UNIVERSIDADE FEDERAL DE SANTA CATARINA
PÓS GRADUAÇÃO EM ENSINO DE MATEMÁTICA

Jocinéia Medeiros

UM BREVE ESTUDO SOBRE MATRIZES ORTOGONAIS E O PROBLEMA DE MÍNIMOS QUADRADOS

Foz do Iguaçu

2010

Jocinéia Medeiros

UM BREVE ESTUDO SOBRE MATRIZES ORTOGONAIS E O PROBLEMA DE MÍNIMOS QUADRADOS

Monografia apresentada ao departamento de Matemática e Física da Universidade Federal de Santa Catarina, como requisito para obtenção do Título de Especialista em Matemática.

Orientador: Fermin S. V. Bazán

Foz do Iguaçu

2010

Jocinéia Medeiros

UM BREVE ESTUDO SOBRE MATRIZES ORTOGONAIS E O PROBLEMA DE MÍNIMOS QUADRADOS

Monografia apresentada ao departamento de Matemática e Física da Universidade Federal de Santa Catarina, como requisito para obtenção do Título de Especialista em Matemática.

Orientador: Fermin S. V. Bazán

Aprovada em: _________ / __________ / __________.

BANCA EXAMINADORA

____

Dr. Fermín S. V. Bazán
(Orientador)

____

Dr. Daniel Norberto kozakevich
(Examinador)

____

Dr. Márcio Rodolfo Fernandes
(Examinador)

Dedico este trabalho de conclusão de curso à minha família, que sempre estiveram ao meu lado em todos os momentos e em especial ao meu irmão Gilson que apesar de não estar dentre nós, permanece vivo em nossos pensamentos.

AGRADECIMENTOS

Agradeço ao Professor Dr. Fermín Bazán pela paciência na orientação, além do incentivo, simpatia, e presteza que tornaram possível a conclusão desta monografia.

À todos os professores pela dedicação, entusiasmo demonstrados durante o curso e que foram muito importantes na minha vida acadêmica.

À Universidade Federal de Santa Catarina por nos proporcionar esta especialização.

À UAB (PTI) pelo espaço Cedido e estrutura excelente.

À Deus, por me proporcionar a vida e me permitir que chegasse até onde estou.

À meus colegas do pólo: João Bonfim, Gilderléia, Maikon, Cristiane, Graziela, Paula, Patrick, Mariana, Valmei, Diego, Ademir e Emanuele pela ajuda durante os grupos de estudos e especialmente ao tutor Gilberto Nunes, pelo incentivo que sempre nos proporcionou durante o curso, nos animando sempre para concluir esta especialização.

À minha família que sempre teve paciência comigo, dedicando carinho e apoio a todo o momento. Não medindo esforços para que eu conseguisse chegar até esta etapa do curso.

E finalmente, ao meu namorado Ailson, pelo incentivo e principalmente pela paciência que sempre teve comigo.

“O principio criador reside na matemática; a sua certeza é absoluta, enquanto se trata de matemática abstrata, mas diminui na razão direta de sua concretização.”

(Albert Einstein)

RESUMO

Neste trabalho são estudados alguns conceitos da Álgebra Linear, para então desenvolver um breve estudo de matrizes ortogonais. Tais matrizes apresentam propriedades de conservar ângulos e módulos, produtos escalares e além disso do ponto de vista numérico, preservam a ordem de grandeza de números durante operações numéricas. No decorrer do trabalho foram contemplados além da fatoração QR, o problema de mínimos quadrados calculados através da equação normal e via fatoração QR. Durante a pesquisa fez-se uma abordagem da reta de melhor ajustes, em que a tal reta é $y=ax+b$ que minimiza a soma dos quadrados dos resíduos.

Palavras-chave: Matrizes ortogonais. Fatoração QR. Método dos Mínimos Quadrados.

LISTA DE FIGURAS

SUMÁRIO

INTRODUÇÃO

Ao se falar em matriz ortogonal, segundo Poole (2006), tal nomenclatura não é muito adequada, um termo melhor seria “matriz ortonormal”, o qual não é utilizado, até porque não tem termo algum para designar uma matriz não quadrada com colunas ortonormais.

Para construir uma base ortogonal (ou ortonormal) para qualquer subespaço de R^n é utilizado um método simples (O Processo de Gram-Schmidt) que nos leva a uma das mais úteis fatorações de Matrizes. Se A é uma matriz mxn com colunas linearmente independentes (exigindo m>n), o processo de Gram-Schmidt produz uma fatoração da matriz A em um produto de uma matriz ortogonal Q por uma matriz triangular superior R, conhecida como a fatoração QR. A fatoração QR tem aplicações em programas de computador para encontrar os autovalores de uma matriz, para resolver sistemas lineares e para encontrar as aproximações por mínimos quadrados,
que discutiremos neste trabalho.

Sistemas impossíveis aparecem muitas vezes em aplicações, embora nem sempre com uma matriz de coeficientes tão grande. Quando o sistema não possui solução, o melhor que podemos fazer é encontrar um x que faça com que Ax fique o mais próximo possível de b. O problema geral de mínimos quadrados é encontrar um x que torne $\|b - Ax\|$ o menor possível. O termo mínimos quadrados vem do fato de que $\|b - Ax\|$ é a raiz quadrada de uma soma de quadrados.

O presente trabalho tem por objetivo fazer um breve estudo à respeito de matrizes ortogonais, visando aplicações em problemas que envolvem o método dos mínimos quadrados. Sua elaboração constitui-se de 4 capítulos e está organizado da seguinte maneira:

No capítulo I são mencionados os conceitos preliminares que serão utilizados nos capítulos seguintes.

No capítulo II apresentamos os conjuntos de propriedades das matrizes ortogonais e dois exemplos típicos de matrizes ortogonais: a matriz reflexão e a matriz rotação.

No capítulo III utilizamos a propriedade do processo de Gram-Schmidt para provar a fatoração QR.

No capítulo IV mostramos o problema de mínimos quadrados, onde o sistema linear $Ax = b$ não possui solução, porque o vetor do lado direito não pertence ao espaço coluna da matriz. Nesses casos, podemos recorrer de b, para isso podemos utilizar as equações normais para x ou fatoração QR.

Finalmente procuramos uma reta de melhor ajuste que se aproxime ao máximo dos pontos dados, ou seja, que possui o mínimo possível de erro.

CAPÍTULO 1

CONCEITOS PRELIMINARES

Neste capítulo apresentaremos alguns conceitos de álgebra linear que serão necessários para a compreensão dos próximos tópicos. Para maiores detalhes em relação aos conceitos indicamos os livros KOLMAN & HILL (2006),LAY(2007), CALLIOLI ET AL (1990) e BOLDRINI ET AL (1980).

1.1 Definição de Matrizes

Uma matriz A_{mxn} é um arranjo retangular de *mn* números reais (ou complexos) distribuídos em *m* linhas horizontais e *n* colunas verticais:

$$A = \begin{array}{c} \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-ésima linha} \\ \\ \\ \end{array} \\ \uparrow j\text{-ésima coluna} \end{array}$$

Figura 1

A i-ésima linha de A é

$$\left[a_{i1}a_{i2}\cdots a\cdots_{in}\right] \quad (1 \le i \le m);$$

e a j-ésima coluna de A é

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad (1 \le j \le n).$$

Dizemos que A é m por n (representado por mxn).

1.2 Operações com Matrizes

1.2.1Adição de Matrizes

Se A=$\left[a_{ij}\right]$e B=$\left[b_{ij}\right]$ são matrizes mxn, então a soma de A e B é a matriz C=$\left[c_{ij}\right]$, mxn, definida por $C_{ij}= a_{ij}+b_{ij}$ $\quad (1 \le i \le m, 1 \le j \le n)$.

Logo, C é obtida pela adição dos elementos correspondentes de A e B.

Se A e B são matrizes mxn, escrevemos A + (-1)B como A – B e chamamos isto de diferença entre A e B.

Observamos que pela maneira com que foi definida, a adição de matrizes tem as mesmas propriedades que a adição de números reais.

Propriedades: Dadas as matrizes A, B e C de mesma ordem mxn, temos:

i) A + B = B + A (comutatividade)
ii) A + (B + C) = (A + B) + C (associatividade)
iii) A + 0 = A, onde 0 denota a matriz nula mxn.

1.2.2 Produtos

1.2.2.1 Multiplicação por um escalar

Se A=$\left[a_{ij}\right]$é uma matriz *mxn* e r é um número real, então a multiplicação por um escalar de A por r, rA, é a matriz B=$\left[b_{ij}\right]$, *mxn*, onde $b_{ij} = ra_{ij}$ $\quad (1 \le i \le m, 1 \le j \le n)$.

Logo, B é obtida pela multiplicação de cada elemento de A por r.

Propriedades: Dadas as matrizes A e B de mesma ordem *mxn* e números r, r_1 e r_2, temos:

i) r(A + B)= rA + rB
ii) $(r_1 + r_2)A = r_1A + r_2A$
iii) 0 . A = 0, isto é se multiplicarmos o número zero por qualquer matriz A, teremos a matriz nula.
iv) $r_1(r_2A) = (r_1r_2)A$

1.2.2.2 Multiplicação de Matrizes

Se A=$\left[a_{ij}\right]$ é uma matriz *mxp* e B=$\left[b_{ij}\right]$ é uma matriz pxn, então o produto de A e B, representado por AB, é a matriz *mxn* C = $\left[c_{ij}\right]$, definida por $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$ = $\sum_{k=1}^{p} a_{ik}b_{kj} \quad (1 \le i \le m, 1 \le j \le n)$.

Propriedades da multiplicação de matrizes

a. Se A, B e C têm os tamanhos apropriados, então A(BC) = (AB)C (propriedade associativa da multiplicação);
b. Se A, B e C têm os tamanhos apropriados, então A(B+C) = AB+AC (propriedade distributiva à esquerda);
c. Se A, B e C têm os tamanhos apropriados, então (A+B)C = AC+BC (propriedade distributiva à direita)

1.3 Alguns Tipos Especiais de Matrizes

1.3.1 Definição de Matriz Quadrada

É aquela cujo número de linhas é igual ao número de colunas ($m = n$). Quando nos referimos a uma matriz quadrada $n \text{ x } n$, podemos dizer que a sua ordem é n ao invés de $n \text{ x } n$.

1.3.2 Definição de Matriz Identidade

É aquela em que $a_{ii} = 1$ e $a_{ij} = 0$, para $i \neq j$. Isto é, é a matriz quadrada de ordem n, em que todos os elementos da diagonal principal são iguais a 1 e os demais elementos são iguais a zero. Representa-se a matriz identidade por I_n.

1.3.3 Definição de Matriz Triangular Superior

É uma matriz quadrada onde todos os elementos abaixo da diagonal são nulos, isto é, se $a_{ij} = 0$, para $i > j$.

1.3.4 Definição de Matriz Simétrica

Uma matriz quadrada A é dita simétrica, se A^T=A equivalentemente, A é simétrica se os elementos simétricos com relação à diagonal principal são iguais, ou seja, aquela onde $a_{ij} = a_{ji}$.

Exemplo:

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

Podemos observar que a parte superior é uma “reflexão” da parte inferior, em relação a diagonal.

1.3.5 Matriz Transposta

Definição: Dada uma matriz *mxn* A, a transposta de A é a matriz *nxm*, denotada por A^T, cujas colunas são formadas com as linhas correspondentes de A. Note que a transposta de um vetor coluna é um vetor linha e vice-versa.

A operação de matriz transposta satisfaz às seguintes propriedades:

Sejam A e B matrizes cujos tipos são apropriados para as seguintes somas e produtos.

a. $(A^T)^T=A$
b. $(A+B)^T=A^T+B^T$
c. Para qualquer escalar r, $(rA)^T= rA^T$
d. $(AB)^T=B^TA^T$

1.4 Produto Interno

Definição: Seja V um espaço vetorial de dimensão finita sobre IR*. Entende-se por produto interno sobre V uma aplicação que transforma cada par ordenado $(u,v)\in VxV$ em um número real, (que indicaremos por $\langle u,v\rangle$) obedecendo às seguinte condições:

a) $\langle u+v,w\rangle=\langle u,w\rangle+\langle v,w\rangle, \forall u,v,w\in V$;

b) $\langle \alpha u,v\rangle=\alpha\langle u,v\rangle\forall\alpha\in IRe\forall u,v\in V$;

c) <u,v> = <v,u>, para todo $u,v\in V$;

d) $\langle u,u\rangle$ é um número real maior que zero para todo vetor $u\neq 0$.

Um espaço vetorial V onde está definido um produto interno é chamado espaço com produto interno (real). Como exemplo de um espaço com produto interno temos o $\mathbb{R}^n$, com $\langle u,v\rangle=x_1y_1+x_2y_2+...+x_ny_n$, onde u = $(x_1, x_2,..., x_n)$ e v = $(y_1, y_2,..., y_n)$.

1.5 Propriedades do produto Escalar

Se u, v, e w são vetores em R^n e c é um escalar, então:

a) $u.u\geq 0; u.u=0$ se, e somente se, $u=0$
b) $u.v=v.u$
c) $(u+v).w=u.w+v.w$
d) $(cu).v=u.(cv)=c(u.v)$

1.6 Norma em $\mathbb{R}^n$

O comprimento (ou norma) de v é o escalar não-negativo $\|v\|$ definido por

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

1.7 Ortogonalidade

Seja V um espaço euclidiano. Dizemos que dois vetores $u, v \in V$ são ortogonais se, e somente se, $\langle u, v\rangle = 0$. Um conjunto S={u_1, ...,u_r}$\subset$V se diz ortonormal se, e somente se, (I) $\|u_i\| = 1$ (i=1, 2,...,r) e (II) dois vetores quaisquer de S, distintos entre si, são ortogonais.

CAPÍTULO 2

MATRIZES ORTOGONAIS

No capítulo II apresentamos o conjunto de propriedades das matrizes ortogonais e dois exemplos típicos: matrizes de reflexão e matrizes de rotação.

Segundo Poole (2006), a nomenclatura Matriz Ortogonal não é muito adequada, pois um termo melhor seria “matriz ortonormal”, o qual não é utilizado, até porque não tem nenhum termo para designar uma matriz não quadrada com colunas ortonormais. Para este capítulo nos baseamos nos trabalhos de POOLE (2006), LIPSCHUTZ (1994), STEINBRUCH (2009).

Definição: Uma matriz real Q é ortogonal se $QQ^T = Q^TQ = I$.

Observe que uma matriz ortogonal Q é necessariamente quadrada e invertível, e sua inversa coincide com a transposta isto é, $Q^{-1} = Q^T$.

Exemplo:

Seja $Q = \frac{1}{9}\begin{pmatrix} 1 & 8 & -4 \\ 4 & -4 & -7 \\ 8 & 1 & 4 \end{pmatrix}$. Então

$$QQ^T = \frac{1}{9}\begin{pmatrix} 1 & 8 & -4 \\ 4 & -4 & -7 \\ 8 & 1 & 4 \end{pmatrix} . \frac{1}{9}\begin{pmatrix} 1 & 4 & 8 \\ 8 & -4 & 1 \\ -4 & -7 & 4 \end{pmatrix} =$$

$$QQ^T = \frac{1}{81}\begin{pmatrix} 1+64+16 & 4-32+28 & 8+8-16 \\ 4-32+28 & 16+16+49 & 32-4-28 \\ 8+8-16 & 32-4-28 & 64+1+16 \end{pmatrix}$$

$$QQ^T = \frac{1}{81}\begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Definição: Se det$|A|$ = 1, a matriz ortogonal A é chamada matriz ortogonal própria e se detA = -1, a matriz ortogonal A é chamada matriz ortogonal imprópria.

Existem matrizes ortogonais que produzem efeitos especiais quando aplicados num vetor. Duas dessas são a Matriz de Reflexão e Matriz de Rotação. As matrizes de transformação que representam rotações do plano sobre a origem são exemplos de matrizes ortogonais próprias e as reflexões do plano com respeito a uma reta pela origem são exemplos de matrizes ortogonais impróprias. Além disso, tanto os vetores linha quanto os vetores coluna de uma matriz de rotação e de uma matriz de reflexão podem ser considerados como um par de vetores unitários ortogonais.

2.1 Matriz Reflexão

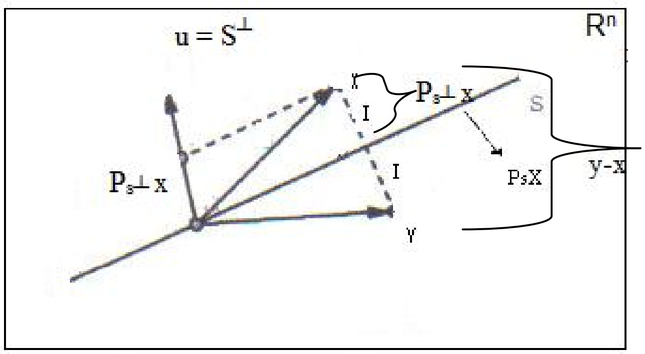

Figura 2

u : unitário

$S = u^T$ (complemento ortogonal)

$S^{\perp} = u$

V= $S \oplus S^{\perp}$

$v \in V$; $v = v_1 + v_2$; $v_1 \in S$ e $v_2 \in S^{\perp}$; $v_1 \perp v_2$

Uma reflexão de Householder é uma matriz da forma H= $I - 2uu^T$.

Temos que

$y - x = -2P_{S^{\perp}}x$

$\Rightarrow y = x - 2uu^T x$

$= (I - 2uu^T)x$

Ao multiplicar uma matriz de Householder por um vetor x é equivalente a refletir o vetor x através do plano perpendicular a u .

As matrizes de Householder são simétricas e ortogonais, como veremos no teorema a seguir:

Teorema: A matriz de Householder é uma Matriz ortogonal.

Demonstração:

Vamos provar que $H^T H = I$ e $HH^T = I$

Observamos que H é simétrica, de fato,

$$H^T = (I - 2uu^T)^T$$

$$= I^T - 2(uu^T)^T$$

$$= I - 2(u^T)^T u^T$$

$$= I - 2uu^T = H$$

Assim, temos

$$H^T H = HH = H^2$$

$$= (I - 2uu^T)^2$$

$$= (I - 2uu^T)(I - 2uu^T)$$

$$= I^2 - 2Iuu^T - 2uu^T I + 4(uu^T)(uu^T)$$

$$= I - 2uu^T - 2uu^T + 4uu^T = \mathrm{I}$$

Exemplo:

Suponha $u^T = \frac{1}{5}(3,4)^T$ calcular a reflexão de Householder.

Seja $H = I - 2uu^T$ então,

$$H = I - 2\left(\frac{3}{5}, \frac{4}{5}\right)\left(\frac{3}{5}, \frac{4}{5}\right)^T$$

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{6}{5} \\ \frac{8}{5} \end{pmatrix}\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \end{pmatrix}$$

$$H=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}\frac{18}{25} & \frac{24}{25}\\ \frac{24}{25} & \frac{32}{25}\end{pmatrix}$$

$$H=\begin{pmatrix}\frac{7}{25} & -\frac{24}{25}\\ -\frac{24}{25} & -\frac{7}{25}\end{pmatrix}\text{logo,}$$

$$H=\frac{1}{25}\begin{pmatrix}7 & -24\\ -24 & -7\end{pmatrix}\ \text{Det (H) = -1}$$

2.2 Matriz Rotação

Em duas dimensões, a rotação pode ser definida por um único ângulo α,θ. Por convenção, ângulos positivos representam rotação no sentido anti-horario.

Rotação é uma operação básica muito usada em aplicações gráficas. Ela faz parte de um grupo maior de transformações chamadas isométricas, que tem a característica de preservarem o comprimento do vetor. Para construir a matriz de rotação considere a figura abaixo:

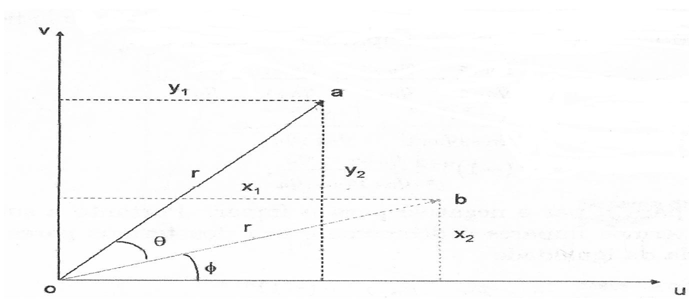

Figura 3

Já sabemos que o tamanho do vetor não se altera. Aqui chamado de r.

$$\cos\phi=\frac{x_1}{r}\longrightarrow x_1=r\cos\phi$$

$$sen\phi=\frac{x_2}{r}\longrightarrow x_2=rsen\phi$$

$$\cos(\phi+\theta)=\frac{y_1}{r}\longrightarrow y_1=r\cos(\phi+\theta)$$

$$sen(\phi+\theta)=\frac{y_2}{r}\longrightarrow y_2=rsen(\phi+\theta)$$

$$y_1 = r\cos(\phi+\theta) \rightarrow y_1 = \overbrace{r\cos\phi}^{x_1}\cos\theta - \overbrace{rsen\phi}^{x_2}\,sen\theta$$

$$y_2 = rsen(\phi+\theta) \rightarrow y_2 = \underbrace{r\cos\phi}_{x_1}\,sen\theta + \underbrace{rsen\phi}_{x_2}\cos\theta \qquad \text{ou,}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -sen\theta \\ sen\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{ou,}$$

Y= R . X

Onde, R é ortogonal.

Para verificar se esta matriz rotação é ortogonal basta multiplicar pela sua transposta, obtendo assim a matriz identidade.

$$\begin{bmatrix} \cos\theta & -sen\theta \\ sen\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & sen\theta \\ -sen\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta + sen^2\theta & 0 \\ 0 & sen^2\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$ e determinante igual a 1.

$D = \begin{bmatrix} \cos\theta & -sen\theta \\ sen\theta & \cos\theta \end{bmatrix}$ Sejam $v = \begin{bmatrix} x \\ y \end{bmatrix}$ e u=D.v= $\begin{bmatrix} x\cos\theta & -ysen\theta \\ xsen\theta & y\cos\theta \end{bmatrix}$

$$\|v\| = \sqrt{x^2+y^2}$$

$$\|u\| = \sqrt{\left(x\cos\theta - ysen\theta\right)^2 + \left(xsen\theta + y\cos\theta\right)^2}$$

$$\|u\| = \sqrt{x^2\cos^2\theta - 2xy\cos\theta sen\theta + y^2sen^2\theta + x^2sen^2\theta + 2xysen\theta\cos\theta + y^2\cos^2\theta}$$

$$\|u\| = \sqrt{x^2(\cos^2\theta + sen^2\theta) + y^2(sen^2\theta + \cos^2\theta)}$$

$\|v\| = \sqrt{x^2+y^2} \rightarrow$ Preserva a norma do vetor.

Exemplo1: qual a rotação do vetor

$\vec{u} = (1,0)$, e $\theta = 30^o$ no sentido anti-horário. $T:\mathbb{R}^2 \rightarrow \mathbb{R}^2$

Multiplicando a matriz rotação pelo vetor coluna obtemos:

$\begin{bmatrix} \cos 30^o & -sen30^o \\ sen30^o & \cos 30^o \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$ que é as coordenadas do vetor que é rotação do vetor $\vec{u} = (1,0)$ de ângulo 30º.

Exemplo 2:

Com $\theta = \dfrac{\pi}{4}$ rad D= $\begin{bmatrix} \cos\theta & -sen\theta \\ sen\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{2}}{2} \\ \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \end{bmatrix}$ v= $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

u= $\begin{bmatrix} \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{2}}{2} \\ \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{\sqrt{2}}{2} \\ \dfrac{\sqrt{2}}{2} \end{bmatrix}$

O vetor u é uma rotação de v, com 45 graus para a direita , têm a mesma norma e as colunas de **D** são ortogonais.

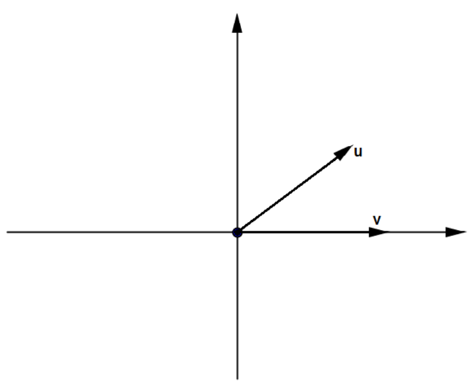

Figura 4

As rotações planas apresentadas anteriormente podem ser generalizadas para dimensões maiores. Para R^3, por exemplo, podemos definir:

$R_1(\theta) \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & sen\theta \\ 0 & -sen\theta & \cos\theta \end{bmatrix}$ Neste caso a rotação é feita apenas no plano YZ, mantendo o eixo X fixado.

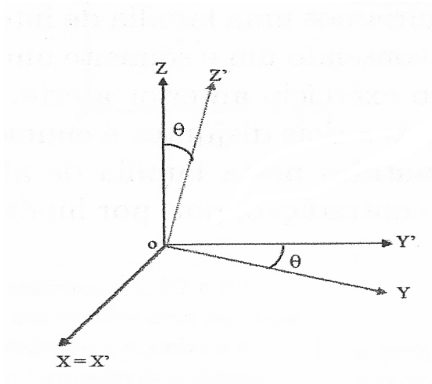

Figura 5

$R_2(\theta)\begin{bmatrix} \cos\theta & 0 & -sen\theta \\ 0 & 1 & 0 \\ sen\theta & 0 & \cos\theta \end{bmatrix}$ Neste caso a rotação é feita apenas no plano XZ, mantendo o eixo Y fixado.

$R_3(\theta)\begin{bmatrix} \cos\theta & sen\theta & 0 \\ -sen\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Neste caso a rotação é feita apenas no plano XY, mantendo o eixo Z fixado.

2.3 Propriedades das Matrizes Ortogonais

As matrizes ortogonais,satisfazem o conjunto das propriedades enunciadas nos teoremas a seguir:

Proposição 2.3.1 Seja Q uma matriz mxn. Suas colunas formam um conjunto ortonormal se, e somente se, $Q^TQ = I_n$.

Demonstração:

Seja q_i a $i-ésima$ coluna de Q (consequentemente a $i-ésima$ linha de Q^T). Como o elemento (i, j) de Q^TQ é produto escalar da $i-ésima$ linha de Q^T pela $j-ésima$ coluna de Q, temos que,

$(Q^TQ)_{i,j} = \langle q_i, q_j \rangle$ (*)

devido a definição de multiplicação de matrizes. De fato, as colunas de Q formam um conjunto ortonormal se e somente se $\left\langle q_i, q_j \right\rangle = \begin{cases} 1 \text{ se i} = j \\ 0 \text{ se i} \neq j \end{cases}$ logo, da equação (*), vale se e somente se $(Q^T Q)_{i,j} = \begin{cases} 1 \text{ se i} = j \\ 0 \text{ se i} \neq j \end{cases}$

Proposição 2.3.2 Uma matriz quadrada Q é ortogonal se e somente se $Q^{-1} = Q^T$.

Demonstração:

Segundo POOLE (2006), no teorema anterior vimos que Q é ortogonal se e somente se $Q^T Q = I$. Isso é verdade se e somente se, Q é invertível. Assim, se multiplicarmos ambos os lados de $Q^T Q = I$ por Q^{-1} pela direita obtemos $Q^T.QQ^{-1} = IQ^{-1} \Rightarrow Q^T I = Q^{-1} \Rightarrow Q^T = Q^{-1}$

Nesse caso, as colunas de Q são ortogonais entre si e formam uma base ortonormal de $\mathbb{R}^n$.

Proposição 2.3.3 Seja Q uma matriz nxn. As seguintes afirmações são equivalentes:

a. Q é ortogonal.

b. $\|Qx\| = \|x\|$ para todo **x** em $\mathbb{R}^n$.

c. $Qx.Qy = x.y$ para todo $\mathbb{R}^n$.

Demonstração: Como as afirmações são equivalentes temos que provar que $(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$. Para provar tais equivalências POOLE (2006) utilizou o fato de que, se **x** e **y** são vetores (coluna) em $\mathbb{R}^n$, então $x.y = x^T y$.

$(a) \Rightarrow (c)$ Se Q é ortogonal $\Rightarrow \left\langle Qx, Qy \right\rangle = \left\langle x, y \right\rangle$ isto é, o produto das imagens é igual ao produto do domínio.

Seja Q ortogonal, então $Q^T Q = I$. De fato, $\left\langle Qx, Qy \right\rangle = (Qx)^T Qy$ então $\left\langle Qx, Qy \right\rangle = x^T Q^T Qy$ como $Q^T Q = I$ temos que $\left\langle Qx, Qy \right\rangle = x^T Iy$ e portanto

$\left\langle Qx, Qy \right\rangle = x^T y$ logo,

$\left\langle Qx, Qy \right\rangle = \left\langle x, y \right\rangle$.

$(c) \Rightarrow (b)$ Se $\left\langle Qx, Qy \right\rangle = \left\langle x, y \right\rangle \forall x, y \Rightarrow \|Qx\| = \|x\|$

Suponha que $\left\langle Qx, Qy \right\rangle = \left\langle x, y \right\rangle \; \forall x, y \in \mathbb{R}^n$. Tomando x=y, por hipótese temos que $\left\langle Qx, Qx \right\rangle = \left\langle x, x \right\rangle$ assim, $\|Qx\| = \sqrt{\left\langle Qx, Qx \right\rangle} = \sqrt{\left\langle x, x \right\rangle} = \|x\|$.

$(b) \Rightarrow (a)$ Considere válida a propriedade (b) e que q_i denota a $i-ésima$ coluna de Q. Então:

a) $\|Q(x+y))\|^2 = \|x+y\|^2$

$= \langle x+y, x+y \rangle$

$= \langle x,x \rangle + 2\langle x,y \rangle + \langle y,y \rangle$

$= \|x\|^2 + 2\langle x,y \rangle + \|y\|^2$

b) $\|Q(x-y)\|^2 = \|x-y\|^2 = \langle x-y, x-y \rangle$

$= \|x\|^2 - 2\langle x,y \rangle + \|y\|^2$

i) (a)-(b)

$\|Q(x+y)\|^2 - \|Q(x-y)\|^2 =$

$\|x\|^2 + 2\langle x,y \rangle + \|y\|^2$ - $\left(\|x\|^2 - 2\langle x,y \rangle + \|y\|^2\right) = 4\langle x,y \rangle$

c) $\|Q(x+y)\|^2 = \|Qx+Qy\|^2$

$= \langle Qx+Qy, Qx+Qy \rangle$

$= \|Qx\|^2 + 2\langle Qx,Qy \rangle + \|Qy\|^2$

d) $\|Q(x-y)\|^2 = \langle Qx-Qy, Qx-Qy \rangle$

$\|Qx\|^2 - 2\langle Qx,Qy \rangle + \|Qy\|^2$

ii) (c)-(d)

$\|Qx\|^2 + 2\langle Qx,Qy \rangle + \|Qy\|^2$ - $\left(\|Qx\|^2 - 2\langle Qx,Qy \rangle + \|Qy\|^2\right) = 4\langle Qx,Qy \rangle$

i = ii

$4\langle x,y \rangle = 4\langle Qx,Qy \rangle$ multiplicando ambos os lados por $\frac{1}{4}$ logo teremos

$\langle x,y \rangle = \langle Qx,Qy \rangle$ $\forall$ x e y em $\mathbb{R}^n$. [Isso mostra que (b) $\Rightarrow$(c)]. Analisando esta propriedade percebemos que preserva o ângulo dos vetores.

Agora, $\langle x, y\rangle = \langle Qx, Qy\rangle$ se x= q_i e y= q_j então, $\langle q_i, q_j\rangle = \langle Qe_i, Qe_j\rangle = \langle e_i, e_j\rangle = \begin{cases} 1 \text{ se i} = j \\ 0 \text{ se i} \neq j \end{cases}$ As colunas de Q formam, portanto, um conjunto ortonormal, e Q é uma matriz ortogonal.

Analisando esta propriedade percebemos que preserva o ângulo dos vetores.

Proposição 2.3.4 Se Q é uma matriz ortogonal, então suas linhas formam um conjunto ortonormal.

Demonstração: Assuma que Q é uma matriz ortogonal, então $Q^{-1} = Q^T$. Assim, $\left(Q^T\right)^{-1} = \left(Q^{-1}\right)^{-1} = Q = \left(Q^T\right)^T$, logo Q^T é uma matriz ortogonal. Então, as colunas de Q^T que são exatamente as linhas de Q formam um conjunto ortonormal.

Proposição 2.3.5 Seja Q uma matriz ortogonal.

a. Q^{-1} é ortogonal.

Demonstração: LIPSCHUTZ (1994)

Temos $Q^T = Q^{-1}$, pois, Q é ortogonal. Assim, como Q é ortogonal se e somente se Q^T é ortogonal logo Q^{-1} é ortogonal.

b. det $|Q| = \pm 1$

Demonstração: STEINBRUCH (2009)

Para provar esta propriedade é suficiente lembrar que Q^T e Q têm o mesmo determinante e que o determinante da identidade é igual a 1. Sendo Q ortogonal, $Q^T Q$ =I. Logo, $\det(Q^T Q) = \det I \Rightarrow \det(Q^T)\ \det(Q) = 1$

Como $\det(Q) = \det(Q^T)$, vem: $(\det Q)^2 = 1$ ou seja, $\det(Q) = +1$ ou $\det(Q) = -1$

c. Se λ é um autovalor de Q, $|\lambda| = 1$.

Demonstração: POOLE (2006) prova esta propriedade utilizando o fato de que uma matriz ortogonal conserva o tamanho do vetor.

Seja λ um autovalor e **v** um autovetor associado a Q, então $Qv = \lambda v$ logo, $\|v\| = \|Qv\| = \|\lambda v\| = |\lambda| \|v\|$ e como $\|v\| \neq 0, \|\lambda v\| = 1$.

d. Se Q_1 e Q_2 são matrizes ortogonais nxn, $Q_1 Q_2$ também é.

Esta propriedade nos mostra que o produto de duas matrizes ortogonais é ortogonal, e para demonstrá-la utilizamos a idéia de LIPSCHUTZ (1994).

Demonstração :

Para provar esta propriedade, lembremos que $(AB)^T = B^T A^T$. Sejam agora Q_1 e Q_2 ortogonais, isto é, $Q_1^T = Q_1^{-1}$ e $Q_2^T = Q_2^{-1}$. Assim, $(Q_1Q_2)(Q_1Q_2)^T = Q_1Q_2Q_2^TQ_1^T =$

Utilizando o fato de $Q_1^T = Q_1^{-1}$e $Q_2^T = Q_2^{-1}$ temos que $Q_1Q_2Q_2^{-1}Q_1^{-1} = Q_1(Q_2Q_2^{-1})Q_1^{-1}$ logo $Q_1IQ_1^{-1} = Q_1Q_1^{-1} = I$ Assim, $(Q_1Q_2)^T = (Q_1Q_2)^{-1}$, e, portanto, Q_1Q_2 é ortogonal.

CAPÍTULO 3

O PROCESSO DE GRAM-SCHMIDT E A FATORAÇÃO QR

No capítulo III descrevemos o processo de ortogonalização de Gram-Schmid e mostramos sua relação com a fatoração QR.

3.1 Projeção Ortogonal

Definição: Seja W um subespaço de R^n e seja $\{u_{1,}...u_k\}$ uma base ortogonal de W. Para um vetor v em R^n, a projeção ortogonal de v em W é definida por $proj_w(v) = \left(\dfrac{u_1.v}{u_1.u_1}\right)u_1 + \cdots + \left(\dfrac{u_k.v}{u_k.u_k}\right)u_k$ A componente ortogonal de v em relação a W é o vetor $perp_w(v) = v - proj_w(v)$.

Como os vetores u_i são ortogonais, a projeção de v sobre W é a soma de suas projeções ortogonais sobre subespaços unidimensionais que são mutuamente ortogonais.

$v, s \subset V$

Ex: $v = \mathbb{R}^3$

$S = Span\{v\}$

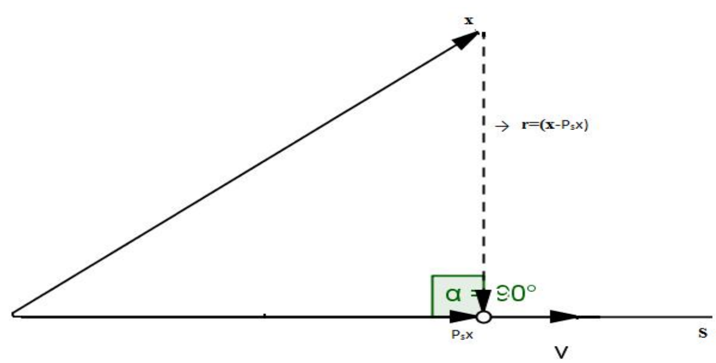

Figura 6

$P_s = ?$

Seja $P_s x = tv$, algum para $t \in \mathbb{R}$

Para determinar t, exigimos que

$$(x - P_s x) \perp S$$

$$\Updownarrow$$

$$(x - P_s x) \perp v$$

$$\Updownarrow$$

$$\langle x - P_s x, v \rangle = 0$$
$$\langle x - tv, v \rangle = 0$$
$$\langle x, v \rangle - t\langle v, v \rangle = 0$$
$$t = \frac{\langle x, v \rangle}{\langle v, v \rangle}$$

Conclusão

$$P_s x = \frac{\langle x, v \rangle}{\langle v, v \rangle} v$$

Se $\langle .,. \rangle$: produto interno em $\mathbb{R}^n$, $\langle x, v \rangle = x^T v \ = v^T x$

$$P_v x = \frac{v^T x}{v^T v} v$$

$$= \frac{\left(vv^T\right)}{v^T v} x$$

$$= \underbrace{\left(\frac{vv^T}{v^T v}\right)}_{\substack{matriz \\ projeção}} x$$

$$= P_s$$

Exemplo numérico

$$S = Span\{v_1, v_2\}$$

$$v_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \qquad P_s =$$

$$v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Note que $\|v_1\| = 1$ e $\|v_2\| = 2$

$$v_1^T v_2 = \frac{1}{\sqrt{3}} . \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} . \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$v_1^T v_2 = \frac{1}{\sqrt{3}} . \frac{1}{\sqrt{2}} \left[1.1 + 1.0 + (-1).1 \right] = 0$$

$$P_s = \left[\underbrace{\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}}_{v_1} \quad \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}}_{v_2} \right] \begin{bmatrix} \overbrace{\frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}}^{v_1^T} \\ \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}}_{v_2^T} \end{bmatrix}$$

calculando obtemos a matriz projeção

$$P_s = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} =$$

$$P_s = \begin{bmatrix} \frac{1}{3} + \frac{1}{2} & \frac{1}{3} + 0 & -\frac{1}{3} + \frac{1}{2} \\ \frac{1}{3} + 0 & \frac{1}{3} + 0 & -\frac{1}{3} + 0 \\ -\frac{1}{3} + \frac{1}{2} & -\frac{1}{3} + 0 & \frac{1}{3} + \frac{1}{2} \end{bmatrix}$$

$$P_s = \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{5}{6} \end{bmatrix}$$

3.2 O PROCESSO GRAM SCHMIDT

Teorema:

Seja $\{x_1, ..., x_k\}$ uma base para um subespaço W de $\mathbb{R}^n$ e defina o seguinte:

$$v_1 = x_1, \qquad W_1 = span(x_1)$$

$$v_2 = x_2 - \left(\frac{v_1 \cdot x_2}{v_1 \cdot v_1}\right)v_1, \quad W_2 = span(x_1, x_2)$$

$$v_3 = x_3 - \left(\frac{v_1 \cdot x_3}{v_1 \cdot v_1}\right)v_1 - \left(\frac{v_2 \cdot x_3}{v_2 \cdot v_2}\right)v_2, \quad W_3 = span(x_1, x_2, x_3)$$

$$\vdots$$

$v_k = x_k - \left(\frac{v_1 \cdot x_k}{v_1 \cdot v_1}\right)v_1 - \left(\frac{v_2 \cdot x_k}{v_2 \cdot v_2}\right)v_2 - \cdots \left(\frac{v_{k-1} \cdot x_k}{v_{k-1} \cdot v_{k-1}}\right)v_{k-1}, \qquad W_k = span(x_1, ..., x_k)$ Então para cada $i = 1, ..., k$, $\{v_1, ..., v_i\}$ é uma base ortogonal de W_i. Em particular $\{v_1, ..., v_k\}$ é uma base de W. Em suma, todo subespaço de $\mathbb{R}^n$ tem uma base ortogonal e fornece um algoritmo para construir uma base dessas. Para normalizarmos os vetores do processo de Gram-Schmidt, basta a cada i substituir v_i pelo vetor unitário $q_i = \left(\frac{v_i}{\|v_i\|}\right)$.

Demonstração: veja POOLE [2006], página 343

Exemplo: Considere a base S=$\{x_1, x_{2,} x_3\}$ para $\mathbb{R}^3$, onde $x_1 = (1,1,1)$, $x_2 = (-1,0,-1)$ e $x_3 = (-1,2,3)$ utilizaremos o processo de Gram schmidt para transformar S em uma base ortonormal para $\mathbb{R}^3$.

Passo1: Seja $v_1 = x_1 = (1,1,1)$.

Passo 2: calculando agora v_2 e v_3:

$$v_2 = x_2 - \left(\frac{x_2 \cdot v_1}{v_1 \cdot v_1}\right)v_1$$

$$v_2 = (-1,0,-1) - \left(\frac{(-1,0,-1)\cdot(1,1,1)}{(1,1,1)\cdot(1,1,1)}\right)\cdot(1,1,1)$$

$$v_2 = (-1,0,-1) - \left(-\frac{2}{3}\right)(1,1,1) =$$

$$v_2 = (-1,0,-1) - \left(-\frac{2}{3}, -\frac{2}{3}, -\frac{2}{3}\right) = \left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$$

$$v_3 = x_3 - \left(\frac{x_3 \cdot v_1}{v_1 \cdot v_1}\right)v_1 - \left(\frac{x_3 \cdot v_2}{v_2 \cdot v_2}\right)v_2$$

$$v_3 = (-1,2,3) - \left(\frac{(-1,2,3)\cdot(1,1,1)}{(1,1,1)\cdot(1,1,1)} \right)\cdot(1,1,1) - \left(\frac{(-1,2,3)\cdot\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)}{\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)\cdot\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)} \right)\cdot\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)$$

$$v_3 = (-1,2,3) - \left(\frac{4}{3}\right)\cdot(1,1,1) - \left(\frac{\frac{2}{3}}{\frac{2}{3}} \right)\cdot\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)$$

$$v_3 = (-1,2,3) - \left(\frac{4}{3},\frac{4}{3},\frac{4}{3}\right) - (1)\cdot\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)$$

$$v_3 = (-1,2,3) - \left(\frac{4}{3},\frac{4}{3},\frac{4}{3}\right) + \left(\frac{1}{3},-\frac{2}{3},\frac{1}{3}\right)$$

$$v_3 = (-2,0,2)$$

Assim, uma base ortogonal para $\mathbb{R}^3$ é $\left\{(1,1,1),\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right),(-2,0,2)\right\}$.

Agora normalizando a base ortogonal

$$q_1 = \frac{v_1}{\|v_1\|} = \frac{(1,1,1)}{\sqrt{1^2+1^2+1^2}} = \frac{(1,1,1)}{\sqrt{3}} = \left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$$

$$q_2 = \frac{v_2}{\|v_2\|} = \frac{\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)}{\sqrt{\left(-\frac{1}{3}\right)^2+\left(\frac{2}{3}\right)^2+\left(-\frac{1}{3}\right)^2}} = \frac{\left(-\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)}{\frac{\sqrt{6}}{3}} = \left(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}\right)$$

$$q_3 = \frac{v_3}{\|v_3\|} = \frac{(-2,0,2)}{\sqrt{(-2)^2+0^2+2^2}} = \left(\frac{-2}{\sqrt{8}},\frac{0}{\sqrt{8}},\frac{2}{\sqrt{8}}\right) = \left(\frac{-1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)$$

Logo obteremos a base ortonormal para $\mathbb{R}^3 \left\{\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right),\left(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}\right),\left(\frac{-1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)\right\}$.

3.3 FATORAÇÃO QR

Teorema: Seja A uma matriz mxn com colunas linearmente independentes. Então A pode ser escrita como: $A = QR$, onde Q é uma matriz mxn com colunas ortonormais e R é uma matriz triangular superior invertível.

Demonstração: Ao aplicar o processo de Gram schimdt (com normalizações) a $x_1,...,x_n$ corresponde a fatorar A.

As colunas de A formam uma base $\{x_1,...,x_n\}$ para coluna de A. Utilizando o processo de Gram schmidt, podemos obter um base ortonormal $\{q_1,...,q_n\}$ para o espaço coluna de A. Seja $Q = [q_1 \quad q_2 \quad ... \quad q_n]$ para $K = 1,...,n$, x_k está em $Span\{x_1,...,x_k\} = Span\{q_1,...,q_n\}$.

Agora, cada um dos vetores x_i pode ser escrito como uma combinação linear dos vetores q, isto é, existem constantes $r_{1k}, r_{2k},..., r_{kk}$ tais que $x_k = r_{1k}q_1 + r_{2k}q_2 + ... + r_{kk}q_k + 0 \cdot q_{k+1} + ... + 0 \cdot q_n$. Podemos supor que os elementos da diagonal de R são positivos. Se $r_{kk} < 0$, basta substituirmos q_k por $-q_k$ e r_{kk} por $-r_{kk}$, ou então simplesmente multiplicar tanto q_k como r_{kk} por -1. Assim, como foi dito anteriormente, x_k é uma combinação linear das colunas de Q usando como coeficientes as coordenadas do vetor

$$r_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

ou seja, $x_k = Qr_k$ para $k = 1,...,n$. Seja $R = [r_1 \quad \cdots \quad r_n]$. Então

$$A = [x_1 \quad \cdots \quad x_n] = [Qr_1 \quad \cdots \quad Qr_n] = QR, \text{ onde } R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \cdots & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

Mostraremos agora que R é invertível. Seja y uma solução para o sistema linear $Ry = 0$. Multiplicando esta equação por Q à esquerda, temos

$Q(Ry) = Q0$

$(QR)y = 0$ usando o fato de $A = QR$

$Ay = 0$

Escrevendo $Ay = 0$ como

$y_1x_1 + y_2x_2 + \cdots + y_nx_n = 0$, onde $y_1, y_2, \ldots, y_n$ são as componentes do vetor y. Como as colunas de A são linearmente independentes, $y_1 = y_2 = \cdots = y_n = 0$, portanto y tem que ser o vetor nulo. Logo R é invertível.

Exemplo:

Encontre a fatoração QR de A= $\begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 2 \\ 1 & -1 & 3 \end{bmatrix}$

As colunas de A são os vetores $x_1, x_{2,} x_3$ do exemplo anterior. Aplicando o processo de Gram Schmidt para essa base iniciando com x_1, encontramos a seguinte base ortonormal para o espaço coluna de A já calculados ainda no exemplo anterior:

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad q_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix}, \quad q_3 = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Utilizando estes vetores ortonormais como colunas de Q temos;

$$Q = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Por construção, as k primeiras colunas de Q formam uma base ortonormal para $Span\{x_1, \ldots x_k\}$. Pela demonstração do teorema anterior A=QR para algum R. Para encontrar R, note que $Q^TQ = I$, já que as colunas de Q são ortonormais. Logo, multiplicando esta equação por Q^T à esquerda, temos

$A = QR$

$Q^TA = Q^TQR$

$Q^T A = (Q^T Q)R$

$Q^T A = IR$

$Q^T A = R$

calculamos

$$R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 2 \\ 1 & -1 & 3 \end{bmatrix}$$ e então obtemos

$$R = \begin{bmatrix} \frac{3}{\sqrt{3}} & -\frac{2}{\sqrt{3}} & \frac{4}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ 0 & 0 & \frac{4}{\sqrt{2}} \end{bmatrix}.$$

CAPÍTULO 4

MÉTODO DOS MÍNIMOS QUADRADOS

No capítulo IV estudamos o problema de mínimos quadrados, onde o sistema linear não possui solução porque o vetor do lado direito não pertence ao espaço coluna da matriz, e mostramos que a solução do problema pode ser encontrado pelas equações normais ou pela fatoração QR.

Para ilustrar a teoria e os métodos estudados, procuramos uma reta de melhor ajuste de um conjunto de pontos dados, e finalizamos com o problema do melhor ajuste quadrático.

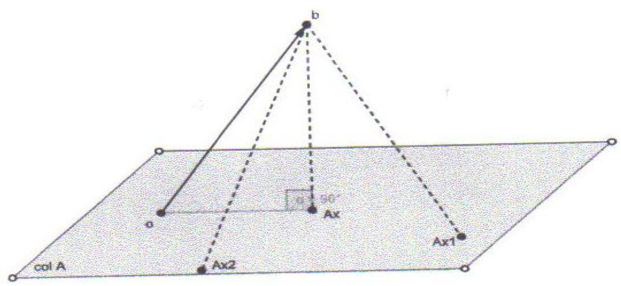

Figura 7

b está mais próximo de Ax do que de Ax1 e Ax2 , isto é para outro qualquer x_n.

4.1 O Teorema da Melhor Aproximação

Se W é um subespaço de dimensão finita de um espaço V com produto interno e se v é um vetor em V, então $proj_W(v)$ é a melhor aproximação para v em W.

Denotaremos $proj_W(v)$ como $\bar{v}$

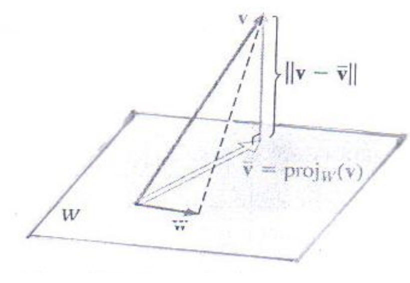

Figura 8

Se $\bar{v} = proj_W(v)$***, então*** $\|v - \bar{v}\| < \|v - w\|$ ***para todo*** $w \neq \bar{v}$

Demonstração: seja w um vetor em W diferente da projeção de v no subespaço W, isto é diferente de $\bar{v}$. Como w e $\bar{v}$ estão ambos em W, $\bar{v}$-w está também em W, e $v-\bar{v} = perp_W(v)$ é ortogonal a $\bar{v}-w$. Aplicando o teorema de Pitágoras temos que

$$\|(v-\bar{v})+(\bar{v}-w)\|^2 = \|v-\bar{v}\|^2 + \|\bar{v}-w\|^2$$

$$\|v-\bar{v}+\bar{v}-w\|^2 = \|v-\bar{v}\|^2 + \|\bar{v}-w\|^2$$

$$\|v-w\|^2 = \|v-\bar{v}\|^2 + \|\bar{v}-w\|^2$$

Como $w \neq \bar{v}$, então $\|\bar{v}-w\|^2$ é positivo e $\|v-\bar{v}\|^2 + \|\bar{v}-w\|^2 > \|v-\bar{v}\|^2$ o que equivale a $\|v-w\|^2 > \|v-\bar{v}\|^2$. Assim, segue que $\bar{v}$ é o vetor em W que minimiza $\|v-w\|^2$ e portanto minimiza $\|v-w\|$.

O teorema acima diz que a distância entre v e w, dada por $\|v-w\|$, pode ser vista como o "erro" em usar w no lugar de v, e que o erro é minimizado quando $w=\bar{v}$, assim, o vetor $\bar{v}$ é chamado de projeção ortogonal de v sobre w, e $\bar{v}$ é chamado de melhor aproximação por elementos de W.

O problema geral de mínimos quadráticos é: determinar $x \in \mathbb{R}^n$ tal que a diferença entre b e Ax, isto é, $\|b-Ax\|$ é mínima.

4.2 Teorema dos Mínimos Quadrados

Seja A uma matriz mxn e seja b em R^m. Então, $Ax=b$ sempre tem pelo menos uma solução por mínimos quadrados x. Além disso,

a) x é uma solução por mínimos quadrados de $Ax=b$ se, e somente se, x é uma solução da equação normal $A^T Ax = A^T b$.

b) A possui colunas linearmente independentes se, e somente se, $A^T A$ é invertível. Nesse caso, a solução por mínimos quadrados de $Ax=b$ é única e é dada por $x=(A^T A)^{-1} A^T b$.

Demonstração:

a) $r = b - Ax$ (resíduo)

Para que $\|b-Ax\|$ seja o menor possível, o vetor resíduo deve ser ortogonal ao espaço coluna.

Se a_i é uma coluna de A, então

$$\Leftrightarrow \begin{cases} a_1^T r = 0 \\ a_2^T r = 0 \\ \vdots \\ a_m^T r = 0 \end{cases}$$

Como cada a_i^T é uma linha de A^T, temos

$\Leftrightarrow A^T r = 0$

$A^T(b - Ax) = 0$

$A^T b - A^T Ax = 0$

$\Leftrightarrow A^T Ax = A^T b$

Logo, uma solução por mínimos quadrados para um sistema de equações lineares $Ax = b$ é um vetor x que minimiza o comprimento do vetor $Ax - b$, isto é que satisfaz as equações $A^T Ax = A^T b$ conhecido como equações normais para x.

b) Note que as n colunas de A são linearmente independentes se, e somente se, posto de A=n. Mas isso é verdade se, e somente se, $A^T A$ é invertível. Podemos provar sua invertibilidade mostrando que o sistema linear $A^T Ax = 0$ tem apenas a solução trivial. Multiplicando ambos os lados de $A^T Ax = 0$ por x^T, temos

$A^T Ax = 0$

$x^T A^T Ax = x^T 0$

$(Ax)^T (Ax) = 0$

$(Ax).(Ax) = 0$

De acordo com a propriedade 1.5 dos conceitos preliminares temos que Ax=0. Isso implica que temos uma combinação linear de colunas de A linearmente independentes que é nula. Logo, x=0. Assim, $A^T A$ é invertível e a equação $A^T Ax = A^T b$ tem uma única solução $x = (A^T A)^{-1} A^T b$.

Neste método de resolução o procedimento para encontrar a aproximação por mínimos quadrados é primeiramente formar a matriz $A^T A$ e $A^T b$ e em seguida resolver o sistema normal $A^T Ax = A^T b$ por escalonamento, isto é, pelo método de Gauss-Jordan.

Exemplo de um aplicação:

Encontre uma solução por mínimos quadráticos do sistema impossível $Ax = b$, onde

$$A=\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix},\; b=\begin{bmatrix}2\\0\\11\end{bmatrix}$$

Solução:

Formaremos as matrizes A^TA e A^Tb

$$A^TA=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix}=\begin{bmatrix}17&1\\1&5\end{bmatrix}$$

$$A^Tb=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}2\\0\\11\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

Agora formamos o sistema normal $A^TAx=A^Tb$

$$\begin{bmatrix}17&1\\1&5\end{bmatrix}x=\begin{bmatrix}19\\11\end{bmatrix}$$

$$\begin{bmatrix}17&1\\1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

Aplicando o método de Gauss-Jordan, e usando o fato de A^TA é invertível temos:

$$\left(A^TA\right)^{-1}=\begin{bmatrix}17&1\\1&5\end{bmatrix}^{-1}=\left[\begin{array}{cc|cc}17&1&1&0\\1&5&0&1\end{array}\right]\cong\begin{array}{c}L'_1=L_1\\L'_2=L_1-17L_2\end{array}\left[\begin{array}{cc|cc}17&1&1&0\\0&-84&1&-17\end{array}\right]\cong$$

$$\begin{array}{c}L''_1=\frac{1}{17}L'_1\\L''_2=-\frac{1}{84}L'_2\end{array}\left[\begin{array}{cc|cc}1&\frac{1}{17}&\frac{1}{17}&0\\0&1&-\frac{1}{84}&\frac{17}{84}\end{array}\right]\cong\begin{array}{c}L'''_1=L''_1+\frac{1}{17}L''_2\\L'''_2=L''_2\end{array}\left[\begin{array}{cc|cc}1&0&\frac{5}{84}&-\frac{1}{84}\\0&1&-\frac{1}{84}&\frac{17}{84}\end{array}\right]$$

Logo $\left(A^TA\right)^{-1}=\frac{1}{84}\begin{bmatrix}5&-1\\-1&17\end{bmatrix}$

Resolvendo $A^TAx=A^Tb$ diretamente:

$$x=(A^TA)^{-1}A^Tb$$

$$=\begin{bmatrix}\frac{5}{84}&\frac{-1}{84}\\\frac{-1}{84}&\frac{17}{84}\end{bmatrix}\begin{bmatrix}19\\11\end{bmatrix}=\begin{bmatrix}\frac{84}{84}\\\frac{168}{84}\end{bmatrix}\text{logo,}$$

$$x=\begin{bmatrix}1\\2\end{bmatrix}$$

Pelo teorema dos mínimos quadrados (b), o sistema normal tem uma única solução, devido às colunas de A serem independentes.

Os componentes de x são os coeficientes da combinação linear das colunas de A

$Ax = b$

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} \cong \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

que ocasionam a melhor aproximação para b.

4.3 Utilizando Fatoração QR Para o Problema de Mínimos Quadrados

Seja A uma matriz mxn com colunas linearmente independentes, e seja b em R^m. Se $A = QR$ é uma fatoração QR de A, então a única solução por mínimos quadrados x de $Ax = b$ é $x = R^{-1}Q^T b$.

Demonstração:

Seja uma fatoração QR de A, $A = QR$. Substituindo A nas equações normais temos:

$A^T Ax = A^T b$

$(QR)^T (QR)x = (QR)^T b$

$R^T Q^T QRx = R^T Q^T b$

$R^T (Q^T Q)Rx = R^T Q^T b$

como as colunas de Q formam um conjunto ortonormal, $Q^T Q = I$

$R^T IRx = R^T Q^T b$

$R^T Rx = R^T Q^T b$ devido R ser invertível, R^T também é uma matriz invertível e logo temos $Rx = Q^T b$ o que é equivalente a $x = R^{-1}Q^T b$. (Sistema alternativo para resolver o problema de mínimos quadrados).

Pelo fato de R ser triangular superior é muito mais rápido e fácil resolver $Rx = Q^T b$ do que usar substituições ou operações elementares para calcular R^{-1} e depois calcular $x = R^{-1}Q^T b$.

Exemplo: Encontre uma solução por mínimos quadráticos do sistema impossível $Ax = b$, onde

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix},\ b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Utilizando a fatoração QR para o problema de mínimos quadrados temos:

$$v_1 = x_1 = (4,0,1)$$

$$v_2 = (0,2,1) - \left(\frac{(0,2,1).(4,0,1)}{(4,0,1).(4,0,1)}\right).(4,0,1)$$

$$v_2 = (0,2,1) - \left(\frac{1}{17}\right).(4,0,1)$$

$$v_2 = (0,2,1) + \left(-\frac{4}{17}, 0, -\frac{1}{17}\right)$$

$$v_2 = \left(-\frac{4}{17}, 2, \frac{16}{17}\right)$$

Base ortogonal $\left\{(4,0,1)\left(-\frac{4}{17}, 2, \frac{16}{17}\right)\right\}$

Normalizando a base ortogonal:

$$q_1 = \frac{v_1}{\|v_1\|} \rightarrow q_1 = \frac{(4,0,1)}{\sqrt{4^2+0^2+1^2}} = \frac{(4,0,1)}{\sqrt{17}} = \left(\frac{4}{17}, 0, \frac{1}{\sqrt{17}}\right)$$

$$q_2 = \frac{v_2}{\|v_2\|} \rightarrow q_2 = \frac{\left(-\frac{4}{17}, 2, \frac{16}{17}\right)}{\sqrt{\left(-\frac{4}{17}\right)^2 + 2^2 + \left(\frac{16}{17}\right)^2}} = \frac{\left(-\frac{4}{17}, 2, \frac{16}{17}\right)}{\sqrt{\frac{16}{289} + 4 + \frac{256}{289}}} = \frac{\left(-\frac{4}{17}, 2, \frac{16}{17}\right)}{\sqrt{\frac{1428}{289}}} =$$

$$\frac{\left(-\frac{4}{17}, 2, \frac{16}{17}\right)}{\frac{\sqrt{1428}}{17}} = \left(-\frac{4}{\sqrt{1428}}, \frac{34}{\sqrt{1428}}, \frac{16}{\sqrt{1428}}\right)$$

$$Q = \begin{bmatrix} q_1 & q_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{17}} & -\frac{4}{\sqrt{1428}} \\ 0 & \frac{34}{\sqrt{1428}} \\ \frac{1}{\sqrt{17}} & \frac{16}{\sqrt{1428}} \end{bmatrix} \text{ matriz ortogonal}$$

Calculando $R = Q^T A$

$$R = \begin{bmatrix} \frac{4}{\sqrt{17}} & 0 & \frac{1}{\sqrt{17}} \\ -\frac{4}{\sqrt{1428}} & \frac{34}{\sqrt{1428}} & \frac{16}{\sqrt{1428}} \end{bmatrix} . \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} \frac{17}{\sqrt{17}} & \frac{1}{\sqrt{17}} \\ 0 & \frac{84}{\sqrt{1428}} \end{bmatrix}$$

A solução de mínimos quadráticos x satisfaz $Rx = Q^T b$, isto é,

$$\begin{bmatrix} \frac{17}{\sqrt{17}} & \frac{1}{\sqrt{17}} \\ 0 & \frac{84}{\sqrt{1428}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{17}} & 0 & \frac{1}{\sqrt{17}} \\ -\frac{4}{\sqrt{1428}} & \frac{34}{\sqrt{1428}} & \frac{16}{\sqrt{1428}} \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} \frac{17}{\sqrt{17}} & \frac{1}{\sqrt{17}} \\ 0 & \frac{84}{\sqrt{1428}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{19}{\sqrt{17}} \\ \frac{168}{\sqrt{1428}} \end{bmatrix}$$

Essa equação é resolvida facilmente e obtemos $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Talvez o fato mais importante do uso da fatoração QR é evitar o cálculo da matriz $A^T A$. Tal cálculo pode produzir dificuldades no cálculo da solução do problema. Isto pode ser ilustrado com a seguinte matriz: $A = \begin{bmatrix} 1+\xi & 1 \\ 1 & 1+\xi \\ 1 & 1 \end{bmatrix}$ e $A^T = \begin{bmatrix} 1+\xi & 1 & 1 \\ 1 & 1+\xi & 1 \end{bmatrix}$;

$$A^T A = \begin{bmatrix} 3+2\xi+\xi^2 & 3+2\xi \\ 3+2\xi & 3+2\xi+\xi^2 \end{bmatrix}$$

Na prática, (na utilização do computador) se ξ é muito pequeno, ξ^2 é representado no computador como nulo, e $A^T A$ torna-se $A^T A = \begin{bmatrix} 3+2\xi & 3+2\xi \\ 3+2\xi & 3+2\xi \end{bmatrix}$ que é singular. Nesse caso, perde-se a unicidade da solução do problema. Isso pode ser evitado calculando QR de A.

4.4 Reta de Melhor Ajuste

Geralmente medimos o valor de y conforme o valor de x dado, e então marcamos no plano cartesiano os pontos (x,y), tentamos relacionar as variáveis x e y que pode ser utilizada para prever valores de y para determinados valores de x. Quando os pontos marcados no gráfico não pertencem exatamente a uma reta, ou seja, quando dados experimentais fornecem pontos que ao serem colocados em um gráfico parecem estar próximo de uma reta costumamos usar o método dos mínimos quadrados para obter a reta que melhor se ajuste aos dados. Calcular a solução de mínimos quadráticos de Ax=b é equivalente a encontrar o vetor x que determina a reta de mínimos quadráticos. Iremos determinar os coeficientes a e b que tornam a reta mais próxima possível dos pontos. A reta de melhor ajustes é a reta y=ax +b que minimiza a soma dos quadrados dos resíduos.

4.5 Erro Quadrático (resíduo)

A diferença entre um valor de y observado e um previsto é chamada de resíduo. Para medir o quanto a reta está próxima dos dados utilizamos somar os quadrados dos resíduos, isto é, o quadrado da distância entre os vetores Ax e b é precisamente a soma dos quadrados dos resíduos. Uma vez calculados e encontrados a e b o objetivo é minimizar o comprimento do resíduo (e), o que é equivalente a encontrar a solução de mínimos quadráticos de Ax=b. Assim, a solução de mínimos quadráticos x é a solução das equações normais. $(A^T A)x = A^T b$

Sendo

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad x = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{e} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

então,

$$A^T Ax = A^T b$$

$$A^T A = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} = \begin{bmatrix} x_1{}^2 + x_2{}^2 + \dots + x_m{}^2 & x_1 + x_2 + \dots + x_m \\ x_1 + x_2 + \dots + x_m & m \end{bmatrix}$$

$$A^T b = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} x_1 . y_1 + x_2 y_2 + \dots + x_m y_m \\ y_1 + y_2 + \dots + y_m \end{bmatrix}$$

Se porventura os pontos estiverem exatamente sobre a reta de mínimos quadrados a equação estaria da forma y=ax+b. Porém, como alguns dos pontos podem não pertencer exatamente a reta, então a equação se apresenta da forma y=Ax+b+e,ou e=y-(Ax+b) em que e é o desvio vertical do ponto dado até a reta de mínimos quadrados (representando os erros na direção y). Para que e seja o menor possível o $\|e\|$ deve estar o mais próximo possível de zero. O vetor erro é apenas b-Ax, devido a forma matricial com que é representada a equação da reta onde os pontos não são colineares, no qual o sistema é indeterminado.Para calcular o resíduo utilizamos $e = \|b - Ax\|$.

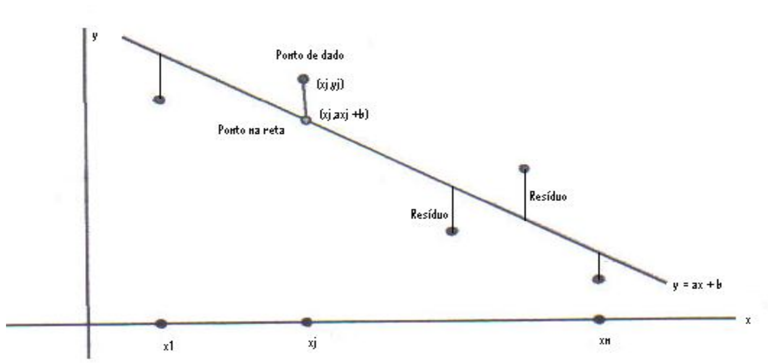

Figura 9

Quando os pontos de dados (x_1,y_1), ..., (x_n,y_n), em um gráfico, não parecem estar próximos de nenhuma reta, pode ser conveniente escolher uma outra relação funcional entre x e y, e então podemos utilizar o método de mínimos quadrados para aproximar pontos dados por outras curvas, além de retas. A solução de mínimos quadrados x é a solução das equações normais: $(A^T A)x = A^T b$

Suponha que queremos aproximar os dados por uma equação da forma $y = ax^2 + bx + c$, então as coordenadas do primeiro ponto (x_1,y_1) satisfazem uma equação da forma : $y_1 = ax_1^2 + bx_1 + c + e_1$, onde e_1 é o erro residual entre o valor observado y_1 e o valor previsto para y, $y_1 = ax_1^2 + bx_1 + c$. Vamos formar uma equação análoga para cada ponto dos dados.

$$
\begin{aligned}
y_1 &= ax_1^2 + bx_1 + c + e_1 \\
y_2 &= ax_2^2 + bx_2 + c + e_2 \\
\vdots & \qquad \vdots \\
y_n & \quad ax_n^2 + bx_n + c + e_n
\end{aligned}
$$

Escrevendo este sistema na forma $y = Ax + e$

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{b} = \underbrace{\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}}_{e}$$

Temos:

$(A^T A)x = A^T b$

$$A^T A = \begin{bmatrix} x_1^2 & x_2^2 & \cdots & x_n^2 \\ x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} =$$

$$\begin{bmatrix} x_1^4 + x_2^4 + \cdots + x_n^4 & x_1^3 + x_2^3 + \cdots + x_n^3 & x_1^2 + x_2^2 + \cdots + x_n^2 \\ x_1^3 + x_2^3 + \cdots + x_n^3 & x_1^2 + x_2^2 + \cdots + x_n^2 & x_1 + x_2 + \cdots + x_n \\ x_1^2 + x_2^2 + \cdots + x_n^2 & x_1 + x_2 + \cdots + x_n & m \end{bmatrix}$$

$$A^T b = A^T b = \begin{bmatrix} x_1^2 & x_2^2 & \cdots & x_n^2 \\ x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1^2.y_1 + x_2^2.y_2 + \cdots + x_n^2.y_n \\ x_1.y_1 + x_2.y_2 + \cdots + x_n.y_n \\ y_1 + y_2 + \cdots + y_n \end{bmatrix}$$

Exemplos:

1) Encontre a reta de mínimos quadrados e o erro quadrático mínimo para os pontos (1,1), (2,2), (3,2), (4,3)

Seja y=ax +b a reta procurada. Substituindo os quatros pontos nessa equação, obtemos:

1=a+b

2=2a+ b   ou

2=3a+b

3=4a+b

$$\begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$ Assim, queremos a solução por mínimos quadrados de Ax=b, onde

$$A=\begin{bmatrix}1&1\\2&1\\3&1\\4&1\end{bmatrix} \text{ e } b=\begin{bmatrix}1\\2\\2\\3\end{bmatrix}$$

Calculando a equação normal $(A^T A)x = A^T b$

$$A^T A=\begin{bmatrix}1^2+2^2+3^2+4^2 & 1+2+3+4\\ 1+2+3+4 & 4\end{bmatrix}$$

$$A^T A=\begin{bmatrix}30&10\\10&4\end{bmatrix}$$

$$A^T b=\begin{bmatrix}1.1+2.2+3.2+4.3\\ 1+2+2+3\end{bmatrix}$$

$$A^T b=\begin{bmatrix}23\\8\end{bmatrix}$$

$$\begin{bmatrix}30&10\\10&4\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}23\\8\end{bmatrix}$$

$$\begin{bmatrix}30&10&\vdots&23\\10&4&\vdots&8\end{bmatrix}\cong \begin{matrix}L_1^{'}=L_1\\ L_2^{'}=3L_2-L_1\end{matrix}\begin{bmatrix}30&10&\vdots&23\\0&2&\vdots&1\end{bmatrix}\cong \begin{matrix}L_1^{''}=L_1^{'}-5L_2^{'}\\ L_2^{''}=L_2^{'}\end{matrix}\begin{bmatrix}30&0&\vdots&18\\0&2&\vdots&1\end{bmatrix}\cong$$

$$\begin{matrix}L_1^{'''}=\frac{1}{30}L_1^{''}\\ L_2^{'''}=\frac{1}{2}L_2^{''}\end{matrix}\begin{bmatrix}1&0&\vdots&\frac{3}{5}\\0&1&\vdots&\frac{1}{2}\end{bmatrix} \text{ logo } x=\begin{bmatrix}\frac{3}{5}\\ \frac{1}{2}\end{bmatrix}$$

e portanto uma equação para a reta de mínimos quadrados com coeficientes $a=\frac{3}{5}$ e $b=\frac{1}{2}$ que minimiza a soma dos quadrados dos resíduos é: $y=\frac{3}{5}x+\frac{1}{2}$.

Como $e=b-Ax$ temos

$$e=\begin{bmatrix}1\\2\\2\\3\end{bmatrix}-\begin{bmatrix}1&1\\2&1\\3&1\\4&1\end{bmatrix}.\begin{bmatrix}\frac{3}{5}\\ \frac{1}{2}\end{bmatrix}=\begin{bmatrix}1\\2\\2\\3\end{bmatrix}-\begin{bmatrix}\frac{11}{10}\\ \frac{17}{10}\\ \frac{23}{10}\\ \frac{29}{10}\end{bmatrix}=\begin{bmatrix}-\frac{1}{10}\\ \frac{3}{10}\\ -\frac{3}{10}\\ \frac{1}{10}\end{bmatrix}$$

Calculando a norma do resíduo temos $\|e\| = \sqrt{\left(-\frac{1}{10}\right)^2 + \left(\frac{3}{10}\right)^2 + \left(-\frac{3}{10}\right)^2 + \left(\frac{1}{10}\right)^2} =$

$\sqrt{\frac{1}{5}} \cong 0.447$

.

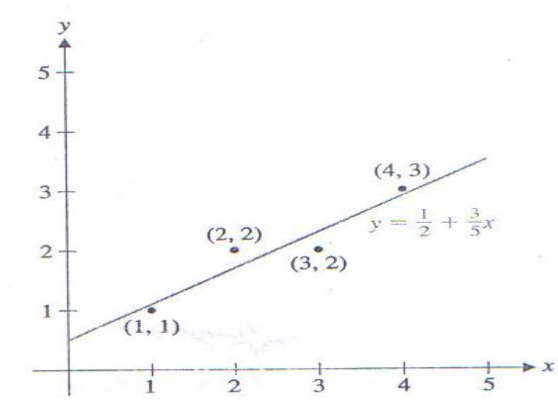

Figura 10

2) Encontre a parábola que tem a melhor aproximação por mínimos quadrados para os pontos (-1,1), (0,-1), (1,0) e (2,2).

Solução: A equação de uma parábola é $y = ax^2 + bx + c$, substituindo os pontos dados nessa equação, obtemos o sistema linear:

$$\begin{aligned} 1 &= a - b + c \\ -1 &= +c \\ 0 &= a + b + c \\ 2 &= 4a + 2b + c \end{aligned}$$

ou

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Assim, queremos a solução por mínimos quadrados de Ax=b, onde

$$A=\begin{bmatrix}1 & -1 & 1\\ 0 & 0 & 1\\ 1 & 1 & 1\\ 4 & 2 & 1\end{bmatrix} \quad \text{e} \quad \text{b}=\begin{bmatrix}1\\ -1\\ 0\\ 2\end{bmatrix}$$

Calculando a equação normal $(A^T A)x = A^T b$

$$A^T A=\begin{bmatrix}(-1)^4+0^4+1^4+2^4 & (-1)^3+0^3+1^3+2^3 & (-1)^2+0^2+1^2+2^2\\ (-1)^3+0^3+1^3+2^3 & (-1)^2+0^2+1^2+2^2 & (-1)+0+1+2\\ (-1)^2+0^2+1^2+2^2 & (-1)+0+1+2 & 4\end{bmatrix}$$

$$A^T A=\begin{bmatrix}18 & 8 & 6\\ 8 & 6 & 2\\ 6 & 2 & 4\end{bmatrix}$$

$$A^T b=\begin{bmatrix}(-1)^2.1+0^2.(-1)+1^2.0+2^2.2\\ (-1).1+0.(-1)+1.0+2.2\\ 1+(-1)+0+2\end{bmatrix}$$

$$A^T b=\begin{bmatrix}9\\ 3\\ 2\end{bmatrix}$$

$$\begin{bmatrix}18 & 8 & 6\\ 8 & 6 & 2\\ 6 & 2 & 4\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}9\\ 3\\ 2\end{bmatrix}$$

$$\begin{bmatrix}18 & 8 & 6 & \vdots & 9\\ 8 & 6 & 2 & \vdots & 3\\ 6 & 2 & 4 & \vdots & 2\end{bmatrix} \cong \begin{matrix}L_1' = \dfrac{L_1}{L_1}\\ L_2' = 6L_2 - 8L_3\\ L_3' = L_1 - 3L_3\end{matrix}\begin{bmatrix}1 & \dfrac{4}{9} & \dfrac{1}{3} & \vdots & \dfrac{1}{2}\\ 0 & 20 & -20 & \vdots & 2\\ 0 & 2 & -6 & \vdots & 3\end{bmatrix} \cong$$

$$\begin{matrix}L_1'' = L_1'\\ L_2'' = \dfrac{L_2'}{20}\\ L_3'' = L_2' - 10L_3'\end{matrix}\begin{bmatrix}1 & \dfrac{4}{9} & \dfrac{1}{3} & \vdots & \dfrac{1}{2}\\ 0 & 1 & -1 & \vdots & \dfrac{1}{10}\\ 0 & 0 & 40 & \vdots & -28\end{bmatrix} \cong \begin{matrix}L_1''' = L_1'' - \dfrac{4}{9}L_2''\\ L_2''' = 40L_2'' + L_3''\\ L_3''' = \dfrac{1}{40}L_3''\end{matrix}\begin{bmatrix}1 & 0 & \dfrac{7}{9} & \vdots & \dfrac{41}{90}\\ 0 & 40 & 0 & \vdots & -24\\ 0 & 0 & 1 & \vdots & -\dfrac{7}{10}\end{bmatrix} \cong$$

$$L_1'''' = L_1''' - \frac{7}{9}L_3''' \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{7}{10} \end{array}\right]$$
$$L_2'''' = \frac{1}{40}L_2'''$$
$$L_3''' = L_3'''$$

Logo x = $\begin{bmatrix} 1 \\ -\frac{3}{5} \\ -\frac{7}{10} \end{bmatrix}$ assim, a parábola de mínimo quadrado tem equação $y = x^2 - \frac{3}{5}x - \frac{7}{10}$.

Como $e = b - Ax$ temos

$$e = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{3}{5} \\ -\frac{7}{10} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{9}{10} \\ -\frac{7}{10} \\ -\frac{3}{10} \\ \frac{21}{10} \end{bmatrix} = \begin{bmatrix} \frac{1}{10} \\ -\frac{3}{10} \\ \frac{3}{10} \\ -\frac{1}{10} \end{bmatrix}$$

Calculando a norma do resíduo temos $\|e\| = \sqrt{\left(\frac{1}{10}\right)^2 + \left(-\frac{3}{10}\right)^2 + \left(\frac{3}{10}\right)^2 + \left(-\frac{1}{10}\right)^2} =$

$\sqrt{\left(\frac{1}{5}\right)} \cong 0.447$

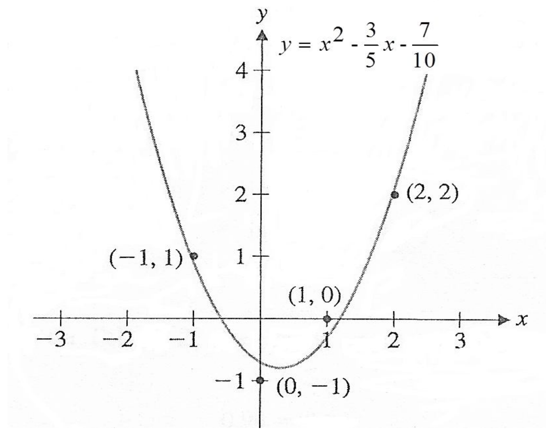

Figura 11

CAPÍTULO 5

CONCLUSÕES

A presente pesquisa foi de grande beneficio para meu conhecimento, pois tive a oportunidade de aprender que matrizes ortogonais apresentam propriedades de conservar ângulos e módulos, além de preservar produtos escalares. Ao se calcular o problema de mínimos quadrados em muitos casos, torna–se mais eficiente utilizar a fatoração QR de A (se as colunas de A forem linearmente independentes) pois as equações normais podem algumas vezes introduzir pequenos erros nos cálculos dos elementos de $A^T A$ que podem causar erros relativamente grandes na solução. Um exemplo disso está na matriz A = $\begin{bmatrix} 1+\xi & 1 \\ 1 & 1+\xi \\ 1 & 1 \end{bmatrix}$.

Quando se calcula $A^T A = \begin{bmatrix} 3+2\xi+\xi^2 & 3+2\xi \\ 3+2\xi & 3+2\xi+\xi^2 \end{bmatrix}$ num programa de computador, se ξ é muito pequeno, ξ^2 é representado como nulo, e $A^T A$ torna-se $A^T A = \begin{bmatrix} 3+2\xi & 3+2\xi \\ 3+2\xi & 3+2\xi \end{bmatrix}$ sendo singular, o que ocasiona a perda da unicidade da solução do problema garantida no teorema dos mínimos quadrados (item 4.2). Na reta de melhor ajuste, os dados experimentais fornecem muitas vezes, pontos que, ao serem colocados em um gráfico, parecem estar próximos de uma reta, temos então de determinar os parâmetros a e b que tornam a reta mais próxima possível dos pontos e para medir o quão próxima está a reta dos dados a melhor escolha é somar os quadrados dos resíduos. Quando os pontos colocados num gráfico não parecem estar próximos de nenhuma reta, pode ser apropriado escolher uma outra relação funcional, tornando-se necessário ajustar os dados a uma curva diferente de uma reta. Assim, propõe-se para futuros trabalhos aprofundar o estudo de matrizes ortogonais e suas aplicações.

REFERÊNCIAS

BOLDRINI, J.L. et. Al; **Álgebra Linear.** 3ª ed. São Paulo: Harbra LTD, 1980.

CALLIOLI, Carlos A. et Al. **Álgebra Linear e Aplicações**. 6ª ed. São Paulo: Atual, 1990.

KOLMAN, Bernard; HILL, David R; tradução: BOSQUILHA, Alessandra. **Introdução à Álgebra Linear:** com aplicações. Rio de Janeiro, RJ: LTC, 2006.

LAY, David C; tradução: CAMELIER, Ricardo; IÓRIO, Valéria de Magalhães. **Álgebra Linear e suas Aplicações.** – 2ª ed. **-** Rio de Janeiro, RJ: LTC, 2007.

LIPSCHUTZ, Seymour; tradução: FARIAS, Alfredo A. **Álgebra Linear:** Teoria e Problemas. – 3ª ed. – São Paulo, SP: Makron Books, 1994.

NADAL, Carlos A. **Sistemas de coordenadas tridimensionais Translação e Rotação de Sistemas** <http://www.cartografica.ufpr.br/docs/sistreftempo/translacao.pdf> Acesso em 19 jan. 2011.

POOLE, David; tradução: MONTEIRO, Martha S. (coord.) [et al] **Álgebra Linear.** São Paulo, SP:Thomson Learning, 2006.

STEINBRUCH. Alfredo; WINTERLE, Paulo. **Álgebra Linear.** 2ª ed. São Paulo: Pearson Makron Books,1987.