UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

CRISLAINE APARECIDA HISSAI MIYASAKI

MÉTODO DOS MÍNIMOS QUADRADOS: ASPECTOS TEÓRICOS E SUAS APLICAÇÕES

MONOGRAFIA DE ESPECIALIZAÇÃO

CAMPO MOURÃO

2010

CRISLAINE APARECIDA HISSAI MIYASAKI

MÉTODO DOS MÍNIMOS QUADRADOS: ASPECTOS TEÓRICOS E SUAS APLICAÇÕES

Monografia apresentada ao Programa de Pós-graduação em Matemática da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Especialista em Ciências” – Área de Concentração: Matemática.

Orientadora: Profa. Msc. Sara Coelho da Silva

CAMPO MOURÃO

2010

TERMO DE APROVAÇÃO

Crislaine Aparecida Hissai Miyasaki

MÉTODO DOS MÍNIMOS QUADRADOS: ASPECTOS TEÓRICOS E SUAS APLICAÇÕES

Monografia apresentada ao Programa de Pós-graduação em Matemática da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de "Especialista em Ciências" – Área de Concentração: Matemática.

Orientador: Prof. Msc. Sara Coelho da Silva.

Prof. Dr. Juan Amadeo Soriano Palomino

Prof. Msc. Wellington Jose Corrêa

Campo Mourão, 2010

Dedico este trabalho à Deus, aos meus pais , familiares e ao meu noivo.

AGRADECIMENTOS

Agradeço á Deus por toda a ajuda, como dar paciência aos meus pais, meus familiares, meu noivo nos momentos em que eu pensava somente neste trabalho e nos meus momentos de preguiça (gerando assim muita preocupação). Agradeço a minha orientadora por ter me atendido em momentos complicados, como no meio da mudança de casa, em período de provas e nas altas horas da madrugada. Agradeço também aos amigos da pós, pois juntos nos meus momentos de crise me deram força para continuar. Agradeço aos professores que com poucas palavras me ajudaram a tomar decisões difíceis e que com muitas palavras, mas muitas mesmo, me ajudaram a adquirir conhecimento e experiências que vou levar para minha vida toda. Para todos vocês digo muito obrigado!

Os números governam o mundo. (Platão)

RESUMO

MIYASAKI, Crislaine Aparecida Hissai. MÉTODO DOS MÍNIMOS QUADRADOS: ASPECTOS TEÓRICOS E SUAS APLICAÇÕES. 37 f. Monografia – Programa de Pós-graduação em Matemática, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2010.

O Método dos Mínimos Quadrados, é uma técnica que procura encontrar o gráfico de melhor ajuste, para um conjunto de pontos dados, portanto o uso deste método contribui para muitas pesquisas, porque é como se pudessemos encontrar uma ordem no caos, podendo assim fazer previsões e estudos sobre vários fenômenos. Este trabalho mostrará a facilidade de desenvolver este método no ajuste de curvas: usando sistemas lineares ou fazendo um tratamento matricial. Na seção **2.1** ilustraremos o ajuste dos dados á uma reta e na seção **2.2** apresentaremos o ajuste dos dados a uma parábola. Finalmente na seção **2.3** generalizaremos o método para um polinômio de grau n. Mostraremos ainda que o método pode ser deduzido matricialmente e apontaremos algumas aplicações do método na modelagem de problemas reais.

Palavras-chave: mínimos quadrados, função , ajuste de curvas, equações , matrizes .

ABSTRACT

MIYASAKI, Crislaine Aparecida Hissai. METHOD OF LEAST SQUARES: THEORETICAL ASPECTS AND ITS APPLICATIONS. 37 f. Monografia – Programa de Pós-graduação em Matemática, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2010.

The Method of Least Squares, is a technique that attempts to find the chart of best fit for a set of data points, so using this method contributes to a lot of research because it is as if we could find order in chaos, and thus can make predictions and studies on various phenomena. This paper will show the ease of developing this method in curve fitting: linear systems or by using a matrix treatment. In section textbf 2.1 illustrate the fit of the data will be a straight section and the textbf 2.2 we present the data fitting to a parabola. Finally in section textbf 2.3 we generalize the method to a polynomial of degree n. We show that the method can still be deducted in matrix and we consider some applications of the method in modeling real problems.

Keywords: least squares, function, curve fitting equations, matrices.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

Na educação básica, no estudo das funções reais, algumas funções são apresentadas como modelos de fenômenos:

- Função Lucro
  O lucro L (em reais) de um estabelecimento comercial pode ser estimado pela lei
  $L(x) = -x^2 + 75x - 4$, sendo x o número de unidades vendidas (IEZZI et al,pg.50,2010)

- Quilômetros de congestionamento
  A lei representa o número de quilômetros de congestionamento, em função da hora do dia (á partir das 12 horas), registrado em uma cidade
  $f(t) = -t^2 + 12t + 20$
  onde $f(t)$ é o número em quilômetros e t é dada pela seguinte convenção: t=0 corresponde às 12 h e assim por diante. (IEZZI et al,pg.102,2010)

- Altura e peso da criança
  A altura e peso da criança á partir da 20^a semana pode ser modelada de acordo com as funções
  $h(t) = 1,5t - 9,4e$
  $p(t) = 3,8t^2 - 72t + 246$
  onde t indica o tempo em semanas, $h(t)$ é a altura em centímetros e $p(t)$ a massa em gramas.(IEZZI et al,pg.114,2010)

No entanto,torna-se necessário responder as seguintes perguntas:

- Como tais funções foram obtidas?

- Como utilizar dados reais de um experimento para criar uma função que modele este experimento?

O método dos mínimos quadrados nos permitirá responder tais perguntas usando recursos teóricos básicos, como resolução de sistemas lineares, funções lineares, quadráticas e exponenciais.

Portanto a utilização do método na modelagem de fenômenos reais pode ser utilizado como um fator de motivação na aprendizagem na educação básica, mostrando a utilidade da matemática na resolução de problemas reais.

Além disso, nos dias atuais as informações de previsão são altamente necessárias seja em estudos sociais, econômicos, ambientais e até mesmo na medicina.

Portanto apresentar este método de ajuste de curvas (que utiliza um número finito de dados amostrais para determinarmos uma curva que melhor se aproxima destes dados e que descreve o comportamento do experimento em questão, podendo ser utilizado para analisar dados futuros), evidenciará uma das grandes aplicabilidades da matemática. O que justifica a nossa motivação para o desenvolvimento deste trabalho.

2 DESENVOLVIMENTO

2.1 AJUSTE PARA A RETA

Vamos descrever o método geral, conhecido como o método dos mínimos quadrados, para determinar a linha reta que, melhor se ajusta a um conjunto de pontos dados. Para ilustrar o princípio do método dos mínimos quadrados, suponhamos, que nos são dados cinco pontos $P_1(x_1,y_1), P_2(x_2,y_2), P_3(x_3,y_3), P_4(x_4,y_4), P_5(x_5,y_5)$ que descrevem a relação entre duas variáveis x e y. Representando graficamente estes pontos, obtemos um gráfico chamado diagrama de dispersão.

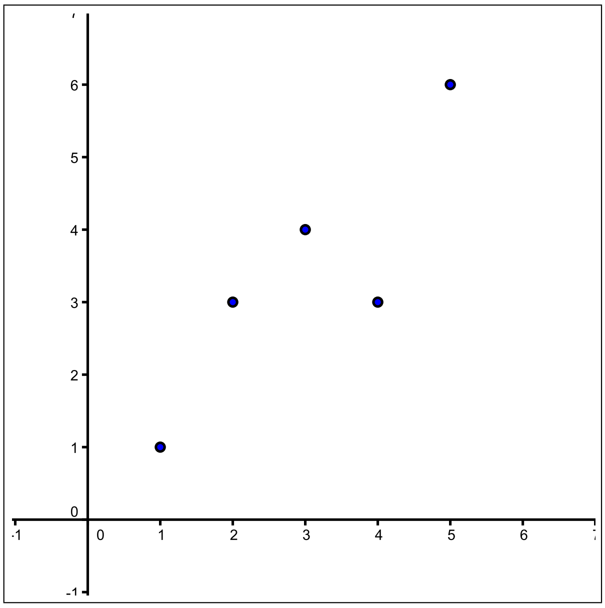

Figura 1: *diagrama de dispersão*

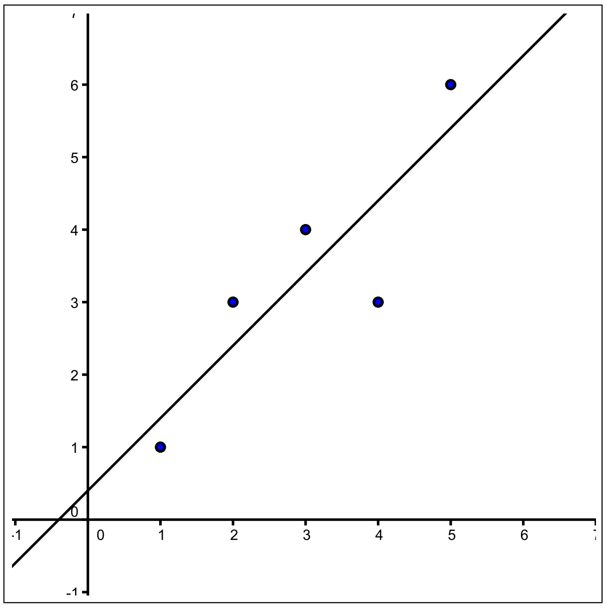

Figura 2: *reta de melhor ajuste*

Mediremos a distância vertical de cada ponto dado até a reta, representando assim os erros na direção y, e então tentaremos escolher a reta que minimize o erro total.

Os erros estão denotados por $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5$, poderemos formar o vetor-erro

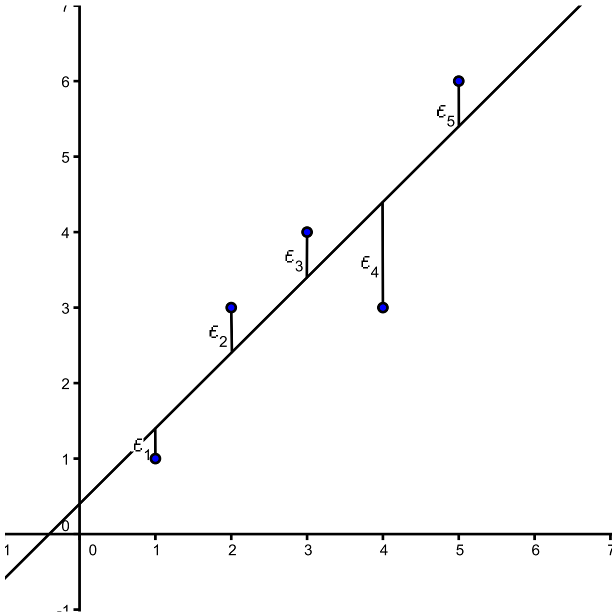

Figura 3: Erro na direção y

$$e = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix}$$

Queremos que e seja o menor possível, então $||e||$ deve estar o mais próximo de zero. Utilizando a norma euclidana obtemos:

$$||e|| = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2} \text{ ou, equivalente a } ||e||^2 = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2$$

O número $||e||$ é chamado de erro quadrático mínimo da aproximação. Sendo assim devemos encontrar a reta que melhor se a ajusta ao conjunto de pontos, suponhamos que a reta seja representada por $y = f(x) = a_0 + a_1 x$, com isso a_0 e a_1 são constantes a serem determinadas. Observe que:

$$\begin{gathered} \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2 = \\ [f(x_1) - y_1]^2 + [f(x_2) - y_2]^2 + [f(x_3) - y_3]^2 + [f(x_4) - y_4]^2 + [f(x_5) - y_5]^2 = \\ (a_0 + a_1 x_1 - y_1)^2 + (a_0 + a_1 x_2 - y_2)^2 + (a_0 + a_1 x_3 - y_3)^2 + (a_0 + a_1 x_4 - y_4)^2 + (a_0 + a_1 x_5 - y_5)^2 \end{gathered}$$

e pode ser vista como uma função que depende de a_0 e a_1.

Assim, o critério dos mínimos quadrados é equivalente a minimizar a função

$$f(a_0,a_1) = (a_0+a_1x_1-y_1)^2+(a_0+a_1x_2-y_2)^2+(a_0+a_1x_3-y_3)^2 \\ +(a_0+a_1x_4-y_4)^2+(a_0+a_1x_5-y_5)^2$$

com relação a a_0 e a_1.

Usando a regra da cadeia

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2(a_0+a_1x_1-y_1)+2(a_0+a_1x_2-y_2)+ \\ &\quad 2(a_0+a_1x_3-y_3)+2(a_0+a_1x_4-y_4)+ \\ &\quad 2(a_0+a_1x_5-y_5)\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2a_0+2a_1x_1-2y_1+2a_0+2a_1x_2-2y_2+ \\ &\quad 2a_0+2a_1x_3-2y_3+2a_0+2a_1x_4-2y_4+ \\ &\quad 2a_0+2a_1x_5-2y_5\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2[a_1(x_1+x_2+x_3+x_4+x_5)+ \\ &\quad 5a_0-(y_1+y_2+y_3+y_4+y_5)]\end{aligned}$$

e

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2(a_0+a_1x_1-y_1)x_1+2(a_0+a_1x_2+a_0-y_2)x_2+ \\ &\quad 2(a_0+a_1x_3-y_3)x_3+2(a_0+a_1x_4-y_4)x_4+ \\ &\quad 2(a_0+a_1x_5-y_5)x_5\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2a_0x_1+2a_1x_1^2-2y_1x_1+2a_0x_2+2a_1x_2^2-2y_2x_2+ \\ &\quad 2a_0x_3+2a_1x_3^2-2y_3x_3+2a_0x_4+2a_1x_4^2-2y_4x_4+ \\ &\quad 2a_0x_5+2a_1x_5^2-2y_5x_5\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2[a_1(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)+ \\ &\quad a_0(x_1+x_2+x_3+x_4+x_5)- \\ &\quad (y_1x_1+y_2x_2+y_3x_3+y_4x_4+y_5x_5)]\end{aligned}$$

Para obter o mínimo de f, façamos:

$$\frac{\partial f}{\partial a_0} = 0 \text{ e } \frac{\partial f}{\partial a_1} = 0$$

Assim,

$$\begin{cases} 5a_0 + a_1(x_1+x_2+x_3+x_4+x_5) & = & (y_1+y_2+y_3+y_4+y_5) \\ a_0(x_1+x_2+x_3+x_4+x_5) + a_1(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2) & = & (y_1x_1+y_2x_2+y_3x_3+y_4x_4+y_5x_5) \end{cases}$$

Logo, é um sistema de duas equações com duas variáveis a_0, a_1

Resolvendo as duas equações simultâneas encontraremos o valor das variáveis a_0 e a_1, obtendo assim a equação da reta de melhor ajuste $y = a_0 + a_1x$. Vamos representar um caso geral, onde a derivação é identica ao caso com cinco pontos.

2.1.1 CASO GERAL PARA A RETA

Suponha que são dados n pontos, $P_1(x_1,y_1), P_2(x_2,y_2), \ldots, P_n(x_n,y_n)$. Assim a reta que melhor se ajusta aos pontos dados é representada por;

$$y = a_0 + a_1x$$

onde a_0 e a_1 são as constantes que satisfazem o sistema;

$$\begin{cases} a_0n & + & a_1(x_1+x_2+\ldots+x_n) & = & (y_1+y_2+\ldots+y_n) \\ a_0(x_1+x_2+\ldots+x_n) & + & a_1(x_1^2+x_2^2+\ldots+x_n^2) & = & (y_1x_1+y_2x_2+\ldots+y_nx_n) \end{cases}$$

equivalente;

$$\begin{cases} a_0n & + & a_1\sum x_i & = & \sum y_i \\ a_0\sum x_i & + & a_1\sum x_i^2 & = & \sum y_ix_i \end{cases}$$

Essas equações são chamadas de equações normais.

Exemplo 2.1 *A tabela mostra a expectativa de vida para pessoas nascidas nos Estados Unidos nos anos dados. Determine a reta de mínimos quadrados para estes dados e utilize para prever a expectativa de vida de alguém nascido em 2020.*

Período	Expectativa de vida
1920	54,1
1930	59,7
1940	62,9
1950	68,2
1960	69,7
1970	70,8
1980	73,7
1990	75,4

Tabela 1: Expectativa de vida

	x_i	y_i	x_i^2	x_iy_i
	0	54,1	0	0
	1	59,7	1	59,7
	2	62,9	4	125,8
	3	68,2	9	204,6
	4	69,7	16	278,8
	5	70,8	25	354
	6	73,7	36	442,2
	7	75,4	49	527,8
Total	28	534,5	140	1992,9

Tabela 2: Cáculos para a tabela 1

Solução:

$$\begin{cases} a_0 n & + & a_1 \sum x_i & = & \sum y_i \\ a_0 \sum x_i & + & a_1 \sum x_i^2 & = & \sum y_i x_i \end{cases}$$

substituindo x_i e y_i temos:

$$\begin{cases} 8a_0 + 28a_1 & = & 534,5 \\ 28a_0 + 140a_1 & = & 1992,9 \end{cases}$$

Resolvendo as duas equações simultâneas encontraremos o valor das variáveis $a_0 = 56,6333$ e $a_1 = 2,9083$, obtendo assim a equação da reta de melhor ajuste $y = 2,9083x + 56,6333$.

Com isso podemos prever a expectativa em 2020, que é representado pelo x=10 na tabela;

$$\begin{aligned} y &= 2,9083x + 56,6333 \\ y &= 2,9083 \cdot 10 + 56,6333 \\ y &\approx 85,7 \end{aligned}$$

Portanto em 2020 a expectativa de vida será aproximadamente de 86 anos.

2.2 AJUSTE PARA A PARÁBOLA

Agora vamos descrever o método dos mínimos quadrados onde a função que melhor se ajusta ao diagrama é a parábola.
Abaixo a ilustração:

Como vimos na ilustração são dados cinco pontos. Portanto encontraremos a distância vertical do ponto até a parábola, que são os erros na direção de y e serão representados por $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5$, e então tentaremos escolher a parábola que minimize esses erros. Logo formamos com os erros o vetor e e queremos que ele seja o menor possível, ou seja, o mais

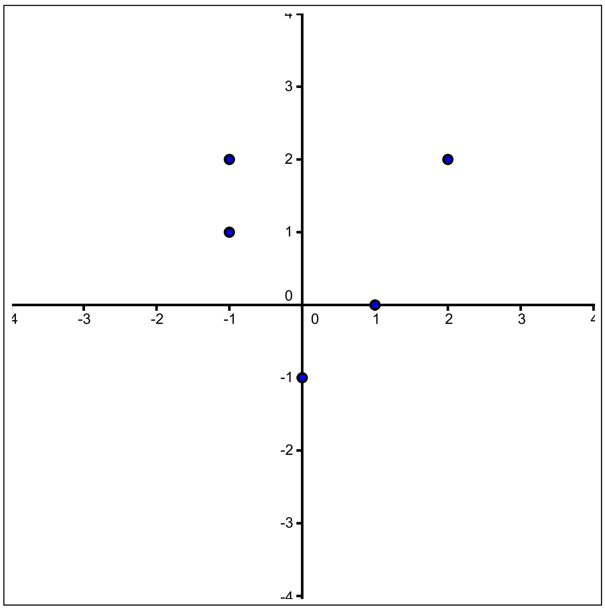

Figura 4: *diagrama de dispersão*

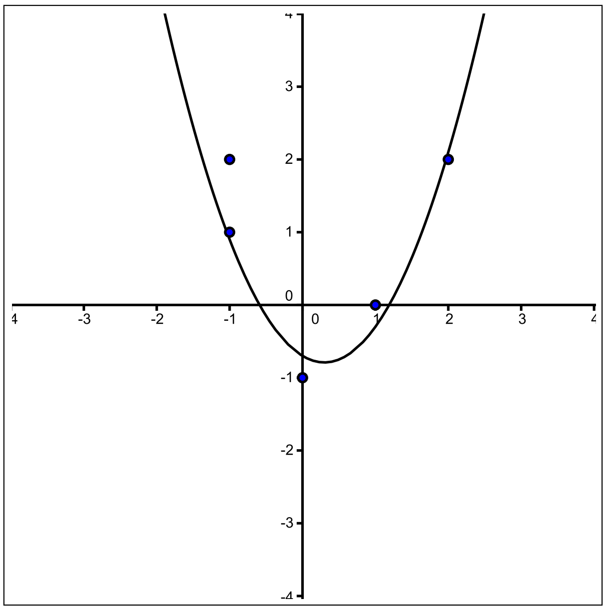

Figura 5: *parábola de melhor ajuste*

próximo de zero.

$$e = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix}$$

Utilizando a norma euclidiana, temos

$$||e|| = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2} \text{ ou, equivalente a } ||e||^2 = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2$$

Sendo assim devemos encontrar a menor soma de quadrados para determinar a parábola que melhor se ajusta aos pontos. A nosssa parábola será representada por $y = f(x) = a_0 + a_1x + a_2x^2$, com isso a_0, a_1 e a_2 são as constantes a serem determinadas. Observe que:

$$\begin{gathered}
\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 + \varepsilon_5^2 = \\
[f(x_1) - y_1]^2 + [f(x_2) - y_2]^2 + [f(x_3) - y_3]^2 + [f(x_4) - y_4]^2 + [f(x_5) - y_5]^2 = \\
(a_0 + a_1x_1 + a_2x_1^2 - y_1)^2 + (a_0 + a_1x_2 + a_2x_2^2 - y_2)^2 + \\
(a_0 + a_1x_3 + a_2x^2 - y_3)^2 + (a_0 + a_1x_4 + a_2x_4^2 - y_4)^2 + \\
(a_0 + a_1x_5 + a_2x_5^2 - y_5)^2
\end{gathered}$$

portanto pode ser vista como uma função que depende de a_0, a_1 e a_2 Assim, o critério dos mínimos quadrados é equivalente a minimizar a função

$$\begin{aligned} f(a_0,a_1,a_2) &= (a_0+a_1x_1+a_2x_1^2-y_1)^2+(a_0+a_1x_2+a_2x_2^2-y_2)^2+ \\ &\quad (a_0+a_1x_3+a_2x_3^2-y_3)^2+(a_0+a_1x_4+a_2x_4^2-y_4)^2+ \\ &\quad (a_0+a_1x_5+a_2x_5^2-y_5)^2 \end{aligned}$$

com relação a a_0, a_1 e a_2.

Usando a regra da cadeia

$$\begin{aligned} \frac{\partial f}{\partial a_0} &= 2(a_0+a_1x_1+a_2x_1^2-y_1)+2(a_0+a_1x_2+a_2x_2^2-y_2)+ \\ &\quad 2(a_0+a_1x_3+a_2x_3^2-y_3)+2(a_0+a_1x_4+a_2x_4^2-y_4)+ \\ &\quad 2(a_0+a_1x_5+a_2x_5^2-y_5) \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a_0} &= 2a_0+2a_1x_1+2a_2x_1^2-2y_1+2a_0+2a_1x_2+2a_2x_2^2-2y_2+ \\ &\quad 2a_0+2a_1x_3+2a_2x_3^2-2y_3+2a_0+2a_1x_4+2a_2x_4^2-2y_4+ \\ &\quad 2a_0+2a_1x_5+2a_2x_5^2-2y_5 \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a_0} &= 2[a_2(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)+a_1(x_1+x_2+x_3+x_4+x_5)+ \\ &\quad 5a_0-(y_1+y_2+y_3+y_4+y_5)] \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a_1} &= 2(a_0+a_1x_1+a_2x_1^2-y_1)(x_1)+2(a_0+a_1x_2+a_2x_2^2-y_2)(x_2)+ \\ &\quad 2(a_0+a_1x_3+a_2x_3^2-y_3)(x_3)+2(a_0+a_1x_4+a_2x_4^2-y_4)(x_4)+ \\ &\quad 2(a_0+a_1x_5+a_2x_5^2-y_5)(x_5) \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a_1} &= 2a_0x_1+2a_1x_1^2+2a_2x_1^3-2y_1x_1+2a_0x_2+2a_1x_2^2+2a_2x_2^3-2y_2x_2+ \\ &\quad 2a_0x_3+2a_1x_3^2+2a_2x_3^3-2y_3x_3+2a_0x_4+2a_1x_4^2+2a_2x_4^3-2y_4x_4+ \\ &\quad 2a_0x_5+2a_1x_5^2+2a_2x_5^3-2y_5x_5 \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a_1} &= 2[a_2(x_1^3+x_2^3+x_3^3+x_4^3+x_5^3)+a_1(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)+ \\ &\quad a_0(x_1+x_2+x_3+x_4+x_5)-(y_1x_1+y_2x_2+y_3x_3+y_4x_4+y_5x_5)] \end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2(a_0+a_1x_1+a_2x_1^2-y_1)(x_1^2)+2(a_0+a_1x_2+a_2x_2^2-y_2)(x_2^2)+\\ &\quad 2(a_0+a_1x_3+a_2x_3^2-y_3)(x_3^2)+2(a_0+a_1x_4+a_2x_4^2-y_4)(x_4^2)+\\ &\quad 2(a_0+a_1x_5+a_2x_5^2-y_5)(x_5^2)\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2a_0x_1^2+2a_1x_1^3+2a_2x_1^4-2y_1x_1^2+2a_0x_2^2+2a_1x_2^3+2a_2x_2^4-2y_2x_2^2+\\ &\quad 2a_0x_3^2+2a_1x_3^3+2a_2x_3^4-2y_3x_3^2+2a_0x_4^2+2a_1x_4^3+2a_2x_4^4-2y_4x_4^2+\\ &\quad 2a_0x_5^2+2a_1x_5^3+2a_2x_5^4-2y_5x_5^2\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2[a_2(x_1^4+x_2^4+x_3^4+x_4^4+x_5^4)+a_1(x_1^3+x_2^3+x_3^3+x_4^3+x_5^3)+\\ &\quad a_0(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)-(y_1x_1^2+y_2x_2^2+y_3x_3^2+y_4x_4^2+y_5x_5^2)]\end{aligned}$$

Para obter o mínimo de f, façamos:

$$\frac{\partial f}{\partial a_0}=0\ ,\ \frac{\partial f}{\partial a_1}=0 \text{ e } \frac{\partial f}{\partial a_2}=0$$

Assim

$$\left\{\begin{array}{rl} 5a_0+a_1(x_1+x_2+x_3+x_4+x_5)+a_2(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2) & = \\ (y_1+y_2+y_3+y_4+y_5) & \\ a_0(x_1+x_2+x_3+x_4+x_5)+a_1(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)+a_2(x_1^3+x_2^3+x_3^3+x_4^3+x_5^3) & = \\ (y_1x_1+y_2x_2+y_3x_3+y_4x_4+y_5x_5) & \\ a_0(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)+a_1(x_1^3+x_2^3+x_3^3+x_4^3+x_5^3)+a_2(x_1^4+x_2^4+x_3^4+x_4^4+x_5^4) & = \\ (y_1x_1^2+y_2x_2^2+y_3x_3^2+y_4x_4^2+y_5x_5^2) & \end{array}\right.$$

é um sistema de três equações com três variáveis a_0, a_1 e a_2. Ao resolver este sistema determina-se os coeficientes a_0, a_1 e a_2 da parábola que melhor se ajusta aos pontos dados.

Vamos representar um caso geral, onde a derivação é identica ao caso com cinco pontos

2.2.1 CASO GERAL PARA A PARÁBOLA

Suponha que são dados n pontos, $P_1(x_1,y_1), P_2(x_2,y_2), \ldots, P_n(x_n,y_n)$. Assim a parábola que melhor se ajusta aos pontos dados é representada por;

$$y = a_0 + a_1x + a_2x^2$$

onde a_0 , a_1 e a_2 são as constantes que satisfazem o sistema;

$$\begin{cases} na_0 + a_1(x_1 + x_2 + \dots + x_n) + a_2(x_1^2 + x_2^2 + \dots + x_n^2) & = \\ (y_1 + y_2 + \dots + y_n) \\ a_0(x_1 + x_2 + \dots + x_n) + a_1(x_1^2 + x_2^2 + \dots + x_n^2) + a_2(x_1^3 + x_2^3 + \dots + x_n^3) & = \\ (y_1x_1 + y_2x_2 + \dots + y_nx_n) \\ a_0(x_1^2 + x_2^2 + \dots + x_n^2) + a_1(x_1^3 + x_2^3 + \dots + x_n^3) + a_2(x_1^4 + x_2^4 + \dots + x_n^4) & = \\ (y_1x_1^2 + y_2x_2^2 + \dots + y_nx_n^2) \end{cases}$$

$$\begin{cases} a_0 n & + & a_1 \sum x_i & + & a_2 \sum x_i^2 & = & \sum y_i \\ a_0 \sum x_i & + & a_1 \sum x_i^2 & + & a_2 \sum x_i^3 & = & \sum y_i x_i \\ a_0 \sum x_i^2 & + & a_1 \sum x_i^3 & + & a_2 \sum x_i^4 & = & \sum y_i x_i^2 \end{cases}$$

Essas equações são chamadas de equações normais.

Exemplo 2.2 *Utilizaremos o método dos mínimos quadrados para ajustar a curva aos dados da população brasileira entre 1872 a 1990, com isso podemos fazer uma previsão da população para 2015.*

Período	População
1872	9,9
1890	14,3
1900	17,4
1920	30,6
1940	41,2
1950	51,9
1960	70,2
1970	93,1
1980	118,6
1990	146,6

Tabela 3: População Brasileira]

Vamos utilizar a função quadrática $y = a_0 + a_1x + a_2x^2$

Solução:

$$\begin{cases} 10a_0 & + & 652a_1 & + & 57056a_2 & = & 593,8 \\ 652a_0 & + & 57056a_1 & + & 5452768a_2 & = & 54472,2 \\ 57056a_0 & + & 5452768a_1 & + & 546557504a_2 & = & 5457415,6 \end{cases}$$

Resolvendo as três equações simultâneas encontraremos o valor das variáveis $a_0 = 15.90326991$; $a_1 = -0.485864853$ *e* $a_2 = 0.013172167$, *obtendo assim a equação da parábola de melhor ajuste* $y = 15.90326991 - 0.485864853x + 0.013172167x^2$.

Com isso podemos prever a população de 2015, que é representado pelo x=143 na tabela; $y = 15.90326991 - 0.485864853x + 0.013172167x^2$

	x_i	y_i	x_i^2	x_i^3	x_i^4	$y_i x_i$	y_i
	0	9,9	0	0	0	0	0
	18	14,3	324	5832	104976	286	5720
	28	17,4	784	21952	614656	522	15660
	48	30,6	2304	110592	5308416	1530	76500
	68	41,2	4624	314432	21381376	2884	201880
	78	51,9	6084	474552	37015056	4152	332160
	88	70,2	7744	681472	59969536	6318	63180
	98	93,1	9604	941192	92236816	9310	931000
	108	118,6	11664	1259712	136048896	13046	1435060
	118	146,6	13924	1643032	193877776	17592	2111040
Total	652	593,8	57056	5452768	546557504	54472,2	5457415,6

Tabela 4: Cáculos para a tabela 3

$y = 15.90326991 - 0.485864853 \cdot 143 + 0.013172167 \cdot (143)^2.$

$y = 15.90326991 - 69,47867398 + 269,357643$

$y \approx 215,78$

Logo em 2015 a previsão da população brasileira é de 215,78 milhões de habitantes

2.3 AJUSTE PARA FUNÇÃO POLINOMIAL

Como vimos na seção **2.1** e **2.2**, o princípio do método dos mínimos quadrados, consiste em encontrar a distância na direção de y. Ou seja encontrar a menor valor para a $\sum[f(x_i) - y_i]^2 = [f(x_1) - y_1]^2 + [f(x_2) - y_2]^2 + [f(x_3) - y_3]^2 + \ldots + [f(x_n) - y_n]^2 =, \forall i = 1, 2, 3, \ldots, n.$
Como a função polinomial será representada por $y = f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p$.

temos:

$$[f(x_1) - y_1]^2 + [f(x_2) - y_2]^2 + \ldots + [f(x_n) - y_n]^2 =$$
$$[a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_p x_1^p - y_1]^2 + [a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_p x_2^p - y_2]^2 +$$
$$\ldots + [a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_p x_n^p - y_n]^2$$

Logo a expressão pode ser vista como uma função que depende de $a_0, a_1, a_2, \ldots, a_p$

$$\begin{aligned} f(a_0, a_1, a_2, \ldots, a_p) &= [a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_p x_1^p - y_1]^2 + \\ &\quad [a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_p x_2^p - y_2]^2 + \ldots + \\ &\quad [a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_p x_n^p - y_n]^2 \end{aligned}$$

Para minimizar a função usamos a regra da cadeia

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2[a_0+a_1x_1+a_2x_1^2+\ldots+a_px_1^p-y_1]+\\ & \quad 2[a_0+a_1x_2+a_2x_2^2+\ldots+a_px_2^p-y_2]+\ldots+\\ & \quad 2[a_0+a_1x_n+a_2x_n^2+\ldots+a_px_n^p-y_n]\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2a_0+2a_1x_1+2a_2x_1^2+\ldots+2a_px_1^p-2y_1+\\ & \quad 2a_0+2a_1x_2+2a_2x_2^2+\ldots+2a_px_2^p-2y_2+\ldots+\\ & \quad 2a_0+2a_1x_n+2a_2x_n^2+\ldots+2a_px_n^p-2y_n\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_0} &= 2[na_0+a_1(x_1+x_2+\ldots+x_n)+\\ & \quad a_2(x_1^2+x_2^2+\ldots+x_n^2)+\ldots+a_p(x_1^p+x_2^p+\ldots+x_n^p)\\ & \quad -(y_1+y_2+\ldots+yn)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2[a_0+a_1x_1+a_2x_1^2+\ldots+a_px_1^p-y_1]x_1+\\ & \quad 2[a_0+a_1x_2+a_2x_2^2+\ldots+a_px_2^p-y_2]x_2+\ldots+\\ & \quad 2[a_0+a_1x_n+a_2x_n^2+\ldots+a_px_n^p-y_n]x_n\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2a_0x_1+2a_1x_1^2+2a_2x_1^3+\ldots+2a_px_1^{p+1}-2y_1x_1+\\ & \quad 2a_0x_2+2a_1x_2^2+2a_2x_2^3+\ldots+2a_px_2^{p+1}-2y_2x_2+\ldots+\\ & \quad 2a_0x_n+2a_1x_n^2+2a_2x_n^3+\ldots+2a_px_n^{p+1}-2y_nx_n\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_1} &= 2[a_0(x_1+x_2+\ldots+x_n)+a_1(x_1^2+x_2^2+\ldots+x_n^2)+\\ & \quad a_2(x_1^3+x_2^3+\ldots+x_n^3)+\ldots+a_p(x_1^{p+1}+a_px_2^{p+1}+\ldots+a_px_n^{p+1})\\ & \quad -(y_1x_1+y_2x_2+\ldots+y_nx_n)\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2[a_0+a_1x_1+a_2x_1^2+\ldots+a_px_1^p-y_1]x_1^2+\\ & \quad 2[a_0+a_1x_2+a_2x_2^2+\ldots+a_px_2^p-y_2]x_2^2+\ldots+\\ & \quad 2[a_0+a_1x_n+a_2x_n^2+\ldots+a_px_n^p-y_n]x_n^2\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2a_0x_1^2+2a_1x_1^3+2a_2x_1^4+\ldots+2a_px_1^{p+2}-2y_1x_1^2+\\ & \quad 2a_0x_2^2+2a_1x_2^3+2a_2x_2^4+\ldots+2a_px_2^{p+2}-2y_2x_2^2+\ldots+\\ & \quad 2a_0x_n^2+2a_1x_n^3+2a_2x_n^4+\ldots+2a_px_n^{p+2}-2y_nx_n^2\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_2} &= 2[a_0(x_1^2+x_2^2+\ldots+x_n^2)+a_1(x_1^3+x_2^3+\ldots+x_n^3)+\\ &a_2(x_1^4+x_2^4+\ldots+x_n^4)+\ldots+a_p(x_1^{p+2}+a_px_2^{p+2}+\ldots+a_px_n^{p+2})\\ &-(y_1x_1^2+y_2x_2^2+\ldots+y_nx_n^2)\end{aligned}$$

.

.

.

$$\begin{aligned}\frac{\partial f}{\partial a_p} &= 2[a_0+a_1x_1+a_2x_1^2+\ldots+a_px_1^p-y_1]x_1^p+\\ &2[a_0+a_1x_2+a_2x_2^2+\ldots+a_px_2^p-y_2]x_2^p+\ldots+\\ &2[a_0+a_1x_n+a_2x_n^2+\ldots+a_px_n^p-y_n]x_n^p\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_p} &= 2a_0x_1^p+2a_1x_1^{p+1}+2a_2x_1^{p+2}+\ldots+2a_px_1^{2p}-2y_1x_1^p+\\ &2a_0x_2^p+2a_1x_2^{p+1}+2a_2x_2^{p+2}+\ldots+2a_px_2^{2p}-2y_2x_2^p+\ldots+\\ &2a_0x_n^p+2a_1x_n^{p+1}+2a_2x_n^{p+2}+\ldots+2a_px_n^{2p}-2y_nx_n^p\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial a_p} &= 2[a_0(x_1^p+x_2^p+\ldots+x_n^p)+a_1(x_1^{p+1}+x_2^{p+1}+\ldots+x_n^{p+1})+\\ &a_2(x_1^{p+2}+x_2^{p+2}+\ldots+x_n^{p+2})+\ldots+a_p(x_1^{2p}+a_px_2^{2p}+\ldots+a_px_n^{2p})\\ &-(y_1x_1^p+y_2x_2^p+\ldots+y_nx_n^p)\end{aligned}$$

Para obter o mínimo de f, façamos:

$\frac{\partial f}{\partial a_0}=0, \frac{\partial f}{\partial a_1}=0 \ldots \frac{\partial f}{\partial a_p}=0$

obtemos, o sistema equivalente;

$$\left\{\begin{array}{c}
na_0+a_1(x_1+x_2+...+x_n)+a_2(x_1^2+x_2^2+...+x_n^2)+...+\\
a_p(x_1^p+x_2^p+...+x_n^p)=(y_1+y_2+...+yn)\\
a_0(x_1+x_2+...+x_n)+a_1(x_1^2+x_2^2+...+x_n^2)+a_2(x_1^3+x_2^3+...+x_n^3)+...+\\
a_p(x_1^{p+1}+a_px_2^{p+1}+...+a_px_n^{p+1})=(y_1x_1+y_2x_2+...+y_nx_n)\\
a_0(x_1^2+x_2^2+...+x_n^2)+a_1(x_1^3+x_2^3+...+x_n^3)+a_2(x_1^4+x_2^4+...+x_n^4)+...+\\
a_p(x_1^{p+2}+a_px_2^{p+2}+...+a_px_n^{p+2})=(y_1x_1^2+y_2x_2^2+...+y_nx_n^2)\\
.\\
.\\
.\\
a_0(x_1^p+x_2^p+...+x_n^p)+a_1(x_1^{p+1}+x_2^{p+1}+...+x_n^{p+1})+a_2(x_1^{p+2}+x_2^{p+2}+...+x_n^{p+2})+...+\\
a_p(x_1^{2p}+a_px_2^{2p}+...+a_px_n^{2p})=(y_1x_1^p+y_2x_2^p+...+y_nx_n^p)
\end{array}\right.$$

$$\left\{\begin{array}{cccccccc}
a_0\sum x_n^0 & + & a_1\sum x_n^1 & + & ... & +a_p\sum x_n^p & = & \sum x_n^0y_n\\
a_0\sum x_n^1 & + & a_1\sum x_n^2 & + & ... & +a_p\sum x_n^{p+1} & = & \sum x_n^1y_n\\
a_0\sum x_n^2 & + & a_1\sum x_n^3 & + & ... & +a_p\sum x_n^{p+2} & = & \sum x_n^1y_n^2\\
 & & & & . & & & .\\
 & & & & . & & & .\\
 & & & & . & & & .\\
a_0\sum x_n^p & + & a_1\sum x_n^{p+1} & + & ... & +a_p\sum x_n^{2p} & = & \sum x_n^py_n
\end{array}\right.$$

é um sistema de $p+1$ equações e $p+1$ variáveis $a_0,a_1,a_2,...,a_p$

Este sistema pode ser escirto da forma matricial:

$$\begin{bmatrix}
\sum x_n^0 & \sum x_n^1 & . \ . \ . & \sum x_n^p\\
\sum x_n^1 & \sum x_n^2 & . \ . \ . & \sum x_n^{p+1}\\
. & . & . & .\\
. & . & . & .\\
. & . & . & .\\
\sum x_n^p & \sum x_n^{p+1} & . \ . \ . & \sum x_n^{2p}
\end{bmatrix}
\begin{bmatrix} a_0\\ a_1\\ .\\ .\\ .\\ a_p\end{bmatrix}
=
\begin{bmatrix} \sum x_n^0y_n\\ \sum x_n^1y_n\\ .\\ .\\ .\\ \sum x_n^py_n\end{bmatrix}$$

2.4 FORMA MATRICIAL

Para desenvolvermos o método dos mínimos quadrados matricialmente será necessário enunciarmos alguns conceitos de Álgebra Linear.

Definição 2.1 *Um espaço vetorial real é um conjunto V, não vazio, munido de duas operações:*

$$\begin{array}{rcl} + : V \times V & \to & V \\ (u,v) & \mapsto & u+v \end{array} \qquad \begin{array}{rcl} \cdot : \mathbb{R} \times V & \to & V \\ (a,v) & \mapsto & av \end{array}$$

satisfazendo as seguintes propriedades, $\forall\ u,v,w \in V$ *e* $a,b \in \mathbb{R}$ *ou* $a,b \in \mathbb{C}$*:*

01) $u+v=v+u$ *;*

02) $u+(v+w)=(u+v)+w$ *;*

03) *Existe um elemento* $0 \in V$*, tal que* $0+u=u$ *(*0 *é o elemento neutro da adição);*

04) *Existe um elemento* $(-v) \in V$*, tal que* $v+(-v)=(-v)+v=0$ *(existência do simétrico);*

05) $1v=v$ *(*1 *é o elemento neutro da multiplicação por escalar);*

06) $a(bv)=(ab)v$*;*

07) $a(v+w)=av+aw$*;*

08) $(a+b)v=av+bv$*.*

Exemplo 2.3 *O* ***espaço euclidiano*** $\mathbb{R}^n = \{(x_1,x_2,\ldots,x_n); x_i \in \mathbb{R}\}$ *(n natural qualquer) é um espaço vetorial real, munido das operações*

$u+v=(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)$ *e* $a.u=(ax_1,ax_2,\ldots,ax_n)\ \forall\ u,v,\in V$ *e* $a \in \mathbb{R}$.

Definição 2.2 *Subespaço Vetorial*

Seja V um espaço vetorial real. Um subconjunto W de V é um ***subespaço vetorial*** *de V se:*

1) *W contém o vetor nulo;*

2) *Se* $u,v \in W$*, então* $u+v \in W$*;*

3) *Se* $u \in W$ *e* $a \in \mathbb{R}$*, então* $au \in W$*.*

Exemplo 2.4 *Seja* $A_{m\times n}$*. O conjunto* $W=\{A.x : x \in \mathbb{R}^n\}$ *é subespaço vetorial de* $\mathbb{R}^n$*, pois:*

$$w_1+w_2 = Ax_1+Ax_2 = A(x_1+x_2) \in W$$

$$aw_1 = aAx_1 = A(ax_1) \in W$$

Observação 2.1 *O subespaço* $W=\{A.x : x \in \mathbb{R}^n\}$ *é dito espaço coluna de A.*

Para justificarmos tal nomenclatura, usaremos um exemplo.

Exemplo 2.5 *Para* $A_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}$ *temos:*

$$w \in W \Leftrightarrow w = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 2x_2 \\ 5x_1 + 10x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 \\ 5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 10x_2 \end{bmatrix} =$$

$$= x_1 \begin{bmatrix} 1 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 10 \end{bmatrix}$$

Portanto, o subespaço W apresentado acima é o subespaço gerado pelas colunas de A.

Definição 2.3 *O produto interno no espaço vetorial* V *é uma função de* $V \times V$ *em* $\mathbb{R}$ *que a todo par de vetores* $(u,v) \in V \times V$ *associa um número real, indicado por* $u.v$ *ou* $< u,v >$, *que satisfaz as seguintes propriedades:*

1) $u.v = v.u$;

2) $u.(v+w) = u.v + v.w$;

3) $(\alpha u).v = \alpha(u.v)$ *para todo real* α;

4) $u.v \geq 0$ *e* $u.u = 0$ *se, e somente se,* $u = 0_V$.

Exemplo 2.6 *Se* $u = (x_1, x_2, \ldots, x_n)$ *e* $v = (y_1, y_2, \ldots, y_n)$ *são* n-*uplas no* $\mathbb{R}^n$ *então temos que o* ***produto interno*** *usual no* $\Re^n$ *de* u *por* v *é dado por*

$$u \cdot v = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Definição 2.4 *Vetores Ortogonais*

Seja V *um espaço vetorial euclidiano. Dizemos que dois vetores* u *e* v *são* ortogonais, *denotamos por* $u \perp v$, *se, e somente se,* $u.v = 0$.

Definição 2.5 *O complemento ortogonal*

Seja V *um espaço vetorial com produto interno. O* ***complemento ortogonal*** *de um conjunto não vazio* $W \in V$ *é o conjunto*

$$W^{\perp} = \{ v \in V :< v, w >= 0, \forall w \in W \}.$$

Teorema 2.1 *Seja V um espaço vetorial munido de produto interno. Se W é subespaço de V, então*

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

Observação 2.2 *Do teorema acima temos que, para cada $u \in V, u = w_1 + w_2$, onde $w_1 \in W$ e $w_2 \in W^{\perp}$. O vetor $w_1 \in W$ é dito* ***projeção de u em*** *W, $proj_W\, u$.O vetor w_2 é chamado de* ***componente de*** *u* ***ortogonal a*** *W e é denotado por $proj_{w^{\perp}} u$. Assim, a fórmula ($u = w_1 + w_2$) no teorema projeção pode ser reformulada como*

$u = proj_W\, u + proj_{w^{\perp}}\, u$

Como

$w_2 = u - w_1$, *decorre*

$proj_w^{\perp} u = u - proj_W\, u$

Exemplo 2.7 *Seja $A_{m\times n}$. O conjunto $N = \{v \in \mathbb{R}^m : A^T.v = 0_{\mathbb{R}^m}\}$ é um subespaço vetorial de $\mathbb{R}^m$ dito* ***espaço nulo de*** *A^T.*

Justificativa*:*

Considere $v_1, v_2 \in N$, ou seja, $A^T(v_1) = A^T(v_2) = 0_{\mathbb{R}^m}$. Assim,

1. $A^T(v_1+v_2) = \underbrace{A^T(v_1)}_{0_{\mathbb{R}^m}} + \underbrace{A^T(v_2)}_{0_{\mathbb{R}^m}} = 0_{\mathbb{R}^m} \Rightarrow v_1 + v_2 \in N$
2. *Dado* $k \in \mathbb{R}, A^T(kv_1) = k.\underbrace{A^T(v_1)}_{0_{\mathbb{R}^m}} = k.0_{\mathbb{R}^m} = 0_{\mathbb{R}^m} \Rightarrow kv_1 \in N$.

Exemplo 2.8 *Para* $A = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}$ *temos* $A^T = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}$. *Portanto, dado $v \in N$ temos:*

$$A^T.v = \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

que é equivalente ao sistema de equações:

$$\begin{cases} y_1 + 5y_2 = 0 \\ 2y_1 + 10y_2 = 0 \end{cases}$$

cuja solução é $v = (-5y_2, y_2)$.

Observação 2.3 *Observe que no exemplo anterior, dado $v \in N$ temos $v = (-5y_2, y_2)$. E portanto, v é ortogonal aos vetores w do subespaço coluna W dado no exemplo* (2.5).

De fato, considerando $w = (1x_1 + 2x_2, 5x_1 + 10x_2) \in W$ e $v = (-5y_2, y_2) \in N$ temos:

$$v.w = 0$$

Esta observação sugere que $N = W^{\perp}$, ou seja, ***o espaço nulo de A^T é o complemento ortogonal do espaço coluna de A.*** *O teorema a seguir generaliza esse resultado.*

Teorema 2.2 *Seja $A_{m\times n}$. Se $W = \{A.x : x \in \mathbb{R}^n\}$ é o espaço coluna de A então:*

$$W^T = \{v \in \mathbb{R}^m : A^T.v = 0_{\mathbb{R}^m}\} = N$$

ou seja, ***o complemento ortogonal de W é o espaço nulo de A^T.***

i) Considere $v \in W^{\perp}$, ou seja, $< w, v >= 0$, para todo $w \in W$

Como $W = w \in Ax : w = x_1C_1 + x_2C_2 + ... + x_nC_n, x_i \in \mathbb{R}$,

em particular temos:

$$< C_1, v >= 0, ..., < C_n, v >= 0$$

Por outro lado,

$$A^T.v = \begin{bmatrix} a_{11} & a_{21} & . & . & . & a_{m1} \\ a_{12} & a_{22} & . & . & . & a_{m2} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ a_{1n} & a_{2n} & . & . & . & a_{mn} \end{bmatrix}_{m\times n} \cdot v_{\mathbb{R}^m} = \begin{bmatrix} < C_1, v > \\ < C_2, v > \\ . \\ . \\ . \\ < C_n, v > \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} \Rightarrow v \in N$$

Portanto $W_{\perp} \subset N$

ii) Considere $v \in N$, ou seja,

$A^T \cdot V = 0_{\mathbb{R}^m}$

$$\begin{bmatrix} < C_1, v > \\ < C_2, v > \\ . \\ . \\ . \\ < C_n, v > \end{bmatrix} = 0_{\mathbb{R}^n} \Rightarrow v \text{ é ortogonal a } C_i, i = 1, 2, \ldots n.$$

Logo considerando $w \in W$ temos:

$$w = x_1C_1 + x_2C_2 + \ldots + x_nC_n, x_i$$
$$< w, v > = < \sum x_iC_i, v > = \sum x_i < C_i, v > = 0$$

$$\textit{ou seja,} v \in W^{\perp}$$

Portanto $N \subset W^{\perp}$

Conclusão: O complemento ortogonal do espaço coluna de A é o espaço nulo de A^T.

Teorema 2.3 *O Teorema da Melhor Aproximação*

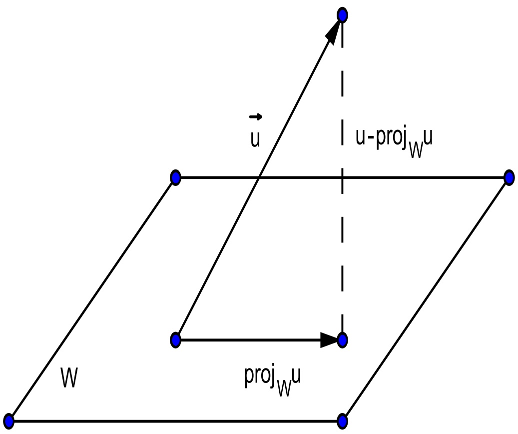

Figura 6: $proj_W v$

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

DEMONSTRAÇÃO: Seja **w** *um vetor em* W *diferente da* $proj_W(v)$*. Então,* $proj_W(v) - w$ *também está em* W*, e* $v - proj_W(v) = per_W(v)$ *é ortogonal a* $proj_W(v) - w$*. Agora, o Teorema de Pitágoras implica que*

$$||v - proj_W(v)||^2 + ||proj_W(v) - w||^2 = ||(v - proj_W(v)) + (proj_W(v) - w)||^2 = ||v - w||^2$$

como ilustrado na figura 6. Entretanto, $||proj_W(v) - w||^2 > 0$*, já que* $w \neq proj_W(v)$*, então*

$$||v - proj_W(v)||^2 < ||v - proj_W(v)||^2 + ||proj_W(v) - w||^2 = ||v - w||^2$$

ou equivalente,

$$||v - proj_W(v)||^2 < ||v - w||^2$$

2.4.1 MÍNIMOS QUADRADOS DE $Ax = b$

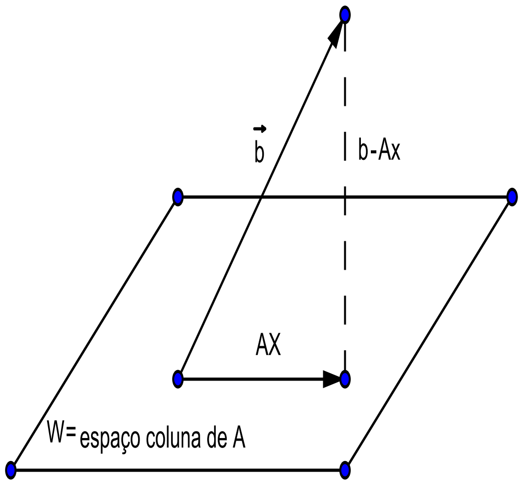

Figura 7: $proj_W b$

Dado um sistema $Ax = b$ de m equações em n variáveis encontre, se posível, 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 $Ax = b$.

Decorre do teorema da Melhor Aproximação (2.3) que o vetor em W mais próximo 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

$$Ax = proj_W b$$
$$b - Ax = b - proj_W b$$

é ortogonal a W. Como W é o espaço coluna de A, segue do teorema 2.1 que $b - Ax$ está no espaço nulo de A^T. Desse modo, uma solução de mínimos quadrados de $Ax = b$ deve satisfazer

$A^T(b - Ax) = 0$, ou ainda $A^T Ax = A^T b$

Este sistema é chamado sistema normal associado a $Ax = b$ e as equações que a compõe são chamadas de equações normais associadas a $Ax = b$. Assim, o problema de encontrar uma solução de mínimos quadrados foi reduzido a encontrar uma solução exata do sistema normal associado.

Observação 2.4 *Ajuste linear de mínimos quadrados*
Digamos que nós queremos ajustar uma reta $y = a_0 + a_1 x$ *aos pontos* $P_1(x_1, y_1), P_2(x_2, y_2), \ldots, P_n(x_n, y_n)$ *se estes pontos fossem colineares, a reta passaria pelos n pontos e então os coeficientes* a_0 *e* a_1*, iriam satisfazer :*

$$\begin{aligned} y_1 &= a_0 + a_1 x_1 \\ y_2 &= a_0 + a_1 x_2 \\ &. \\ &. \\ &. \\ y_n &= a_0 + a_1 x_n \end{aligned}$$

Nós podemos escrever este sistema em forma matricial como:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & \\ . & \\ . & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}$$

ou, mais compactante como $Ax = b$ *onde*

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & \\ . & \\ . & \\ 1 & x_n \end{bmatrix} X = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}$$

Exemplo 2.9 *Vamos utilizar o método dos mínimos quadrados para encontrar a reta que melhor se ajusta a tabela de dados, referente a taxa bruta de natalidade no Brasil do ano 1950 à 2000.*

Período	Taxa bruta de natalidade
1950	43,50
1960	44,00
1970	37,70
1980	31,87
1990	23,72
2000	21,06

Tabela 5: Taxa bruta de natalidade

Solução: $X = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 43,50 \\ 44,00 \\ 37,70 \\ 31,87 \\ 23,72 \\ 21,06 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ 1 & x_5 \\ 1 & x_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$$

Dadas as matrizes substituimos $A^T AX = A^T Y$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 43,50 \\ 44,00 \\ 37,70 \\ 31,87 \\ 23,72 \\ 21,06 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 15 \\ 15 & 55 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 201,85 \\ 415,19 \end{bmatrix}$$

Resolvendo encontraremos o valor das variáveis $a_1 = -5,1$ *e* $a_0 = 46,42$*, obtendo assim a* *equação da reta de melhor ajuste* $y = -5,1x + 46,42$.

Abaixo visualisação do gráfico

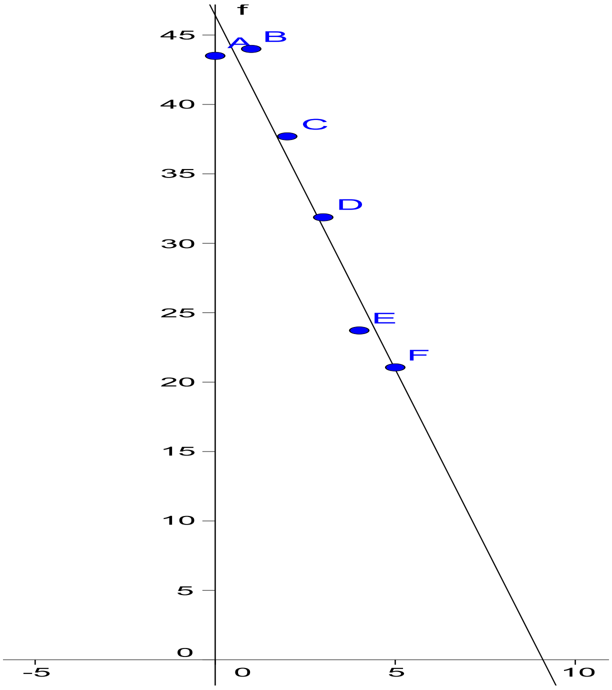

Figura 8: reta da natalidade

Com isso podemos fazer a previsão da natalidade no brasil no ano de 2015, que é representado por x=6,5

$y = -5,1x + 46,42$

$y = -5,1 \cdot 6,5 + 46,42$

$y = 13,27$

3 CONSIDERAÇÕES FINAIS

Neste trabalho apresentamos o ajuste de pontos à retas, parábolas e funções polinomiais. Caso for uma curva exponencial que melhor se ajusta aos pontos dados, ou seja $f(x) = a_0e^{a_1x}$ é só aplicar ln na equação e usar ou o método da equação ou matrizes. Abaixo mostraremos um exemplo:

Exemplo 3.1 *Utilizaremos o método dos mínimos quadrados para ajustar a curva exponencial aos dados da emissão de CO_2. O banco de daods do Carbon Dioxide Information Analysis Center (CDIAC) e de Oak Ridge National Laboratory (ORNL) fornecem os seguintes dados sobre a emissão do carbono CO_2.*

Período	$Quantidade(CO_2)$
1800	$5,0x10^9$
1850	$1,0x10^10$
1900	$1,5x10^10$
1950	$2,0x10^10$
2000	$2,5x10^10$

Tabela 6: $Quantidade(CO_2)$

Vamos utilizar a função exponencial $y = a_0e^{a_1x}$ que não é linear. Aplicamos ln de ambos os lados das equações

$$\begin{aligned} lny &= lna_0e^{a_1x} \\ lny &= lna_0 + lne^{a_1x} \\ lny &= lna_0 + a_1xlne \\ lny &= lna_0 + a_1x \end{aligned}$$

Assim devemos ajustar os novos dados (x_i, lny_i) uma reta. Segue daqui que devemos determinar as constantes lna_0 e a_1 os novos dados são

x	y	ln (y)
Período	$Quantidade(CO_2)$	$ln(Quantidade(CO_2))$
1,00	$5,0x10^9$	22,33270375
1,25	$1,0x10^{10}$	23,02585093
1,50	$1,5x10^{10}$	23,43131604
1,75	$2,0x10^{10}$	23,71899811
2,00	$2,5x10^{10}$	23,94214166

Tabela 7: Cáculos para a tabela 6

$$\textit{Logo as matrizes são: } A = \begin{bmatrix} 1 & 1,00 \\ 1 & 1,25 \\ 1 & 1,50 \\ 1 & 1,75 \\ 1 & 2,00 \end{bmatrix} X = \begin{bmatrix} lna_0 \\ a_1 \end{bmatrix} lny = \begin{bmatrix} 22,33270375 \\ 23,02585093 \\ 23,43131604 \\ 23,71899811 \\ 23,94214166 \end{bmatrix}$$

Resolvendo o sistema $A^TAX = A^TY$, temos:

$$X = \begin{bmatrix} 20,94298828 \\ 1,5648092 \end{bmatrix}$$

$$\begin{aligned} lny &= lna_0 + a_1x \\ lny &= 20,94298828 + 1,5648092x \\ e^{lny} &= e^{20,94298828+1,5648092x} \\ y &= e^{20,94298828+1,5648092x} \end{aligned}$$

Aplicando o ponto 2010,onde x=2,05 na função temos

$$\begin{aligned} y &= e^{20,94298828+1,5648092x} \\ f(2.05) &= e^{20,94298828+1,5648092\cdot 2,05} \\ f(2.05) &= e^{24,15084714} \\ f(2.05) &\approx 3,08 \times 10^{10} \end{aligned}$$

Exemplo 3.2 *A disponibilidade de água potável no mundo vem caindo gradualmente. A tabela, com dados da América Latina baseados na realidade, apresenta a quantidade de água por habitante medida em 1000 m^3. Determine a curva da forma $y = a_0e^{a_1x}$ que melhor se ajusta aos dados. Veja a tabela:*

Vamos utilizar a função exponencial $y = a_0e^{a_1x}$ que não é linear. Aplicamos ln de ambos

Período	Disponibilidade da água
1950	105
1960	80,2
1970	61,7
1980	48,8
2000	28,3

Tabela 8: Disponibilidade da água

os lados das equações

$$\begin{aligned} lny &= lna_0 e^{a_1 x} \\ lny &= lna_0 + lne^{a_1 x} \\ lny &= lna_0 + a_1 x lne \\ lny &= lna_0 + a_1 x \end{aligned}$$

Assim devemos ajustar os novos dados (x_i, lny_i) *uma reta. Segue daqui que devemos determinar as constantes* lna_0 *e* a_1 *os novos dados são*

x	y	ln (y)
1,00	105	4,6539603501
1,25	80,2	4,3845235148
1,50	61,7	4,1222839309
1,75	48,8	3,8877303128
2,00	28,3	3,3428618046

Tabela 9: Cáculos para a tabela 8

Logo as matrizes são: $A = \begin{bmatrix} 1 & 1,00 \\ 1 & 1,25 \\ 1 & 1,50 \\ 1 & 1,75 \\ 1 & 2,25 \end{bmatrix}$ $X = \begin{bmatrix} lna_0 \\ a_1 \end{bmatrix}$ $lny = \begin{bmatrix} 4,6539603501 \\ 4,3845235148 \\ 4,1222839309 \\ 3,8877303128 \\ 3,3428618046 \end{bmatrix}$

Resolvendo o sistema $A^T AX = A^T Y$*, temos:*

$$X = \begin{bmatrix} 5,6929532610 \\ -1,0417298570 \end{bmatrix}$$

$$\begin{aligned} lny &= lna_0 + a_1 x \\ lny &= 5,6929532610 - 1,0417298570x \\ e^{lny} &= e^{5,6929532610 - 1,0417298570x} \\ y &= e^{5,6929532610 - 1,0417298570x} \end{aligned}$$

Aplicando o ponto 2020, onde x=3 na função temos

$$\begin{aligned} y &= e^{5,6929532610-1,0417298570\cdot 3} \\ f(3) &= e^{2,56776369} \\ f(3) &\approx 13,04 \end{aligned}$$

REFERÊNCIAS

POOLE, David. Álgebra linear. São Paulo, SP: Thomson Learning: Cengage Learning, 2004.

LEITHOLD, Louis. O cálculo com geometria analítica. São Paulo, SP: HARBRA, 1977 2 v.

ANTON, Howard; RORRES, Chris. Álgebra linear com aplicações. 8. ed. Porto Alegre: Bookman, 2001.

TAN, S. T. Matemática aplicada à administração e economia. São Paulo: Pioneira Thomson Learning, 2005.

BOLDRINI, José Luiz et al. Álgebra linear. 3. ed. ampl. e rev. São Paulo, SP: HARBRA, c 1986.

IEZZI, Gelson. Matemática ciências e aplicações. São Paulo, 2010.