor da ação de A sobre os vetores do círculo unitário que é representado pelo maior semieixo da cônica.

Neste caso obtemos que $\|A\|_2 = \sqrt{\frac{1}{2}(3 + \sqrt{5})}$.

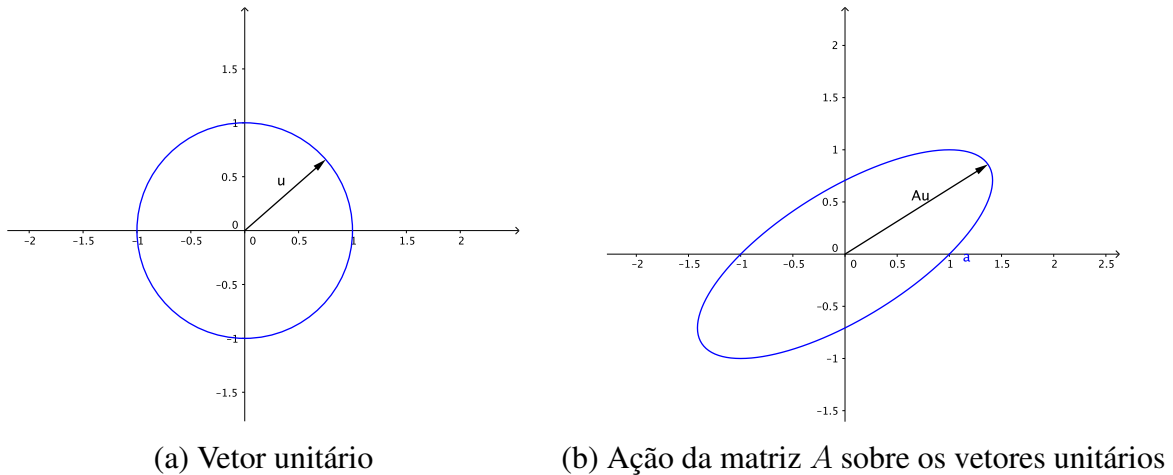


Figura 6 – Ilustração da norma 2 de matrizes

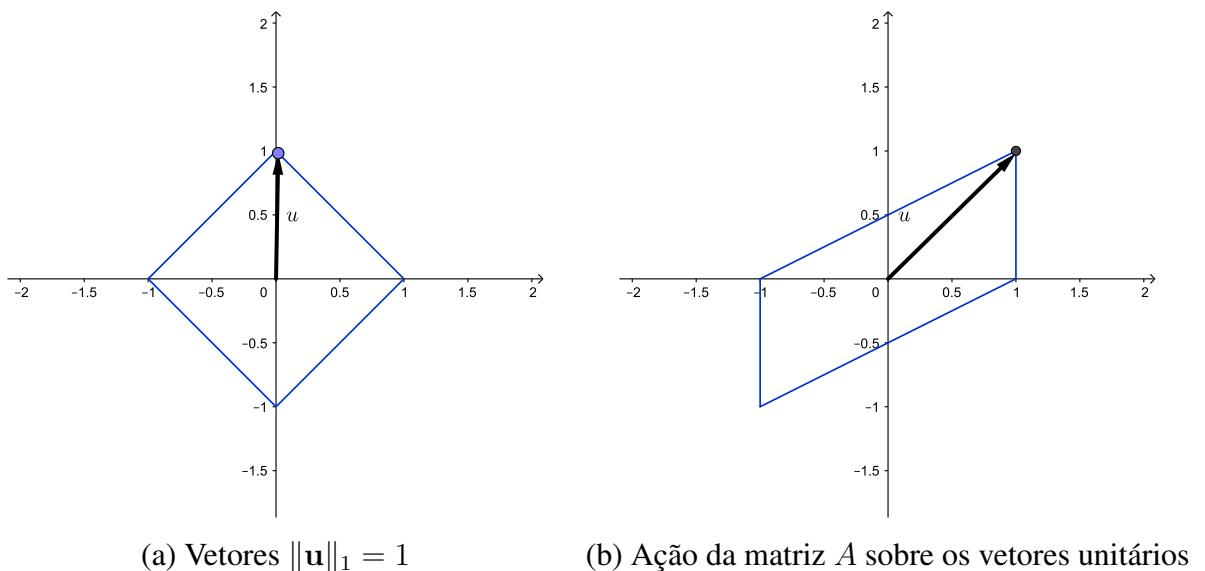


Figura 7 – Ilustração da norma 1 de matrizes

Exemplo 1.31. Considere, ainda a matriz A do exemplo anterior, e vejamos a ilustração da norma 1 da matriz, ou seja,

$$\|A\|_1 = \sup_{\|\mathbf{u}\|_1=1} \|A\mathbf{u}\|_1. \quad (1.32)$$

Observe que se $\|\mathbf{u}\|_1 = 1$ então o vetor \mathbf{u} está definido conforme ilustrado na Figura 7 (a). Ao calcular $A\mathbf{u}$ obtemos um paralelogramo conforme ilustrado na Figura 7 (b). Logo $\|A\|_1$ representa, neste caso, o valor da maior semidiagonal do paralelogramo. Neste caso obtemos que $\|A\mathbf{u}\|_1 = \sqrt{2}$.

1.4 ORTOGONALIDADE

Nesta seção serão apresentados os principais conceitos envolvendo ortogonalidade de vetores, de conjuntos e de subespaços vetoriais. Além disso será definido complemento ortogonal

de um subespaço vetorial.

Definição 1.32. Dados, V um espaço vetorial, \mathbf{u} e \mathbf{v} vetores de V , diz-se que \mathbf{u} e \mathbf{v} são ortogonais se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Definição 1.33. Um conjunto $W = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$ é dito ortogonal quando seus elementos são ortogonais dois a dois, isto é:

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \quad \forall i, j = 1, \dots, n, \quad i \neq j \quad (1.33)$$

Exemplo 1.34. O conjunto $A = \{(1, -1, 0)^T, (1, 1, 0)^T, (0, 0, 1)^T\}$ é um conjunto ortogonal. De fato, basta observar que

$$\begin{aligned} \langle (1, -1, 0)^T, (1, 1, 0)^T \rangle &= 1 \cdot 1 + (-1) \cdot 1 + 0 \cdot 0 = 0 \\ \langle (1, -1, 0)^T, (0, 0, 1)^T \rangle &= 1 \cdot 0 + (-1) \cdot 0 + 0 \cdot 1 = 0 \\ \langle (1, 1, 0)^T, (0, 0, 1)^T \rangle &= 1 \cdot 0 + 1 \cdot 0 + 0 \cdot 1 = 0. \end{aligned}$$

O conceito de ortogonalidade também pode ser estendido para subespaços vetoriais.

Definição 1.35. Sejam, V um espaço vetorial, W_1 e W_2 subespaços de V . W_1 e W_2 são ditos ortogonais se todo vetor $\mathbf{v} \in W_1$ é ortogonal a todo vetor $\mathbf{w} \in W_2$. Neste caso denota-se por $W_1 \perp W_2$.

Observe que se W_1 e W_2 são ortogonais então $W_1 \cap W_2 = \{\mathbf{0}\}$.

Os próximos resultados estabelecem a ortogonalidade entre núcleo e imagem de uma matriz A .

Proposição 1.36. Dada uma matriz $A_{m \times n}$, tem-se que o espaço linha de A é ortogonal ao $N(A)$, ou seja, $L(A) \perp N(A)$.

Demonstração. Seja $\mathbf{u} \in L(A)$ de acordo com (1.8), $\mathbf{u} = \sum_{i=1}^m \alpha_i \bar{\mathbf{a}}_i$ em que $\bar{\mathbf{a}}_i$ denota as linhas da matriz A . Agora dado $\mathbf{v} \in N(A)$ tem-se que $\langle \bar{\mathbf{a}}_i, \mathbf{v} \rangle = 0$ para $i = 1, \dots, m$. Portanto

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \left\langle \sum_{i=1}^m \alpha_i \bar{\mathbf{a}}_i, \mathbf{v} \right\rangle \\ &= \sum_{i=1}^m \alpha_i \langle \bar{\mathbf{a}}_i, \mathbf{v} \rangle \\ &= 0. \end{aligned}$$

□

Proposição 1.37. Dada uma matriz $A_{m \times n}$, tem-se que o espaço coluna de A é ortogonal ao núcleo de A^T , ou seja, $C(A) \perp N(A^T)$.

Demonstração. Seja $\mathbf{u} \in C(A)$ de acordo com (1.5), $\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{a}_i$ em que \mathbf{a}_i denota as colunas da matriz A . Agora dado $\mathbf{v} \in N(A^T)$ tem-se que $\langle \mathbf{a}_i, \mathbf{v} \rangle = 0$ para $i = 1, \dots, n$. Portanto

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \left\langle \sum_{i=1}^n \alpha_i \mathbf{a}_i, \mathbf{v} \right\rangle \\ &= \sum_{i=1}^n \alpha_i \langle \mathbf{a}_i, \mathbf{v} \rangle \\ &= 0. \end{aligned}$$

□

Exemplo 1.38. Considere a matriz A definida por:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

O $N(A)$ é dado por todo vetor $\mathbf{x} = (x_1, x_2)^T$ tal que

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Assim tem-se que $N(A)$ é todo vetor da forma $\mathbf{x} = (-2t, t)^T$, com $t \in \mathbb{R}$. Além disso note que os vetores linha de A são ortogonais ao núcleo. De fato

$$\begin{aligned} \langle (1, 2), (-2t, t) \rangle &= 1 \cdot (-2t) + 2 \cdot t = 0, \\ \langle (2, 4), (-2t, t) \rangle &= 2 \cdot (-2t) + 4 \cdot t = 0. \end{aligned}$$

Assim vemos que o espaço linha de A é ortogonal a $N(A)$.

Exemplo 1.39. Dada a matriz A definida por:

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}. \tag{1.34}$$

O $N(A^T)$ é todo vetor da forma $\mathbf{x} = (x_1, x_2)^T$

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Assim temos que o $N(A^T)$ é formado por todo vetor da forma $(-2x_2, x_2)^T$, com $x_2 \in \mathbb{R}$. Além disso note que os vetores coluna de A são perpendiculares ao $N(A^T)$. De fato:

$$\begin{aligned} \langle (1, 2), (-2x_2, x_2) \rangle &= 1 \cdot (-2x_2) + 2 \cdot x_2 = 0, \\ \langle (3, 6), (-2x_2, x_2) \rangle &= 3 \cdot (-2x_2) + 6 \cdot x_2 = 0. \end{aligned}$$

Assim vemos que o espaço coluna de A é ortogonal a $N(A^T)$.

Definição 1.40. Sejam V um espaço vetorial e S um subespaço de V . O complemento ortogonal de S , denotado por S^\perp é definido por

$$S^\perp = \{ \mathbf{u} \in V \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0, \forall \mathbf{v} \in S \}. \quad (1.35)$$

Observe que S^\perp também é um subespaço de V . De fato, seja $\mathbf{u}_1, \mathbf{u}_2 \in S^\perp$, ou seja, $\langle \mathbf{u}_1, \mathbf{v} \rangle = 0$ e $\langle \mathbf{u}_2, \mathbf{v} \rangle = 0$ para $\mathbf{v} \in S$, assim tem-se que

$$\begin{aligned} \langle \mathbf{u}_1 + \alpha \mathbf{u}_2, \mathbf{v} \rangle &= \langle \mathbf{u}_1, \mathbf{v} \rangle + \alpha \langle \mathbf{u}_2, \mathbf{v} \rangle \\ &= 0. \end{aligned}$$

Exemplo 1.41. Seja $S = \{ \mathbf{v} \in \mathbb{R}^3 \mid \mathbf{v} = (3\alpha, \alpha, 0) \}$ com $\alpha \in \mathbb{R}$. O complemento ortogonal de S é o conjunto formado por todos os vetores $\mathbf{u} = (x, y, z)$ tal que $\langle (x, y, z), (3\alpha, \alpha, 0) \rangle = 0$. Assim $S^\perp = \{ (x, -3x, z) \mid x, z \in \mathbb{R} \}$

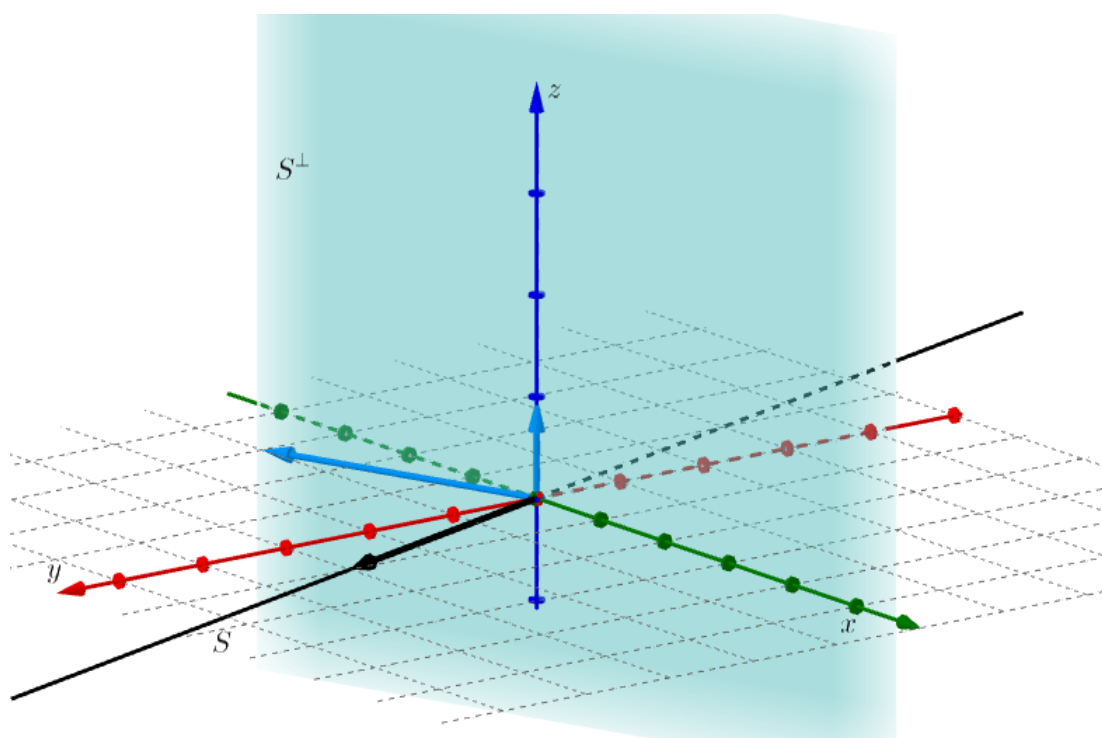


Figura 8 – Representação dos conjuntos S e S^\perp

2 SISTEMAS SOBREDETERMINADOS

Um sistema de equações lineares é um conjunto finito de equações lineares aplicadas a um conjunto, também finito, de variáveis. O objetivo deste capítulo é explorar um método para determinar uma solução para sistemas lineares sobredeterminados, ou seja, sistemas lineares em que o número de equações é maior que o número de incógnitas. Para isso serão apresentadas as principais definições relacionadas a projeção ortogonal e de matriz de projeção ortogonal para, por fim, aplicar na resolução de sistemas sobredeterminados através do método dos mínimos quadrados.

2.1 PROJEÇÃO ORTOGONAL

Nesta seção serão apresentadas as ideias que envolvem projeção de um vetor sobre outro vetor e projeção de um vetor sobre um subespaço de \mathbb{R}^n . Posteriormente serão associados estes resultados com matrizes de projeção, apresentando algumas propriedades importantes.

A partir deste momento, por simplicidade, será utilizada norma 2 e o produto interno usual de \mathbb{R}^n , salvo se menção em contrário.

2.1.1 PROJEÇÃO DE UM VETOR SOBRE OUTRO VETOR

Projetar um vetor u sobre um vetor v é determinar um vetor p pertencente à reta que tem como vetor diretor v através de alguma relação, e é esta relação na qual temos interesse.

Afim de ilustrar geometricamente a construção de uma projeção considere o seguinte problema: Dados os vetores u e v em \mathbb{R}^2 , encontrar um vetor $p = \alpha v$ com $\alpha \in \mathbb{R}$ tal que $u - p$ seja ortogonal a v , como ilustra a Figura 9.

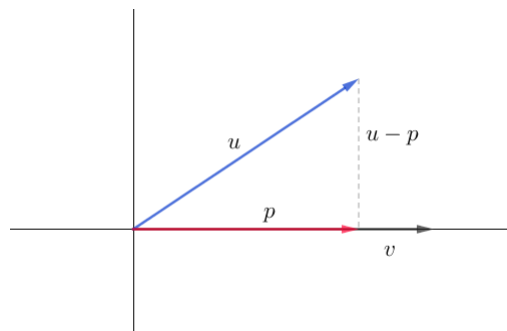


Figura 9 – Projeção do vetor u sobre o vetor v

Para determinar o valor de α observe que $u - p$ é ortogonal a v , desta forma tem-se que:

$$0 = \langle \mathbf{u} - \mathbf{p}, \mathbf{v} \rangle \quad (2.1)$$

$$= \langle \mathbf{u}, \mathbf{v} \rangle - \alpha \langle \mathbf{v}, \mathbf{v} \rangle. \quad (2.2)$$

Como $\mathbf{v} \neq \mathbf{0}$ segue que $\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$ e portanto a projeção de \mathbf{u} sobre \mathbf{v} é dada por

$$\mathbf{p} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}. \quad (2.3)$$

Exemplo 2.1. Dados $\mathbf{u} = (2, 5)$ e $\mathbf{v} = (6, 4)$, a projeção ortogonal \mathbf{p} de \mathbf{u} sobre \mathbf{v} é dada por:

$$\mathbf{p} = \frac{\langle (2, 5), (6, 4) \rangle}{\langle (6, 4), (6, 4) \rangle} = \frac{8}{13} (6, 4). \quad (2.4)$$

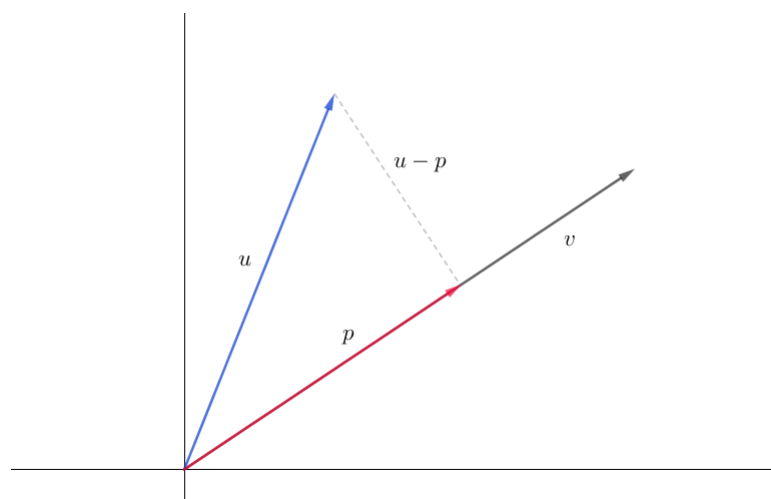


Figura 10 – Projeção de \mathbf{u} sobre \mathbf{v}

2.1.2 PROJEÇÃO DE UM VETOR SOBRE UM SUBESPAÇO

O objetivo desta Seção é estabelecer a projeção de um vetor sobre um subespaço de forma similar o que já foi descrito anteriormente. Para ilustrar esta situação considere o seguinte problema: Dados o vetor $\mathbf{u} = (1, 1, 1)^T$ e S um subespaço gerado pelos vetores $\mathbf{v}_1 = (1, 2, 0)^T$ e $\mathbf{v}_2 = (3, -1, 1)^T$ encontrar a projeção ortogonal \mathbf{p} de \mathbf{u} sobre S .

Agora, como $\mathbf{p} \in S$ então \mathbf{p} pode ser escrito como

$$\mathbf{p} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \quad (2.5)$$

com $\alpha_i \in \mathbb{R}$, $i = 1, 2$.

Observe que, para determinar a projeção \mathbf{p} basta determinar os valores de $\alpha_i \in \mathbb{R}$, $i = 1, 2$. Para isso, note que $\mathbf{u} - \mathbf{p}$ é ortogonal a S , então segue que

$$\begin{cases} \langle \mathbf{u} - \mathbf{p}, \mathbf{v}_1 \rangle = 0 \\ \langle \mathbf{u} - \mathbf{p}, \mathbf{v}_2 \rangle = 0. \end{cases} \quad (2.6)$$

Substituindo (2.5) no sistema (2.6) tem-se

$$\begin{cases} \langle \mathbf{u}, \mathbf{v}_1 \rangle - \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle - \alpha_2 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 0 \\ \langle \mathbf{u}, \mathbf{v}_2 \rangle - \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle - \alpha_2 \langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 0. \end{cases} \quad (2.7)$$

Tal sistema pode ser escrito na seguinte forma matricial

$$\begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_1 \rangle \\ \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ \langle \mathbf{u}, \mathbf{v}_2 \rangle \end{pmatrix}. \quad (2.8)$$

Pode-se provar que como os vetores \mathbf{v}_1 e \mathbf{v}_2 são linearmente independentes a matriz A dada pelo lado esquerdo de (2.8) é não singular portanto, o sistema (2.8) tem única solução e é dada por

$$\mathbf{p} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ \langle \mathbf{u}, \mathbf{v}_2 \rangle \end{pmatrix}. \quad (2.9)$$

Aplicando os valores de \mathbf{u} , \mathbf{v}_1 e \mathbf{v}_2 obtêm-se

$$\mathbf{p} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 1 & 11 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 3 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 11 \\ 8 \\ 2 \end{pmatrix}. \quad (2.10)$$

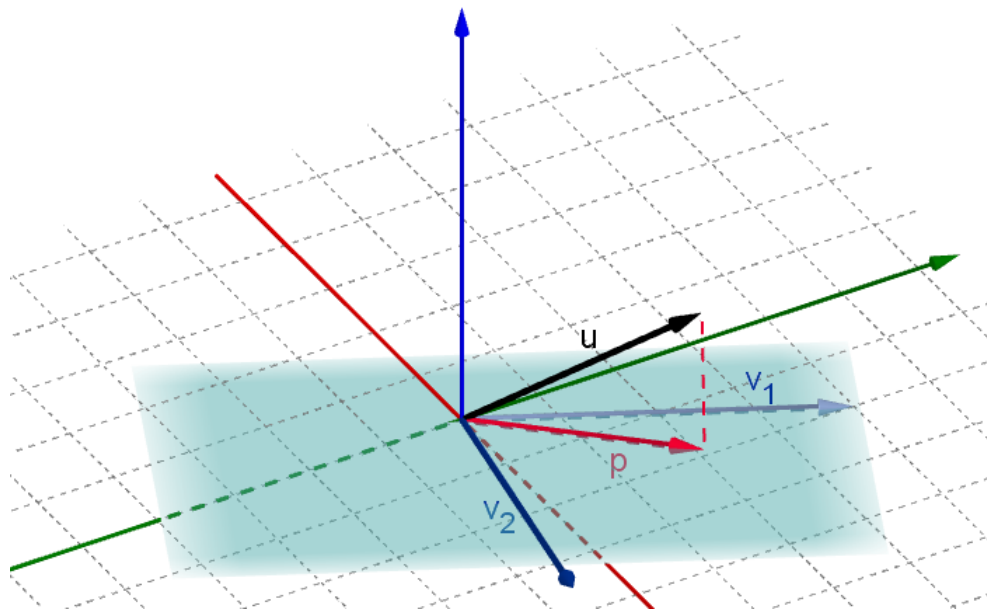


Figura 11 – Projeção do vetor \mathbf{u} sobre S

Note que se adotarmos uma base ortonormal para S , o cálculo da projeção do vetor \mathbf{u} sobre S se torna mais simples. De fato, se $\{\mathbf{v}_1, \mathbf{v}_2\}$ é ortonormal, tem-se que $A = I$ e portanto a projeção (2.9) fica da forma:

$$\mathbf{p} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ \langle \mathbf{u}, \mathbf{v}_2 \rangle \end{pmatrix}, \quad (2.11)$$

ou seja,

$$\mathbf{p} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2. \quad (2.12)$$

Exemplo 2.2. Seja $\mathbf{u} = (1, 1, 1)^T$ e S o subspaço gerado por $\mathbf{v}_1 = \left(\frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5}, 0\right)^T$ e $\mathbf{v}_2 = \left(-\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5}, 0\right)^T$ (note que $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 0$, $\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 1$ e $\langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 1$). De acordo com (2.12) a projeção de \mathbf{u} sobre S é dado por:

$$\mathbf{p} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \begin{pmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \\ 0 \end{pmatrix} + \langle \mathbf{u}, \mathbf{v}_2 \rangle \begin{pmatrix} -\frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \\ 0 \end{pmatrix} \quad (2.13)$$

$$= \frac{3\sqrt{5}}{5} \begin{pmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \\ 0 \end{pmatrix} + \frac{-\sqrt{5}}{5} \begin{pmatrix} -\frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \\ 0 \end{pmatrix} \quad (2.14)$$

$$= \begin{pmatrix} \frac{11}{5} \\ \frac{2}{5} \\ 0 \end{pmatrix}. \quad (2.15)$$

Baseado nos exemplos apresentados, vamos estabelecer a projeção de um vetor $\mathbf{u} \in \mathbb{R}^n$ sobre um subspaço n -dimensional S gerado por $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, vetores linearmente independentes. Neste caso a projeção ortogonal é dada por,

$$\mathbf{p} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n. \quad (2.16)$$

Como o vetor $\mathbf{u} - \mathbf{p}$ é ortogonal a S , então:

$$\langle \mathbf{u} - \mathbf{p}, \mathbf{v}_i \rangle = 0 \quad \forall i = 1, \dots, n, \quad (2.17)$$

ou ainda

$$\begin{cases} \langle \mathbf{u} - (\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n), \mathbf{v}_1 \rangle = 0 \\ \vdots \\ \langle \mathbf{u} - (\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n), \mathbf{v}_n \rangle = 0. \end{cases} \quad (2.18)$$

O sistema (2.18) também pode ser escrito na forma matricial.

$$\begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \dots & \langle \mathbf{v}_n, \mathbf{v}_1 \rangle \\ \vdots & & \vdots \\ \langle \mathbf{v}_1, \mathbf{v}_n \rangle & \dots & \langle \mathbf{v}_n, \mathbf{v}_n \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{v}_n \rangle \end{pmatrix}. \quad (2.19)$$

O sistema (2.19) é chamado de **sistema normal**. É possível provar que a matriz do sistema normal é simétrica e não singular, se as colunas da matriz A são L.I., portanto, neste caso, o sistema admite única solução.

Afim de caracterizar a solução do sistema observe que tomando A a matriz $m \times n$ cujas colunas são formadas pelos vetores \mathbf{v}_i , com $i = 1, \dots, n$ tem-se que:

$$A^T A = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_n, \mathbf{v}_1 \rangle \\ \vdots & & \vdots \\ \langle \mathbf{v}_1, \mathbf{v}_n \rangle & \cdots & \langle \mathbf{v}_n, \mathbf{v}_n \rangle \end{pmatrix} \quad (2.20)$$

e

$$A^T \mathbf{u} = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \mathbf{u} \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{v}_n \rangle \end{pmatrix}. \quad (2.21)$$

Desta forma, tomando $\alpha = (\alpha_1, \dots, \alpha_n)^T$, o sistema (2.28) é equivalente à

$$A^T A \alpha = A^T \mathbf{u}. \quad (2.22)$$

Como os vetores $\mathbf{v}_1, \dots, \mathbf{v}_n$ são L.I. segue que $A^T A$ é não singular e portanto a solução de (2.22) é dada por

$$\alpha = (A^T A)^{-1} A^T \mathbf{u}. \quad (2.23)$$

Com isso tem-se que, a projeção ortogonal \mathbf{p} é da forma

$$\mathbf{p} = A(A^T A)^{-1} A^T \mathbf{u}. \quad (2.24)$$

A matriz $P = A(A^T A)^{-1} A^T$ é chamada de **matriz de projeção ortogonal** e assim a projeção ortogonal toma a forma simplificada $\mathbf{p} = P\mathbf{u}$.

Note que, para fazer a projeção de qualquer vetor $\mathbf{v} \in \mathbb{R}^n$ sobre S basta multiplicar a matriz de projeção P por \mathbf{v} . Além disso, a matriz P satisfaz as seguintes propriedades.

1.

$$\begin{aligned} P^2 &= (A(A^T A)^{-1} A^T)(A(A^T A)^{-1} A^T) \\ &= A(A^T A)^{-1} \underbrace{(A^T A)(A^T A)^{-1}}_I A^T \\ &= A(A^T A)^{-1} A^T \\ &= P. \end{aligned}$$

2.

$$\begin{aligned}
P^T &= (A(A^T A)^{-1} A^T)^T \\
&= (A^T)^T (A(A^T A)^{-1})^T \\
&= A((A^T A)^T)^{-1} A^T \quad (A^T A \text{ é simétrica}) \\
&= A(A^T A)^{-1} A^T \\
&= P.
\end{aligned}$$

Estas duas propriedades caracterizam a projeção ortogonal.

O próximo exemplo tem por objetivo caracterizar as projeções ortogonais em \mathbb{R}^2 , para isso tomaremos uma matriz simétrica A em \mathbb{R}^2 e estabeleceremos condições para que $A^2 = A$.

Exemplo 2.3. Seja $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$, logo $A^2 = A$ se e somente se

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad (2.25)$$

ou seja,

$$a_{11}^2 + a_{12}^2 = a_{11} \quad (2.26)$$

$$a_{12} \cdot a_{11} + a_{22} \cdot a_{12} = a_{12} \quad (2.27)$$

$$a_{12}^2 + a_{22}^2 = a_{22}. \quad (2.28)$$

De (2.27) tem-se que $a_{12}(a_{11} + a_{22} - 1) = 0$ logo $a_{12} = 0$ ou $a_{11} + a_{22} = 1$.

1. Se $a_{12} = 0$ substituindo em (2.26) tem-se que $a_{11} = 0$ ou $a_{11} = 1$. Do mesmo modo substituindo em (2.28) tem-se que $a_{22} = 0$ ou $a_{22} = 1$. Com isso obtemos as seguintes matrizes de projeção ortogonal:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.29)$$

2. Se $a_{11} + a_{22} = 1$ substituindo em (2.26) e (2.28) tem-se que $a_{22}^2 - a_{22} + a_{12}^2 = 0$, logo

$$a_{22} = \frac{1 \pm \sqrt{1 - 4 \cdot a_{12}^2}}{2}. \quad (2.30)$$

Observe que (2.30) apenas faz sentido somente se

$$-\frac{1}{2} \leq a_{12} \leq \frac{1}{2}. \quad (2.31)$$

Portanto se

$$a) \text{ Se } a_{22} = \frac{1 + \sqrt{1 - 4 \cdot a_{12}^2}}{2} \text{ então}$$

$$a_{11} = \frac{1 - \sqrt{1 - 4 \cdot a_{12}^2}}{2}; \quad (2.32)$$

$$b) \text{ Se } a_{22} = \frac{1 - \sqrt{1 - 4 \cdot a_{12}^2}}{2} \text{ então}$$

$$a_{11} = \frac{1 + \sqrt{1 - 4 \cdot a_{12}^2}}{2}. \quad (2.33)$$

Com isso tem-se que em \mathbb{R}^2 uma matriz de projeção ortogonal devem ter seus coeficientes satisfazendo (2.30), (2.31), (2.32) e (2.33).

As matrizes abaixo são exemplos de matrizes de projeção ortogonal no caso em que $a_{12} = \frac{1}{2}$ e $a_{12} = -\frac{1}{2}$ respectivamente.

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \quad (2.34)$$

Dentre os assuntos abordados neste capítulo, a projeção ortogonal é de fundamental importância para este trabalho. Na próxima Seção aplicaremos o estudo sobre projeção ortogonal na resolução de sistemas sobredeterminados.

2.2 MÍNIMOS QUADRADOS

No início da Seção 2.1.2 vimos um exemplo em que o objetivo era encontrar a projeção ortogonal de um vetor \mathbf{v} sobre um subespaço S . Em termos de sistemas lineares o problema é equivalente a encontrar $\mathbf{x} \in \mathbb{R}^2$ tal que

$$\underbrace{\begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 1 \end{pmatrix}}_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_b. \quad (2.35)$$

Observe que $\rho(A) = \dim(C(A)) = 2$, logo não é possível encontrar $\mathbf{x} \in \mathbb{R}^2$ tal que $\mathbf{b} \in \mathbb{R}^3$ pertença a $C(A)$. Este tipo de situação caracteriza os sistemas sobredeterminados. Neste caso o que se pretende é encontrar $\mathbf{x}^* \in C(A)$ de forma que $A\mathbf{x}^*$ esteja o mais próximo de \mathbf{b} possível, ou seja, resolver o problema

$$\min_{\mathbf{x} \in C(A)} \|A\mathbf{x} - \mathbf{b}\|. \quad (2.36)$$

Vimos na Seção 2.1.2 que a projeção ortogonal de \mathbf{b} sobre $C(A)$ resolve o problema (2.35). O próximo teorema garante que a projeção ortogonal resolve o problema (2.36).

Teorema 2.4. (Teorema da Aproximação) A projeção ortogonal $P\mathbf{b} \in C(A)$ é a mais próxima de \mathbf{b} do que qualquer outro elemento de $C(A)$.

Demonstração. Seja $P\mathbf{b}$ a projeção ortogonal de \mathbf{b} sobre $C(A)$ e $\mathbf{y} \in C(A)$ um elemento qualquer. Logo, como $P\mathbf{b} - \mathbf{y} \in C(A)$ e $\mathbf{b} - P\mathbf{b} \perp C(A)$

$$\begin{aligned}\|\mathbf{b} - \mathbf{y}\|^2 &= \|\mathbf{b} - P\mathbf{b} + P\mathbf{b} - \mathbf{y}\|^2 \\ &= \|\mathbf{b} - P\mathbf{b}\|^2 + \|P\mathbf{b} - \mathbf{y}\|^2 \\ &\geq \|\mathbf{b} - P\mathbf{b}\|^2.\end{aligned}$$

Portanto $P\mathbf{b}$ resolve (2.36). □

Como a projeção ortogonal resolve o problema de minimização e na Seção 2.1.2 resolvemos o problema (2.35) por meio das equações normais é de se esperar que para resolver (2.36) basta resolvermos um sistema de equações normais, é o que afirma o próximo resultado.

Teorema 2.5. Seja A uma matriz $m \times n$. Então \mathbf{x}^* resolve (2.36) se e somente se \mathbf{x}^* resolve

$$A^T A \mathbf{x} = A^T \mathbf{b}. \quad (2.37)$$

Demonstração. (\Rightarrow) Seja \mathbf{x}^* solução de (2.36), como $A\mathbf{x}^* \in C(A)$ e $\mathbf{b} - A\mathbf{x}^* \perp C(A)$ segue que $\mathbf{b} - A\mathbf{x}^* \in C(A)^\perp = N(A^T)$, logo $A^T(\mathbf{b} - A\mathbf{x}^*) = \mathbf{0}$, ou seja, \mathbf{x}^* resolve (2.37).

(\Leftarrow) Seja \mathbf{x}^* a solução de (2.37), se mostrarmos que $A\mathbf{x}^*$ é a projeção ortogonal de \mathbf{b} sobre $C(A)$ segue do Teorema 2.4 que \mathbf{x}^* resolve (2.36). Para isso defina $r(\mathbf{x}) = \mathbf{b} - A\mathbf{x}$, logo

$$A^T r(\mathbf{x}^*) = A^T(\mathbf{b} - A\mathbf{x}^*) = \mathbf{0},$$

ou seja, $\mathbf{b} - A\mathbf{x}^* \in N(A^T) = C(A)^\perp$. Como $A\mathbf{x}^* \in C(A)$ e $\mathbf{b} - A\mathbf{x}^* \perp C(A)$ segue que $A\mathbf{x}^*$ é a projeção ortogonal de \mathbf{b} sobre $C(A)$ e portanto, pelo Teorema 2.4, resolve (2.36). □

Para garantir que (2.36) tenha uma única solução basta garantir que a matriz $A^T A$ seja não singular.

Teorema 2.6. Se A é uma matriz $m \times n$ com $m > n$ e o posto de A é n então $A^T A$ é não singular.

Demonstração. Seja A uma matriz $m \times n$ com $m > n$ de posto n , logo $N(A) = \mathbf{0}$, mostremos que $N(A) = N(A^T A)$. De fato, seja $\mathbf{x} \in N(A)$, logo $A\mathbf{x} = \mathbf{0}$ e assim $A^T A\mathbf{x} = \mathbf{0}$, ou seja, $\mathbf{x} \in N(A^T A)$.

Agora, dado $\mathbf{x} \in N(A^T A)$ tem-se que $A^T(A\mathbf{x}) = \mathbf{0}$, logo $A\mathbf{x} \in N(A^T) = C(A)^\perp$. Com isso temos que $A\mathbf{x} \in C(A)^\perp \cap C(A) = \{\mathbf{0}\}$, portanto $\mathbf{x} \in N(A)$. □

Baseado nos resultados desta seção temos que dado um sistema $Ax = \mathbf{b}$ com $A_{m \times n}$ e $m > n$ se posto de A é n o problema normal (2.37) tem única solução e é dada por

$$\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b}. \quad (2.38)$$

O próximo exemplo ilustra os resultados apresentados ao longo desta Seção.

Exemplo 2.7. *Considere o seguinte sistema*

$$\begin{pmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 10 \end{pmatrix}. \quad (2.39)$$

Denotando por A a matriz do sistema (2.39), temos que $\rho(A) = 3$, e portanto o sistema normal

$$\begin{pmatrix} 34 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 44 \\ 20 \\ 14 \end{pmatrix} \quad (2.40)$$

possui única solução e é dado por

$$\begin{aligned} \mathbf{x}^* &= \begin{pmatrix} 34 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 44 \\ 20 \\ 14 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}. \end{aligned}$$

Neste caso $\|Ax^* - \mathbf{b}\| = 3.1622777$.

Vimos ao longo deste capítulo que o método dos mínimos quadrados é uma importante ferramenta para resolver sistemas sobredeterminados. Foram estabelecidos resultados que garantem que, dado um sistema $Ax = \mathbf{b}$ com $A_{m \times n}$ e $m > n$, a projeção ortogonal sobre o subespaço gerado pelas colunas de A é a melhor aproximação para este problema, no sentido de minimizar $\|Ax - \mathbf{b}\|$.

No próximo capítulo, aplicaremos os resultados aqui apresentados num problema de geoposicionamento, que é modelado por um sistema sobredeterminado.

3 APLICAÇÃO

O objetivo deste capítulo é aplicar os assuntos vistos anteriormente em um problema de geoposicionamento por satélite, onde o receptor observa mais do que quatro satélites. Todos os resultados e conceitos aqui apresentados foram consultados em (AKALA et al., 2011), (RAHEMI et al., 2014), (SAUER, 2012), (MOSAVI et al., 2014) e (OSZCZAK, 2014).

3.1 SISTEMA DE GEOPOSICIONAMENTO POR SATÉLITE

GPS é a sigla para *Global Positioning System*, que significa Sistema de Posicionamento Global. Este é um sistema de posicionamento de satélites usado para determinar a posição de um receptor em qualquer lugar sobre a Terra com uma grande precisão.

O projeto GPS foi projetado para uso militar, mas em 1980, por decisão do então presidente dos Estados Unidos Ronald Reagan, foi implementado para uso civil. Ainda assim, foi implantado um erro proposital no sistema GPS de uso civil. Isso se deu devido ao medo do então governo americano de que nações inimigas se utilizassem deste sistema para realizar atentados contra a nação americana.

Enquanto os receptores de uso militar possuem precisão de 1 metro, o de uso civil possuía uma margem de erro de 100 a 140 m devido a um processo de deterioração da precisão das informações dos satélites. Mas este processo foi abolido a 0h do dia 2 de maio de 2000, fazendo com que a precisão do GPS de uso civil obtivesse uma melhora de até dez vezes na sua precisão, fazendo com que a margem de erro variasse de 15 a 100 metros.

A precisão no receptor de uso militar se deve ao fato da existência de um relógio atômico em cada satélite, propiciando uma medição de tempo mais precisa. Este é o sistema de medição de tempo com maior precisão atualmente.

Para o correto funcionamento do GPS, é necessário o uso de três segmentos: espacial, de controle e o receptor. O segmento espacial é composto pelos satélites que orbitam sobre o globo. O de controle são as estações de controle ou de monitoramento dos satélites. E o receptor GPS são os aparelhos que mostram a posição do usuário naquele instante.

3.2 DETERMINANDO A DISTÂNCIA ENTRE O SATÉLITE E O RECEPTOR

Para o cálculo da distância entre o satélite e o receptor, o satélite transmite um longo sinal digital e, neste mesmo instante, o receptor gera o mesmo código. No momento em que o sinal chega ao receptor, existe uma defasagem em relação ao sinal gerado pelo receptor. Esta

diferença é igual ao tempo de trânsito do sinal. Então o satélite calcula a distância utilizando o fato de que:

$$d = ct, \quad (3.1)$$

onde d é a distância do satélite até o receptor, c é a velocidade da luz no vácuo que é dada por $c \approx 299792.458 \text{ km/s}$ e t é o tempo que este sinal demora para transitar entre o satélite e o receptor. A medida desta distância é chamada de *pseudo-distância*. Por conta de possíveis erros na medição do tempo medido pelos receptores, a pseudo-distância não representa a distância real entre o satélite e o receptor. Para minimizar os erros, os relógios dos receptores são constantemente reiniciados e sincronizados com o relógio atômico dos satélites.

As informações utilizadas para determinar a posição do receptor são as posições dos satélites e as distâncias dos satélite até o receptor. Se o receptor capta o sinal de apenas um satélite percebe-se que todas as possíveis posições para o receptor estão na superfície de uma esfera de centro (a, b, c) , onde a, b, c são as coordenadas cartesianas da posição do satélite, e raio d , ou seja, as coordenadas do receptor deve satisfazer:

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = (ct)^2. \quad (3.2)$$

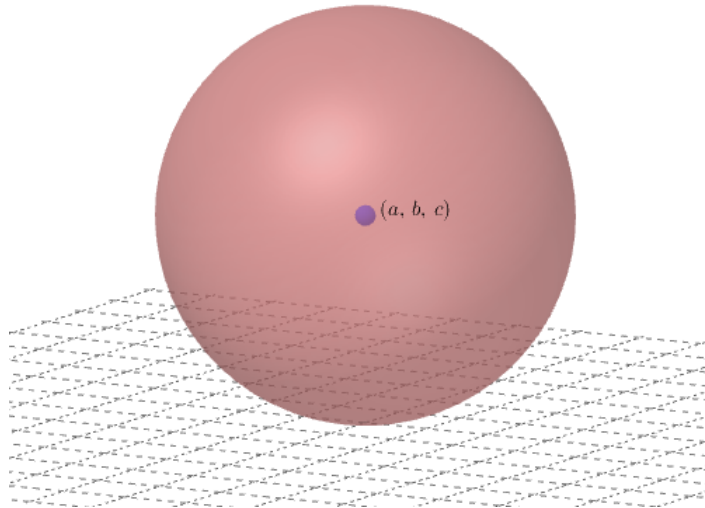


Figura 12 – Representação da superfície esférica de centro (a, b, c)

Se o receptor capta o sinal de dois satélites tem-se que a posição do receptor está contida na interseção das duas esferas cujos centros são os satélites. Esta interseção é não vazia, já que contém pelo menos um ponto em comum, que é a posição do receptor. Caso as esferas sejam tangentes então o ponto de interseção é a posição do receptor (Figura 13 (a)), caso sejam secantes a posição do receptor esta contido em uma circunferência de possibilidades (Figura 13 (b)).

Se o receptor capta o sinal de três satélites a posição do receptor está contida na interseção de três esferas. Se considerar que as três esferas são secantes duas a duas, sua interseção será

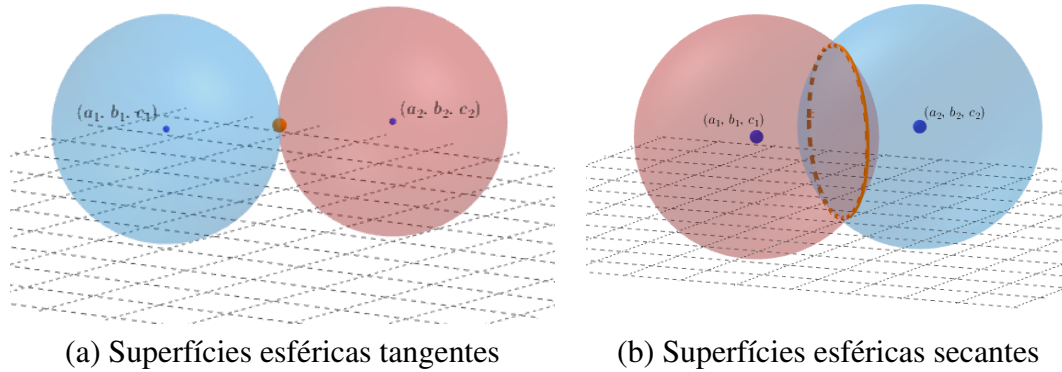


Figura 13 – Representação da interseção de duas superfícies esféricas de centros (a_1, b_1, c_1) e (a_2, b_2, c_2)

o conjunto de dois pontos (Figura 14). Caso duas sejam tangentes, a interseção será um único ponto, que neste caso será a posição do receptor.

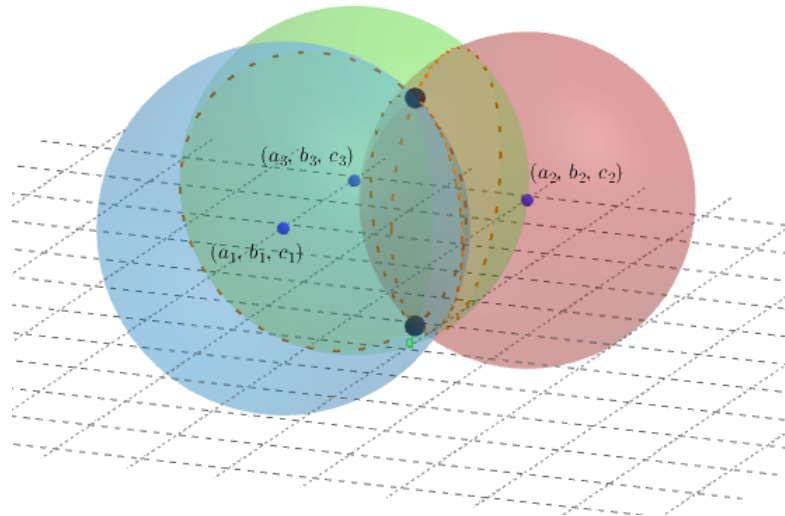


Figura 14 – Representação da interseção de três superfícies esféricas de centros (a_1, b_1, c_1) , (a_2, b_2, c_2) , e (a_3, b_3, c_3)

Considere portanto um sistema de informações coletados de n satélites de posições $S_i = (x_i, y_i, z_i)$ e $d_i = ct_i$ com $i = 1, \dots, n$. Neste caso o sistema de equações associado é dado por,

$$\begin{cases} (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = (ct_1)^2 \\ (x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = (ct_2)^2 \\ \vdots = \vdots \\ (x - x_n)^2 + (y - y_n)^2 + (z - z_n)^2 = (ct_n)^2. \end{cases} \quad (3.3)$$

Sendo assim para determinar a posição de um receptor que recebe o sinal de n satélites devemos resolver o seguinte sistema:

$$\begin{cases} x^2 + y^2 + z^2 - 2xx_1 - 2yy_1 - 2zz_1 + x_1^2 + y_1^2 + z_1^2 - (ct_1)^2 = 0 \\ x^2 + y^2 + z^2 - 2xx_2 - 2yy_2 - 2zz_2 + x_2^2 + y_2^2 + z_2^2 - (ct_2)^2 = 0 \\ \vdots \\ x^2 + y^2 + z^2 - 2xx_n - 2yy_n - 2zz_n + x_n^2 + y_n^2 + z_n^2 - (ct_n)^2 = 0. \end{cases} \quad (3.4)$$

Afim de transformar o sistema (3.4) em um sistema linear, subtraímos a primeira equação das $n - 1$ equações restantes dando origem a um sistema de $(n - 1) \times 3$.

$$\begin{pmatrix} 2(x_1 - x_2) & 2(y_1 - y_2) & 2(z_1 - z_2) \\ 2(x_1 - x_3) & 2(y_1 - y_3) & 2(z_1 - z_3) \\ \vdots & \vdots & \vdots \\ 2(x_1 - x_n) & 2(y_1 - y_n) & 2(z_1 - z_n) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1^2 - x_2^2 + y_1^2 - y_2^2 + z_1^2 - z_2^2 - c^2(t_1^2 - t_2^2) \\ x_1^2 - x_3^2 + y_1^2 - y_3^2 + z_1^2 - z_3^2 - c^2(t_1^2 - t_3^2) \\ \vdots \\ x_1^2 - x_n^2 + y_1^2 - y_n^2 + z_1^2 - z_n^2 - c^2(t_1^2 - t_n^2) \end{pmatrix}. \quad (3.5)$$

Como o sistema (3.5) possui mais equações que incógnitas resolveremos este sistema usando a teoria desenvolvida no Capítulo 2.

A seguir são apresentados exemplos que ilustram toda a teoria desenvolvida ao longo deste trabalho.

Exemplo 3.1. ((SAUER, 2012), pag. 240) Considere um barco que está em uma posição desconhecida que recebe simultaneamente sinais de quatro satélites indicando as suas posições e o tempo como mostrado na tabela abaixo.

Satélite	Posição(km)	Tempo(s)
1	(15600, 7540, 20140)	0,070
2	(18760, 27050, 18610)	0,072
3	(17610, 14630, 13480)	0,076
4	(19170, 610, 18390)	0,073

Tabela 1 – Dados dos satélites

Conhecido o tempo, determina-se a distância pela equação (3.1). Neste caso o sistema linearizado (3.5) é da forma

$$\begin{pmatrix} -6320 & -39020 & 3060 \\ -4020 & -14180 & 13320 \\ -7140 & 13860 & 3500 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8.48121 \times 10^7 \\ 3.0264 \times 10^8 \\ 1.05984 \times 10^8 \end{pmatrix}. \quad (3.6)$$

O sistema normal associado é dado por

$$\begin{pmatrix} 107082400 & 204649600 & -97875600 \\ 204649600 & 1915732400 & -259768800 \\ -97875600 & -259768800 & 199036000 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2.50935 \times 10^{12} \\ -6.13187 \times 10^{12} \\ 4.66164 \times 10^{12} \end{pmatrix}. \quad (3.7)$$

Logo as coordenadas da posição do receptor são $\mathbf{x} = (-3888.88, 159.374, 21716.7)$ e o erro $\|\mathbf{b} - A\mathbf{x}\|_2 = 1.47514 \times 10^{-7}$. A Figura 15 ilustra a interseção das esferas duas a duas e a solução do sistema.

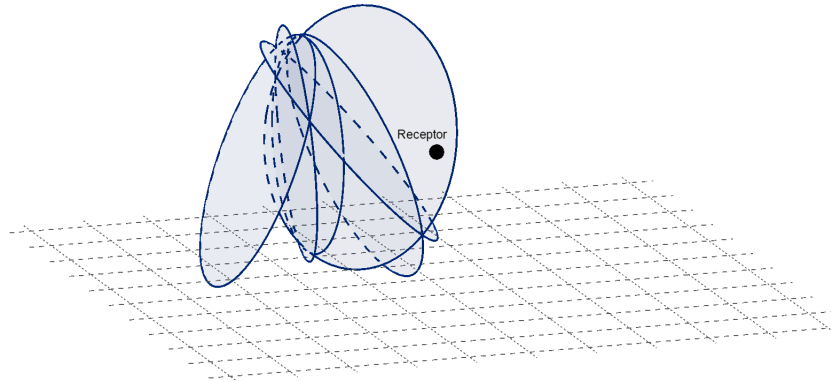


Figura 15 – Localização do receptor com 4 satélites

A seguir são apresentados dois exemplos com um número maior de satélites. Estes exemplos foram extraídos de (OSZCZAK, 2014).

Exemplo 3.2. Considere um sistema de GPS em que o receptor capta o sinal de 5 satélites cujas posições em metros são:

$$\begin{aligned} x_1 &= 28573624.909, & y_1 &= 176258.719, & z_1 &= 475886.493 \\ x_2 &= 20534972.474, & y_2 &= 3620869.695, & z_2 &= 20821515.054 \\ x_3 &= 13834909.426, & y_3 &= 9331764.237, & z_3 &= 24705373.313 \\ x_4 &= -18325015.195, & y_4 &= 12831313.778, & z_4 &= 20831862.073 \\ x_5 &= -11441576.697, & y_5 &= 19817392.158, & z_5 &= 15998439.113 \end{aligned}$$

e suas respectivas pseudo-distâncias, também em metros são:

$$\begin{aligned} d_1 &= 25573786.094 \\ d_2 &= 23269991.712 \\ d_3 &= 23527045.278 \\ d_4 &= 29205487.559 \\ d_5 &= 26129807.790 \end{aligned}$$

Assim o sistema linearizado (3.5) é da forma

$$\begin{pmatrix} 1.60773 \times 10^7 & -6.88922 \times 10^6 & -4.06913 \times 10^7 \\ 2.94774 \times 10^7 & -1.8311 \times 10^7 & -4.8459 \times 10^7 \\ 9.37973 \times 10^7 & -2.53101 \times 10^7 & -4.0712 \times 10^7 \\ 8.00304 \times 10^7 & -3.92823 \times 10^7 & -3.10451 \times 10^7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -5.45835 \times 10^{14} \\ -7.10626 \times 10^{14} \\ -2.34798 \times 10^{14} \\ -2.26975 \times 10^{14} \end{pmatrix}.$$

Neste caso, o sistema normal associado é dado por

$$\begin{pmatrix} 1.63302 \times 10^{16} & -6.16832 \times 10^{15} & -8.38587 \times 10^{15} \\ -6.16832 \times 10^{15} & 2.56645 \times 10^{15} & 3.41761 \times 10^{15} \\ -8.38587 \times 10^{15} & 3.41761 \times 10^{15} & 6.62531 \times 10^{15} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -6.99113 \times 10^{22} \\ 3.16315 \times 10^{22} \\ 7.32525 \times 10^{22} \end{pmatrix},$$

e portanto as coordenadas da posição do receptor são $(5.83074 \times 10^6, 5.71024 \times 10^6, 1.5491 \times 10^7)$. Neste caso o erro do problema de mínimos quadrados é 4.27462×10^{13} . Note que o erro é consideravelmente grande. Este fato ocorre por se tratar de um exemplo que não necessariamente representa uma situação real.

Exemplo 3.3. Um receptor capta o sinal de 10 satélites de posições

$$\begin{aligned} x_1 &= 28573843.196, & y_1 &= 186705.396, & z_1 &= 458504.029 \\ x_2 &= -13737297.587, & y_2 &= 23793697.380, & z_2 &= 440829.364 \\ x_3 &= 135280.549, & y_3 &= 9472446.041, & z_3 &= 23550389.315 \\ x_4 &= -17629491.025, & y_4 &= 10178391.389, & z_4 &= 20326540.307 \\ x_5 &= 21444538.037, & y_5 &= 9999752.312, & z_5 &= 16543394.085 \\ x_6 &= -8952698.519, & y_6 &= 24597337.024, & z_6 &= 12187985.352 \\ x_7 &= 13576242.929, & y_7 &= 20905580.826, & z_7 &= 11605617.387 \\ x_8 &= 2107612.980, & y_8 &= 24090126.595, & z_8 &= 19555410.293 \\ x_9 &= -10553478.506, & y_9 &= 4921167.847, & z_9 &= 26114803.717 \\ x_{10} &= -290863.203, & y_{10} &= 5550000.536, & z_{10} &= 26104633.518, \end{aligned}$$

e suas respectivas pseudo-distâncias

$$\begin{aligned} d_1 &= 25449152.282, & d_2 &= 28710125.200, & d_3 &= 22512803.080, \\ d_4 &= 27609639.021, & d_5 &= 22920682.547, & d_6 &= 27338791.883, \\ d_7 &= 22881688.771, & d_8 &= 26984600.739, & d_9 &= 25643828.772, & d_{10} &= 21830588.390. \end{aligned}$$

Aplicando os dados acima descritos, o sistema linearizado 9×3 obtido é

$$\begin{pmatrix} 8.46223 \times 10^7 & -4.7214 \times 10^7 & 35349.3 \\ 5.68771 \times 10^7 & -1.85715 \times 10^7 & -4.61838 \times 10^7 \\ 9.24067 \times 10^7 & -1.99834 \times 10^7 & -3.97361 \times 10^7 \\ 1.42586 \times 10^7 & -1.96261 \times 10^7 & -3.21698 \times 10^7 \\ 7.50531 \times 10^7 & -4.88213 \times 10^7 & -2.3459 \times 10^7 \\ 2.99952 \times 10^7 & -4.14378 \times 10^7 & -2.22942 \times 10^7 \\ 5.29325 \times 10^7 & -4.78068 \times 10^7 & -3.81938 \times 10^7 \\ 7.82546 \times 10^7 & -9.46892 \times 10^6 & -5.13126 \times 10^7 \\ 5.77294 \times 10^7 & -1.07266 \times 10^7 & -5.12923 \times 10^7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1.76628 \times 10^{14} \\ -6.95244 \times 10^{14} \\ -2.98325 \times 10^{14} \\ -3.95775 \times 10^{14} \\ -4.85866 \times 10^{13} \\ -2.58568 \times 10^{14} \\ -3.01695 \times 10^{14} \\ -6.71826 \times 10^{14} \\ -8.52326 \times 10^{14} \end{pmatrix},$$

e o sistema normal associado é da forma

$$\begin{pmatrix} 3.79292 \times 10^{16} & -1.5976 \times 10^{16} & -1.8182 \times 10^{16} \\ -1.5976 \times 10^{16} & 9.9494 \times 10^{15} & 7.21257 \times 10^{15} \\ -1.8182 \times 10^{16} & 7.21257 \times 10^{15} & 1.25168 \times 10^{16} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1.86957 \times 10^{23} \\ 6.13151 \times 10^{22} \\ 1.5332 \times 10^{23} \end{pmatrix}.$$

Logo as coordenadas da posição do receptor são

$$(1.60315 \times 10^6, -3.14448 \times 10^6, 1.63898 \times 10^7). \quad (3.8)$$

Antes de finalizar o capítulo é necessário fazer algumas observações. Nos exemplos anteriores não foi considerado que o sistema possa ter imprecisões ao tentar calcular a distância entre o receptor e o satélite. Ao levar esta possibilidade em consideração pode acontecer do receptor não estar contido na superfície da esfera obtida do sinal de um satélite, conforme ilustrado na Figura 16 (a). Este mesmo fato pode ocorrer para mais de um satélite conforme Figura 16 (b). Neste caso quanto mais satélites maior é a precisão obtida.

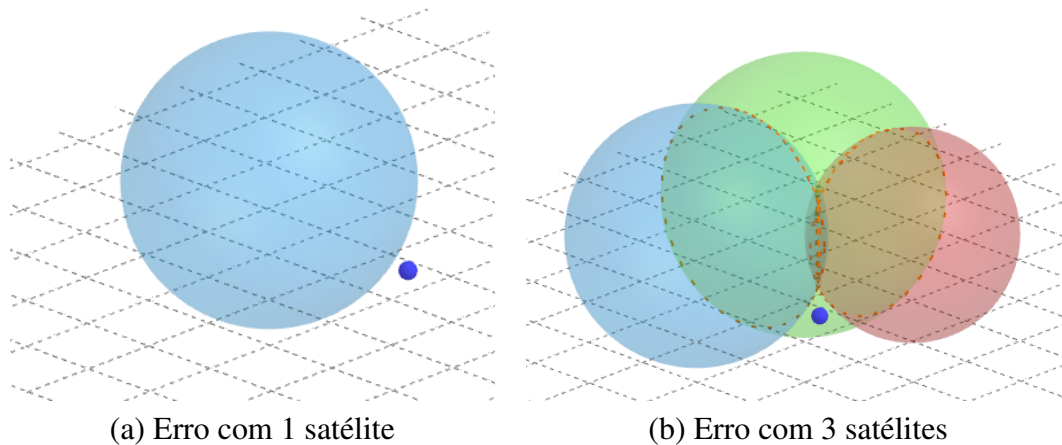


Figura 16 – Representação da posição real do receptor

Alguns fatores podem ocasionar perda de precisão, são eles:

- (i) Atrasos da Ionosfera e Troposfera: A velocidade do sinal não é precisamente a velocidade da luz. O sinal passa por 100 km de Ionosfera e 10 km de Troposfera que possuem propriedades eletromagnéticas que podem afetar a velocidade de transmissão. O atraso do sinal ocorrido nesta parte da atmosfera é compensado por um modelo matemático incluído no receptor, porém por fazer apenas uma média dos atrasos, deixa-se de utilizar dados exatos perdendo assim um pouco da precisão.
- (ii) Sinais “multi-path”: quando o sinal é refletido por construções como prédios e também formações rochosas, o tempo de propagação do sinal é alterado causando imprecisão.

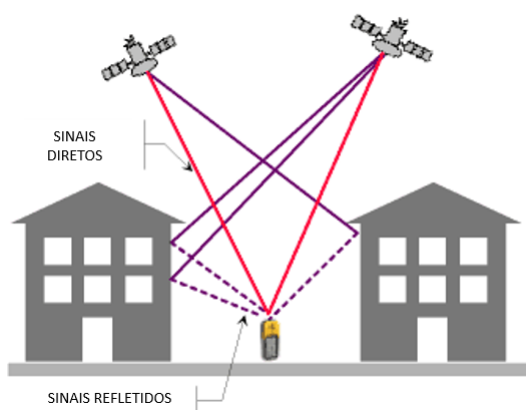


Figura 17 – Representação dos sinais “multi-path”

- (iii) Erros no relógio do receptor: Não é possível ter um relógio atômico em cada receptor, então eles usam um cristal de quartzo comum que não tem a mesma precisão mas é reiniciado em sincronia com quatro ou mais satélites. Mesmo assim podem ocorrer variações no relógio do receptor, causando assim perda de precisão.
- (iv) Números de satélites visíveis: Quanto maior for o número de satélites visíveis maior é a precisão. Não se consegue um bom funcionamento do sistema dentro de locais fechados, embaixo da água ou da terra.
- (v) Geometria dos satélites/sombra: Existe uma geometria entre a posição dos satélites que favorece a decodificação da posição do receptor. Se houver satélites muito próximos uns dos outros ou alinhados, o resultado é uma precisão ruim.

Baseado nisso, existem trabalhos em que a possibilidade de erro é incluída no modelo, como por exemplo (STRANG; BORRE, 1997), (OSZCZAK, 2014). No entanto este tipo de situação não foi abordado neste trabalho por necessitar de outras ferramentas que não foram aqui desenvolvidas.

4 CONCLUSÃO

Os sistemas lineares são empregados na resolução de vários tipos de problemas dentro da Matemática. Um caso particular, que foi o objeto de estudo deste trabalho, são os sistemas lineares sobredeterminados, aqueles em que o número de equações é maior do que o número de incógnitas.

Para resolver este tipo de sistema foi apresentado o método dos mínimos quadrados. Para dar clareza ao método fez-se necessário o estudo de conceitos de álgebra linear como espaços e subespaços vetoriais, base e dimensão de um subespaço vetorial, norma de vetores e de matrizes e por fim, ortogonalidade de vetores, de conjuntos e de subespaços vetoriais.

Posteriormente foi apresentado o método dos mínimos quadrados explorando várias ideias de projeção ortogonal culminando em uma sistemática de resolução de sistemas sobredeterminados.

O trabalho foi finalizado com uma apresentação de um problema de geoposicionamento em que o objetivo era determinar a posição de um receptor em um sistema de coordenadas cartesianas. Este tipo de problema foi modelado por meio de um sistema sobredeterminado de equações lineares, que foi resolvido com o uso de mínimos quadrados.

Para que este trabalho fosse desenvolvido, foi necessário uma série de conhecimentos que foram adquiridos ao longo deste processo tais como: aprimorar a experiência com a pesquisa acadêmica, utilização de editores de texto, como o uso do *Latex*, o uso de *softwares* matemáticos como o Geogebra e Wolfram Mathematica. Isso tudo somado contribuiu para uma formação mais completa.

REFERÊNCIAS

- AKALA, A. O. et al. Determined optimization technique for solving over-determined linear systems. **Latin-American Journal of Physics Education**, 2011.
- GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)**. EUA: Johns Hopkins University Press, 1996. ISBN 0-8018-5414-8.
- LIMA, E. L. **Álgebra Linear**. Rio de Janeiro: IMPA, 2009.
- MEYER, C. D. **Matrix Analysis and Applied Linear Algebra**. EUA: SIAM, 2000.
- MOSAVI, M. R. et al. Least squares techniques for GPS receivers positioning filter using pseudo-range and carrier phase measurements. **Iran University of Science and Technology**, 2014. ISSN 2383-3890.
- OSZCZAK, B. GNSS position algorithms using methods of reference point indications. **Artificial Satellites**, v. 49, n. 1, p. 21–32, 2014.
- RAHEMI, N. et al. Accurate solution of navigation equations in GPS receivers for very high velocities using pseudorange measurements. **Iran University of Science and Technology**, p. 8, 2014. Disponível em: <<http://dx.doi.org/10.1155/2014/435891>>.
- SAUER, T. **Numerical Analysis**. 2nd. edition. ed. Boston: Pearson Education, 2012.
- STRANG, G. **Introduction to Linear Algebra, Fourth Edition**. EUA: Wellesley Cambridge Press, 2009. ISBN 978-0-9802327-1-4.
- STRANG, G.; BORRE, K. **Linear Algebra, geodesy and GPS**. EUA: Wellesley Cambridge Press, 1997.
- TREFETHEN, L. N.; III, D. B. **Numerical Linear Algebra**. Philadelphia: SIAM: Society for Industrial and Applied Mathematics, 1997. ISBN 0-89871-361-7.