

# Álgebra Linear – Semana 13

20 de Junho de 2017

## Conteúdo

1	Método dos Mínimos Quadrados	1
2	Fatoração QR e Mínimos Quadrados	4
3	Regressão Linear Simples	5
4	Regressão não linear	8
5	Regressão linear múltipla	9

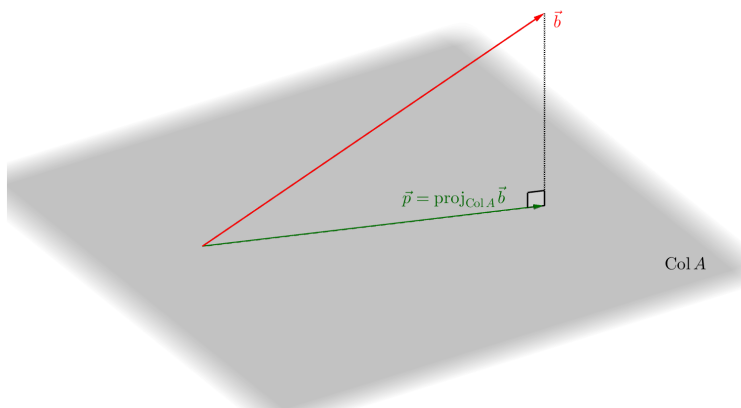
## 1 Método dos Mínimos Quadrados

Sistemas lineares aparecem como modelos matemáticos de vários fenômenos e em várias situações. Acontece que alguns sistemas simplesmente não possuem soluções e ficamos sem saber como proceder. O **Método dos Mínimos Quadrados** é uma técnica que nos permite, de forma aproximada, retirar alguma informação desses sistemas impossíveis. A terminologia se deve ao fato de que, como veremos, este método minimiza a soma dos quadrados dos erros obtidos na aproximação.

Como sabemos, resolver o sistema linear

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

consiste em encontrar um vetor  $\vec{x}$  que satisfaça a esta equação. Na terminologia que construímos ao longo do curso, isto significa que  $\vec{b}$  pertence ao espaço coluna da matriz  $A$ , isto é,  $\vec{b} \in \text{Col } A$ . Desta forma, não é possível resolver o sistema quando  $\vec{b} \notin \text{Col } A$ .



O Método dos Mínimos Quadrados consiste em olhar para o vetor  $\vec{p} = \text{proj}_{\text{Col } A} \vec{b}$  e resolver o sistema linear associado

$$A\vec{x} = \vec{p}.$$

Esta solução de  $A\vec{x} = \vec{p}$  é chamada de **solução de mínimos quadrados**. A ideia é (ver figura) considerar o vetor em  $\text{Col } A$  que é “mais próximo” de  $\vec{b}$  e cujo sistema linear associado possua solução. Neste contexto, estamos pensando em mais próximo no sentido que

$$\|\vec{b} - \text{proj}_{\text{Col } A} \vec{b}\| \leq \|\vec{b} - \vec{c}\|, \text{ para qualquer } \vec{c} \in \text{Col } A,$$

isto é, no sentido de ser a projeção a melhor aproximação de  $\vec{b}$  no espaço coluna de  $A$ . Escrevendo os vetores em coordenadas:

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ e } \text{proj}_{\text{Col } A} \vec{b} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix},$$

temos que o valor

$$\|\vec{b} - \text{proj}_{\text{Col } A} \vec{b}\| = \sum_{i=1}^n (b_i - p_i)^2$$

é chamado de **erro da aproximação**. Logo, como anunciado, a soma dos quadrados dos erros obtidos cada componente é o mínimo possível.

**Exemplo 1.** Considere o sistema linear

$$\begin{cases} x_1 + x_2 - 2x_3 = 3 \\ x_1 - 2x_3 = 2 \\ x_2 + 2x_3 = 0 \\ x_1 + x_2 - 2x_3 = 0 \end{cases} \iff A\vec{x} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} = \vec{b}.$$

Este sistema é claramente impossível, pois a primeira equação é inconsistente com a última. De forma mais automática, um passo no escalonamento da matriz aumentada associada revelaria a mesma conclusão:

$$[A \mid \vec{b}] = \begin{bmatrix} 1 & 1 & -2 & 3 \\ 1 & 0 & -2 & 2 \\ 0 & 1 & 2 & 0 \\ -1 & -1 & 2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 \\ 1 & 0 & -2 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Vamos tentar encontrar, conforme descrito acima, uma solução de mínimos quadrados. Passos a serem realizados:

- Para calcular a projeção de  $\vec{b}$  sobre  $\text{Col } A$ , precisamos, em primeiro lugar, obter uma base ortogonal do espaço  $\text{Col } A$ ; isto pode ser feito pelo Processo de Gram-Schmidt;
- Calcular a projeção  $\vec{p} = \text{proj}_{\text{Col } A} \vec{b}$ ;
- Resolver o sistema linear  $A\vec{x} = \vec{p}$ .

Embora seja possível realizar estes três passos, temos de convir que seria muito trabalhoso. Por isto, vamos analisar um pouco mais da teoria para encontrar um método mais eficaz.  $\triangleleft$

Suponhamos que:

- $A$  é uma matriz  $m \times n$ ,
- $\vec{x} \in \mathbb{R}^n$ ,
- $\vec{b} \in \mathbb{R}^m$  e

- o sistema  $A\vec{x} = \vec{b}$  não possui solução.

A solução de mínimos quadrados é a solução de  $A\vec{x} = \vec{p}$ , onde  $\vec{p} = \text{proj}_{\text{Col } A} \vec{b}$ . Observamos, por outro lado, que, sendo  $\vec{p}$  a projeção sobre o espaço coluna de  $A$ , o vetor  $\vec{b} - \vec{p}$  deve ser ortogonal a todos os elementos de  $\text{Col } A$ . Em particular, se escrevermos  $A$  em termos de suas colunas

$$A = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & \cdots & | \end{bmatrix}, \quad \left( \text{que, em particular, implica } A^T = \begin{bmatrix} - & \vec{a}_1 & - \\ - & \vec{a}_2 & - \\ & \vdots & \\ - & \vec{a}_n & - \end{bmatrix} \right)$$

devemos ter

$$\vec{a}_j \cdot (\vec{b} - \vec{p}) = 0, \text{ para todo } j \iff A^T(\vec{b} - \vec{p}) = \vec{0} \iff A^T\vec{b} = A^T\vec{p} = A^T A\vec{x}.$$

Esta última linha é válida porque o produto escalar  $\vec{a}_j \cdot (\vec{b} - \vec{p})$  é justamente a entrada  $j$  do vetor obtido ao multiplicar  $A^T(\vec{b} - \vec{p})$ .

Concluimos, desta forma, que, se  $\vec{x}$  é uma solução de mínimos quadrados, ou seja, se  $A\vec{x} = \vec{p}$ , então necessariamente

$$\boxed{A^T A\vec{x} = A^T \vec{b}.} \quad (1)$$

**Exercício 1 (Teórico).** Justifique que o conjunto de soluções do sistema (1) coincide com o conjunto das soluções de mínimos quadrados do sistema  $A\vec{x} = \vec{b}$ .

Tudo isto implica que podemos utilizar a equação (1) para encontrar as soluções de mínimos quadrados de  $A\vec{x} = \vec{b}$ .

**Exemplo 2 (De volta ao Exemplo 1).** Calculamos

$$A^T A = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ -2 & -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -6 \\ 2 & 3 & -2 \\ -6 & -2 & 16 \end{bmatrix}$$

e

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ -2 & -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -10 \end{bmatrix}$$

Para resolvermos o sistema  $A^T A\vec{x} = A^T \vec{b}$ , vamos escalonar a matriz aumentada associada:

$$\begin{bmatrix} 3 & 2 & -6 & 5 \\ 2 & 3 & -2 & 3 \\ -6 & -2 & 16 & -10 \end{bmatrix} \sim \begin{bmatrix} 3 & 2 & -6 & 5 \\ 0 & 5 & 6 & -1 \\ 0 & 2 & 4 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 2 & -6 & 5 \\ 0 & 5 & 6 & -1 \\ 0 & 0 & 4 & 1 \end{bmatrix}$$

Logo,

$$\begin{cases} 3x_1 + 2x_2 - 6x_3 = 5 \\ 5x_2 + 6x_3 = -1 \\ 4x_3 = 1 \end{cases}.$$

Assim,

$$x_3 = \frac{1}{4} \implies 5x_2 + 6 \cdot \frac{1}{4} = -1 \implies x_2 = -\frac{1}{2} \implies 3x_1 - 1 - \frac{3}{2} = 5 \implies x_1 = \frac{3}{2}.$$

Portanto, uma solução de mínimos quadrados é

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -1/2 \\ 1/4 \end{bmatrix}.$$

Observe como isto é muito menos trabalhoso do que o método que havíamos esquematizado quando discutimos este sistema no Exemplo 1.  $\triangleleft$

**Exercício 2.** Colocar em prática o método discutido no Exemplo 1 e comparar o resultado com o que obtivemos no Exemplo 2.

**Observação 3.** Caso a matriz  $A$  já seja uma matriz ortogonal, podemos calcular projeções diretamente, pois uma base ortogonal para  $\text{Col } A$  já está disponível desde o começo. Neste caso, ambos os métodos exigiriam um trabalho parecido.

Nossa aproximação linear foi encontrada de modo a minimizar o erro quadrático. Para encontrar o erro, deveríamos calcular

$$\|\vec{b} - \text{proj}_{\text{Col } A} \vec{b}\|.$$

No entanto, resolvemos de outra maneira e não calculamos a projeção. Observando que a solução de mínimos quadrados satisfaz  $A\vec{x} = \text{proj}_{\text{Col } A} \vec{b}$ , podemos calcular o erro por

$$\text{erro} = \|\vec{b} - A\vec{x}\|,$$

onde  $\vec{x}$  é a solução de mínimos quadrados.

No Exemplo 2 acima, podemos calcular o erro desta forma. De fato,

$$\begin{aligned} A\vec{x} &= \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3/2 \\ -1/2 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ 0 \\ -1/2 \end{bmatrix} \implies \vec{b} - A\vec{x} = \begin{bmatrix} 5/2 \\ 1 \\ 0 \\ 1/2 \end{bmatrix} \\ \implies \text{erro} &= \|\vec{b} - A\vec{x}\| = \sqrt{\frac{25}{4} + 1 + \frac{1}{4}} = \frac{\sqrt{30}}{2}. \end{aligned}$$

## 2 Fatoração QR e Mínimos Quadrados

Vimos que o processo de Gram-Schmidt nos permite escrever uma matriz  $A$  cujas linhas são linearmente independentes na forma  $A = QR$ , onde  $Q$  é uma matriz ortogonal,  $R$  é triangular superior invertível, e as colunas de  $Q$  formam uma base ortonormal para o espaço-coluna de  $A$ . Chamamos esta fatoração de  $A$  de fatoração QR.

A fatoração QR tem uma importante aplicação na solução de problemas em mínimos quadrados. Ocorre que, em muitos problemas, a matriz  $A^T A$  é uma matriz *mal-condicionada*, ou seja uma matriz que é muito sensível aos erros de aproximação cometidos em cálculos numéricos usualmente executados no tratamento de problemas aplicados. Os exemplos que virão a seguir nos mostram situações em que há uma grande variação de magnitude entre as entradas da matriz  $A^T A$ , e isto costuma ser fonte de instabilidade numérica.

Para isso, a fatoração QR oferece uma solução alternativa em mínimos quadrados para um sistema linear  $Ax = b$ : Se  $A = QR$ , então

$$A^T = R^T Q^T$$

e

$$A^T A = R^T Q^T Q R = R^T R,$$

onde usamos o fato de que  $Q^{-1} = Q^T$ . Assim as equações normais  $A^T A\vec{x} = A^T \vec{b}$  se reduzem a

$$R^T R\vec{x} = R^T Q^T \vec{b}.$$

Ainda, como  $R$  é invertível podemos eliminar (cancelar)  $R^T$  na equação acima, obtendo o seguinte:

**Teorema 4.** Se  $A = QR$  com  $Q$  ortogonal e  $R$  triangular superior invertível, então a solução em mínimos quadrados do sistema  $Ax = b$  é única e dada pela solução do sistema

$$R\vec{x} = Q^T \vec{b}.$$

Note que o fato de  $R$  ser triangular superior torna a resolução do sistema acima pouco custosa do ponto de vista computacional.

Alternativamente temos uma expressão explícita para solução em mínimos quadrados:

$$\vec{x} = R^{-1}Q^T\vec{b}.$$

É importante notar que existem algoritmos para fatoração QR que são *numericamente estáveis*, ou seja pouco suscetíveis a erros de aproximação ou arredondamento. Tais algoritmos são diferentes do processo de Gram Schmidt usual.

### 3 Regressão Linear Simples

Vamos apresentar uma aplicação de algumas das técnicas da seção anterior à Estatística ou Econometria. Uma **regressão linear simples** é uma equação, tipicamente da forma

$$y = a + bx,$$

para estimar os valores  $(x, y)$  apenas conhecendo alguns valores específicos  $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$ . A ideia é tentar capturar como que mudanças na variável independente  $x$  afetam a variável dependente  $y$ .

**Exemplo 5.**<sup>1</sup> Com o intuito de analisar se é razoável supor que há uma relação linear entre a idade de um motorista e quão longe ele consegue ver, uma empresa (Last Resource, Inc., Bellefonte, PA) coletou dados de 30 motoristas. Para simplificar as nossas contas, vamos listar abaixo *apenas alguns* destes dados.

Idade	Distância (em m)
20	590
32	410
41	460
49	380
66	350

Podemos pensar em  $y$  como a distância e em  $x$  como a idade. Gostaríamos de achar uma relação linear da forma

$$y = a + bx.$$

Desta forma, os dados obtidos implicam que

$$\begin{cases} b + 20a = 590 \\ b + 32a = 410 \\ b + 41a = 460 \\ b + 49a = 380 \\ b + 66a = 350 \end{cases} \iff \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix}$$

Ora, dificilmente um sistema linear com duas incógnitas e cinco equações terá solução (só terá solução se todos os pontos do conjunto de dados estiverem alinhados em uma reta perfeita!). Vamos procurar por uma solução de mínimos quadrados. Isto é **regressão linear simples**.

Denotando

$$A = \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \quad \text{e} \quad \vec{b} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix}$$

<sup>1</sup>Exemplo adaptado de <https://onlinecourses.science.psu.edu/stat501/node/257>

precisamos calcular

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \end{bmatrix} \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} = \begin{bmatrix} 5 & 208 \\ 208 & 9832 \end{bmatrix}$$

e

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \end{bmatrix} \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} = \begin{bmatrix} 2190 \\ 85500 \end{bmatrix}.$$

Em seguida, a matriz associada aumentada pode ser reduzida por escalonamento:

$$[A^T A \mid \vec{b}] = \begin{bmatrix} 5 & 208 & 2190 \\ 208 & 9832 & 85500 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 468510/737 \\ 0 & 1 & -7005/1474 \end{bmatrix} \implies \begin{cases} a \simeq 635.7 \\ b \simeq -4.75 \end{cases}.$$

Os números são feios, mas as contas feitas foram as que sempre fizemos ao escalonar uma matriz até sua forma escalonada reduzida.

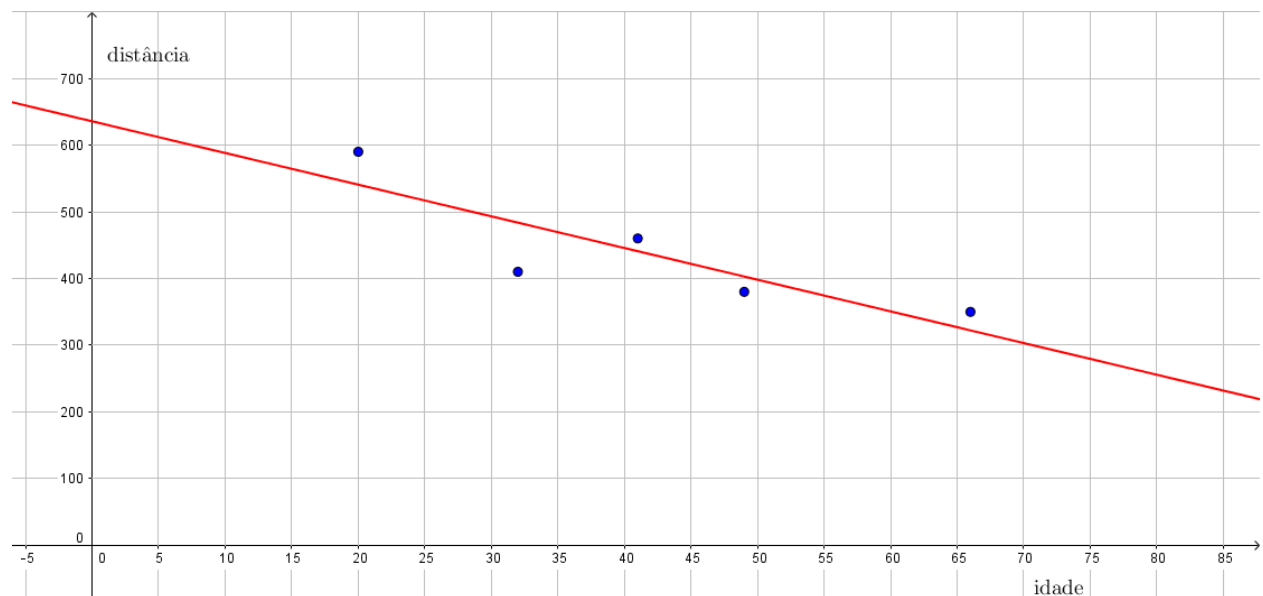
A conclusão é que a reta de mínimos quadrados que melhor aproxima os nossos dados é a reta

$$y = a + bx = 635.7 - 4.75x.$$

O erro de mínimos quadrados nesta aproximação (ou norma do resíduo) pode ser calculado como

$$A\vec{x} = \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \begin{bmatrix} 635.7 \\ 4.75 \end{bmatrix} \simeq \begin{bmatrix} 730.7 \\ 787.7 \\ 830.45 \\ 868.45 \\ 949.2 \end{bmatrix} \simeq \text{proj}_{\text{Col } A} \vec{b} \implies \text{erro} = \|\vec{b} - A\vec{x}\| \simeq 947.3$$

Na figura abaixo, mostramos um esboço da reta que melhor aproxima os dados deste exemplo. ◁



De uma maneira geral, para encontrar uma **reta de melhor ajuste** a uma quantidade de pontos (dados coletados de algum problema)

$$(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k),$$

devemos procurar pela solução  $(a, b)$  de mínimos quadrados do sistema linear

$$\begin{cases} a + bx_1 = y_1 \\ a + bx_2 = y_2 \\ \vdots \\ a + bx_k = y_k \end{cases} \rightsquigarrow \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_k \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}.$$

Vejamos um outro exemplo:

**Exemplo 6.** Encontrar a reta que melhor se ajusta aos pontos do plano

$$(1, 2), (2, 3), (2, 4), (3, 2), (4, 3), (4, 4), (5, 3), (5, 5), (6, 4).$$

Como vimos acima, podemos considerar

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 4 \\ 1 & 5 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \text{ e } \vec{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 2 \\ 3 \\ 4 \\ 3 \\ 5 \\ 4 \end{bmatrix}$$

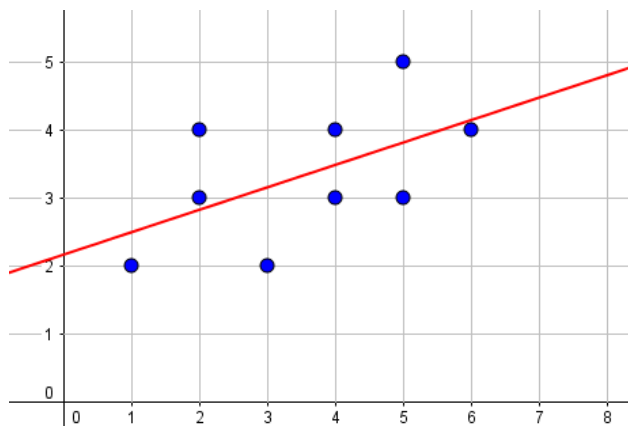
Assim,

$$A^T A = \begin{bmatrix} 9 & 32 \\ 32 & 136 \end{bmatrix} \text{ e } A^T \vec{b} = \begin{bmatrix} 30 \\ 114 \end{bmatrix}.$$

A solução de mínimos quadrados, que é a solução de  $A^T A \vec{x} = A^T \vec{b}$ , é dada por

$$b = 54/25 = 2.16, a = 33/100 = 0.33.$$

A reta procurada é portanto  $y = 0.33x + 2.16$ .



## 4 Regressão não linear

Podemos também procurar (em contraste com a seção anterior) por funções não lineares que se ajustem a um conjunto de pontos. Por exemplo, dado um certo conjunto de pontos

$$(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k),$$

vamos procurar por uma parábola

$$y = a + bx + cx^2$$

que melhor se ajuste a este conjunto de pontos, no sentido de que a soma dos quadrados dos erros seja a menor possível. Isto corresponde a procurar pela solução  $(a, b, c)$  de mínimos quadrados do seguinte sistema linear

$$\begin{cases} a + bx_1 + cx_1^2 = y_1 \\ a + bx_2 + cx_2^2 = y_2 \\ \vdots \\ a + bx_k + cx_k^2 = y_k \end{cases} \iff \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_k & x_k^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}.$$

Vejamos como ficaria a parábola que se ajusta aos dados do Exemplo 5:

**Exemplo 7.** Nossos pontos no Exemplo 5 são

$$(20, 590), (32, 410), (41, 460), (49, 380), (66, 350).$$

Para encontrar a parábola de melhor ajuste como acima, devemos procurar pela solução de mínimos quadrados do sistema

$$\begin{bmatrix} 1 & 20 & 400 \\ 1 & 32 & 1024 \\ 1 & 41 & 1681 \\ 1 & 49 & 2401 \\ 1 & 66 & 4356 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} \iff A\vec{x} = \vec{b}.$$

Calculamos (com o auxílio de uma calculadora)

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \\ 400 & 1024 & 1681 & 2401 & 4356 \end{bmatrix} = \begin{bmatrix} 5 & 208 & 9862 \\ 208 & 9862 & 514834 \\ 9862 & 514834 & 28773874 \end{bmatrix},$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \\ 400 & 1024 & 1681 & 2401 & 4356 \end{bmatrix} \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} = \begin{bmatrix} 2190 \\ 85500 \\ 3866080 \end{bmatrix}$$

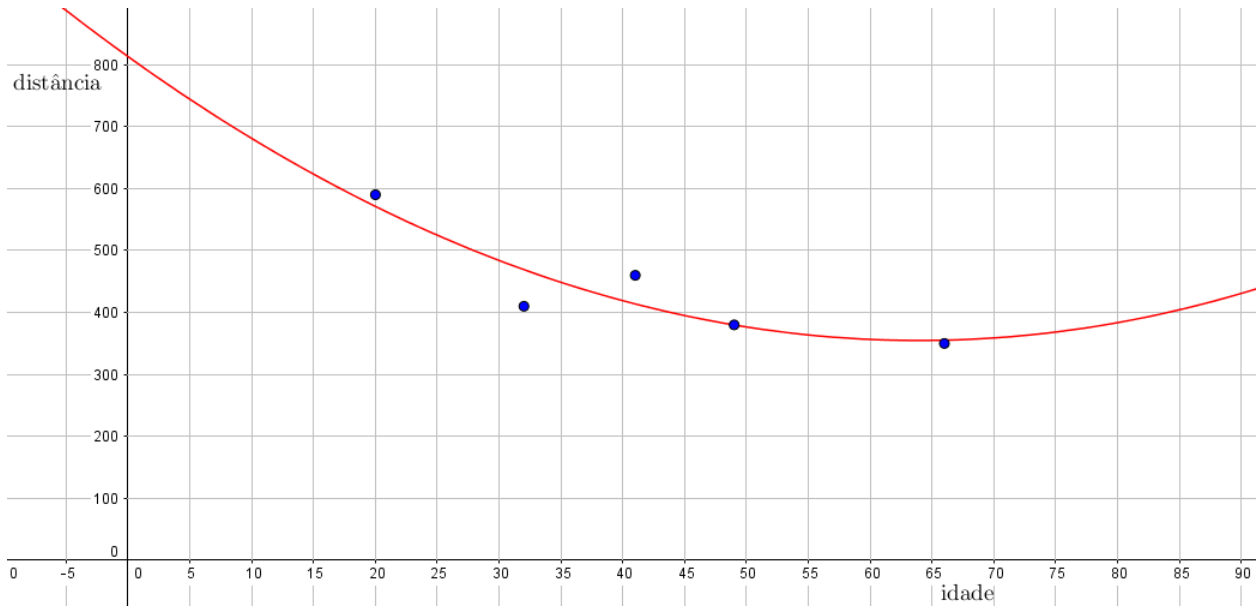
e resolvemos (escalonando, por exemplo, com o auxílio de um computador)

$$\begin{bmatrix} 5 & 208 & 9862 & 2190 \\ 208 & 9862 & 514834 & 85500 \\ 9862 & 514834 & 28773874 & 3866080 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 28482529840/35036713 \\ 0 & 1 & 0 & -1505841055/105110139 \\ 0 & 0 & 1 & 11779375/105110139 \end{bmatrix}.$$

Logo, aproximando estas frações, obtemos

$$\begin{cases} a \simeq 812.934 \\ b \simeq -14.326 \\ c \simeq 0.112 \end{cases}$$





e a parábola desejada é, aproximadamente

$$y = 812.934 - 14.326x + 0.112x^2.$$

Observamos que, apesar de o erro quadrático ser provavelmente menor, esta parábola se torna crescente a partir de um certo momento, fazendo com que o modelo não seja tão razoável para idades maiores. Por exemplo, é pouco provável que (na média), pessoas com 95 anos vejam a uma distância maior do que pessoas de 65 anos, como sugere a parábola acima.

## 5 Regressão linear múltipla

Uma **regressão linear múltipla** é o problema análogo à regressão linear simples no caso em que a variável dependente pode depender de mais fatores independentes. Tipicamente, queremos encontrar uma equação afim que melhor se ajusta a alguns dados conhecidos.

No caso especial em que  $y$  depende de apenas outros dois fatores, escreveremos

$$z = a + bx + cy,$$

que, geometricamente, representa um plano no espaço tridimensional  $\mathbb{R}^3$ . Seja agora um conjunto de dados:

$$(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_k, y_k, z_k).$$

Queremos encontrar coeficientes  $(a, b, c)$  que satisfaçam:

$$\begin{cases} a + bx_1 + cy_1 = z_1 \\ a + bx_2 + cy_2 = z_2 \\ \vdots \\ a + bx_k + cy_k = z_k \end{cases} \iff \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_k & y_k \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}.$$

Quanto maior o número de dados, mais provável de o sistema ser impossível (podemos fazer a analogia geométrica de que por três pontos não colineares no espaço passa um único plano; se aumentarmos o número de pontos, mais difícil que haja um plano contendo todos eles). Por isto, procuramos por uma solução de mínimos quadrados.

**Exemplo 8.** Uma pesquisa com 214 mulheres na em uma universidade americana<sup>2</sup> coletou informações sobre a altura das participantes, assim como a altura de seus pais. Abaixo, listamos *apenas alguns destes dados*, para que nossas contas não fiquem tão extensas. Fizemos também uma mudança de unidades nas alturas (de polegadas) para centímetros,

Altura	Altura Mãe	Altura Pai
152	155	165
162	155	160
165	170	173
170	163	183
173	168	183
183	165	183

Queremos encontrar uma solução de mínimos quadrados para o sistema linear

$$\begin{bmatrix} 1 & 155 & 165 \\ 1 & 155 & 160 \\ 1 & 170 & 173 \\ 1 & 163 & 183 \\ 1 & 168 & 183 \\ 1 & 165 & 183 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 152 \\ 162 \\ 165 \\ 170 \\ 173 \\ 183 \end{bmatrix} \rightsquigarrow A\vec{x} = \vec{b}.$$

Calculamos

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 155 & 155 & 170 & 163 & 168 & 165 \\ 165 & 160 & 173 & 183 & 183 & 183 \end{bmatrix} \begin{bmatrix} 1 & 155 & 165 \\ 1 & 155 & 160 \\ 1 & 170 & 173 \\ 1 & 163 & 183 \\ 1 & 168 & 183 \\ 1 & 165 & 183 \end{bmatrix} = \begin{bmatrix} 6 & 976 & 1047 \\ 976 & 158968 & 170553 \\ 1047 & 170553 & 183221 \end{bmatrix}$$

e

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 155 & 155 & 170 & 163 & 168 & 165 \\ 165 & 160 & 173 & 183 & 183 & 183 \end{bmatrix} \begin{bmatrix} 152 \\ 162 \\ 165 \\ 170 \\ 173 \\ 183 \end{bmatrix} = \begin{bmatrix} 1005 \\ 163689 \\ 175803 \end{bmatrix}.$$

Por escalonamento,

$$\begin{bmatrix} 6 & 976 & 1047 & 1005 \\ 976 & 158968 & 170553 & 163689 \\ 1047 & 170553 & 183221 & 175803 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2154573/145769 \\ 0 & 1 & 0 & 14475/145769 \\ 0 & 0 & 1 & 114081/145769 \end{bmatrix} \Rightarrow \begin{cases} a \simeq 14.781 \\ b \simeq 0.099 \\ c \simeq 0.783 \end{cases}$$

A equação de melhor ajuste procurada é, portanto, aproximadamente,

$$z \simeq 14.781 + 0.099x + 0.783y.$$

Tente calcular sua altura  $z$  a partir da altura de sua mãe  $x$  e de seu pai  $y$ . O teste deve funcionar melhor para mulheres!

<sup>2</sup>Referência: <https://onlinecourses.science.psu.edu/stat501/node/292>

