Universidade Federal de Santa Catarina – UFSC

Centro de Ciências Físicas e Matemáticas

Departamento de Matemática

Curso de Especialização em Matemática - Formação do Professor de Matemática na modalidade a Distância

AJUSTE DE CURVAS POR MÍNIMOS QUADRADOS

Ademir de Lima

Florianópolis, Santa Catarina, Março de 2011

Universidade Federal de Santa Catarina – UFSC

Centro de Ciências Físicas e Matemáticas

Departamento de Matemática

Curso de Especialização em Matemática - Formação do Professor de Matemática na modalidade a Distância

AJUSTE DE CURVAS POR MÍNIMOS QUADRADOS

Ademir de Lima

Monografia submetida à Comissão de avaliação do curso de Especialização em Matemática-Formação do professor em cumprimento parcial para a obtenção do título de Especialista em Matemática.

Florianópolis, Santa Catarina, Março de 2011

Universidade Federal de Santa Catarina – UFSC
Centro de Ciências Físicas e Matemáticas
Departamento de Matemática
Curso de Especialização em Matemática - Formação do Professor de Matemática na modalidade a Distância

AJUSTE DE CURVAS POR MÍNIMOS QUADRADOS

Ademir de Lima

Monografia submetida à Comissão de avaliação do curso de Especialização em Matemática-Formação do professor em cumprimento parcial para a obtenção do título de Especialista em Matemática.

Aprovado em 25/03/11

Banca Examinadora

____

Prof. Dr. Roberto Corrêa da Silva (Orientador)

____

Prof. Dr. Marcio Fernandes (Examinador)

____

Prof. Drª. Silvia Martini de Holanda Janesch (Examinadora)

Dedico este trabalho a minha esposa Viviani e a minha filha Allana Vitória pelo incentivo e apoio nos momentos de maior dificuldade e principalmente a mim, pelo empenho e dedicação.

Agradecimentos

Agradeço a Deus, pela vida e pela oportunidade de chegar até aqui com forças para buscar novos desafios.

Agradeço o apoio incondicional de todos que me auxiliaram na conclusão desse trabalho em especial ao meu orientador Prof. Dr. Roberto Corrêa da Silva e ao tutor do Pólo de Foz do Iguaçu Gilberto por depositar em mim sua confiança e ter-me apoiado em momentos difíceis.

E também aos meus colegas do Pólo de Foz do Iguaçu pela força e incentivo, pois sem a união de todos nos grupos de estudos nada disso seria possível.

“É muito melhor arriscar coisas grandiosas, alcançar triunfos e glórias, mesmo espondo-se a derrota, do que formar fila com os pobres de espírito que nem gozam muito e nem sofrem muito, porque vivem nessa penumbra cinzenta que nem conhece vitória e nem derrota”
(Theodore Roosevelt)

Resumo

Com esse trabalho é possível observar, como obter-se estimativas para os valores de uma função em pontos não tabulados, quando são fornecidos os dados experimentais que geralmente formam um sistema inconsistente. A razão provável de esse sistema ser inconsistente e nenhuma função se ajustar precisamente aos dados, é a presença de erros, ocasionados por falhas de observações ou por equipamentos com defeitos. Resolveremos sistemas inconsistente utilizando o *Método dos Mínimos Quadrados* (MMQ), e que também tal método nos permite resolver problemas de ajuste de curvas, ou seja, "ajustar" a melhor curva polinomial de aproximação, quando o erro envolvido for a soma dos quadrados das diferenças entre os valores de y na curva de aproximação e os valores de y dados. Assim o *Método dos Mínimos Quadrados* é o procedimento mais adequado para determinar melhores aproximações e fazer *ajuste de curvas*.

Palavras-chaves: Sistema Inconsistente, Método dos Mínimos Quadrados, Melhores Aproximações, Ajuste de Curvas

Sumário

Índice de Figuras

Introdução

Responder as perguntas, “Como a partir de um conjunto de pontos dados num plano cartesiano, encontrar uma função que passe, ou mais se aproxime desses pontos?”, ou ainda “Como encontrar uma solução aproximada para um sistema incompatível, já que não *existe* solução?”, foi o que motivou-me a desenvolver o trabalho sobre *Ajuste de Curvas por Mínimos Quadrados.*

Em 1809, Carl Friedrich Gauss (1777-1855) publicou um artigo demonstrando que a *melhor maneira de determinar um parâmetro desconhecido de uma equação de condições é minimizando a soma dos quadrados dos resíduos*, sendo posteriormente denominado de *Mínimos Quadrados* por Adrien-Marie Legendre (1752-1833). Em abril de 1810, Pierre-Simon Laplace (1749-1827) apresenta a generalização a problemas com vários parâmetros desconhecidos.

No capítulo 1 trataremos da fundamentação teórica, falaremos sobre Espaços e Subespaços Vetoriais e suas propriedades, Espaço-solução de um Sistema Linear, Base e Dimensão de um espaço vetorial e então, Produto Interno, onde trataremos de distância entre vetores, Bases Ortogonais e Ortonormais e finalmente falaremos sobre o Processo de Gram-Schmidt que é um método para encontrar uma base ortonormal a partir de uma base qualquer de um espaço vetorial.

No capítulo 2, poderemos observar que Projeção Ortogonal, é parte importante do trabalho, pois quando tratamos de encontrar a solução de um sistema $Ax = b$, por Mínimos Quadrados, estamos procurando um vetor $\|Ax - b\|$ que seja ortogonal a b. Abordaremos Complemento Ortogonal e também verificaremos a relação fundamental entre o espaço-nulo e o espaço-linha de uma matriz.

Capítulo 3, trataremos nesse capítulo de sistemas incompatíveis e como projeções ortogonais podem ser usadas para resolver problemas de aproximações. Temos por objetivo encontrar solução desses sistemas utilizando o Método dos Mínimos Quadrados, de modo que, a diferença entre a solução por mínimos quadrados e a solução “exata” do sistema seja a menor possível.

No capítulo 4, mostraremos como fazer ajuste de curvas a um conjunto de pontos discretos no plano cartesiano utilizando o método dos Mínimos Quadrados. A aproximação por mínimos quadrados consiste em, dados um “tipo de função” e um conjunto de pontos, encontrar entre estas funções, a que melhor se aproxima do conjunto de pontos. Após ter utilizado esse método para fazer o ajuste dos pontos a um polinômio de grau um, podemos generalizar esse método para encontrar polinômios de grau $n \geq 2$, e ainda para sistemas não lineares.

1. ESPAÇOS VETORIAIS E PRODUTO INTERNO

Abordaremos nesse capítulo definições sobre espaço vetorial, subespaço vetorial, base e dimensão, por ser necessário para um bom entendimento do trabalho. Contudo estaremos interessados em formalizar os conceitos de produto interno e a partir dessa formalização definir as noções de comprimento, distância e de ângulos em *espaço vetoriais*. Com isso teremos processos para que possamos medir num *espaço vetorial*, da mesma forma pela qual se mede no plano $\mathbb{R}^2$.

1.1- Espaços Vetoriais

Seja um conjunto V, não-vazio, sobre o qual estão definidas as operações adição e multiplicação por escalar, isto é:

$\forall u, v \in V, u + v \in V$

$\forall \alpha \in \mathbb{R}, \forall u \in V, \alpha u \in V$

O conjunto V com essas duas operações é chamado de *espaço vetorial real* (ou espaço vetorial sobre $\mathbb{R}$) se forem verificados os seguintes axiomas:
Para diferenciar do número real zero (0) utilizaremos o símbolo (o) para o vetor nulo.

A) Em relação à adição:

A_1) $(u+v)+w = u+(v+w), \forall u, v, w \in V$ (associatividade)

A_2) $u+v = v+u, \forall u, v \in V$, (comutatividade)

A_3) $\exists o \in V; \forall u \in V,\ u + o = u \in V$, (existência de vetor nulo)

A_4) $\forall u \in V, \exists (-u) \in V; u + (-u) = o$, (existência de vetor oposto)

M) Em relação à multiplicação por escalar:

M_1) $(\alpha\beta)u = \alpha(\beta u)$, (associatividade)

M_2) $(\alpha+\beta)u = \alpha u + \beta u$, (distributiva a esquerda por um vetor)

M_3) $\alpha(u+v) = \alpha u + \alpha v$, (distributiva a esquerda por um escalar)

M_4) $1u = u$, (existência de elemento neutro)

Para $\forall u, v \in V$ e $\forall \alpha, \beta \in \mathbb{R}$.

Exemplo 1.1: O conjunto $V = \mathbb{R}^2 = \{(x, y) / x, y \in \mathbb{R}\}$ é um espaço vetorial com as operações de adição e multiplicação por um número real assim definidas: $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ e $\alpha(x, y) = (\alpha x, \alpha y)$. Essas são as operações usuais de adição e multiplicação por escalar.

Para verificarmos os oito axiomas de espaço vetorial, consideremos $u = (x_1, y_1)$, $v = (x_2, y_2)$ e $w = (x_3, y_3)$. Tem-se:

A1) $$\begin{aligned}(u+v)+w &= ((x_1, y_1) + (x_2, y_2)) + (x_3, y_3)\\ &= ((x_1 + x_2, y_1 + y_2)) + (x_3, y_3)\\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3)\\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3))\\ &= (x_1, y_1) + ((x_2 + x_3, y_2 + y_3))\\ &= (x_1, y_1) + ((x_2, y_2) + (x_3, y_3))\\ &= u + (v + w)\end{aligned}$$

A2) $$\begin{aligned}u + v &= (x_1, y_1) + (x_2 + y_2)\\ &= (x_1 + x_2, y_1 + y_2)\\ &= (x_2 + x_1, y_2 + y_1)\\ &= (x_2, y_2) + (x_1, y_1)\\ &= v + u\end{aligned}$$

A3) $$\exists o = (0,0) \in \mathbb{R}^2, \forall u \in \mathbb{R}^2;\ \begin{aligned}[t] u + o &= (x_1, y_1) + (0,0)\\ &= (x_1 + 0, y_1 + 0)\\ &= (x_1, y_1)\\ &= u\end{aligned}$$

A4) $$\forall u = (x_1, y_1) \in \mathbb{R}^2, \exists(-u) = (-x_1, -y_1) \in \mathbb{R}^2;\ \begin{aligned}[t] u + (-u) &= (x_1, y_1) + (-x_1, -y_1)\\ &= (x_1 - x_1, y_1 - y_1)\\ &= (0,0) = o\end{aligned}$$

M1) $\alpha\beta(u) = \alpha\beta(x_1, y_1)$

$$= ((\alpha\beta)x_1, (\alpha\beta)y_1)$$

$$= (\alpha(\beta x_1), \alpha(\beta y_1))$$

$$= \alpha(\beta x_1, \beta y_1)$$

$$= \alpha(\beta u)$$

M2) $(\alpha + \beta)u = (\alpha + \beta)(x_1, y_1)$

$$= ((\alpha + \beta)x_1, (\alpha + \beta)y_1)$$

$$= (\alpha x_1 + \beta x_1, \alpha y_1 + \beta y_1)$$

$$= (\alpha x_1, \alpha y_1) + (\beta x_1, \beta y_1)$$

$$= \alpha(x_1, y_1) + \beta(x_1, y_1)$$

$$= \alpha u + \beta u$$

M3) $\alpha(u + v) = \alpha((x_1, y_1) + (x_2, y_2))$

$$= \alpha(x_1 + x_2, y_1 + y_2)$$

$$= (\alpha(x_1 + x_2), \alpha(y_1 + y_2))$$

$$= (\alpha x_1 + \alpha x_2, \alpha y_1 + \alpha y_2)$$

$$= (\alpha x_1, \alpha y_1) + (\alpha x_2, \alpha y_2)$$

$$= \alpha(x_1, y_1) + \alpha(x_2, y_2)$$

$$= \alpha u + \alpha v$$

M4) $1u = 1(x_1, y_1)$

$$= (1x_1, 1y_1)$$

$$= (x_1, y_1)$$

$$= u$$

Exemplo 1.2: Os conjuntos $\mathbb{R}^3, \mathbb{R}^4, ..., \mathbb{R}^n$ são espaços vetoriais com as operações de adição e de multiplicação por escalar usuais. A verificação dos conjuntos acima citados é de forma análoga ao que foi feita para o $\mathbb{R}^2$.

1.2- Subespaços Vetoriais

Definição 1.2.1: Dado um espaço vetorial V, um subconjunto W, não vazio será um *subespaço vetorial de* V se:

i) Para quaisquer u e $v \in W$ tivermos $u+v \in W$.

ii) Para quaisquer $\alpha \in \mathbb{R}, u \in W$ tivermos $\alpha u \in W$.

Podemos fazer três observações:

a) As condições da definição acima garantem que ao operarmos em W (soma e multiplicação por escalar), não obteremos um vetor fora de W. Isto é suficiente para afirmar que W é ele próprio um espaço vetorial, pois assim as operações ficam bem definidas e, além disso, não precisamos verificar as propriedades de $i)$ a $viii)$ de espaço vetorial, por que elas são validas em V que contém W.

b) Qualquer subespaço W de V precisa necessariamente conter o vetor nulo (por causa da condição $ii)$ quando $\alpha = 0$).

c) Todo espaço vetorial admite pelo menos dois subespaços (que são chamados de subespaços triviais), o conjunto formado apenas pelo vetor nulo e o próprio espaço vetorial.

Às vezes, é necessário detectar, dentro de um espaço vetorial V, subconjuntos W que sejam eles próprios espaços vetoriais "menores". Tais conjuntos serão chamados de subespaços de V. Isto acontece, por exemplo, em $V = \mathbb{R}^2$, o plano, onde W é uma reta deste plano, que passa pela origem.

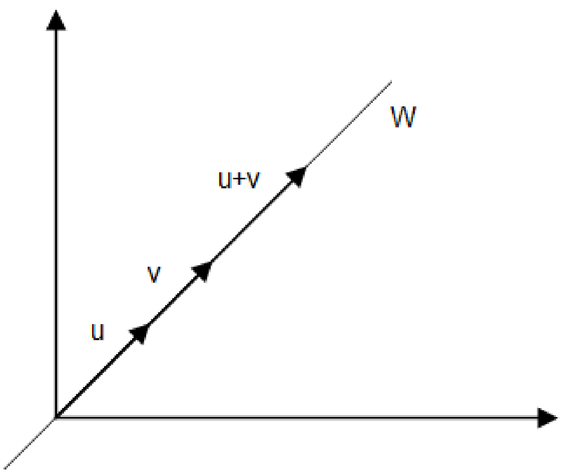

Figura 1: W é um subespaço vetorial de $\mathbb{R}^2$.

Veja que a reta W, funciona sozinha como espaço vetorial, pois, ao somarmos dois vetores de W e se multiplicarmos um vetor em W por um número, o vetor resultante ainda estará em W. Isto é, o subconjunto W é “fechado” em relação á soma de vetores e multiplicação deste por um escalar. Estas são as condições exigidas para que um subconjunto W de um espaço vetorial V seja um subespaço.

Exemplo 1.2.1: Sejam $V = \mathbb{R}^2$ e $S = \{(x, y) \in \mathbb{R}^2 \,/\, y = 2x\}$ ou $S = \{(x, 2x); x \in \mathbb{R}\}$, isto é, S é o conjunto dos vetores do plano que têm a segunda componente igual ao dobro da primeira.

Temos que $S \neq \phi$, pois $(0,0) \in S$.

Verificaremos agora as condições i) e ii).

Consideremos $u = (x_1, 2x_1) \in S$ e $v = (x_2, 2x_2) \in S$.

i) $u + v = (x_1, 2x_1) + (x_2, 2x_2) = (x_1 + x_2, 2x_1 + 2x_2) = (x_1 + x_2, 2(x_1 + x_2) \in S$, pois a segunda componente de $u + v$ é igual ao dobro da primeira.

ii) $\alpha u = \alpha(x_1, 2x_1) = (\alpha x_1, 2\alpha x_1) = (\alpha x_1, 2(\alpha x_1)) \in S$, pois a segunda componente é igual ao dobro da primeira.

Portanto S é um subespaço de $\mathbb{R}^2$.

Exemplo 1.2.2: $V = \mathbb{R}^3$ e $W \subset V$, um plano passando pela origem.

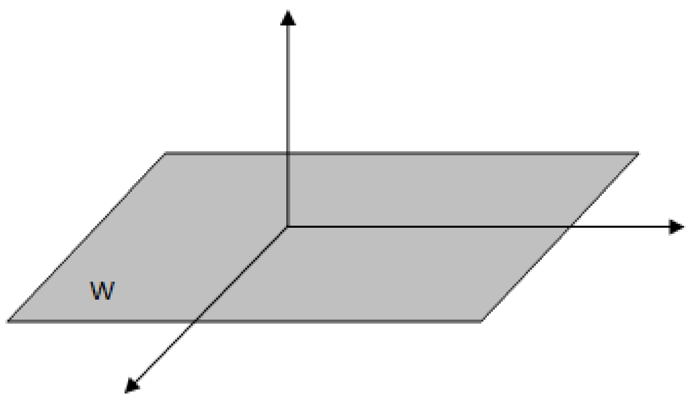

Figura 2: $V = \mathbb{R}^3$ e $W \subset V$ é um subespaço.

Podemos observar geometricamente a validade de i) e ii). Temos que se W não passasse pela origem, não seria um subespaço (pois pela condição ii) qualquer subespaço precisa necessariamente conter o vetor nulo, no caso de $\alpha = 0$) e ainda, os únicos subespaços do $\mathbb{R}^3$ são a origem, as retas e planos que passam pela origem, e o próprio $\mathbb{R}^3$.

Assim podemos ter a falsa impressão de que o espaço vetorial $\mathbb{R}^2$ é um subespaço de $\mathbb{R}^3$, contudo $\mathbb{R}^2$ não é nem mesmo um subconjunto de $\mathbb{R}^3$. (Os vetores do $\mathbb{R}^3$ têm todos três componentes enquanto os vetores do $\mathbb{R}^2$ têm apenas duas).

O conjunto $W = \{(x, y, 0); x, y \in \mathbb{R}\}$ é um subconjunto do $\mathbb{R}^3$ que se "parece" e "age" como o $\mathbb{R}^2$, apesar de ser diferente do $\mathbb{R}^2$. W é um subespaço do $\mathbb{R}^3$, pois o vetor nulo está em W e também é fechado em relação a soma de vetores e à multiplicação por escalar, porque estas operações em vetores de W, sempre produzem vetores cujas terceiras componentes são iguais a zero (e, assim, pertencem a W). Portanto W é um subespaço de $\mathbb{R}^3$.

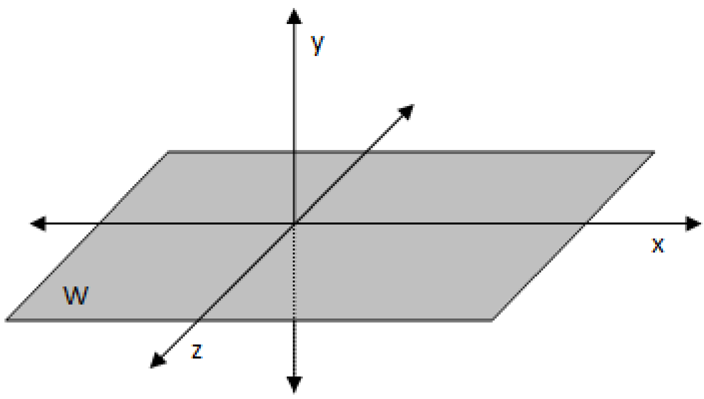

Figura 3: Um plano como subespaço do $\mathbb{R}^3$.

Se $Ax = b$ é um sistema de equações lineares, então cada vetor x que satisfaz esta equação é chamado um *vetor-solução* do sistema. O teorema a seguir mostra que os vetores solução de um sistema linear homogêneo formam um espaço vetorial. Chamamos o conjunto de todas as soluções deste sistema de *espaço-solução*.

Teorema 1.2.2: Se $Ax = 0$ é um sistema linear homogêneo de m equações em n incógnitas, então o conjunto dos vetores-solução é um subespaço vetorial de $\mathbb{R}^n$.

Demonstração: Consideremos W o conjunto dos vetores-solução de $Ax = 0$. Temos que em W existe pelo menos um vetor nulo (o) por se tratar de um sistema linear homogêneo. Para mostrar que W é um subespaço vetorial de $\mathbb{R}^n$, precisamos mostrar que W é fechado na adição e na multiplicação por escalar, para isso consideremos x e $x' \in W$, vetores-solução e $k \in \mathbb{R}$ um escalar qualquer.

De fato:

i) Se $Ax = 0$ e $Ax' = 0$, então, $A(x + x') = Ax + Ax' = 0 + 0 = 0$.

ii) Temos que, $A(kx) = k(Ax) = k0 = 0$.

Desse modo como x e $x' \in W$ e $k \in \mathbb{R}$ um escalar qualquer, mostramos que W é fechado na adição e na multiplicação por escalar e portanto W é um subespaço vetorial de $\mathbb{R}^n$.

■

Exemplo 1.2.3: Considere o sistema linear com três equações e três incógnitas $\begin{cases} x - 2y + 3z = 0 \\ 2x - 4y + 6z = 0 \\ 3x - 6y + 9z = 0 \end{cases}$, temos que o conjunto solução do sistema é um subespaço de $\mathbb{R}^3$, veja:

Transformando num sistema matricial temos, $\begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, onde:

$\begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix} \overset{\substack{2L_1+L_2 \to L_2 \\ 3L_1+L_3 \to L_3}}{\sim} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x - 2y + 3z = 0$ ou $x = 2y - 3z$, isto é, a solução do sistema é a equação de um plano que passa pela origem com vetor normal $n = (1, -2, 3)$. Segue do teorema 1.2.2 que o conjunto solução do sistema se trata de um subespaço vetorial de $\mathbb{R}^3$.

Temos que em um sistema homogêneo de três incógnitas as soluções formam subespaços vetoriais de $\mathbb{R}^3$. Geometricamente isto significa que cada espaço-solução deve ser uma reta que passa pela origem, um plano que passa pela origem, somente a origem ou todo o $\mathbb{R}^3$.

Definição 1.2.3: Sejam os vetores $v_1, v_2, ..., v_n$ do espaço vetorial V e os escalares $a_1, a_2, ..., a_n$. Qualquer vetor $v \in V$ da forma $v = a_1v_1 + a_2v_2 + ... + a_nv_n$ é uma *combinação linear* dos vetores $v_1, v_2, ..., v_n$.

Seja V um espaço vetorial. Consideremos um subconjunto $A = \{v_1, v_2, ..., v_n\} \subset V$, $A \neq \phi$.

O conjunto S de todos os vetores de V que são combinações lineares dos vetores de A é um subespaço vetorial de V.

Então podemos escrever $u = a_1v_1 + a_2v_2 + ... + a_nv_n$ e $v = b_1v_1 + b_2v_2 + ... + b_nv_n$, como sendo dois vetores de S, assim:

$$u + v = (a_1v_1 + a_2v_2 + ... + a_nv_n) + (b_1v_1 + b_2v_2 + ... + b_nv_n)$$

$$u + v = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + ... + (a_n + b_n)v_n \in S \quad \text{e}$$

$$\alpha u = (\alpha a_1)v_1 + (\alpha a_2)v_2 + ... + (\alpha a_n)v_n \in S$$

Assim, podemos concluir que S é um subespaço vetorial de V. De fato, o subespaço S é:

$$S = \{v \in V \,/\, v = a_1v_1 + a_2v_2 + ... + a_nv_n, \quad a_1, a_2, ..., a_n \in \mathbb{R}\}.$$

O subespaço S diz-se gerado pelos vetores $v_1, v_2, ..., v_n$, ou gerado pelo conjunto A, e representado por:

$S = [v_1, v_2, ..., v_n]$ ou $S = G(A)$.

Os vetores $v_1, v_2, ..., v_n$ são chamados geradores do subespaço S, enquanto A é o *conjunto gerador* de S. Para o caso particular de A somente conter o vetor nulo, define-se $[A] = \{o\}$.

Exemplo 1.2.4: Os vetores $i = (1,0)$ e $j = (0,1)$ geram o espaço vetorial $\mathbb{R}^2$, pois qualquer $(x, y) \in \mathbb{R}^2$ é combinação linear de i e j.

$$(x, y) = xi + yj = x(1,0) + y(0,1) = (x,0) + (0, y) = (x, y)$$

1.3- Base de um espaço vetorial

Definição 1.3.1: Se $S=\{v_1,v_2,...,v_n\}$ é um conjunto não-vazios de vetores, então a equação vetorial $a_1v_1+a_2v_2+...+a_nv_n=0$, tem pelo menos uma solução a saber, $a_1=0$, $a_2=0$,..., $a_n=0$. Se esta é a única solução, então o conjunto de vetores é *linearmente independente*. Se existem outras soluções, então S é um conjunto *linearmente dependente*.

Definição 1.3.2: Um conjunto $B=\{v_1,...,v_n\}\subset V$ é uma base do espaço vetorial V se:

i) B é Linearmente Independente (LI);

ii) B gera V.

Exemplo 1.3.1: $B=\{(1,1),(-1,0)\}$ é base do $\mathbb{R}^2$.

De fato:

i) B é LI, pois $a(1,1)+b(-1,1)=(0,0)$ implica em $\begin{cases} a-b=0 \\ a=0 \end{cases}$, assim $a=b=0$.

ii) B gera $\mathbb{R}^2$, pois para todo $(x,y)\in\mathbb{R}^2$, tem-se:

$(x,y)=a(1,1)+b(-1,0) \quad \Rightarrow \quad \begin{cases} a-b=x \\ a=y \end{cases}$, donde $a=y$ e $b=y-x$, o que implica em $(x,y)=y(1,1)+(y-x)(-1,0)$.

1.4- Dimensão de um Espaço Vetorial

Definição 1.4.1: A *Dimensão* de um espaço vetorial de dimensão finita V é definida como o número de vetores de uma base de V e denotada por $\dim(V)$. Além disso, definimos o espaço vetorial nulo como tendo dimensão zero.

Assim se V possui uma base com n vetores, então V tem dimensão n e anota-se $\dim V=n$.

Exemplo 1.4.1: Seja $W=\{(x,y,z)\in\mathbb{R}^3 / 2x-3y+4z=0\}$ um subespaço vetorial de $V=\mathbb{R}^3$, vamos encontrar uma base e determinar a sua dimensão.

Para encontrar a base, vamos encontrar um conjunto gerador de W e verificar se é L I.

De fato;

$2x-3y+4z=0 \Rightarrow 3y=2x+4z \Rightarrow y=\frac{2}{3}x+\frac{4}{3}z$, assim podemos escrever W como $W=\{(x,\frac{2}{3}x+\frac{4}{3}z,z)\in \mathbb{R}^3\}$

$$W=\{(x,\frac{2}{3}x,0)+(0,\frac{4}{3}z,z)\in \mathbb{R}^3\}$$

$$W=\{x(1,\frac{2}{3},0)+z(0,\frac{4}{3},1)\in \mathbb{R}^3\}$$

$$W=[(1,\frac{2}{3},0)+(0,\frac{4}{3},1)]=B$$

Logo B é o conjunto gerador para S. Falta verificar se B é L I.

$$a(1,\frac{2}{3},0)+(0,\frac{4}{3},1)=(0,0,0)$$

$$\Rightarrow \begin{cases} a=0 \\ \frac{2}{3}a+\frac{4}{3}a=0 \\ b=0 \end{cases} \Rightarrow \quad a=b=0$$, como o sistema tem solução única é L I.

Portanto B é base para W e ainda como a base possui dois vetores, isso implica que $\dim W=2$.

1.5- Produto Interno

Definição 1.5.1: Um *produto interno* num espaço vetorial V, é uma função que, a cada par de vetores u e v em V associa um número real $<u,v>$ de tal maneira que os seguintes axiomas são satisfeitos por quaisquer vetores u,v e w de V e qualquer escalar α.

i) $<u,v>=<v,u>\forall u,v\in V$ (axioma de simetria)

ii) $<u+v,w>=<u,w>+<v,w>\forall u,v,w\in V$ (axioma de aditividade)

iii) $<\alpha u,v>=\alpha<u,v>\forall \alpha\in R$ e $u,v\in V$ (axioma de homogeneidade)

iv) $<v,v>\geq 0$ e $<v,v>=0 \Leftrightarrow v=o$ (axioma de positividade)

Um espaço vetorial real com produto interno é chamado *espaço com produto interno real*.

Obs.: O produto interno euclidiano para o $\mathbb{R}^n$ deve ser entendido como produto interno usual para o espaço $\mathbb{R}^n$.

Exemplo 1.5.1: O produto escalar usual de vetores do espaço $\mathbb{R}^3$ para $v=(x_1,x_2,x_3)$ e $w=(y_1,y_2,y_3)$ é $<v,w>=x_1y_1+x_2y_2+x_3y_3$. De modo análogo, define-se o que chamamos de *produto interno usual* para o espaço $\mathbb{R}^n$.

De fato:

Dados $v=(x_1,x_2,...,x_n)$ e $w=(y_1,y_2,...,y_n)$ $\Rightarrow <v,w>=x_1y_1+x_2y_2+...+x_ny_n$

Exemplo 1.5.2: Sejam $V=P_2$, $p=a_2x^2+a_1x+a_0$ e $q=b_2x^2+b_1x+b_0$ vetores quaisquer de P_2. A fórmula $<p,q>=a_2b_2+a_1b_1+a_0b_0$ define um produto interno em P_2. Por exemplo, se:

$p=2x^2-6x+7$ e $q=3x^2+3x+5$, então, $<p,q>=2.3-6.3+7.5=23$

Podemos notar que $<p,q>=a_2b_2+a_1b_1$, não define sobre V, um produto interno, pois existem polinômios $p\in V$ tais que $<p,p>=0$ sem que $p=0$. Por exemplo, se $p=0x^2+0x+5$. Isso acarreta uma falha no axioma *iv.*

O produto interno euclidiano é o mais importante dos produtos internos, contudo existem outros produtos internos relevantes como o *Produto interno euclidiano ponderado*. Mais precisamente, se $w_1,w_2,...,w_n$, são números reais positivos, que designaremos por *pesos* e se $u=(x_1,x_2,...,x_n)$ e $v=(y_1,y_2,...,y_n)$ são vetores do $\mathbb{R}^n$, então:

$$<u,v>=w_1x_1y_1+w_2x_2y_2+...+w_nx_ny_n,$$

define um produto interno no $\mathbb{R}^n$, denominado de *Produto interno euclidiano ponderado com pesos* $w_1,w_2,...,w_n$.

Para entendermos melhor uma maneira pela qual pode surgir um produto euclidiano ponderado, suponha que um experimento físico possa produzir qualquer um de n valores numéricos $x_1,x_2,...,x_n$ e que m repetições do experimento fornecem esses valores com várias frequências, ou seja, que x_1 ocorre f_1 vezes, x_2 ocorre f_2 vezes e assim por diante. Como há um total de m repetições do experimento, temos $f_1 + f_2 + ... + f_n = m$.

Assim a média aritmética dos valores observados ($\bar{x}$) é;

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + ... + f_n x_n}{f_1 + f_2 + ... + f_n} = \frac{1}{m}(f_1 x_1 + f_2 x_2 + ... + f_n x_n)$$

Se escrevermos $f = (f_1, f_2, ..., f_n)$, $x = (x_1, x_2, ..., x_n)$ e $w_1 = w_2 = ... = w_n = \frac{1}{m}$, então a fórmula da média aritmética pode ser escrita como o produto interno ponderado;

$$\bar{x} = < f, x > = w_1 f_1 x_1 + w_2 f_2 x_2 + ... + w_n f_n x_n$$

Exemplo 1.5.3: Sejam $u = (x_1, x_2)$ e $v = (y_1, y_2)$ vetores do $\mathbb{R}^2$. O produto interno euclidiano ponderado $< u, v > = 3x_1 y_1 + 2x_2 y_2$, satisfaz os quatro axiomas de produto interno.

De fato:

i) $< u, v > = < (x_1, x_2), (y_1, y_2) > = 3x_1 y_1 + 2x_2 y_2 = 2x_2 y_2 + 3x_1 y_2 = < v, u >$

ii) Se $w = (z_1, z_2)$, então,

$$\begin{aligned} < u + v, w > &= < (x_1, x_2) + (y_1, y_2), (z_1, z_2) > \\ &= 3(x_1 + y_1).z_1 + 2(x_2 + y_2).z_2 \\ &= 3x_1 z_1 + 3y_1 z_1 + 2x_2 z_2 + 2y_2 z_2 \\ &= (3x_1 z_1 + 2x_2 z_2) + (3y_1 z_1 + 2y_2 z_2) \\ &= < u, w > + < x, w > \end{aligned}$$

iii) Seja $\alpha \in \mathbb{R}$, então:

$$\begin{aligned} < \alpha u, v > &= < \alpha(x_1, x_2), (y_1, y_2) > \\ &= 3(\alpha x_1) y_1 + 2(\alpha x_2) y_2 \\ &= \alpha(3x_1 y_1) + \alpha(2x_2 y_2) \\ &= \alpha(3x_1 y_1 + 2x_2 y_2) \\ &= \alpha < u, v > \end{aligned}$$

iv) $<v,v>=3y_1y_1+2y_2y_2=3y_1^2+2y_2^2$, assim $<v,v>\geq 0$ para qualquer $v\in\mathbb{R}^2$ e também, $<v,v>=3y_1^2+2y_2^2=0$ se, e somente se $y_1=y_2=0$ o que vai implicar em $v=o$, o que satisfaz o axioma iv.

Vamos formalizar agora a noção de comprimento e distância em espaços com produto interno

Definição 1.5.2: Se V é um espaço com produto interno, então a *norma* (ou *comprimento*) de um vetor $v\in V$ é denotada por $\|v\|$ e é definida por

$$\|v\|=<v,v>^{1/2}$$

A distância entre dois vetores u e v é denotada por $d(u,v)$ e é definida por

$$d(u,v)=\|u-v\|$$

Se $\|v\|=1$, isto é, $<v,v>^{1/2}=1$, v é chamado de *vetor unitário*. Dizemos nesse caso que v está *normalizado*.

Observe que todo vetor não nulo $v\in V$ pode ser *normalizado*, basta multiplicar v por $\dfrac{1}{\|v\|}$, assim teremos um vetor u tal que $u=\dfrac{v}{\|v\|}$. Consideremos por exemplo, $V=\mathbb{R}^3$ e $<\,,\,>$ produto interno usual, então se $v=(x_1,x_2,x_3)\in\mathbb{R}^3$, $\|v\|=\sqrt{<v,v>}=\sqrt{x_1^2+x_2^2+x_3^2}$ que é o comprimento do vetor v. Assim para $v=(1,2-1)$ teremos o vetor normalizado

$$u=\frac{v}{\|v\|}=\frac{(1,2,-1)}{\sqrt{1^2+2^2+(-1)^2}}=\left(\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{-1}{\sqrt{6}}\right).$$

Exemplo 1.5.4: De maneira análoga podemos estender para o $\mathbb{R}^n$. Considere $u=(x_1,x_2,...,x_n)$ e $v=(y_1,y_2,...,y_n)$ dois vetores do $\mathbb{R}^n$ com produto interno euclidiano.O comprimento do vetor u será:

$$\|u\|=<u,u>^{1/2}=\sqrt{x_1^2+x_2^2+...+x_n^2} \text{ e}$$

$$d(u,v)=\|u-v\|=<(u-v).(u-v)>^{1/2}=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+...+(x_n-y_n)^2}\ .$$

Exemplo 1.5.5: É importante salientar que norma e distância dependem do produto interno que se está sendo usado. Se o produto interno for mudado, então o comprimento e a distância também mudam. Vejamos para o caso de $u=(3,4)$ e $v=(2,-1)$ do $\mathbb{R}^2$ com o produto interno euclidiano nós temos:

$$\|u\|=\sqrt{3^2+4^2}=\sqrt{25}=5 \text{ e}$$

$$d(u,v)=\sqrt{(3-2)^2+(4+1)^2}=\sqrt{1+25}=\sqrt{26}$$

Agora, se mudarmos para o produto interno euclidiano ponderado $<u,v>=3x_1y_1+2x_2y_2$, então teremos:

$$\|u\|=<u,u>^{1/2}=\sqrt{3.3.3+2.4.4}=\sqrt{27+32}=\sqrt{59} \text{ e}$$

$$d(u,v)=\|u-v\|=<(u-v),(u-v)>^{1/2}=<(1,5).(1,5)>^{1/2}=\sqrt{3.1.1+2.5.5}=\sqrt{3+50}=\sqrt{53}.$$

Se V é um espaço com produto interno, então o conjunto dos pontos de V que satisfazem $\|u\|=1$ é chamado de *esfera unitária*, ou então, *círculo unitário* de V. No $\mathbb{R}^2$ e $\mathbb{R}^3$ estes são os pontos que distam uma unidade da origem.

Exemplo 1.5.6: a) Vejamos como é o esboço gráfico de um círculo unitário num sistema de coordenadas xy em $\mathbb{R}^2$, usando o produto interno euclidiano $<u,v>=x_1y_1+x_2y_2$.

Se $u=(x,y)$ então, $\|u\|=<u,u>^{1/2}=\sqrt{x^2+y^2}$, de modo que a equação do círculo unitário é $\sqrt{x^2+y^2}=1 \quad \Rightarrow \quad x^2+y^2=1$. Como é de se esperar o gráfico desta equação é um círculo de raio 1 centrado na origem (figura 4).

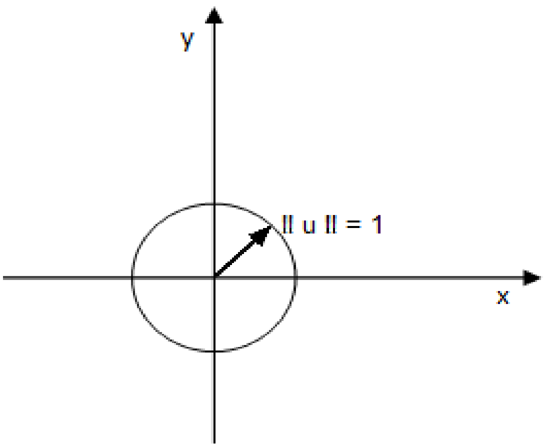

Figura 4: O círculo unitário com a norma euclidiana $\|u\|=\sqrt{x^2+y^2}$.

b) Agora se mudarmos para o produto interno euclidiano ponderado $< u,v >= \frac{x_1 y_1}{9} + \frac{x_2 y_2}{4}$ o esboço do círculo unitário no sistema de coordenadas xy em $\mathbb{R}^2$ será:

Se $u = (x, y)$ então $\|u\| = < u,u >^{1/2} = \sqrt{\frac{x^2}{9} + \frac{y^2}{4}} = 1$, elevando ao quadrado os dois membros da equação temos, $\frac{x^2}{9} + \frac{y^2}{4} = 1$. O gráfico desta equação é uma elipse mostrada na figura 5.

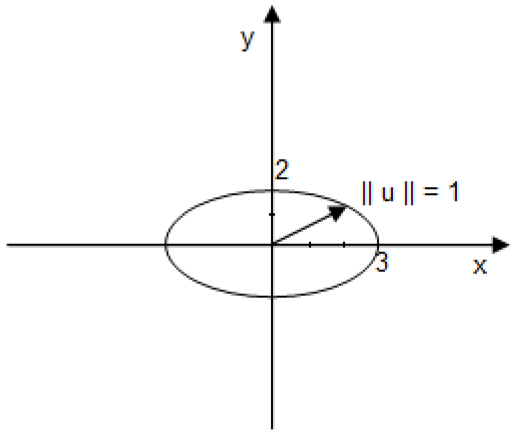

Figura 5: O círculo Unitário com norma $\|u\| = \sqrt{\frac{x^2}{9} + \frac{y^2}{4}}$.

Teorema 1.5.3: Se u, v e w são vetores em um espaço com produto interno real e α é um escalar qualquer então são validas as seguintes propriedades:

a) $< 0,v >=< v,0 >= 0$

b) $< u,v+w >=< u,v > + < u,w >$

c) $< u,\alpha v >= \alpha < u,v >$

d) $< u-v,w >=< u,w > - < v,w >$

e) $< u,v-w >=< u,v > - < u,w >$

Demonstração:

a) $< 0,v >=< 0v,v > \overset{iii)}{=} 0 < v,v >= 0$, como $< v,0 >=< 0,v >$ então $< u,0 >= 0$

b) $< u,v+w > \overset{i)}{=} < v+w,u > \overset{ii)}{=} < v,u > + < w,u > \overset{i)}{=} < u,v > + < u,w >$

c) $< u,\alpha v > \overset{i)}{=} < \alpha v,u > \overset{iii)}{=} \alpha < v,u > \overset{i)}{=} \alpha < u,v >$

d) $< u-v,w > \overset{ii)}{=} < u+(-v),w > = < u,w > + < (-v),w > = < u,w > - < v,w >$

e) $< u,v-w > = < u,v+(-w) > = < u,v > + < u,(-w) > = < u,v > - < u,w >$

■

Teorema 1.5.4: Seja V um espaço vetorial com produto interno. Para quaisquer v, w em V e $\alpha \in \mathbb{R}$.

i) $\|v\| \geq 0$ e $\|v\| = 0$ se, e somente se, $v = o$.

ii) $\|\alpha v\| = |\alpha|.\|v\|$

iii) $|< v,w >| \leq \|v\|\,\|w\|$ (*desigualdade de Cauchy-Schwarz*)

iv) $\|v+w\| \leq \|v\| + \|w\|$ (*desigualdade triangular*)

Demonstração:

i) Por definição temos $\|v\| \geq 0$. Por outro lado,

$$\|v\| = 0 \Leftrightarrow < v,v >^{1/2} = 0 \Leftrightarrow \left((< v,v >)^{1/2}\right)^2 = (0)^2 \Leftrightarrow < v,v > = 0 \Leftrightarrow v = o.$$

ii) $\|\alpha v\| = \sqrt{< \alpha v, \alpha v >} = \sqrt{\alpha^2 < v,v >} = \sqrt{\alpha^2}\sqrt{< v,v >} = |\alpha|.\|v\|$.

iii) Sejam v e w em V com $v \neq 0$. (Para $v = o$ vale a igualdade $|< v,w >| = \|v\|\,\|w\| = 0$).

De fato, $\|v\| = \sqrt{< v,v >} \Rightarrow \|v\|^2 = < v,v >$, assim por definição para qualquer $\alpha \in \mathbb{R}$ temos $\|\alpha v + w\| \geq 0$, ou seja;

$$\begin{aligned}\|\alpha v + w\|^2 &= < (\alpha v + w),(\alpha v + w) > \\ &= < v,v > \alpha^2 + 2 < v,w > \alpha + < w,w > \geq 0\end{aligned}$$

Temos então um trinômio do 2º grau que deve ser positivo para qualquer valor de α. Como o coeficiente $< v,v >$ de α^2 é sempre positivo $(v \neq o)$, o discriminante Δ deve ser negativo: $\Delta = (2 < v,w >)^2 - 4 < v,v >< w,w > \leq 0$.

Isso significa que,

$$4 < v,w >^2 - 4\|v\|^2 \|w\|^2 \le 0$$
$$\Rightarrow \quad 4 < v,w >^2 \le 4\|v\|^2 \|w\|^2$$
$$\Rightarrow \quad < v,w >^2 \le \|v\|^2 \|w\|^2$$
$$\Rightarrow \quad \sqrt{< v,w >^2} \le \sqrt{\|v\|^2 \|w\|^2}$$

E portanto, $|< v,w >| \le \|v\|.\|w\|$, que é o que queríamos provar.

iv) $\|v+w\|^2 =< v+w, v+w > =< v,v > + < v,w > + < w,v > + < w,w >$

$$= \|v\|^2 + \|w\|^2 + 2 < v,w > \le \|v\|^2 + \|w\|^2 + 2\|v\|\,\|w\| = (\|v\| + \|w\|)^2$$

Logo, $\|v+w\|^2 \le (\|v\| + \|w\|)^2$ que extraindo a raiz quadrada da ambos os lados podemos concluir que $\|v+w\| \le \|v\| + \|w\|$.

■

A desigualdade de Cauchy-Schwarz nos dá a possibilidade de definir ângulos entre dois vetores em um espaço vetorial V, munido de um produto interno arbitrário. Suponha v e w são vetores não nulos do espaço vetorial V. Se dividirmos ambos os lados da fórmula $< v,w >^2 \le \|v\|^2 \|w\|^2$ por $\|v\|^2 \|w\|^2$, (observamos que a desigualdade de Cauchy-Schwarz pode ser escrita sob a forma de $< v,w >^2 \le \|v\|^2 \|w\|^2$). Assim:

$$\frac{< v,w >^2}{\|v\|^2 \|w\|^2} \le \frac{\|v\|^2 \|w\|^2}{\|v\|^2 \|w\|^2} \Leftrightarrow \frac{< v,w >^2}{\|v\|^2 \|w\|^2} \le 1 \Leftrightarrow \left|\frac{< v,w >}{\|v\|\,\|w\|}\right| \le 1 \Leftrightarrow -1 \le \frac{< v,w >}{\|v\|\,\|w\|} \le 1$$

Ocorre que se θ é o ângulo cuja medida em radianos varia de 0 a π, então $\cos\theta$ toma qualquer valor entre -1 e 1, inclusive os extremos, uma única vez.

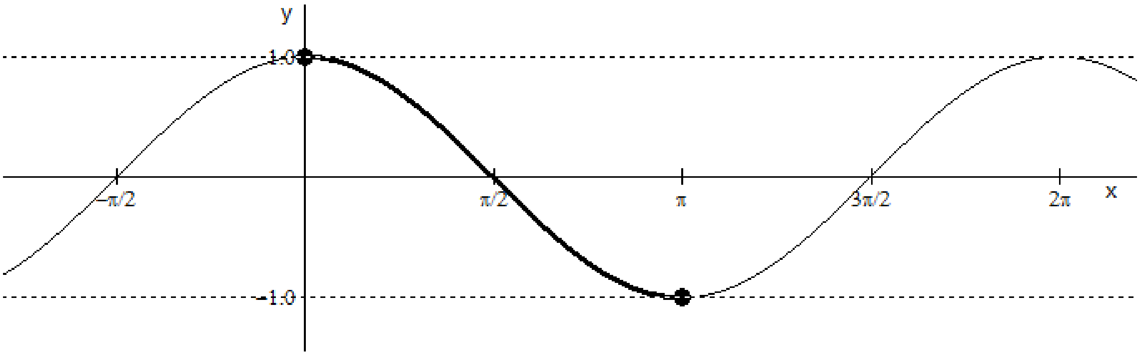

Figura 6: Se $0 \le \theta \le \pi$ então, $\cos\theta \in [-1,1]$.

Seguindo esse raciocínio temos que por $-1 \leq \frac{<v,w>}{\|v\|\,\|w\|} \leq 1$, existe um único ângulo θ tal que:

$$\cos\theta = \frac{<v,w>}{\|v\|\,\|w\|} \text{ e } 0 \leq \theta \leq \pi$$

Dessa maneira definimos θ como sendo o ângulo entre v e w.

Exemplo 1.5.7: Tomando o produto interno euclidiano em $\mathbb{R}^3$ os vetores $v=(1,2,3)$ e $w=(3,2,1)$, vamos calcular o ângulo θ entre eles.

De fato:

$\|v\| = \sqrt{1^2+2^2+3^2} = \sqrt{15}$, $\|w\| = \sqrt{3^2+2^2+3^2} = \sqrt{15}$ e $<v,w>= 1.3+2.2+3.1=10$ assim, $\cos\theta = \frac{<v,w>}{\|v\|\,\|w\|} = \frac{10}{\sqrt{15}.\sqrt{15}} = \frac{10}{15} = \frac{2}{3}$ o que implica, $\theta = arc\cos\frac{2}{3} \cong 48{,}19°$.

Com exceção do $\mathbb{R}^2$ e $\mathbb{R}^3$, há pouca necessidade de se calcular ângulos entre dois vetores, Contudo é de suma importância, verificar se dois ângulos são *ortogonais* entre si, ou seja, se o ângulo formado por eles é $\theta = \pi/2$.

Consideremos v e $w \in V$ dois vetores não nulos e θ o ângulo entre eles, então $\cos\theta = \frac{<v,w>}{\|v\|\,\|w\|} = 0$, se e somente se, $<v,w>=0$. Isso sugere a próxima definição.

Definição 1.5.3: Dois vetores v e w de um espaço vetorial com produto interno são chamados *ortogonais* se $<v,w>=0$, e é representado por $v \perp w$.

Utilizando a definição acima podemos provar que o Teorema de Pitágoras é válido para qualquer espaço com produto interno.

Teorema 1.5.6: Se v e w são vetores ortogonais em um espaço com produto interno, então:

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2$$

Demonstração:

A ortogonalidade de v e w implica em $<v,w>=0$, assim:

$$\begin{aligned}\|v+w\|^2 &= <v+w, v+w> \\ &= <v,v> + <v,w> + <w,v> + <w,w> \\ &= \|v\|^2 + 2<v,w> + \|w\|^2 \\ &= \|v\|^2 + \|w\|^2\end{aligned}$$

■

Exemplo 1.5.8: Seja $V=\mathbb{R}^2$ um espaço vetorial euclidiano em relação ao produto interno $<(x_1,y_1),(x_2,y_2)>=x_1x_2+2x_2y_2$. Em relação a esse produto interno os vetores $v=(-3,2)$ e $w=(4,3)$ são ortogonais, pois:

$$<v,w>=<(-3,2),(4,3)>=-3.4+2.2.3=-12+12=0$$

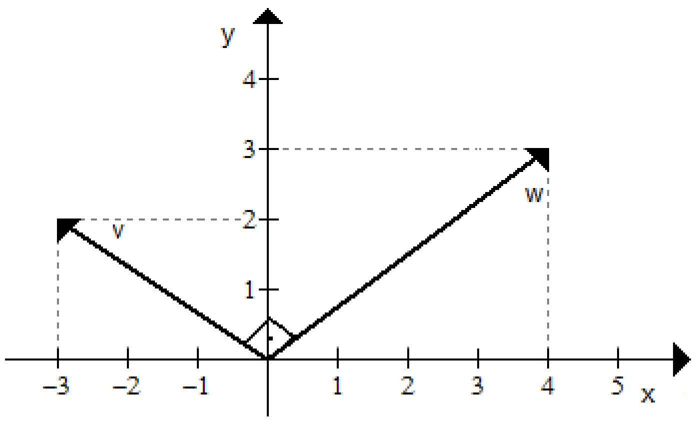

Figura 7: $v \perp w$.

Exemplo 1.5.9: Sejam $V=M_2$ as matrizes quadradas de ordem 2 reais e o produto interno dado pela expressão:

$$\left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right\rangle = ae+2bf+3cg+dh$$

Vamos calcular o ângulo entre as matrizes $\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}$ e $\begin{bmatrix} 1 & 4 \\ -1 & 0 \end{bmatrix}$, segundo este produto interno. Então:

$$\left\langle \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}, \begin{bmatrix} 1 & 4 \\ -1 & 0 \end{bmatrix} \right\rangle = 2.1+2.3.4+3.1.(-1)+(-2).0=2+24-3+0=23$$

$$\left\|\begin{bmatrix}2 & 3\\1 & -2\end{bmatrix}\right\| = \sqrt{\left\langle\begin{bmatrix}2 & 3\\1 & -2\end{bmatrix},\begin{bmatrix}2 & 3\\1 & -2\end{bmatrix}\right\rangle} = \sqrt{2.2+2.3^2+3.1^2+(-2)^2} = \sqrt{4+18+3+4} = \sqrt{29}$$

$$\left\|\begin{bmatrix}1 & 4\\-1 & 0\end{bmatrix}\right\| = \sqrt{\left\langle\begin{bmatrix}1 & 4\\-1 & 0\end{bmatrix},\begin{bmatrix}1 & 4\\-1 & -0\end{bmatrix}\right\rangle} = \sqrt{1.1+2.4^2+3.(-1)^2+0^2} = \sqrt{1+32+3} = \sqrt{36} = 6$$

$$\cos\theta = \frac{\left\langle\begin{bmatrix}2 & 3\\1 & -2\end{bmatrix},\begin{bmatrix}1 & 4\\-1 & 0\end{bmatrix}\right\rangle}{\left\|\begin{matrix}2 & 3\\1 & -2\end{matrix}\right\|\left\|\begin{matrix}1 & 4\\-1 & 0\end{matrix}\right\|} = \frac{23}{\sqrt{29}.6} = \frac{23}{6.\sqrt{29}} \Rightarrow \theta = arc\cos\frac{23}{6.\sqrt{29}} \cong 44{,}6°$$

1.6- Bases Ortonormais

Definição 1.6.1: Uma base $B = \{v_1, v_2, ..., v_n\}$ de um espaço vetorial euclidiano V é *ortonormal* se B é ortogonal e todos os seus vetores são unitários, isto é:

$$< v_i, v_j >= \begin{cases}0, & se \quad i \neq j\\1, & se \quad i = j\end{cases}$$

Exemplo 1.6.1: Em relação ao produto interno usual o conjunto $B = \{(1,0),(0,1)\}$ é uma base ortonormal do $\mathbb{R}^2$, pois:

$$< v_1, v_2 >=< (1,0),(0,1) >= 1.0 + 0.1 = 0$$

$$\|v_1\| = \|(1,0)\| =< (1,0),(1,0) >^{1/2} = \sqrt{1^2+0^2} = \sqrt{1} = 1$$

$$\|v_2\| = \|(0,1)\| =< (0,1),(0,1) >^{1/2} = \sqrt{0^2+1^2} = \sqrt{1} = 1$$

Exemplo 1.6.2: Seja o conjunto $B = \{(\frac{\sqrt{5}}{3}, \frac{2}{3}), (-\frac{2}{3}, \frac{\sqrt{5}}{3})\}$, podemos verificar que também é uma base ortonormal em relação ao produto interno euclidiano, pois:

$$< v_1, v_2 >=< (\frac{\sqrt{5}}{3}, \frac{2}{3}), (-\frac{2}{3}, \frac{\sqrt{5}}{3}) >= -\frac{\sqrt{5}}{3}.\frac{2}{3} + \frac{2}{3}.\frac{\sqrt{5}}{3} = -\frac{2\sqrt{5}}{9} + \frac{2\sqrt{3}}{9} = 0$$

$$\|v_1\| = \left\|(\frac{\sqrt{5}}{3}, \frac{2}{3})\right\| =< (\frac{\sqrt{5}}{3}, \frac{2}{3}) >^{1/2} = \sqrt{(\frac{\sqrt{5}}{3})^2 + (\frac{2}{3})^2} = \sqrt{\frac{5}{9}+\frac{4}{9}} = \sqrt{1} = 1$$

$$\|v_2\| = \left\|(-\frac{2}{3}, \frac{\sqrt{5}}{3})\right\| =< -\frac{2}{3}, \frac{\sqrt{5}}{3} >^{1/2} = \sqrt{(-\frac{2}{3})^2 + (\frac{\sqrt{5}}{3})^2} = \sqrt{\frac{4}{9}+\frac{5}{9}} = \sqrt{1} = 1$$

Exemplo 1.6.3: De maneira análoga ao exemplo 1.6.2, no espaço euclidiano $\mathbb{R}^3$ o conjunto $B=\{(1,0,0),(0,1,0),(0,0,1)\}$ é ortonormal. Vejamos por exemplo a norma de $\|(1,0,0)\| = <(1,0,0)>^{1/2} = \sqrt{1^2+0^2+0^2} = \sqrt{1} = 1$ e o produto interno $<(1,0,0),(0,1,0)> = 1.0+0.1+0.0 = 0$.

Em geral, para todo $n \geq 2$ o conjunto $B=\{(1,0,...,0),(0,1,...,0),...,(0,0,...,1)\}$ é ortonormal no espaço euclidiano $\mathbb{R}^n$. A verificação é análoga ao exemplo 1.6.2.

Já vimos que se $\|v\|=1$, isto é, $<v,v>^{1/2}=1$, v é chamado de *vetor unitário*. Dizemos nesse caso que v está *normalizado*. E ainda que todo vetor não nulo $v \in V$ pode ser *normalizado*, basta multiplicar v por $\frac{1}{\|v\|}$, assim teremos um vetor u tal que $u=\frac{v}{\|v\|}$.

Assim, uma base ortonormal sempre pode ser obtida de uma base ortogonal, normalizando cada vetor.

Exemplo 1.6.4: Seja a base ortogonal $B=\{v_1,v_2,v_3\}$, onde $v_1=(1,-1,3)$, $v_2=(2,2,0)$ e $v_3=(-1,1,\frac{2}{3})$ em relação ao produto interno euclidiano. Normalizando cada vetor, teremos:

$$u_1=\frac{v_1}{\|v_1\|}=\frac{(1,-1,3)}{\sqrt{1^2+(-1)^2+3^2}}=\frac{(1,-1,3)}{\sqrt{11}}=(\frac{1}{\sqrt{11}},-\frac{1}{\sqrt{11}},\frac{3}{\sqrt{11}})$$

$$u_2=\frac{v_2}{\|v_2\|}=\frac{(2,2,0)}{\sqrt{2^2+2^2+0^2}}=\frac{(2,2,0)}{\sqrt{8}}=(\frac{2}{\sqrt{8}},\frac{2}{\sqrt{8}},0)$$

$$u_3=\frac{v_3}{\|v_3\|}=\frac{(-1,1,\frac{2}{3})}{\sqrt{(-1)^2+1^2+(\frac{2}{3})^2}}=\frac{(-1,1,\frac{2}{3})}{\sqrt{\frac{22}{9}}}=(-\frac{3}{\sqrt{22}},\frac{3}{\sqrt{22}},\frac{2}{\sqrt{22}})$$

Agora verificaremos se $B^{'} = \{u_1, u_2, u_3\}$, onde $u_1 = (\frac{1}{\sqrt{11}}, -\frac{1}{\sqrt{11}}, \frac{3}{\sqrt{11}})$, $u_2 = (\frac{2}{\sqrt{8}}, \frac{2}{\sqrt{8}}, 0)$ e $u_3 = (-\frac{3}{\sqrt{22}}, \frac{3}{\sqrt{22}}, \frac{2}{\sqrt{22}})$, é uma base ortonormal.

De fato:

i) Verificação da ortogonalidade entre os vetores.

$$< u_1, u_2 >=< (\frac{1}{\sqrt{11}}, -\frac{1}{\sqrt{11}}, \frac{3}{\sqrt{11}}), (\frac{2}{\sqrt{8}}, \frac{2}{\sqrt{8}}, 0) >= \frac{1}{\sqrt{11}}.\frac{2}{\sqrt{8}} - \frac{1}{\sqrt{11}}.\frac{2}{\sqrt{8}} + \frac{3}{\sqrt{11}}.0$$

$$= \frac{2}{\sqrt{88}} - \frac{2}{\sqrt{88}} + 0 = 0$$

$$< u_1, u_3 >=< (\frac{1}{\sqrt{11}}, -\frac{1}{\sqrt{11}}, \frac{3}{\sqrt{11}}), (-\frac{3}{\sqrt{22}}, \frac{3}{\sqrt{22}}, \frac{2}{\sqrt{22}}) >= -\frac{1}{\sqrt{11}}.\frac{3}{\sqrt{22}} - \frac{1}{\sqrt{11}}.\frac{3}{\sqrt{22}} + \frac{3}{\sqrt{11}}.\frac{2}{\sqrt{22}}$$

$$= -\frac{3}{\sqrt{242}} - \frac{3}{\sqrt{242}} + \frac{6}{\sqrt{242}} = 0$$

$$< u_2, u_3 >=< (\frac{2}{\sqrt{8}}, \frac{2}{\sqrt{8}}, 0), (-\frac{3}{\sqrt{22}}, \frac{3}{\sqrt{22}}, \frac{2}{\sqrt{22}}) >= -\frac{2}{\sqrt{8}}.\frac{3}{\sqrt{22}} + \frac{2}{\sqrt{8}}.\frac{3}{\sqrt{22}} + 0.\frac{2}{\sqrt{22}}$$

$$= -\frac{6}{\sqrt{176}} + \frac{6}{\sqrt{176}} + 0 = 0$$

ii) Verificando se os vetores são unitários.

$$\|u_1\| =< u_1, u_1 >^{1/2} = \sqrt{(\frac{1}{\sqrt{11}})^2 + (-\frac{1}{\sqrt{11}})^2 + (\frac{3}{\sqrt{11}})^2} = \sqrt{\frac{1}{11} + \frac{1}{11} + \frac{9}{11}} = \sqrt{1} = 1$$

$$\|u_2\| =< u_2, u_2 >^{1/2} = \sqrt{(\frac{2}{\sqrt{8}})^2 + (\frac{2}{\sqrt{8}})^2 + (0)^2} = \sqrt{\frac{4}{8} + \frac{4}{8} + 0} = \sqrt{1} = 1$$

$$\|u_3\| =< u_3, u_3 >^{1/2} = \sqrt{(-\frac{3}{\sqrt{22}})^2 + (\frac{3}{\sqrt{22}})^2 + (\frac{2}{\sqrt{22}})^2} = \sqrt{\frac{9}{22} + \frac{9}{22} + \frac{4}{22}} = \sqrt{1} = 1$$

Portanto $B^{'} = \{u_1, u_2, u_3\}$ é uma base ortonormal.

Exemplo 1.6.5: Vamos construir a partir do vetor $v_1 = (1,2,-1)$ uma base ortogonal do $\mathbb{R}^3$ relativamente ao produto interno usual e obter, a partir dela uma base ortonormal.

De fato;

Seja $B = \{v_1, v_2, v_3\}$ a base ortogonal a ser determinada.

Seja $v_2 = (x, y, z)$, temos que v_2 deve ser ortogonal a $v_1 = (1,2,-1)$, então,

$< v_1, v_2 >= 0$,

$\Rightarrow < (1,2,-1),(x,y,z) >= 0$

$\Rightarrow \; x + 2y - z = 0$

$\Rightarrow \; x = -2y + z$

Assim existem infinitos vetores ortogonais a v_1 da forma $(-2y + z, y, z) / y, z \in \mathbb{R}$.

Fazendo $y = 0$ e $z = 1$ obtemos o vetor $v_2 = (1,0,1)$ que é ortogonal a v_1, pois $< v_1, v_2 >= 0$.

Ainda necessitamos de mais um vetor para obtermos uma base ortogonal. Seja $v_3 = (a,b,c)$ de maneira que $v_3 \perp v_1$ e $v_3 \perp v_2$, isto é;

$$\begin{cases} < v_3, v_1 >= 0 \\ < v_3, v_2 >= 0 \end{cases}$$

$$\Rightarrow \begin{cases} < (a,b,c),(1,2,-1) >= 0 \\ < (a,b,c),(1,0,1) >= 0 \end{cases}$$

$$\Rightarrow \begin{cases} a + 2b - c = 0 \\ a + c = 0 \end{cases}$$

$\Rightarrow a = -b = -c$

Portanto os vetores ortogonais a v_1 e v_2 são da forma $(-c, c, c) / c \in \mathbb{R}$.

Fazendo $c = 1$ obtemos o vetor $v_3 = (-1,1,1)$, logo:

$B = \{(1,2,-1),(1,0,1),(-1,1,1)\}$ é uma base ortogonal do $\mathbb{R}^3$ em relação ao vetor v_1.

Agora normalizando cada vetor de $B = \{(1,2,-1),(1,0,1),(-1,1,1)\}$, obteremos uma base ortonormal $B^{'} = \{u_1, u_2, u_3\}$

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{(1,2,-1)}{\|(1,2,-1)\|} = \frac{(1,2,-1)}{<(1,2,-1),(1,2,-1)>^{1/2}} = \frac{(1,2,-1)}{\sqrt{1^2+2^2+(-1)^2}} = (\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}})$$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{(1,0,1)}{\|(1,0,1)\|} = \frac{(1,0,1)}{<(1,0,1),(1,0,1)>^{1/2}} = \frac{(1,0,1)}{\sqrt{1^2+0^2+1^2}} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$$

$$u_3 = \frac{v_3}{\|v_3\|} = \frac{(-1,1,1)}{\|(-1,1,1)\|} = \frac{(-1,1,1)}{<(-1,1,1),(-1,1,1)>^{1/2}} = \frac{(-1,1,1)}{\sqrt{(-1)^2+1^2+1^2}} = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$$

Assim $B' = \{u_1, u_2, u_3\}$ é uma base ortonormal.

Obs.: Esse exemplo tem infinitas soluções, apresentamos apenas uma delas. No próximo tópico será inserido o *processo de ortogonalização de Gram-Schmidt,* que serve para encontrar uma base ortonormal a partir de qualquer base, sendo ela ortogonal ou não.

1.7- Processo de ortogonalização de Gram-Schmidt

Temos que a partir de uma base qualquer de um espaço vetorial, podemos encontrar uma base ortonormal. Para facilitar o entendimento vamos utilizar esse processo para encontrar uma base ortonormal a partir de uma base $\beta = \{v_1, v_2\}$. Mais adiante vamos generalizar essa ortogonalização para uma base $\beta = \{v_1, v_2, ..., v_n\}$.

Considere $v_1' = v_1$, precisamos encontrar a partir de v_2 um novo vetor v_2' ortogonal a v_1', ou seja, $<v_2', v_1'> = 0$.

Para isso tomamos $v_2' = v_2 - cv_1'$, com $c \in \mathbb{R}$ tal que $<v_2', v_1'> = 0$, substituindo v_2' por $v_2 - cv_1'$ temos:

$$< v_2 - cv_1', v_1' >= 0$$

$$\Rightarrow < v_2, v_1' > - < cv_1', v_1' >= 0$$

$$\Rightarrow < v_2, v_1' > -c < v_1', v_1' >= 0$$

$$\Rightarrow < v_2, v_1' >= c < v_1', v_1' >$$

$$\Rightarrow c = \frac{< v_2, v_1' >}{< v_1', v_1' >}$$

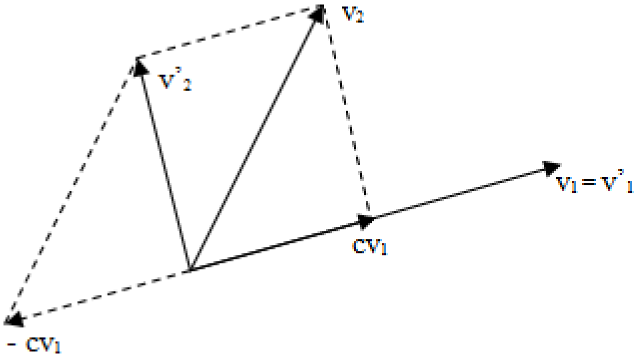

Figura 8: Processo de ortogonalização de Gram-Schmidt.

Assim temos que $v_1' = v_1$ e $v_2' = v_2 - cv_1'$ o que implica em $v_2' = v_2 - \frac{< v_2, v_1' >}{< v_1', v_1' >}.v_1'$.

Observe que v_2' foi obtido de v_2, subtraindo deste a projeção do vetor v_2 na direção de v_1', $cv_1 = \frac{< v_2, v_1' >}{< v_1', v_1' >}.v_1'$ e que v_1' e v_2' são vetores ortogonais não nulos. Podemos então normalizá-los $u_1 = \frac{v_1'}{\|v_1'\|}$ e $u_2 = \frac{v_2'}{\|v_2'\|}$, obtendo uma base ortonormal $B' = \{u_1, u_2\}$.

Exemplo 1.7.1: Seja $\beta = \{(4,2),(2,2)\}$, uma base do $\mathbb{R}^2$. Vamos obter a partir de β uma base ortonormal em relação ao produto interno usual.

Sejam $v_1 = (4,2)$ e $v_2 = (2,2)$.

$v_1' = v_1 = (4,2)$

$v_2' = v_2 - cv_1'$

Como já vimos anteriormente a condição de que v_2' seja ortogonal a v_1' vai implicar em $c = \frac{< v_2, v_1' >}{< v_1', v_1' >}$, e portanto;

$$v_2' = v_2 - \frac{< v_2, v_1' >}{< v_1', v_1' >}.v_1' = (2,2) - \frac{< (2,2),(4,2) >}{< (4,2),(4,2) >}.(4,2) = (2,2) - \frac{8+4}{16+4}.(4,2)$$

$$= (2,\ 2) - \frac{3}{5}.(4,\ 2) = (2,\ 2) - (\frac{12}{5},\ \frac{6}{5}) = (-\frac{2}{5},\ \frac{4}{5})$$

Normalizando estes vetores temos,

$$u_1 = \frac{v_1'}{\|v_1'\|} = \frac{(4,2)}{\|(4,2)\|} = \frac{(4,2)}{<(4,2),(4,2)>^{1/2}} = \frac{(4,2)}{\sqrt{4^2+2^2}} = (\frac{4}{\sqrt{20}}, \frac{2}{\sqrt{20}})$$

$$u_2 = \frac{v_2'}{\|v_2'\|} = \frac{(-\frac{2}{5}, \frac{4}{5})}{\|(-\frac{2}{5}, \frac{4}{5})\|} = \frac{(-\frac{2}{5}, \frac{4}{5})}{<(-\frac{2}{5}, \frac{4}{5})>^{1/2}} = \frac{(-\frac{2}{5}, \frac{4}{5})}{\sqrt{(-\frac{2}{5})^2 + (\frac{4}{5})^2}} =$$

$$= \frac{(-\frac{2}{5}, \frac{4}{5})}{\sqrt{\frac{20}{25}}} = \frac{(-\frac{2}{5}, \frac{4}{5})}{\frac{\sqrt{20}}{5}} = (-\frac{2}{\sqrt{20}}, \frac{4}{\sqrt{20}})$$

Então, $\beta' = \{u_1, u_2\}$ é uma base ortonormal.

Como já havíamos visto o processo de ortogonalização de dois vetores pode ser generalizado para uma base $\beta = \{v_1, v_2, ..., v_n\}$. Tomemos como no caso anterior $v_1' = v_1$ e $v_2' = v_2 - cv_1'$ onde $c = \frac{<v_2, v_1'>}{<v_1', v_1'>}$, então v_1' é ortogonal a v_2'. Queremos encontrar agora um vetor v_3' que seja ortogonal ao mesmo tempo a v_1' e v_2', como no caso anterior, vamos estabelecer que $v_3' = v_3 - mv_2' - kv_1'$ e determinar os valores de m e k de modo que $<v_3', v_1'> = 0$ e $<v_3', v_2'> = 0$. Desenvolvendo essas duas condições temos;

$$<v_3', v_1'> = 0 \Leftrightarrow <v_3' - mv_2' - kv_1', v_1'> = 0 \Leftrightarrow <v_3', v_1'> - m<v_2', v_1'> - k<v_1', v_1'> = 0$$

Assim como $<v_2', v_1'> = 0$ temos, $<v_3', v_2'> = 0$ se, e somente se $k = \frac{<v_3, v_1'>}{<v_1', v_1'>}$.

Da mesma forma $<v_3', v_2'> = 0$, se e somente se, $m = \frac{<v_3, v_2'>}{<v_2', v_2'>}$.

E, portanto $v_3' = v_3 - \frac{<v_3, v_2'>}{<v_2', v_2'>}.v_2' - \frac{<v_3, v_1'>}{<v_1', v_1'>}.v_1'$.

Observe que v_3' é obtido de v_3 subtraindo-se duas projeções sobre v_1' e v_2'.

Procedendo de maneira análoga, obtemos os vetores $v_4,...,v_n$.

Assim a partir de uma base $\beta=\{v_1,v_2,...,v_n\}$ de um espaço vetorial V, construímos a base ortogonal $\{v_1^{'},...,v_n^{'}\}$ dada por;

$$v_1^{'}=v_1$$

$$v_2^{'}=v_2-\frac{<v_2,v_1^{'}>}{<v_1^{'},v_1^{'}>}.v_1^{'}$$

$$v_3^{'}=v_3-\frac{<v_3,v_2^{'}>}{<v_2^{'},v_2^{'}>}.v_2^{'}-\frac{<v_3,v_1^{'}>}{<v_1^{'},v_1^{'}>}.v_1^{'}$$

$$\vdots$$

$$v_n^{'}=v_n-\frac{<v_n,v_{n-1}^{'}>}{<v_{n-1}^{'},v_{n-1}^{'}>}.v_n^{'}-\frac{<v_n,v_{n-2}^{'}>}{<v_{n-2}^{'},v_{n-2}^{'}>}.v_{n-2}^{'}-...-\frac{<v_n,v_1^{'}>}{<v_1^{'},v_1^{'}>}.v_1^{'}$$

Este procedimento é conhecido como *processo de ortogonalização de Gram-Schmidt*.

Sendo assim se precisarmos construir uma base ortonormal, será preciso apenas fazer a normalização dos vetores v_i, ou seja, tomando $u_i=\frac{v_i^{'}}{\left\|v_i^{'}\right\|}$, obtemos a base $\{u_1^{'},...,u_n^{'}\}$ de vetores ortonormais.

Exemplo 1.7.2: Seja $\beta=\{(1,1,1),(0,2,1),(0,0,1)\}$ uma base do $\mathbb{R}^3$. Vamos obter a partir de β uma base ortonormal em relação ao produto interno usual.

De fato;

Sejam $v_1=(1,1,1)$, $v_2=(0,2,1)$ e $v_3=(0,0,1)$

$$v_1^{'}=v_1=(1,1,1)$$

$$v_2^{'}=v_2-\frac{<v_2,v_1^{'}>}{<v_1^{'},v_1^{'}>}.v_1^{'}=(0,2,1)-\frac{<(0,2,1),(1,1,1)>}{<(1,1,1),(1,1,1)>}.(1,1,1)$$

$$=(0,2,1)-\frac{0+2+1}{1+1+1}.(1,1,1)=(0,2,1)-1(1,1,1)=(-1,1,0)$$

$$v_3^{'}=v_3-\frac{<v_3,v_2^{'}>}{<v_2^{'},v_2^{'}>}.v_2^{'}-\frac{<v_3,v_1^{'}>}{<v_1^{'},v_1^{'}>}.v_1^{'}$$

$$= (0,0,1) - \frac{<(0,0,1),(-1,1,0)>}{<(-1,1,0),(-1,1,0)>}.(-1,1,0) - \frac{<(0,0,1),(1,1,1)>}{<(1,1,1),(1,1,1)>}.(1,1,1)$$

$$= (0,0,1) - \frac{0+0+0}{1+1+0}.(-1,1,0) - \frac{0+0+1}{1+1+1}(1,1,1)$$

$$= (0,0,1) - 0(-1,1,0) - \frac{1}{3}(1,1,1)$$

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

Agora falta-nos apenas normalizar os vetores

$$u_1 = \frac{v_1^{'}}{\|v_1^{'}\|} = \frac{(1,1,1)}{<(1,1,1),(1,1,1)>^{1/2}} = \frac{(1,1,1)}{\sqrt{1+1+1}} = \frac{(1,1,1)}{\sqrt{3}} = (\frac{1}{3},\frac{1}{3},\frac{1}{3})$$

$$u_2 = \frac{v_2^{'}}{\|v_2^{'}\|} = \frac{(-1,1,0)}{<(-1,1,0),(-1,1,0)>^{1/2}} = \frac{(-1,1,0)}{\sqrt{1+1+0}} = \frac{(-1,1,0)}{\sqrt{2}} = (\frac{-1}{\sqrt{2},},\frac{1}{\sqrt{2}},0)$$

$$u_3 = \frac{v_3^{'}}{\|v_3^{'}\|} = \frac{(-\frac{1}{3},-\frac{1}{3},\frac{2}{3})}{<(-\frac{1}{3},-\frac{1}{3},\frac{2}{3}),(-\frac{1}{3},-\frac{1}{3},\frac{2}{3})>^{1/2}} = \frac{(-\frac{1}{3},-\frac{1}{3},\frac{2}{3})}{\sqrt{\frac{1}{9}+\frac{1}{9}+\frac{4}{9}}} = \frac{(-\frac{1}{3},-\frac{1}{3},\frac{2}{3})}{\sqrt{\frac{6}{9}}}$$

$$= \frac{(-\frac{1}{3},-\frac{1}{3},\frac{2}{3})}{\frac{\sqrt{6}}{3}} = (-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}})$$

E portanto $\beta = \{u_1, u_2, u_3\}$ é uma base ortonormal.

2. Projeção Ortogonal

Neste capítulo mostraremos a importância de projeção ortogonal, pois permite encontrar um vetor que tem a menor distância de um vetor considerado, critério que servirá de base para o *Método dos Mínimos Quadrados* que abordaremos no capítulo 3. Projeção ortogonal também é utilizada para construção de bases ortogonais e ortonormais de espaços com produto interno.

.

2.1- Complemento Ortogonal

Seja V um espaço vetorial euclidiano e W um subespaço vetorial de V. Um vetor $u \in V$ é dito ortogonal a W se é ortogonal a cada vetor de W, e o conjunto de todos os vetores de V que são ortogonais a W é chamado de *complemento ortogonal* de W e representado por $W^{\perp}$.

Assim;

$$W^{\perp} = \{u \in V / < u, v >= 0, \forall v \in W\}$$

Exemplo 2.1: Considere V um plano passando pela origem do $\mathbb{R}^3$ com o produto interno euclidiano, então, todos os vetores ortogonais a V, formam a reta T passando pela origem, que é o conjunto dos vetores ortogonais a V. Sendo assim podemos dizer que um é *complemento ortogonal* do outro.

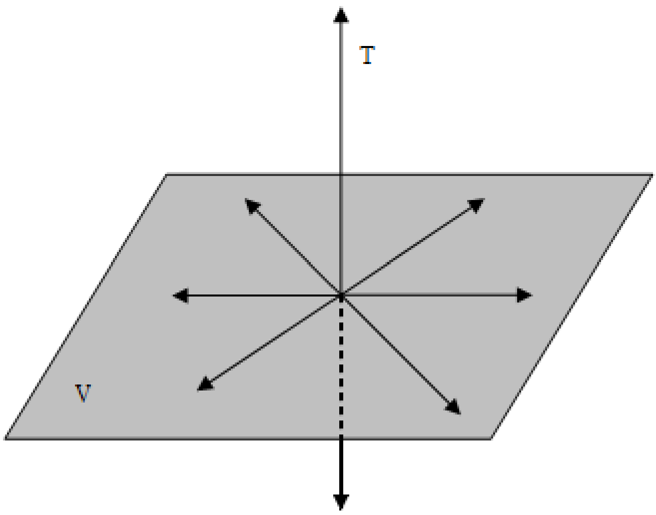

Figura 9: A reta T é o conjunto dos vetores ortogonais a V

Teorema 2.1.1: Seja V um espaço vetorial euclidiano e W um subespaço vetorial de V, de dimensão finita então:

i) $W^{\perp}$ é um subespaço de V.

ii) O único vetor comum a $W^{\perp}$ e W é $\{o\}$.

iii) O complemento ortogonal de $W^{\perp}$ é W, ou seja $(W^{\perp})^{\perp} = W$.

Demonstração:

i) $W^{\perp}$ é um subespaço de V.

Temos que $<o,w>=0$ para cada vetor $w \in W$, assim verificamos que $W^{\perp}$ contém pelo menos o vetor nulo. De fato, seja u e $v \in W^{\perp}$, para qualquer $w \in W$ pela definição de complemento ortogonal $<u,w>=0$ e $<v,w>=0$ então $<u,w>+<v,w> \overset{ii)}{=} <u+v,w>=0$ o que implica em $u+v \in W^{\perp}$. Considere $k \in \mathbb{R}$, u e $v \in W^{\perp}$, temos que $<ku,v> \overset{iii)}{=} k<u,v>=k0=0 \in W^{\perp}$. Assim mostramos que, a soma de dois vetores de $W^{\perp}$ é ortogonal a cada vetor de W e que a multiplicação de um vetor em $W^{\perp}$ por um escalar é ortogonal a cada vetor de W. Portanto $W^{\perp}$ é um subespaço de V.

ii) Se $W \neq \{o\}$ para qualquer $v \in (W \cap W^{\perp})$ então, $<v,w>=0 \overset{iv)}{\Rightarrow} v=o$ o que mostra que $W \cap W^{\perp} = \{o\}$.

iii) Temos que se $w \in W$ e $u \in W^{\perp}$, então $<u,w>=0$, assim $w \in (W^{\perp})^{\perp}$. Portanto, $W \subset (W^{\perp})^{\perp}$. Seja um vetor v tal que $v = w+u$ onde $w \in W$ e $u \in W^{\perp}$, dessa maneira temos que w é ortogonal a u, assim

$$0=<u,v>=<u,(w+u)>=<u,w>+<u,u>=0+<u,u>=<u,u>$$

ou seja $<u,u>=0$ o que implica que $u=o$, então $v=w$, logo $v \in W$ e portanto segue que $(W^{\perp})^{\perp} \subset W$. E assim concluímos que $(W^{\perp})^{\perp} = W$.

■

Exemplo 2.2: Seja $V=\mathbb{R}^3$, com o produto interno usual e $W=\{(0,\ 0,z)\ /\ z\in\mathbb{R}\}$. Então: $W^{\perp}=\{(x,y,0)\ /\ x,y\in\mathbb{R}\}$, Vejamos como é sua representação geométrica.

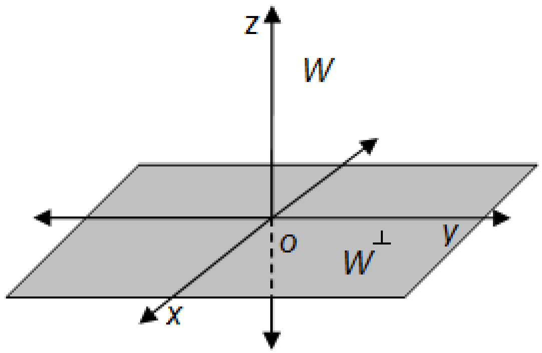

Figura 10: Rep. Geométrica de $W^{\perp}=\{(x,y,0)\ /\ x,y\in\mathbb{R}\}$.

Exemplo 2.3: Seja o produto interno usual no $\mathbb{R}^4$ e o subespaço de dimensão 2, $W=[(1,1,0,-1),(1,-2,1,0)]$, vamos determinar $W^{\perp}$ e uma base ortonormal de $W^{\perp}$.

Um vetor $v=(x,y,z,t)\in W^{\perp}$ se:

$$\begin{cases} <(x,y,z,t),(1,1,0,-1)>=0 \\ <(x,y,z,t),(1,-2,1,0)>=0 \end{cases} \Rightarrow \begin{cases} x+y-t=0 & \Rightarrow t=x+y \\ x-2y+z=0 & \Rightarrow z=2y-x \end{cases}$$ então a solução do sistema será $t=x+y$ e $z=2y-x$. Isso implica que,

$$W^{\perp}=\{(x,y,2y-x,x+y)\}\ /\ x,y\in\mathbb{R}\}$$

Assim uma base de $W^{\perp}$ será, $B=\{(1,0,-1,1),(0,1,2,1)\}$ pois,

$<(1,1,0,-1),(1,0,-1,1)>=1+0+0-1=0$, $<(1,1,0,-1),(0,1,2,1)>=0+1+0-1=0$

$<(1,-2,1,0),(1,0,-1,1)>=1+0-1+0=0$ e $<(1,-2,1,0),(0,1,2,1)>=0-2+2+0=0$

Aplicaremos agora o *processo de Gram-Schmidt* a base B para encontrar a base ortonormal $B'=\{u_1,u_2\}$.

Consideremos $v_1'=v_1=(1,0,-1,1)$ e $v_2=(0,1,2,1)$

$$v_2'=v_2-\frac{<v_2,v_1'>}{<v_1',v_1'>}.v_1'=(0,1,2,1)-\frac{<(0,1,2,1),(1,0,-1,1)>}{<(1,0,-1,1),(1,0,-1,1)>}.(1,0,-1,1)$$

$$= (0,1,2,1) - \frac{-2+1}{1+1+1}.(1,0,-1,1) = (0,1,2,1) + \frac{1}{3}.(1,0,-1,1) = (\frac{1}{3},1,\frac{5}{3},\frac{4}{3})$$

Agora falta-nos apenas normalizar os vetores

$$u_1 = \frac{v_1^{'}}{\|v_1^{'}\|} = \frac{(1,0,-1,1)}{<(1,0,-1,1),(1,0,-1,1)>^{1/2}} = \frac{(1,0,-1,1)}{\sqrt{1+1+1}} = \frac{(1,0,-1,1)}{\sqrt{3}} = (\frac{1}{\sqrt{3}},0,-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})$$

$$u_2 = \frac{v_2^{'}}{\|v_2^{'}\|} = \frac{(\frac{1}{3},1,\frac{5}{3},\frac{4}{3})}{<(\frac{1}{3},1,\frac{5}{3},\frac{4}{3}),(\frac{1}{3},1,\frac{5}{3},\frac{4}{3})>^{1/2}} = \frac{(\frac{1}{3},1,\frac{5}{3},\frac{4}{3})}{\sqrt{\frac{1}{9}+1+\frac{25}{9}+\frac{16}{9}}} = \frac{(\frac{1}{3},1,\frac{5}{3},\frac{4}{3})}{\frac{\sqrt{51}}{3}} =$$

$$= (\frac{1}{\sqrt{51}},\frac{3}{\sqrt{51}},\frac{5}{\sqrt{51}},\frac{4}{\sqrt{51}})$$

E, portanto $B^{'} = \{u_1,u_2\}$ é uma base ortonormal de $W^{\perp}$.

2.2- Espaço-Linha, Espaço-Coluna e Espaço-Nulo

O teorema 2.2.3, fornece uma relação fundamental entre o espaço-nulo e o espaço-linha de uma matriz. Antes, vamos definir o que é espaço-nulo, espaço-linha e espaço-coluna.

Temos que os três espaços citados acima são espaços vetoriais importantes associados a matrizes.

Definição 2.2.1: Para uma matriz $m \times n$

$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, os vetores $\begin{matrix} r_1 = (a_{11},a_{12},...,a_{1n}) \\ r_2 = (a_{21},a_{22},...,a_{2n}) \\ \vdots \qquad \vdots \\ r_m = (a_{n1},a_{n2},...,a_{mn}) \end{matrix}$ em $\mathbb{R}^n$ formados pelas linhas de A, são chamados os *vetores linhas* de A e os vetores $c_1 = (a_{11},a_{21},...,a_{m1})$, $c_2 = (a_{12},a_{22},...,a_{m2})$ ... $c_n = (a_{1n},a_{2n},...,a_{mn})$, em $\mathbb{R}^m$ formados pelas colunas de A são chamados de *vetores coluna* de A.

Definição 2.2.2: Se A é uma matriz de ordem $m \times n$, então o subespaço de $\mathbb{R}^n$ gerado pelos vetores-linha de A é chamado *espaço-linha* de A e o subespaço de $\mathbb{R}^m$ gerado pelos vetores-coluna de A é chamado de *espaço-coluna* de A. O espaço-solução do sistema homogêneo $Ax = 0$ que é um subespaço do $\mathbb{R}^n$ é chamado de *espaço-nulo* de A.

Exemplo 2.2.1: Considere $Ax = b$, o seguinte sistema linear $\begin{cases} 2x + y - z = 5 \\ x - 2y + 3z = -6 \\ 3x + z = 2 \end{cases}$, mostraremos que b está no espaço-coluna de A e que b pode ser expresso como combinação linear dos vetores-coluna de A.

De fato, escrevendo o sistema na forma de expressão matricial temos;

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & -2 & 3 \\ 3 & 0 & 1 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 2 \end{bmatrix}$$

Onde, a solução do sistema é $x = 1$, $y = 2$ e $z = -1$, como o sistema é consistente, b está no espaço coluna de A e ainda o sistema linear $Ax = b$ pode ser escrito como uma combinação linear deste vetor solução;

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + 2\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 2 \end{bmatrix}$$

Teorema 2.2.3: Se A é uma matriz de ordem $m \times n$, então;

i) o espaço-nulo de A e o espaço-linha de A são complementos ortogonais em $\mathbb{R}^n$ com relação ao produto interno euclidiano.

ii) O espaço-nulo de A^T e o espaço-coluna de A são complementos ortogonais em $\mathbb{R}^m$ com relação ao produto interno euclidiano.

Veremos através do exemplo a seguir, a idéia da demonstração deste teorema. Nós queremos mostrar que o complemento ortogonal do espaço-linha de A é o espaço-nulo de A. Para fazer isso, nós precisamos mostrar que se um vetor v pertence ao espaço-nulo, então v é ortogonal a cada vetor do espaço-linha.

Exemplo 2.2.2: Consideremos o sistema linear, $\begin{cases} 2x-3y-z=0 \\ 4x-y+z=0 \\ -5y-3z=0 \end{cases}$, Mostraremos que o espaço-nulo de A e o espaço-linha de A são complementos ortogonais em $\mathbb{R}^n$ com relação ao produto interno euclidiano.

Transformando o sistema numa expressão matricial e usando o método de Gauss para resolvê-lo temos;

$$\begin{bmatrix} 2 & -3 & -1 \\ 4 & -1 & 1 \\ 0 & -5 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 2 & -3 & -1 & 0 \\ 4 & -1 & 1 & 0 \\ 0 & -5 & -3 & 0 \end{array}\right]^{2L_1-L_2\to L_2} \sim \left[\begin{array}{ccc|c} 2 & -3 & -1 & 0 \\ 0 & -5 & -3 & 0 \\ 0 & -5 & -3 & 0 \end{array}\right]^{L_2-L_3\to L_3} \sim \left[\begin{array}{ccc|c} 2 & -3 & -1 & 0 \\ 0 & -5 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$, assim;

$$\begin{cases} 2x-3y-z=0 \\ -5y-3z=0 \end{cases} \Rightarrow \quad z=-\frac{5}{3}y \Rightarrow \quad 2x-3y+\frac{5}{3}y=0$$

$$6x-9y+5y=0$$
$$6x-4y=0$$
$$x=\frac{2}{3}y$$

Desse modo a solução do sistema será $\{(\frac{2}{3}y, y, -\frac{5}{3}y) / y \in \mathbb{R}\}$

De fato;

Seja $r_1=(2,-3,-1)$, $r_2=(4,-1,1)$ e $r_3=(0,-5,-3)$ o espaço-linha de A e $\{(\frac{2}{3}y, y, -\frac{5}{3}y) / y \in \mathbb{R}\}$ o espaço-nulo de A, então;

$$< r_1, (\frac{2}{3}y, y, -\frac{5}{3}y) >= 0$$

$$< (2,-3,-1), (\frac{2}{3}y, y, -\frac{5}{3}y) >= 0$$

$$\frac{4}{3}y-3y+\frac{5}{3}y=0$$

$$0=0$$

$$< r_2, (\frac{2}{3}y, y, -\frac{5}{3}y) >= 0$$

$$< (4,-1,1), (\frac{2}{3}, y, -\frac{5}{3}y) >= 0$$

$$\frac{8}{3}y - y - \frac{5}{3}y = 0$$

$$0 = 0$$

$$< r_3, (\frac{2}{3}y, y, -\frac{5}{3}y) >= 0$$

$$< (0,-5,-3), (\frac{2}{3}y, y, -\frac{5}{3}y) >= 0$$

$$-5y + 5y = 0$$

$$0 = 0$$

Portanto o espaço-nulo e o espaço-linha são complementos ortogonais.

2.3- Projeções Ortogonais

É relativamente fácil visualizar geometricamente que no $\mathbb{R}^2$ ou $\mathbb{R}^3$, com o produto interno euclidiano, que se W é uma reta ou um plano que passa pela origem então cada vetor u do espaço pode ser escrito como uma soma $u = w_1 + w_2$ onde $w_1 \in W$ e $w_2 \in W^{\perp}$.

Para o caso de W ser uma reta.

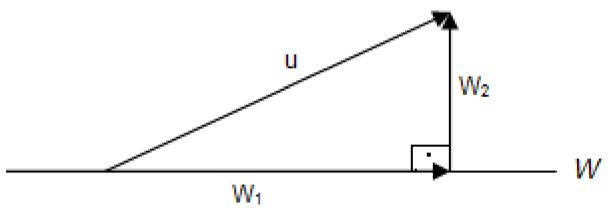

Figura 11: $u = w_1 + w_2$ onde $w_1 \in W$ e $w_2 \in W^{\perp}$.

Para o caso de W ser um plano

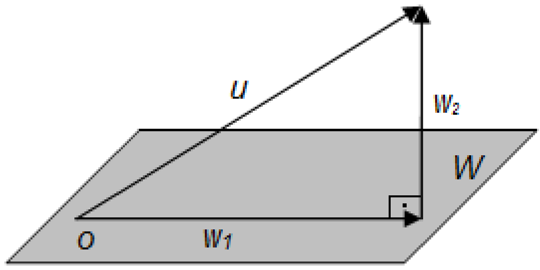

Figura 12: $u = w_1 + w_2$ onde $w_1 \in W$ e $w_2 \in W^{\perp}$.

Teorema 2.3.1: Se W é um subespaço de um espaço vetorial V de dimensão finita com produto interno, então cada vetor u de V pode ser expresso precisamente de uma única maneira como $u = w_1 + w_2$, de onde $w_1 \in W$ e $w_2 \in W^{\perp}$.

Demonstração:

Temos por hipótese que W é um subespaço vetorial de V, assim pelo processo de Gram-Schmidt existe uma base ortonormal $\{v_1, v_2, ..., v_n\}$ para W.

De fato, precisamos mostrar que $u = w_1 + w_2$ com $w_1 \in W$ e $w_2 \in W^{\perp}$, e é único para qualquer $u \in V$.

Dado $u \in V$, tomando $w_1 = < u, v_1 > v_1 + < u, v_2 > v_2 + ... + < u, v_n > v_n$ e $w_2 = u - w_1$. Então $w_1 + w_2 = w_1 + (u - w_1) = u \Rightarrow u = w_1 + w_2$, agora falta mostrar que $w_1 \in W$ e $w_2 \in W^{\perp}$. Contudo temos que $w_1 \in W$, pois é combinação linear da base de W. Para mostrarmos que $w_2 \in W^{\perp}$, basta verificarmos que $< w_2, w >= 0$, para qualquer $w \in W$.

Temos que w pode ser escrito como combinação linear de W, assim $w = k_1 v_1 + k_2 v_2 + ... + k_n v_n$ e lembrando também que $w_2 = u - w_1$, logo;

$$< w_2, w > = < u - w_1, w > = < u, w > - < w_1, w > \qquad (1)$$

De $< u, w >$, temos $< u, w > = < u, k_1 v_1 + k_2 v_2 + ... + k_n v_n >$

$$= < u, k_1 v_1 > + < u, k_2 v_2 > + ... + < u, k_n v_n >$$

$$= k_1 < u, v_1 > + k_2 < u, v_2 > + ... + k_n < u, v_n >$$

e ainda $< w_1, w > = < w_1, k_1 v_1 > + < w_1, k_2 v_2 > + ... + < w_1, k_n v_n >$

$$= < w_1, v_1 > k_1 + < w_1, v_2 > k_2 + ... + < w_1, v_n > k_n$$

$$= k_1 < w_1, v_1 > + k_2 < w_1, v_2 > + ... + k_n < w_1, v_n >$$

temos que $w_1 = u - w_2$, assim,

$$< w_1, w > = < u - w_2, v_1 > k_1 + < u - w_2, v_2 > k_2 + ... + < u - w_2, v_n > k_n$$

$$= (< u, v_1 > - < w_2, v_1 >)k_1 + (< u, v_2 > - < w_2, v_2 >)k_2 + ... + (< u, v_n > - < w_2, v_n >)k_n$$

onde $< w_2, v_i > = 0$, para todo $i = 1, 2, 3, ..., n$ então,

$< w_1, w > = k_1 < u, v_1 > + k_2 < u, v_2 > + ... + k_n < u, v_n >$, o que podemos concluir que $< u, w > = < w_1, w >$ e assim por (1) temos que $< w_2, w > = 0$, que é o que queríamos provar.

Para provar a unicidade $u = w_1 + w_2$ com $w_1 \in W$ e $w_2 \in W^{\perp}$, vamos supor que podemos escrever $u = w_1^{'} + w_2^{'}$ onde $w_1^{'} \in W$ e $w_2^{'} \in W^{\perp}$, então;

$$u - u = (w_1^{'} + w_2^{'}) - (w_1 + w_2)$$

$$0 = (w_1^{'} - w_1) + (w_2^{'} - w_2)$$

$$w_1 - w_1^{'} = w_2^{'} - w_2 \qquad (2)$$

Como w_2 e $w_2^{'}$ são ortogonais a W, sua diferença também é ortogonal a W, pois $\forall w \in W$ temos, $< w, w_2^{'} - w_2 > = < w, w_2^{'} > - < w, w_2 > = 0 - 0 = 0$.

Contudo, $w_2^{'}$ e w_2 também pertencem a W, pois $w_1 - w_1^{'}$ está em W, pela igualdade (2), assim podemos concluir que $w_2^{'} - w_2$ é ortogonal a si mesmo, ou seja, $< w_2^{'} - w_2, w_2^{'} - w_2 > = 0$.

Isso quer dizer que $w_2^{'} - w_2 = 0$ o que implica que $w_2^{'} = w_2$ e ainda por (2) $w_1^{'} = w_1$. Portanto u pode ser expresso de uma única maneira.

■

Temos que $w_1 \in W$ é a projeção de u em W que é denotado por $proj_W u$. O vetor w_2 é chamado de *componente de* u *ortogonal a* W que é denotado por $proj_{W^{\perp}} u$.

Assim, podemos escrever a fórmula do *teorema 2.3.1* da seguinte forma:

$$u = w_1 + w_2 \Rightarrow w_2 = u - w_1 \Rightarrow proj_{W^{\perp}} u = u - proj_W u \Rightarrow u = proj_W u + (u - proj_W u)$$

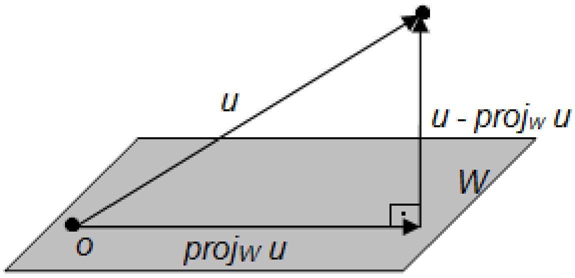

Figura 13: $u = w_1 + w_2$ pode ser escrito como $u = proj_W u + (u - proj_W u)$.

Teorema 2.3.2: Seja W um subespaço de dimensão finita de um espaço vetorial V, com o produto interno.

i) Se $\{v_1, v_2, ..., v_n\}$ é uma base ortonormal de W e u um vetor qualquer de V, então;

$$proj_W u = < u, v_1 > v_1 + < u, v_2 > v_2 + ... + < u, v_n > v_n$$

ii) Se $\{v_1, v_2, ..., v_n\}$ é uma base ortogonal de W e u é um vetor qualquer de V, então;

$$proj_W u = \frac{< u, v_1 >}{\|v_1\|^2} v_1 + \frac{< u, v_2 >}{\|v_2\|^2} v_2 + ... + \frac{< u, v_n >}{\|v_n\|^2} v_n$$

Demonstração

i) Consideremos $u \in V$ e também $V = W + W^{\perp}$, então podemos escrever $u = v + v'$, onde $v \in W$ e $v' \in W^{\perp}$. Por definição de projeção ortogonal de u sobre W temos $proj_W u = v$. Como $v \in W$, v pode ser escrito como combinação linear da base ortonormal de W, $\{v_1, v_2, ..., v_n\}$, assim $proj_W u = a_1 v_1 + a_2 v_2 + ... + a_n v_n$.

Para mostrarmos quem são os coeficientes a_i devemos verificar que $< u, v_i > = a_i$. Para isso considere $u = a_1 v_1 + a_2 v_2 + ... + a_n v_n + b_1 v_{n+1} + ... + b_m v_m$ onde, $\{v_1, v_2, ..., v_n, v_{n+1}, ..., v_m\}$ é uma base ortonormal de V. Então,

$$< u, v_i > = < a_1 v_1 + a_2 v_2 + ... + a_n v_n + b_1 v_{n+1} + ... + b_m v_m, v_i >$$

$$= a_1 < v_1, v_i > + a_2 < v_2, v_i > + ... + a_n < v_n, v_i > + b_1 < v_{n+1}, v_i > + ... + b_m < v_m, v_i >$$

Por hipótese temos que $\{v_1, v_2, ..., v_n\}$ é uma base ortonormal para W, logo os produtos internos serão 0 ou 1, assim, $< v_i, v_i >= 1$ e $< v_i, v_j >= 0$ para $i \neq j$, de 1 até n.

Para produtos internos de n até m será 0, pois é o produto interno de um vetor de W com outro vetor no seu complemento ortogonal $W^{\perp}$. Assim de $proj_W u = a_1 v_1 + a_2 v_2 + ... + a_n v_n$ temos, $proj_W u =< u, v_1 > v_1 + < u, v_2 > v_2 + ... + < u, v_n > v_n$, que é o que queríamos provar.

ii) Temos por hipótese que $\{v_1, v_2, ..., v_n\}$ é uma base ortogonal de W e u um vetor de V. Podemos encontrar uma base ortonormal a partir de uma base ortogonal normalizando cada um de seus vetores, assim $\left\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, ..., \frac{v_n}{\|v_n\|}\right\}$ é uma base ortonormal. Utilizando a parte *i)* desse mesmo teorema, temos;

$proj_W u =< u, \frac{v_1}{\|v_1\|} > \frac{v_1}{\|v_1\|} + < u, \frac{v_2}{\|v_2\|} > \frac{v_2}{\|v_2\|} + ... + < u, \frac{v_n}{\|v_n\|} > \frac{v_n}{\|v_n\|}$ e ainda pela parte *c)* do *teorema 1.5.3* o lado direito da igualdade pode ser reescrito como,

$proj_W u =< u, v_1 > \frac{1}{\|v_1\|}.\frac{v_1}{\|v_1\|} + < u, v_2 > \frac{1}{\|v_2\|}.\frac{v_2}{\|v_2\|} + ... + < u, v_n > \frac{1}{\|v_n\|}.\frac{v_n}{\|v_n\|}$, assim,

$proj_W u = \frac{< u, v_1 >}{\|v_1\|^2} v_1 + \frac{< u, v_2 >}{\|v_2\|^2} v_2 + ... + \frac{< u, v_n >}{\|v_n\|^2} v_n$, que é o que queríamos provar.

■

3- Problemas dos Mínimos Quadrados

Trataremos nesse capítulo de sistemas incompatíveis e como projeções ortogonais podem ser usadas para resolver problemas de aproximações. Tais sistemas podem ter sua origem em algum experimento físico/científico onde por, talvez, na coleta dos dados ou ainda por erros de arredondamentos produz um sistema sem soluções. Temos por objetivo encontrar solução(ões) desses sistemas utilizando o Método dos Mínimos Quadrados, de modo que, a diferença entre a solução por mínimos quadrados e a solução “exata” do sistema seja a menor possível.

3.1- Projeções ortogonais vistas como aproximações

Tendo em vista que a menor distância entre um ponto P a uma reta r ou a um plano α é uma reta perpendicular a reta r ou ao plano α passando por P, consideremos um ponto P no espaço $\mathbb{R}^3$ e W um plano pela origem, então baixando uma perpendicular de P a W, obtemos o ponto Q de W mais próximo de P. Escrevendo $u = \overrightarrow{OP}$, a distância entre P e W é dada por $u = \|u - proj_W u\|$

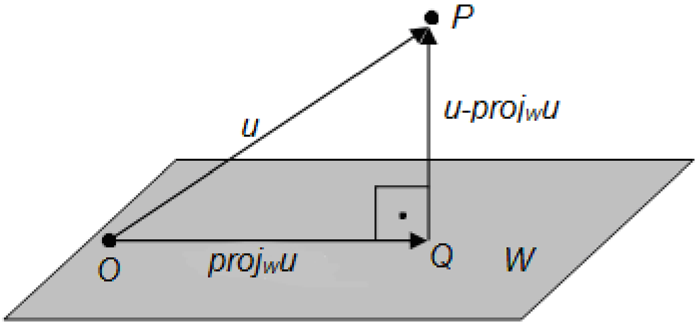

Figura 14: Q é o ponto de W mais próximo de P.

Isso quer dizer que para qualquer vetor $w \in W$, onde $w = proj_W u$, é a menor distância entre P e W, ou seja, minimiza $\|u - w\|$.

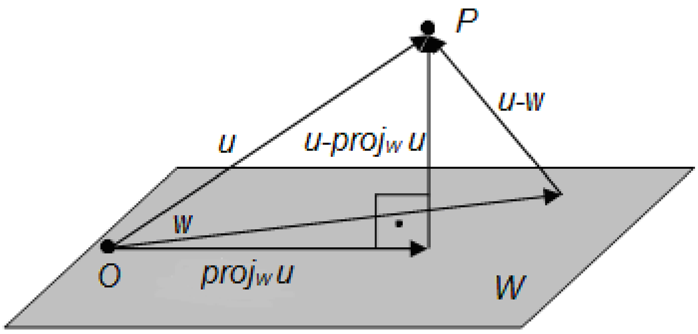

Figura 15: $\|u-w\|$ é minimizado por $w = proj_W u$.

Exemplo 3.1.1: Seja W um subespaço do $\mathbb{R}^3$ com base ortonormal $\{w_1, w_2\}$, onde $w_1 = (\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3})$ e $w_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ e o produto interno usual em $\mathbb{R}^3$, vamos encontrar a projeção ortogonal de $v = (2,1,3)$ sobre W e o vetor u que é ortogonal a todo vetor em W.

Temos que pelo teorema 2.3.2 *i)*,

$$w = proj_W v = < v, w_1 > w_1 + < v, w_2 > w_2$$

$$= < (2,1,3), (\frac{2}{3}. -\frac{1}{3}, -\frac{2}{3}) > (\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}) + < (2,1,3), (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) > (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$$

$$= (\frac{4}{3} - \frac{1}{3} - 2).(\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}) + (\frac{2}{\sqrt{2}} + 0 + \frac{3}{\sqrt{2}}).(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$$

$$= -1.(\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}) + \frac{5}{\sqrt{2}}(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$$

$$= (-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}) + (\frac{5}{2}, 0, \frac{5}{2})$$

$$= (\frac{11}{6}, \frac{1}{3}, \frac{19}{6})$$

e

$$u = v - w = (2,1,3) - (\frac{11}{6}, \frac{1}{3}, \frac{19}{6}) = (\frac{1}{6}, \frac{2}{3}, -\frac{1}{6})$$

Assim podemos observar que a distância de v ao plano W é dado pelo comprimento do vetor $u = v - w$, ou seja, $\|v - proj_W u\|$.

Então a distância entre v e o plano W será dado por:

$$\|u - proj_W u\| = < (\frac{1}{6}, \frac{2}{3}, -\frac{1}{6}), (\frac{1}{6}, \frac{2}{3}, -\frac{1}{6}) >^{1/2} = \sqrt{\frac{1}{36} + \frac{4}{9} + \frac{1}{36}} = \sqrt{\frac{1}{2}}$$

3.2- Teorema da melhor aproximação

Teorema 3.2.1: Seja W um subespaço de dimensão finita de um espaço vetorial V, com produto interno euclidiano, se u é um vetor de V, então $proj_W u$ é a melhor aproximação de u em W, no seguinte sentido:

$\|u - proj_W u\| < \|u - w\|$, para cada vetor $w \in W$ diferente de $proj_W u$.

Demonstração:

Seja w um vetor qualquer de W, então,

$u - w = (u - proj_W u) + (proj_W u - w)$.

Como w e $proj_W u$ estão em W, $proj_W u - w$ está em W e $u - proj_W u$ é ortogonal a W. Sendo assim $proj_W u - w$ e $u - proj_W u$ são ortogonais, logo

$$\|u - w\|^2 = (< u - w, u - w >^{1/2})^2$$
$$= < (u - proj_W u) + (proj_W u - w), (u - proj_W u) + (proj_W u - w) >$$
$$= \|u - proj_W u\|^2 + \|proj_W u - w\|^2$$

Se $w \neq proj_W u$ então $\|proj_W u - w\|^2$ é positivo e

$$\|u - w\|^2 > \|u - proj_W u\|^2 \quad \Leftrightarrow \quad \|u - w\| > \|u - proj_W u\|.$$

Assim, segue que $proj_W u$ é o vetor em W que minimiza $\|u - w\|^2$ e portanto, $\|u - w\|$.

■

No exemplo 3.1.1, $u - proj_W u = (\frac{11}{6}, \frac{1}{3}, \frac{19}{6})$ é o vetor em W mais próximo de $v = (2,1,3)$.

Os sistemas incompatíveis aparecem em várias situações e devemos encontrar um método de lidar com eles. Sistemas incompatíveis de equações lineares também são importantes em aplicações físicas. É comum que em alguns experimentos físicos leve um sistema $Ax = b$, (que na teoria deveria ser consistente), a uma inconsistência, e a causa poderia ser, por exemplo, erros na coleta dos dados ou ainda equipamentos com defeitos, etc. Então a solução seria encontrar um x que faça com que Ax fique tão próximo da solução b quanto possível, ou seja, no sentido que minimiza o valor de $\|Ax-b\|$ em relação ao produto interno euclidiano. A quantidade $\|Ax-b\|$ pode ser vista como uma medida do "erro" que resulta por considerar x uma solução aproximada de $Ax=b$.

Podemos pensar em Ax como uma aproximação de b. Se o sistema for consistente, e x é uma solução do sistema então $\|Ax-b\|=0$, o que implica em dizer que o erro é zero.

3.3- Problemas dos mínimos quadrados

Dado um sistema $Ax=b$ de m equações e n variáveis, encontre se possível, um vetor x que minimiza $\|Ax-b\|$ em relação ao produto interno euclidiano de $\mathbb{R}^m$. Um tal vetor é chamado uma solução de mínimos quadrados de $Ax=b$.

Vamos considerar $e=Ax-b$ como sendo o vetor erro resultado da aproximação do vetor x do sistema $Ax=b$. Sendo $e=(e_1,e_{2,\dots},e_n)$ e considerando a equação $e=Ax-b$ temos, $e=Ax-b \Rightarrow \|e\|=\|Ax-b\|$, assim desenvolvendo apenas o primeiro membro encontramos a origem do termo *mínimos quadrados*.

$$\|e\|=<e,e>^{1/2}$$

$$\|e\|=\sqrt{e_1^2+e_2^2+\dots+e_n^2}$$

$$\|e\|^2=e_1^2+e_2^2+\dots+e_n^2$$

Uma das características de um problema de mínimos quadrados é que, independentemente do vetor x selecionado, o vetor Ax pertence ao espaço-coluna de A. Para cada matriz coluna x de tamanho $n\times 1$, o produto Ax é uma combinação linear dos vetores coluna de A. Assim como x varia sobre $\mathbb{R}^n$, o vetor Ax varia sobre todas possíveis combinações lineares do vetor-coluna de A, ou seja, o vetor Ax varia sobre todo espaço-coluna W.

Geometricamente, resolver um problema de mínimos quadrados significa encontrar um vetor x de $\mathbb{R}^n$ tal que Ax é o vetor em W mais próximo de b.

Pelo *teorema 3.2.1*, o vetor em W que mais se aproxima de b é a projeção ortogonal de b em W. Assim para um vetor x ser uma solução de mínimos quadrados de $Ax = b$, este vetor deve satisfazer

$$Ax = proj_W b.$$

Veja a figura 16, $proj_W b$ é o vetor que mais se aproxima de b. (É importante lembrar que se b está em W, então b é da forma Ax para algum x, e tal x é a solução de mínimos quadrados).

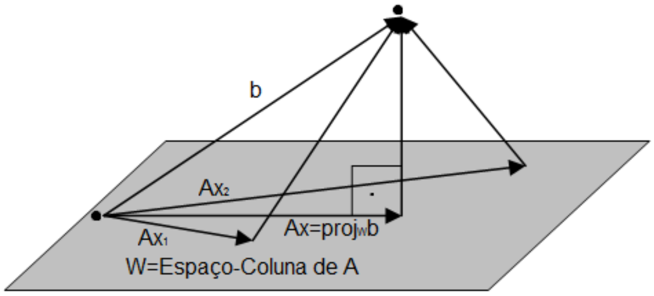

Figura 16: b está mais próximo de $Ax = proj_W b$ do que de Ax_i para outro x_i qualquer.

3.4- Solução de Mínimos Quadrados

Uma maneira de encontrar a soluções de mínimos quadrados seria encontrar o vetor $proj_W b$ e posteriormente resolver a equação $Ax = proj_W b$, no entanto há outra maneira mais prática de resolver utilizando alguns resultados já descritos.

Pelo teorema 2.3.1 e figura 13 do capítulo 2, temos que $b - Ax = b - proj_W b$, é ortogonal a W. Como W é o espaço coluna de A e $b - Ax$ é ortogonal a W, pelo teorema 2.2.3, $b - Ax$ está no espaço-nulo de A^T. Desse modo uma solução de mínimos quadrados de $Ax = b$ deve satisfazer,

$$A^T(b - Ax) = 0 \quad \Leftrightarrow \quad A^T Ax = A^T b$$

Esse sistema é denominado de *sistema normal* associado a $Ax = b$ e as equações que o compõem são chamadas de *equações normais* associadas a $Ax = b$.

Sendo assim, o problema de encontrar uma solução de mínimos quadrados foi reduzido a encontrar uma solução exata do sistema normal associado.

Com base no que acabamos de ver podemos enunciar o próximo teorema.

Teorema 3.4.1: Para qualquer sistema linear $Ax = b$, o sistema normal associado $A^T Ax = A^T b$ é consistente e todas as soluções do sistema normal são *soluções de mínimos quadrados* de $Ax = b$. Além disso, se W é um espaço-coluna de A e x é qualquer solução de mínimos quadrados de $Ax = b$, então a projeção ortogonal de b em W é $proj_W b = Ax$

Demonstração

O sistema normal $A^T Ax = A^T b$ é consistente, pois é satisfeito por uma solução de mínimos quadrados de $Ax = b$, como visto anteriormente. Fica fácil verificar que, se W é um espaço-coluna de A e x é qualquer solução de mínimos quadrados de $Ax = b$, então a projeção ortogonal de b em W é $proj_W b = Ax$, veja a figura 16.

■

3.5- Resolução de Sistemas Lineares Inconsistente por Mínimos Quadrados

Os exemplos a seguir são de sistemas inconsistentes e serão resolvidos utilizando o processo descrito no teorema 3.4.1.

Exemplo 3.5.1: Vamos determinar a solução de mínimos quadrados do sistema linear $Ax = b$ dado por $\begin{cases} x + 2y = 5 \\ 2x - y = -3 \\ 2x - 3y = 4 \end{cases}$ e encontrar a projeção ortogonal de b no espaço-coluna de A.

De fato;

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 2 & -3 \end{bmatrix} \text{ e } b = \begin{bmatrix} 5 \\ -3 \\ 4 \end{bmatrix}$$

Verifica-se que os vetores-colunas de A são L I, logo existirá apenas uma única solução de mínimos quadrados. Então,

$$A^T A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} 9 & -6 \\ -6 & 14 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & -3 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

E, portanto, o sistema normal $A^T Ax = A^T b$, neste caso é,

$$\begin{bmatrix} 9 & -6 \\ -6 & 14 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

Resolvendo esse sistema obtemos a solução de mínimos quadrados

$x = \dfrac{104}{162}$ e $y = \dfrac{51}{162}$

Pelo teorema 3.3.1 temos que $proj_W b = Ax$, ou seja, a projeção de b no espaço-coluna A será dado por;

$$proj_W b = Ax = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 104/162 \\ 51/162 \end{bmatrix} = \begin{bmatrix} 206/162 \\ 157/162 \\ 46/162 \end{bmatrix}$$

Exemplo 3.5.2: Utilizar o método dos mínimos quadrados para resolver o sistema $Ax = b$ dado por $\begin{cases} x + y - 3z = 1 \\ 5 + 2y + z = 2 \\ 9x + 3y + 5z = 5 \end{cases}$ de modo que o "vetor erro" seja mínimo, após calcularemos o vetor erro, isto é, $e = \|Ax - b\|$.

Transformando o sistema numa expressão matricial e usando o método de Gauss para resolvê-lo temos,

$$\begin{bmatrix} 1 & 1 & -3 \\ 5 & 2 & 1 \\ 9 & 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

$\left[\begin{array}{ccc|c} 1 & 1 & -3 & 1 \\ 5 & 2 & 1 & 2 \\ 9 & 3 & 5 & 5 \end{array}\right]^{\substack{9L_1 - L_3 \to L_3 \\ 5L_1 - L_2 \to L_2}} \sim \left[\begin{array}{ccc|c} 1 & 1 & -3 & 3 \\ 0 & 3 & -16 & 3 \\ 0 & 6 & -32 & 4 \end{array}\right]^{2L_2 - L_3 \to L_3} \sim \left[\begin{array}{ccc|c} 1 & 1 & -3 & 3 \\ 0 & 3 & -16 & 3 \\ 0 & 0 & 0 & 2 \end{array}\right]$, o que implica em um sistema incompatível, então vamos encontrar uma solução de mínimos quadrados para o sistema. Sendo assim;

Seja $A = \begin{bmatrix} 1 & 1 & -3 \\ 5 & 2 & 1 \\ 9 & 3 & 5 \end{bmatrix}$ e $b = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$, vamos calcular $A^T A$

$$A^T A = \begin{bmatrix} 1 & 5 & 9 \\ 1 & 2 & 3 \\ -3 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & -3 \\ 5 & 2 & 1 \\ 9 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 1+25+81 & 1+10+27 & -3+5+45 \\ 1+10+27 & 1+4+9 & -3+2+15 \\ -3+5+45 & -3+2+15 & 9+1+25 \end{bmatrix} =$$

$$= \begin{bmatrix} 107 & 38 & 47 \\ 38 & 14 & 14 \\ 47 & 14 & 35 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 5 & 9 \\ 1 & 2 & 3 \\ -3 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+10+45 \\ 1+4+15 \\ -3+2+25 \end{bmatrix} = \begin{bmatrix} 56 \\ 20 \\ 24 \end{bmatrix}$$

Assim o sistema normal $A^T Ax = A^T b$, será

$$\begin{bmatrix} 107 & 38 & 47 \\ 38 & 14 & 14 \\ 47 & 14 & 35 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 56 \\ 20 \\ 24 \end{bmatrix}$$

Resolvendo esse sistema pelo método de Gauss

$$\left[\begin{array}{ccc|c} 107 & 38 & 47 & 56 \\ 38 & 14 & 14 & 20 \\ 47 & 14 & 35 & 24 \end{array}\right]^{\substack{47L_1 - 107L_3 \to L_3 \\ 38L_1 - 107L_2 \to L_2}} \sim \left[\begin{array}{ccc|c} 107 & 38 & 47 & 56 \\ 0 & -54 & 288 & -12 \\ 0 & 288 & -1536 & 64 \end{array}\right]^{288L_2 + 54L_3 \to L_3} \sim$$

$$\sim \left[\begin{array}{ccc|c} 107 & 38 & 47 & 56 \\ 0 & -54 & -33 & -12 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Assim,

$-54y - 33z = -12$

$54y = 12 - 33z$

$y = \dfrac{12 - 33z}{54}$

Então para $z = 0$ teremos

$$-54y-33z=-12$$
$$54y=12$$
$$y=\frac{12}{54}$$
$$y=\frac{2}{9}$$

e

$$107x-38y+47z=56$$
$$107x+38.\frac{2}{9}=56$$
$$107x=56-\frac{76}{9}$$
$$x=\frac{428}{963}$$
$$x=\frac{4}{9}$$

Isso implica que a solução do sistema utilizando o método dos mínimos quadrados é, $x=\frac{4}{9}$ $y=\frac{2}{9}$ e $z=0$

Pelo teorema 3.3.1 temos que $proj_W b=Ax$, ou seja, a projeção de b no espaço-coluna A será dado por;

$$proj_W b=Ax=\begin{bmatrix}1&1&-3\\5&2&1\\9&3&5\end{bmatrix}\begin{bmatrix}4/9\\2/9\\0\end{bmatrix}=\begin{bmatrix}4/9+2/9+0\\20/9+4/9+0\\36/9+6/9+0\end{bmatrix}=\begin{bmatrix}2/3\\8/3\\14/3\end{bmatrix}$$

O vetor $Ax=\begin{bmatrix}2/3\\8/3\\14/3\end{bmatrix}$ é a melhor aproximação de b, vamos calcular o erro fazendo,

$$\|Ax-b\|=<\left(\ (\frac{2}{3},\frac{8}{3},\frac{14}{3})-(1,2,5)\ \right),\left(\ (\frac{2}{3},\frac{8}{3},\frac{14}{3})-(1,2,5)\ \right)>^{1/2}=\sqrt{(-\frac{1}{3})^2+(\frac{2}{3})^2+(-\frac{1}{3})^2}=\sqrt{\frac{2}{3}}$$

4- Ajustes de Curvas por Mínimos Quadrados

No capítulo anterior vimos que em determinadas situações nos deparamos com sistemas inconsistentes e que através do método dos mínimos quadrados podemos encontrar a melhor aproximação possível de modo que diferença entre a solução por mínimos quadrados e a solução “exata” do sistema seja a menor possível.

Um fato que atrai pesquisadores aplicados das mais diversas áreas é a possibilidade de obter uma função real que passe nos pontos ou pelo menos passe próximo dos pontos dados. Geralmente, medimos o valor de y para um valor de x dados, então, marcamos no plano cartesiano os pontos (x, y). Com base no gráfico resultante tentamos encontrar uma relação entre as variáveis x e y que pode então, ser utilizada para prever novos valores de y para determinados valores de x.

A aproximação por mínimos quadrados consiste em, dados um “tipo de função” e um conjunto de pontos, encontrar entre estas funções, a que melhor se aproxima do conjunto de pontos.

4.1- A reta ajustada por mínimos quadrados

Vamos considerar a seguinte situação.

Exemplo 4.1.1: Em determinada marca de um sabonete líquido, a quantidade de hidratante presente no produto é controlada pela quantidade de perfume utilizado no processo por litro. Recolhidas algumas amostras e feita a análise foi construída a seguinte tabela.

Perfume (gramas por litro)	3	4	5	6	7	8	9
Hidratante (gramas por litro)	4,5	5,5	5,7	6,6	7,0	8,5	8,7

Os pontos dessa tabela estão marcados no plano cartesiano, figura 17.

Vamos supor que a quantidade de perfume em relação a quantidade de hidratante seja dada por uma equação linear, ou seja, por uma função do primeiro grau cujo gráfico é uma linha reta. Mas se fizéssemos novamente os experimentos, valores ligeiramente diferentes de hidratante seriam encontrados para a mesma quantidade de perfume, pois todas as medidas estão sujeitas a erros experimentais. Dessa maneira os pontos marcados no plano não pertencem exatamente à reta. Vamos usar o método de mínimos quadrados para encontrar a reta que melhor se ajuste aos dados coletados nos experimentos. Essa reta é chamada de *reta de mínimos quadrados.*

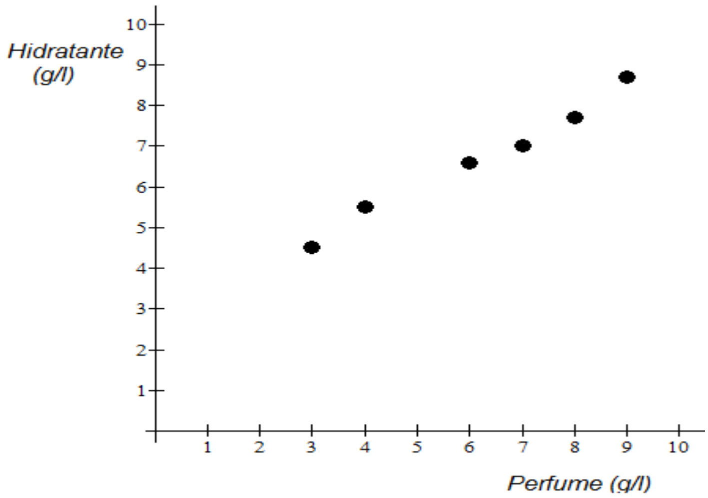

Figura 17: Relação de hidratante em relação a perfume.

Vamos supor que sejam dados n pontos $(x_1, y_1),(x_2, y_2),...,(x_n, y_n)$ com no mínimo dois dos x_i distintos. Queremos encontrar uma reta de mínimos quadrados $y = b_1x + b_0$, que melhor se ajuste aos pontos dados.

Se os pontos (x_i, y_i) com $i = 1,2,...,n$ estiverem exatamente sobre a reta de mínimos quadrados, teríamos, $y_i = b_1x_i + b_0$.

Contudo em alguns destes pontos podem não pertencer exatamente a reta, então;

$$y_i = b_1x_i + b_0 + e_i \text{, com } i = 1,2,...,n$$

Onde e_i é o desvio vertical do ponto (x_i, y_i) à reta de mínimos quadrados. Na figura 18 mostramos quatro pontos de dados $(x_1, y_1),(x_2, y_2),(x_3, y_3),(x_4, y_4)$ e seus desvios correspondentes e_1, e_2, e_3, e_4 da reta de mínimos quadrados.

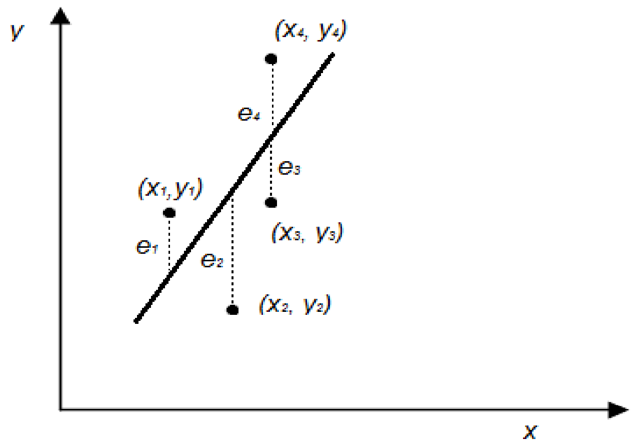

Figura 18: Pontos dados e seus desvios correspondentes em relação a reta de mínimos quadrados

Se fizermos

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\ A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix},\ x = \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} \text{ e } e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

então podemos escrever as equações como uma única equação matricial

$$b = Ax + e$$

Temos que, geralmente o sistema linear $Ax = b$ é inconsistente, então vamos encontrar uma aproximação por mínimos quadrados utilizando o teorema 3.4.1 ($A^T Ax = A^T b$).

De fato;

$$A^T A = \begin{bmatrix} 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \\ 9 & 1 \end{bmatrix} = \begin{bmatrix} 280 & 42 \\ 42 & 7 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4,5 \\ 5,5 \\ 5,7 \\ 6,6 \\ 7,0 \\ 7,7 \\ 8,5 \end{bmatrix} = \begin{bmatrix} 290,7 \\ 45,5 \end{bmatrix}$$

E, portanto, o sistema normal $A^T Ax = A^T b$, neste caso é

$$\begin{bmatrix} 280 & 42 \\ 42 & 7 \end{bmatrix} \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} 290,7 \\ 45,5 \end{bmatrix}$$

Resolvendo esse sistema utilizando o método de Gauss temos.

$$\left[\begin{array}{cc|c} 280 & 42 & 290,7 \\ 42 & 7 & 45,5 \end{array}\right]^{42L_1 - 280L_2 \to L_2} \sim \left[\begin{array}{cc|c} 280 & 42 & 290,7 \\ 0 & -196 & -530,6 \end{array}\right]$$

$$-196b_0 = -530,6$$
$$b_0 = \frac{530,6}{196}$$
$$b_0 = 2,7071$$

$$280b_1 + 42.2,7071 = 290,7$$
$$280b_1 = 176,30$$
$$b_1 = 0,6296$$

Então $x = \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} 0,6296 \\ 2,7071 \end{bmatrix}$

Portanto, uma equação para a reta de mínimos quadrados é $y = 0,6296x + 2,7071$, onde x é a quantidade de perfume e y é a quantidade de hidratante.

Veja como ficou a reta de mínimos quadrados na figura a seguir.

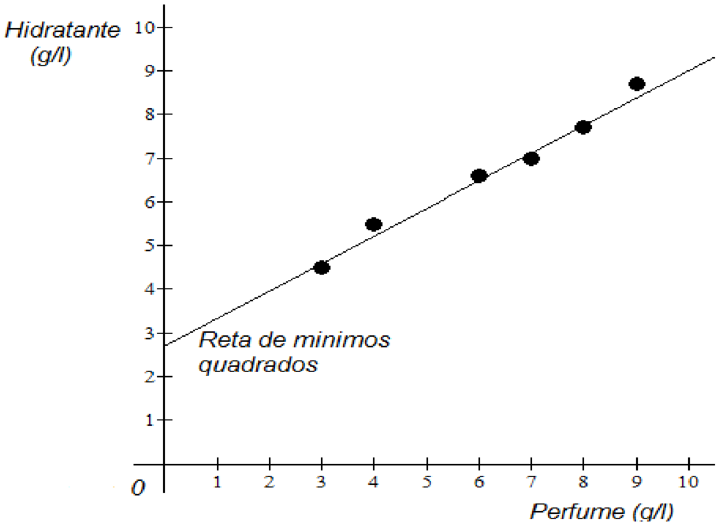

Figura 19: Reta de mínimos quadrados.

Vamos agora calcular o vetor erro comparando as duas tabelas, os valores tabelados originais e os valores tabelados “ajustados”, após ter encontrado a função da *reta de mínimos quadrados*.

Perfume (gramas por litro)	3	4	5	6	7	8	9
Hidratante (gramas por litro)	4,5	5,5	5,7	6,6	7,0	8,5	8,7

Perfume (gramas por litro)	3	4	5	6	7	8	9
Hidratante (gramas por litro)	4,5959	5,5225	5,8551	6,4847	7,1143	7,7439	8,3735

Assim, $\|Ax-b\| = \|e\| = (e_1^2+e_2^2+e_3^2+e_4^2+e_5^2+e_6^2+e_7^2)^{1/2}$ onde

$e_1^2 = (4{,}5-4{,}5959)^2 = 0{,}0091$

$e_2^2 = (5{,}5-5{,}2255)^2 = 0{,}7535$

$e_3^2 = (5{,}7-5{,}8551)^2 = 0{,}0240$

$e_4^2 = (6{,}6-6{,}4847)^2 = 0{,}0132$

$e_5^2 = (7{,}0-7{,}1143)^2 = 0{,}0130$

$e_6^2 = (8{,}5-7{,}4939)^2 = 0{,}5716$

$e_7^2 = (8{,}7-8{,}3735)^2 = 0{,}1066$ então,

$$\|e\| = (0{,}0091 + 0{,}7535 + 0{,}0240 + 0{,}0132 + 0{,}0130 + 0{,}5716 + 0{,}1066)^{1/2} \cong 1{,}2210$$

4.2- O polinômio ajustado pelos mínimos quadrados

Dependendo dos dados fornecidos é preciso encontrar um polinômio de grau predefinido que melhor se ajuste aos pontos. O método apresentado para a obtenção da reta de mínimos quadrados pode ser generalizado para encontrar polinômios de grau $n \geq 2$.

Considere que sejam dados n pontos $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. Queremos encontrar um modelo matemático do tipo

$$y = a_m x^m + a_{m-1} x^{m-1} + ... + a_1 x + a_0 + e_i \;,\; m \leq n-1$$

que melhor se ajuste a esses dados.

Fazendo,

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \; A = \begin{bmatrix} x_1^m & x_1^{m-1} & \ldots & x_1^2 & x_1 & 1 \\ x_2^m & x_2^{m-1} & \cdots & x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^m & x_n^{m-1} & \cdots & x_n^2 & x_n & 1 \end{bmatrix}, \; x = \begin{bmatrix} a_m \\ a_{m-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} \text{ e } e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

podemos escrever as n equações na forma de equação matricial $b = Ax + e$.

De maneira análoga a obtenção da reta por mínimos quadrados, uma solução para o sistema normal $A^T Ax = A^T b$, é uma aproximação por mínimos quadrados para $Ax = b$. Com esta solução garantiremos que $\|e\| = \|Ax - b\|$ é mínimo.

Para encontrarmos o polinômio de mínimos quadrados $y = a_m x^m + a_{m-1} x^{m-1} + ... + a_1 x + a_0$, que melhor se ajuste aos dados $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, onde $m \leq n-1$ e pelo menos $m+1$ dos x_i são distintos, é o seguinte:

Forme $b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $A = \begin{bmatrix} x_1^m & x_1^{m-1} & \ldots & x_1^2 & x_1 & 1 \\ x_2^m & x_2^{m-1} & \cdots & x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^m & x_n^{m-1} & \cdots & x_n^2 & x_n & 1 \end{bmatrix}$, $x = \begin{bmatrix} a_m \\ a_{m-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix}$ e em seguida resolva o sistema normal $A^T Ax = A^T b$ pelo método de Gauss.

Exemplo 4.2.1: considerando a tabela abaixo vamos determinar a equação que melhor se ajusta aos dados.

x	-1,0	-0,75	-0,6	-0,5	-0,3	0,0	0,2	0,4	0,5	0,7	1,0
y	2,0	1,153	0,45	0,4	0,5	0,0	0,2	0,6	0,512	1,2	2,05

Primeiramente vamos inserir os pontos dados no plano cartesiano para observar o comportamento de seu "traçado".

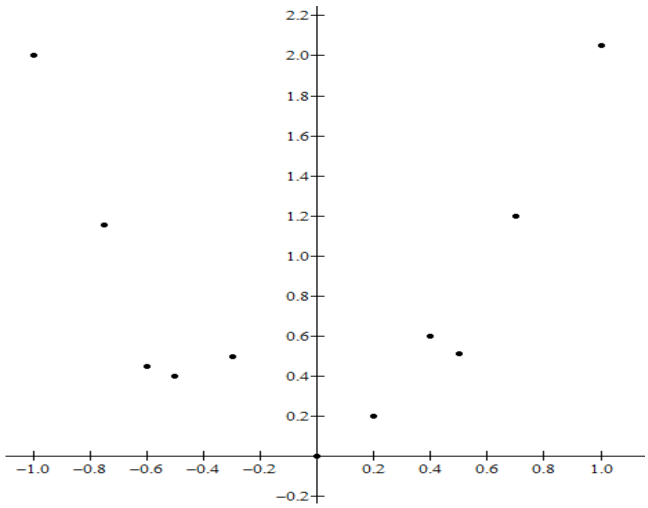

Figura 20: Gráfico de dispersão dos valores tabelados.

O gráfico de dispersão nos sugere uma aproximação a um polinômio do segundo grau, ou seja, devemos aproximar a um polinômio do tipo $y = a_2x^2 + a_1x + a_0$.

Obs.: Se utilizarmos um polinômio de grau maior que 2, é provável que encontraríamos uma curva que "melhor" se ajustaria, mas dependendo da aplicação desejada não compensa trabalhar com um polinômio de grau maior que 2, pois as contas seriam muito maiores.

Assim, utilizando o procedimento citado anteriormente temos;

$$A=\begin{bmatrix} 1 & -1 & 1 \\ 0{,}5675 & -0{,}75 & 1 \\ 0{,}36 & -0{,}6 & 1 \\ 0{,}25 & -0{,}5 & 1 \\ 0{,}09 & -0{,}3 & 1 \\ 0{,}0 & 0{,}0 & 1 \\ 0{,}04 & 0{,}2 & 1 \\ 0{,}16 & 0{,}4 & 1 \\ 0{,}25 & 0{,}5 & 1 \\ 0{,}49 & 0{,}7 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad x=\begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} \quad b=\begin{bmatrix} 2{,}0 \\ 1{,}153 \\ 0{,}45 \\ 0{,}4 \\ 0{,}5 \\ 0{,}0 \\ 0{,}2 \\ 0{,}6 \\ 0{,}512 \\ 1{,}2 \\ 2{,}05 \end{bmatrix}$$

Vamos calcular, $A^T A =$

$$=\begin{bmatrix} 1 & 0{,}5675 & 0{,}36 & 0{,}25 & 0{,}09 & 0{,}0 & 0{,}04 & 0{,}16 & 0{,}25 & 0{,}49 & 1 \\ -1 & -0{,}75 & -0{,}6 & -0{,}5 & -0{,}3 & 0{,}0 & 0{,}2 & 0{,}4 & 0{,}5 & 0{,}7 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ 0{,}5675 & -0{,}75 & 1 \\ 0{,}36 & -0{,}6 & 1 \\ 0{,}25 & -0{,}5 & 1 \\ 0{,}09 & -0{,}3 & 1 \\ 0{,}0 & 0{,}0 & 1 \\ 0{,}04 & 0{,}2 & 1 \\ 0{,}16 & 0{,}4 & 1 \\ 0{,}25 & 0{,}5 & 1 \\ 0{,}49 & 0{,}7 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$A^T A=\begin{bmatrix} 2{,}85205625 & -0{,}253625 & 4{,}2075 \\ -0{,}253625 & 4{,}2075 & -0{,}35 \\ 4{,}2075 & -0{,}35 & 11 \end{bmatrix}$$

Vamos calcular também $A^T b =$

$$= \begin{bmatrix} 1 & 0{,}5675 & 0{,}36 & 0{,}25 & 0{,}09 & 0{,}0 & 0{,}04 & 0{,}16 & 0{,}25 & 0{,}49 & 1 \\ -1 & -0{,}75 & -0{,}6 & -0{,}5 & -0{,}3 & 0{,}0 & 0{,}2 & 0{,}4 & 0{,}5 & 0{,}7 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2{,}0 \\ 1{,}153 \\ 0{,}45 \\ 0{,}4 \\ 0{,}5 \\ 0{,}0 \\ 0{,}2 \\ 0{,}6 \\ 0{,}512 \\ 1{,}2 \\ 2{,}05 \end{bmatrix}$$

$$\Rightarrow \; A^T b = \begin{bmatrix} 5{,}8313 \\ -0{,}05875 \\ 9{,}065 \end{bmatrix}$$

E, portanto o sistema normal $A^T Ax = A^T b$, neste caso é

$$\begin{bmatrix} 2{,}85205625 & -0{,}253625 & 4{,}2075 \\ -0{,}253625 & 4{,}2075 & -0{,}35 \\ 4{,}2075 & -0{,}35 & 11 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 5{,}8313 \\ -2{,}1087 \\ 9{,}065 \end{bmatrix}$$

Para efeito de diminuir os cálculos utilizaremos apenas 4 casas decimais.

Resolvendo esse sistema utilizando o método de Gauss temos.

$$\left[\begin{array}{ccc|c} 2{,}8520 & -0{,}2536 & 4{,}2075 & 5{,}8313 \\ -0{,}2536 & 4{,}2075 & -0{,}35 & -0{,}0587 \\ 4{,}2075 & -0{,}35 & 11 & 9{,}065 \end{array}\right]^{\substack{4,2075L_1 - 2,8520L_3 \to L_3 \\ -0,2536L_1 - 2,8520L_2 \to L_2}} \sim$$

$$\sim \left[\begin{array}{ccc|c} 2{,}8520 & -0{,}2536 & 4{,}2075 & 5{,}8313 \\ 0 & -11{,}9354 & -0{,}0688 & -1{,}3114 \\ 0 & -0{,}0688 & -13{,}6689 & -1{,}3181 \end{array}\right]^{-0,0688L_2 + 11,9354L_3 \to L_3} \sim$$

$$\sim \left[\begin{array}{ccc|c} 2,8520 & -0,2536 & 4,2075 & 5,8313 \\ 0 & -11,9354 & -0,0688 & -1,3114 \\ 0 & 0 & -163,1390 & -15,6418 \end{array}\right]$$

Logo;

$$-163,1390a_0 = -15,6418 \qquad -11,9354a_1 = -1,3114 + (-0,0688).0,0958$$

$$a_0 = \frac{15,6418}{163,1390} \qquad a_1 = \frac{-1,3179}{-11,9354}$$

$$a_0 = 0,0963 \qquad a_1 = 0,1078$$

$$2,8250a_2 - 0,2536a_1 + 4,2075a_0 = 5,8313$$

$$2,8250a_2 - 0,028 + 0,4030 = 5,8313$$

$$a_2 = \frac{5,4563}{2,8563}$$

$$a_2 = 1,9138$$

Assim,

$$Ax = \begin{bmatrix} 1,9138 \\ 0,1078 \\ 0,0963 \end{bmatrix}$$

Portanto obtemos o modelo do polinômio do 2º grau

$$y = 1,9138x^2 + 0,1078x + 0,0963$$

Observe como é o comportamento da curva ajustada por mínimos quadrados aos pontos dados.

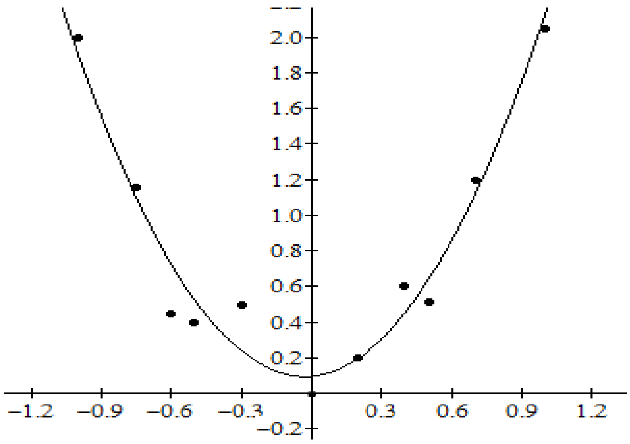

Figura 21: Função do 2º grau ajustado por mínimos quadrados.

Em determinados momentos é apropriado supor que os dados estejam relacionados exponencialmente, isso requer que a função de aproximação seja do tipo $y = be^{ax}$, com a e $b \in R$.

Geralmente o método utilizado quando suspeita-se que os dados estejam relacionados exponencialmente é “linearizar” a função $y = be^{ax}$, assim,

$$y = be^{ax}$$

$$\ln y = \ln be^{ax}$$

$$\ln y = \ln b + \ln e^{ax}$$

$$\ln y = \ln b + ax \ln e$$

$$\ln y = \ln b + ax$$

Em seguida deve-se ajustar a imagem da função, y_i para cada x_i dado, ou seja, deve-se ajustar a tabela para a nova função linear obtida. No entanto a aproximação obtida dessa maneira não é a aproximação por mínimos quadrados para os dados iniciais. Essa aproximação, em alguns casos, difere da aproximação por mínimos quadrados do problema original.

Exemplo 4.2.2: Consideremos a tabela a seguir.

x	-2	-1	0	1	2	3	4	5	6
f(x)	10	8	4	3,5	2,1	1	0,5	0,3	0,2

Vamos construir o gráfico de dispersão e verificar qual é a “função que melhor se ajusta” aos dados.

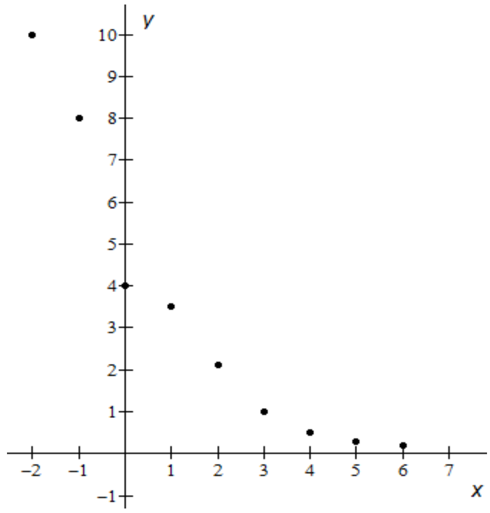

Figura 22: Gráfico de dispersão dos dados.

Temos que pela posição dos pontos no gráfico a "função que melhor se ajusta" é uma função exponencial do tipo $y = be^{-ax}$ que *não* é linear, assim temos que "linearizar" a função e após aplicar os mesmos procedimentos já vistos.

De fato $y = be^{-ax}$ é equivalente a $\ln y = \ln b - ax$, considere $\ln y = g(x) = \alpha_2 - \alpha_1 x$ agora vamos ajustar os dados iniciais para a nova função linear.

x	-2	-1	0	1	2	3	4	5	6
ln y	2,30	2,08	1,38	1,25	0,74	0	-0,69	-1,20	-1,60

Vamos encontrar uma aproximação por mínimos quadrados utilizando o teorema 3.4.1 $(A^T Ax = A^T b)$.

De fato,

$$A = \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \end{bmatrix} \quad x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \text{ e } b = \begin{bmatrix} 2{,}30 \\ 2{,}08 \\ 1{,}38 \\ 1{,}25 \\ 0{,}74 \\ 0 \\ -0{,}69 \\ -1{,}20 \\ -1{,}60 \end{bmatrix}$$

Vamos calcular $A^T A$

$$A^T A = \begin{bmatrix} -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 96 & 18 \\ 18 & 9 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2,30 \\ 2,08 \\ 1,38 \\ 1,25 \\ 0,74 \\ 0 \\ -0,69 \\ -1,20 \\ -1,60 \end{bmatrix} = \begin{bmatrix} -22,31 \\ 4,53 \end{bmatrix}$$

E, portanto, o sistema normal $A^T Ax = A^T b$, neste caso é

$$\begin{bmatrix} 96 & 18 \\ 18 & 9 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -22,31 \\ 4,53 \end{bmatrix}$$

Resolvendo esse sistema utilizando o método de Gauss temos.

$$\left[\begin{array}{cc|c} 96 & 18 & -22,04 \\ 18 & 9 & 4,53 \end{array}\right]^{18L_1 - 96L_2 \to L_2} \sim \left[\begin{array}{cc|c} 96 & 18 & -22,31 \\ 0 & -540 & -831,16 \end{array}\right]$$

$$-540\alpha_2 = -831,6 \qquad 96\alpha_1 + 18.1,54 = -22,04$$

$$\alpha_2 = \frac{831,6}{540} \qquad \alpha_1 = \frac{-49,76}{96}$$

$$\alpha_2 = 1,54 \qquad \alpha_1 = -0,52$$

Entretanto, $g(x) = \ln y = \ln b - ax \quad \Rightarrow \quad g(x) = \alpha_2 - \alpha_1 x$ o que implica em

$\ln b = \alpha_2 \quad \Rightarrow \quad \ln b = 1,54 \quad \Rightarrow \quad b = e^{1,54} \quad \Rightarrow \quad b = 4,66$

$a = -\alpha_1 \quad \Rightarrow \quad a = -(-0,52) \quad \Rightarrow \quad a = 0,52$

Então a função procurada é $g(x) = 4,66e^{-0,52x}$.

Veja como ficou a curva que representa a função $g(x) = 4,66e^{-0,52x}$ acima no plano cartesiano.

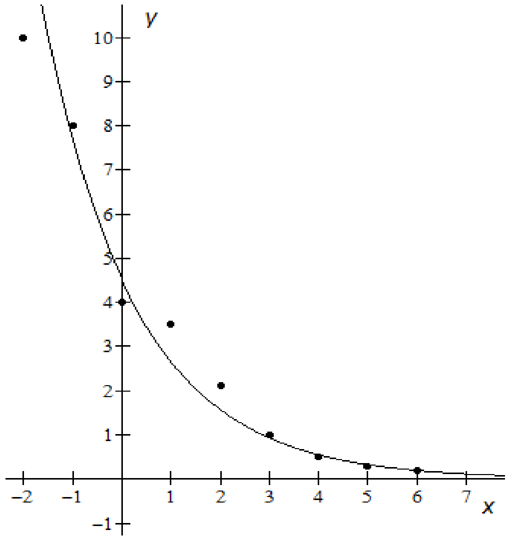

Figura 23: Função exponencial ajustado por Mínimos Quadrados.

Considerações Finais

Nesse trabalho mostramos como encontrar soluções aproximadas para sistemas lineares inconsistentes através do Método dos Mínimos Quadrados, utilizando projeções ortogonais para uma melhor aproximação. O problema de encontrar uma solução de mínimos quadrados foi reduzido a encontrar uma solução exata do sistema normal associado $A^T Ax = A^T b$. Este sistema normal associado é consistente para qualquer sistema linear e todas as soluções do sistema normal são soluções de mínimos quadrados de $Ax = b$.

Observamos que em algumas situações é preciso encontrar uma curva que “melhor se ajuste” aos dados tabulados para se fazer uma estimativa para os valores de uma função em pontos não tabulados. E a abordagem mais acertada é encontrar a melhor “curva” de aproximação resolvendo o sistema normal associado $A^T Ax = A^T b$, mesmo que ela não coincida precisamente com os dados em nenhum ponto.

Referências Bibliográficas

[1] Anton, H. & Rorres, C. – “Álgebra Linear com Aplicações”, Tradução: Claus Ivo Doering, 8ª edição – Porto Alegre, Editora: Bookman, (2001).

[2] Arruda, D. Nunes – “Um Modelo de Mínimos Quadrados para a Audição Humana” Trabalho de Conclusão de Curso, Universidade Federal de Minas Gerais, UFMG (2005).

[3] Boldrini, J. Luiz: Costa, S. Rodrigues: Figueiredo, V. Lúcia & Wetzler G. Henry – “Álgebra Linear”, 3ª edição – Depto de Matemática da Universidade Estadual de Campinas – UNICAMP, Editora: Harbra Ltda, (2009).

[4] Burden, Richard L. & Faires, J. Douglas – “Análise Numérica”, Tradução: All Tasks; Revisão técnica Helena Castro – São Paulo, Editora: Cengage Learning, pag. 459-478, (2008).

[5] Callioli, Carlos A.: Domingues, Higino H. & Costa, Roberto C. F. – “Álgebra Linear e Aplicações”, 6ª edição, São Paulo, Editora: Atual, (1990).

[6] Eves, Howard – “Introdução a História de Matemática” / Howard Eves; Tradução: Higino H. Domingues – Campinas, SP: Editora da UNICAMP, 2004.

[7] HTTP://astro.if.ufrgs.br/minq/minq.htm – acessado dia 01/12/2010 às 18:00 hs

[8] HTTP://pessoal.sercomtel.com.br/matematica/superior/mmq/mmq.htm-acessado 01/12/2010 às 19:00 hs.

[9] Kolman, Bernard & Hill, David R. – “Introdução a Álgebra Linear com Aplicações”, Tradução: Alessandra Bosquilha, Rio de Janeiro, Editora: LTC, (2006).

[10] Lay, David C. – “Álgebra Linear e suas Aplicações”; Tradução: Ricardo Camelier, Valéria de Magalhães Iório, 2ª edição, Rio de Janeiro, Editora: LTC, (2007).

[11] www.mat.ufmg.br/aplicacoes/quadrados_minimos.pdf.

[12] www.labspot.ufsc.br/~campagno/.../aula15_Ajuste_curvas.

[13] www.galileu.esalq.usp.br/mostra_topico.php?cod=39