

## Aproximação usando o método dos mínimos quadrados

Luiz Daniel Gonçalves<sup>1</sup>  
Gilcélia Regiane de Souza<sup>2</sup>

**Resumo:** A aproximação de uma função pode ser desejável por diversos motivos, incluindo eficiência computacional, considerações teóricas sobre a grandeza física representada pela função, ou pelo fato dela ser conhecida apenas parcialmente. Espera-se que as funções aproximadoras pertençam a uma classe de funções mais simples e sejam mais fáceis de calcular e manipular, por derivadas, integrais, etc.

O objetivo deste trabalho é fazer um estudo das aproximações de funções e/ou dados através do Método dos Quadrados Mínimos. O método tem como base, aproximar funções quaisquer por uma combinação linear de funções preestabelecidas. O trabalho é finalizado com uma proposta de aplicação didática para alunos do Ensino Médio, seguida de alguns exemplos desse conteúdo em sala de aula.

**Palavras-chave:** Estudo das Aproximações, Mínimos Quadrados, Ajuste de Dados, Função Aproximada, Aproximação discreta.

## 1 Introdução

Problemas relacionados a aproximação de funções são de fundamental importância em diversas áreas de conhecimento, em particular, nas áreas experimentais principalmente quando deseja-se ajustar dados. No presente trabalho, as funções são aproximadas utilizando-se o Método dos Mínimos Quadrados (M.M.Q.). Este é um bom método, pois é simples e a qualidade dos resultados não é afetada pelo grau de incerteza dos experimentos. Esta última característica evidencia o método para problemas práticos.

No presente estudo, aborda-se um pouco de Álgebra Linear, com os principais temas pertinentes à aproximação, tais como produto interno, norma de um vetor, espaço e subespaço vetorial. Abordamos também a teoria da aproximação e o critério de aproximação escolhido, os mínimos quadrados.

Após o estudo do método, apresenta-se alguns exemplos como ilustração. Finaliza-se o trabalho com uma aplicação no ensino médio, propondo uma forma didática. Apresenta-se também algumas razões, do porque da apresentação, para alunos do ensino médio, do método dos Mínimos Quadrados, para aproximação de funções. Observa-se que este pode

---

<sup>1</sup>Aluno de Mestrado Profissional em Matemática, Turma 2012  
Instituição: Universidade Federal de São João del-Rei - UFSJ / Campus Alto Paraopeba - CAP  
E-mail: luizdaniel@unifemm.edu.br

<sup>2</sup>Orientadora do Trabalho de Conclusão de Curso  
Departamento de Física e Matemática - Defim, UFSJ/CAP  
E-mail: gilcelia@ufsj.edu.br

proporcionar uma enorme motivação aos estudos e o ganho no desenvolvimento matemático dos mesmos.

Podemos estimular os estudantes informando-os que ao se determinar uma função aproximada, é possível estimar o número de habitantes de uma cidade, pesos de materiais, pessoas, ou ainda número de bactérias, etc, em qualquer período que se desejar.

## 2 Um pouco de Álgebra Linear

Nesta seção encontra-se os conceitos de Álgebra Linear utilizados neste trabalho, mais detalhes sobre a teoria de Álgebra Linear podem ser obtidos, por exemplo, nos livros [1, 5, 9]

Seja  $\mathcal{A}$  um conjunto arbitrário de objetos sobre o qual definem-se duas operações, adição:

$$(u, v) \in \mathcal{A} \times \mathcal{A} \rightarrow u + v \in \mathcal{A},$$

e multiplicação escalar (real):

$$(\alpha, u) \in \mathbb{R} \times \mathcal{A} \rightarrow \alpha u \in \mathcal{A}.$$

Então,  $\mathcal{A}$  é dito *espaço vetorial* se todos os objetos  $u, v$  e  $w$  em  $\mathcal{A}$  e todos os escalares  $\alpha, \beta$  satisfazem as seguintes propriedades:

- $A_1$ )  $u + v = v + u$ , para todo  $u, v \in \mathcal{A}$ ,
- $A_2$ )  $(u + v) + w = u + (v + w)$ , para todo  $u, v, w \in \mathcal{A}$ ,
- $A_3$ ) Existe um vetor nulo  $0 \in \mathcal{A}$   $u + 0 = 0 + u = u$ , para todo  $u \in \mathcal{A}$ ,
- $A_4$ ) Para todo  $u \in \mathcal{A}$ , existe  $(-u) \in \mathcal{A}$   $u + (-u) = 0$
- $M_1$ )  $\alpha(u + v) = \alpha u + \alpha v$ , para todo  $\alpha \in \mathbb{R}$ , para todo  $u, v \in \mathcal{A}$ ,
- $M_2$ )  $(\alpha + \beta)u = \alpha u + \beta u$ , para todo  $\alpha, \beta \in \mathbb{R}$ , para todo  $u \in \mathcal{A}$ ,
- $M_3$ )  $(\alpha \beta)u = \alpha(\beta u)$ , para todo  $\alpha, \beta \in \mathbb{R}$ , para todo  $u \in \mathcal{A}$
- $M_4$ )  $1.u = u$ , para todo  $u \in \mathcal{A}$ .

Cada objeto em  $\mathcal{A}$  recebe o nome de vetor.

Chama-se *combinação linear* dos vetores  $v_1, v_2, \dots, v_n$  o vetor que pode ser escrito na forma

$$v = a_1 v_1 + a_2 v_2 + \dots + a_n v_n$$

sendo  $a_1, a_2, \dots, a_n$  escalares.

Um conjunto de vetores  $\{v_1, v_2, \dots, v_n\}$  é *linearmente independente* (L.I.), se a equação  $a_1 v_1 + a_2 v_2 + \dots + a_n v_n = 0$  admite somente a solução  $a_1 = a_2 = \dots = a_n = 0$ . No caso, em que exista algum  $a_i \neq 0$ , tem-se que  $\{v_1, v_2, \dots, v_n\}$  é *linearmente dependente* (L.D.).

Se cada vetor no espaço vetorial  $\mathcal{A}$  for uma combinação linear de  $\{v_1, v_2, \dots, v_n\}$ , o espaço vetorial  $\mathcal{A}$  é *gerado* por estes vetores. Seja  $S = \{v_1, v_2, \dots, v_n\}$ , então  $S$  é um *conjunto gerador*. Se  $S$  for um conjunto L.I., este será uma *base* para  $\mathcal{A}$ , uma vez que  $S$  gera  $\mathcal{A}$ . Além disso, a *dimensão* de  $\mathcal{A}$  é o número de elementos da base, ou seja, o número de vetores de  $S$ , notação:  $\dim \mathcal{A} = n$ .

Um produto interno sobre  $\mathcal{A}$  é uma função

$$\langle , \rangle : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$$

que satisfaz as seguintes propriedades:

- (i)  $\langle u, v \rangle = \langle v, u \rangle$
- (ii)  $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- (iii)  $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
- (iv)  $\langle u, u \rangle \geq 0$  e  $\langle u, u \rangle = 0$  se, e somente se,  $u = 0$

Abaixo encontra-se dois exemplo de produto interno:

1.

$$\langle u, v \rangle = u^\top v = \sum_{i=1}^n u_i v_i.$$

o produto interno entre vetores.

2.

$$\langle u, v \rangle = \int_a^b f(x)g(x)dx$$

o produto interno entre funções, neste caso,  $f(x)$  e  $g(x)$  são conhecidas e integráveis em  $(a, b)$ .

A *norma* de um vetor  $u$ , em símbolo  $\|u\|$ , é uma função  $n : \mathcal{A} \rightarrow \mathbb{R}$ , satisfazendo as seguintes condições:

- $N_1$ )  $\|u\| \geq 0$  e  $\|u\| = 0$  se, e somente se,  $u = 0$  (vetor nulo),
- $N_2$ )  $\|\lambda u\| = |\lambda| \|u\|$  para todo escalar  $\lambda$ ,
- $N_3$ )  $\|u + v\| \leq \|u\| + \|v\|$  (desigualdade triangular).

Um *espaço vetorial normado* é um espaço vetorial onde está definida uma norma.

Seja  $\mathcal{A}$  um espaço vetorial de dimensão  $n$ . Os vetores  $u_1, u_2, \dots, u_n$  formam uma *base ortonormal* de  $\mathcal{A}$  se eles forem vetores ortonormais, ou seja, se:

$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \text{ com } i, j \in \{1, \dots, n\}$$

Lembrando, uma sequência de vetores é ortonormal se cada um dos seus elementos tem norma um e se o produto interno de dois distintos é nulo.

Seja  $\mathcal{A}$  um espaço vetorial. Dados os vetores  $u$  e  $v \in \mathcal{A}$ , define-se distância entre  $u$  e  $v$ , em símbolo  $d(u, v)$ , o comprimento do vetor  $u - v$ , ou seja:

$$d(u, v) = \|u - v\| \rightarrow d(u, v) = \sqrt{\langle u - v, u - v \rangle}$$

Observa-se, ainda que uma aplicação  $d : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$  satisfaz as seguintes condições:

- $c_1$ )  $d(u, v) \geq 0$  e  $d(u, v) = 0$  se e somente se  $u = v$ ,
- $c_2$ )  $d(u, v) = d(v, u)$ , para todo  $v, u \in \mathcal{A}$ ,
- $c_3$ )  $d(u, v) \leq d(u, w) + d(w, v)$ , para todo  $u, v, w \in \mathcal{A}$ .

Seja  $\mathcal{A}$  um espaço vetorial sobre  $\mathbb{R}$ . Um *subespaço vetorial* de  $\mathcal{A}$  é um subconjunto  $W \subset \mathcal{A}$ , tal que:

- (i)  $W \neq \emptyset$ , ou  $0 \in W$  (vetor nulo);
- (ii) Para todo  $u, v \in W$ ,  $u + v \in W$
- (iii) Para todo  $\alpha \in \mathbb{R}$ , e  $\forall u \in W$ ,  $\alpha(u) \in W$

Seja  $\mathcal{R}$  um subespaço do espaço vetorial  $\mathcal{A}$  e considere  $v$  um vetor de  $\mathcal{A}$  que não pertence ao subespaço  $\mathcal{R}$ . Então o vetor  $v$  pode ser escrito de forma única como

$$v = w_1 + w_2$$

sendo que  $w_1$  é paralelo ao subespaço  $\mathcal{R}$  e  $w_2$  é ortogonal ao subespaço  $\mathcal{R}$ , veja a Figura 1.

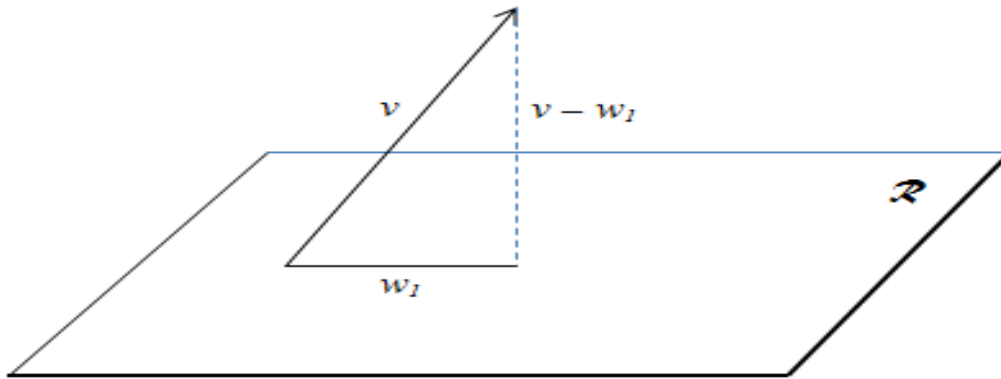


Figura 1: Projeção de  $v$  sobre  $\mathcal{R}$ .

O vetor  $w_1$ , que na verdade pertence à  $\mathcal{R}$ , é chamado de projeção de  $v$  sobre o subespaço  $\mathcal{R}$ , também denotado por  $\text{proj}_{\mathcal{R}}v$ . Considere  $v_1, v_2, \dots, v_r$  uma base ortogonal do subespaço  $\mathcal{R}$ , a *projeção de  $v$  sobre  $\mathcal{R}$*  é definida com a soma das projeções de  $v$  sobre cada vetor da base

$$\text{proj}_{\mathcal{R}}v = \text{proj}_{v_1}v + \text{proj}_{v_2}v + \dots + \text{proj}_{v_r}v, \quad (1)$$

lembre-se  $\text{proj}_u v = \frac{\langle v, u \rangle}{\|u\|^2} u$ .

Tem-se assim, um vetor  $w_1 \in \mathcal{R}$  tal que  $v - w_1$  é ortogonal a todo vetor de  $\mathcal{R}$ .

Considere  $a_1, a_2, \dots, a_r$  uma base de  $\mathcal{R}$ . Como  $w_1 \in \mathcal{R}$ , então  $w_1$  pode ser escrito como combinação linear dos vetores da base de  $\mathcal{R}$ , ou seja:

$$w_1 = \gamma_1 a_1 + \gamma_2 a_2 + \dots + \gamma_r a_r.$$

Deseja-se determinar as coordenadas  $\gamma_1, \gamma_2, \dots, \gamma_r$  de  $w_1$ .

Uma vez que  $v - w_1$  é ortogonal a todo vetor de  $\mathcal{R}$ , então  $v - w_1$  é ortogonal a todos os vetores de uma base de  $\mathcal{R}$ , logo

$$\langle v - w_1, a_j \rangle = 0 \quad \text{para} \quad j = 1, 2, \dots, r,$$

isto é,

$$\langle v - (\gamma_1 a_1 + \gamma_2 a_2 + \dots + \gamma_r a_r), a_j \rangle = 0 \quad j = 1, 2, \dots, r.$$

O que nos fornece:

$$\gamma_1 \langle a_1, a_j \rangle + \gamma_2 \langle a_2, a_j \rangle + \dots + \gamma_r \langle a_r, a_j \rangle = \langle v, a_j \rangle \quad j = 1, 2, \dots, r.$$

Para encontrar as coordenadas de  $w_1$  na base  $\{a_1, a_2, \dots, a_r\}$ , deve-se resolver o sistema de equações lineares:

$$\begin{pmatrix} \langle a_1, a_1 \rangle & \langle a_2, a_1 \rangle & \dots & \langle a_r, a_1 \rangle \\ \langle a_1, a_2 \rangle & \langle a_2, a_2 \rangle & \dots & \langle a_r, a_2 \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle a_1, a_r \rangle & \langle a_2, a_r \rangle & \dots & \langle a_r, a_r \rangle \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_r \end{pmatrix} = \begin{pmatrix} \langle v, a_1 \rangle \\ \langle v, a_2 \rangle \\ \vdots \\ \langle v, a_r \rangle \end{pmatrix},$$

O sistema acima é chamado de *sistema linear normal* e tem uma única solução, uma vez que a matriz dos coeficientes é simétrica positiva definida. O fato da matriz ser simétrica, vem da propriedade do produto interno ( $\langle a_i, a_j \rangle = \langle a_j, a_i \rangle$ ) e é positiva definida, pois cada elemento da base é linearmente independente. Lembre-se, uma matriz simétrica positiva definida tem todos seus autovalores positivos, assim o determinante é não nulo, pois o determinante é o produto dos autovalores.

**Teorema 2.1** *Se  $\mathcal{R}$  for um subespaço de dimensão finita de um espaço vetorial com produto interno sobre  $\mathcal{A}$  e se  $v$  for um vetor pertencente a  $\mathcal{A}$ , então  $\text{proj}_{\mathcal{R}} v$  será a melhor aproximação para  $v$  por  $\mathcal{R}$  no sentido que*

$$\|v - \text{proj}_{\mathcal{R}} v\| < \|v - w\| \quad (2)$$

para todo vetor  $w \in \mathcal{R}$ , diferente de  $\text{proj}_{\mathcal{R}} v$ .

A demonstração encontra-se em [9].

Por último, observa-se o seguinte a respeito das matrizes. Sejam  $A$  uma matriz  $m \times n$  e  $e_j$  o vetor canônico, ou seja, o vetor cuja posição  $j$  tem o valor um e demais posições valor zero. Observa-se que  $Ae_j$  é a  $j$ -ésima coluna de  $A$ , isto é,  $Ae_j = A_{*j} = A_{ij}$  com  $i = 1, 2, \dots, m$ .

Se  $A = [a_{ij}(t)]$  é uma matriz cujas entradas são funções de uma variável  $t$ , a derivada de  $A$  com relação a  $t$  é definida como a matriz das derivadas, isto é:

$$\frac{\partial A}{\partial t} = \left[ \frac{\partial a_{ij}}{\partial t} \right].$$

Mostra-se que

$$\frac{\partial AB}{\partial t} = \frac{\partial A}{\partial t} B + A \frac{\partial B}{\partial t}.$$

De fato, sejam  $A = [a_{ij}(t)]_{m \times n}$  e  $B = [b_{ij}(t)]_{n \times p}$  duas matrizes. Sabemos que o produto  $AB$  é definido como a matriz  $C = [c_{ij}(t)]_{m \times p}$ , tal que

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} \quad \text{para todo} \quad 1 \leq i \leq m$$

Derivando esta expressão, encontra-se:

$$\frac{\partial}{\partial t} \left[ \sum_{k=1}^m a_{ik}(t) b_{kj}(t) \right] = \sum_{k=1}^m \frac{\partial}{\partial t} [a_{ik}(t)] b_{kj}(t) + \sum_{k=1}^m a_{ik}(t) \frac{\partial}{\partial t} [b_{kj}(t)]$$

que é igual a

$$\frac{\partial AB}{\partial t} = \frac{\partial A}{\partial t} B + A \frac{\partial B}{\partial t}.$$

### 3 Teoria de Aproximação

O *espaço de funções a serem aproximadas*  $\mathcal{F}$  é um espaço vetorial de funções normado. Por exemplo,  $\mathcal{F}$  pode ser o espaço das funções contínuas. O *espaço de funções aproximantes*  $\mathcal{A}$  é um subespaço vetorial de dimensão finita  $n$  do conjunto de funções. Isto é, a função aproximadora  $s$  tem sempre a forma

$$s(x) = \sum_{j=1}^n c_j \phi_j(x), \quad (3)$$

para todo  $x$  em  $\mathbb{D}$ ; sendo que  $(\phi_1, \phi_2, \dots, \phi_n)$  é uma base do espaço  $\mathcal{A}$  (uma *base de aproximação*), cujos elementos são funções. Nota-se que o mesmo espaço de aproximação  $\mathcal{A}$  tem uma infinidade de bases de aproximação.

O problema de aproximação é determinar o vetor de coeficientes  $c = (c_1, \dots, c_n)$  desta combinação, de modo a satisfazer os critérios desejados.

Dada a função  $f \in \mathcal{F}$  e uma aproximação  $s \in \mathcal{A}$ , define-se o *resíduo* como a diferença

$$r = f - s.$$

Esta função pertence ao *espaço dos resíduos*.

O *critério de aproximação* é uma regra que permite dizer se uma determinada função  $s \in \mathcal{A}$  é uma “boa aproximação” para uma dada função  $f \in \mathcal{F}$ . Neste trabalho estuda-se um critério de aproximação muito usado, que é o critério da *aproximação ótima*, segundo o qual  $s$  deve minimizar a norma do resíduo.

### 4 Míminos quadrados

Considere  $f$  uma função no espaço das funções a serem aproximadas  $\mathcal{F}$ ,  $s(x)$  uma função no espaço das funções aproximadoras  $\mathcal{A}$ , os espaços das funções aproximadoras é sempre constituído pelas combinações lineares de uma *base de aproximação*  $\phi_1, \phi_2, \dots, \phi_n$ ; isto é, a função aproximadora tem a forma

$$s(x) = \sum_{j=1}^n c_j \phi_j(x). \quad (4)$$

sendo que  $c_1, c_2, \dots, c_n$  são os coeficientes a determinar, e cada  $\phi$  é uma função. A partir da escolha de um critério de aproximação obtem-se  $f \approx s$ .

Neste trabalho é utilizado o método dos mínimos quadrados, como critério de aproximação. Este é um bom método, devido a sua simplicidade e ao fato de que a qualidade dos resultados não é afetada pelo grau de incerteza dos experimentos. Esta última característica evidência o método para problemas práticos. Em muitas situações, conhece-se valores em determinados pontos, ou seja, tem-se  $(x_i, y_i)$ , onde cada  $y_i$  pode ser obtido, por exemplo, experimentalmente. Suponha que deseja-se obter valores para certos pontos desconhecidos, uma maneira de se resolver isto é procurar a expressão analítica de uma curva  $y = s(x)$  que melhor se ajusta a esse conjunto de pontos.

Suponha que os dados serão aproximados por uma função  $s$  do tipo

$$f(x_i) \approx s(x_i) = c_1 \phi_1(x_1) + c_2 \phi_2(x_2) + \dots + c_n \phi_n(x_m) \quad (5)$$

sendo as funções  $\phi_1, \phi_2, \dots, \phi_n$  preestabelecidas. O resíduo pontual é determinado por

$$\begin{aligned} r(x_i) &= f(x_i) - s(x_i) \\ &= f(x_i) - \sum_{j=1}^n c_j \phi_j(x_i), \end{aligned} \quad (6)$$

para  $i = 1, 2, \dots, m$ . Chamando  $r = (r(x_1), r(x_2), \dots, r(x_m))$ , observe que  $\langle r, r \rangle = \sum_{i=1}^m r_i^2$ .

Uma vez que, foi escolhido o método dos mínimos quadrados, tem-se que o critério de aproximação é minimizar a soma dos quadrados dos resíduos pontuais.

Note que

$$f(x_i) - \sum_{j=1}^n c_j \phi_j(x_i) \quad \text{para } i = 1, 2, \dots, m, \quad (7)$$

em que o número de pontos deve ser maior que o números das funções preestabelecidas. Pode-se reescrever como

$$\begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_m) \end{bmatrix} - \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \dots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \dots & \phi_n(x_m) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

Ou seja, na forma matricial,  $b - Ac$ , em que  $A$  é uma matriz  $m \times n$ ,  $c$  é uma matriz coluna  $n \times 1$  e  $b$  é uma matriz coluna  $m \times 1$ , o resíduo pode ser reescrito como

$$r = b - Ac,$$

então deve-se encontrar  $c \in \mathbb{R}^n$  tal que  $\|r\|_2^2$  seja mínima. Observe que,

$$\|r\|_2^2 = \|b - Ac\|_2^2 = \langle b - Ac, b - Ac \rangle = \langle b, b \rangle - 2\langle Ac, b \rangle + \langle Ac, Ac \rangle,$$

isto é,

$$\|r\|_2^2 = b^\top b - 2c^\top A^\top b + c^\top A^\top A c. \quad (8)$$

Defina uma função  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  sendo

$$h(c) = h(c_1, c_2, \dots, c_n) = b^\top b - 2c^\top A^\top b + c^\top A^\top A c,$$

deseja-se encontrar  $c$  que minimize a expressão (8) para isto, deve-se encontrar o mínimo da função  $h$  (observa-se que  $h$  tem origem em uma expressão quadrática, assim um ponto crítico de uma função com esta característica é um ponto de mínimo). Logo, deve-se diferenciar a função  $h$  em relação a cada  $c_i$ . Ou seja,

$$\frac{\partial h}{\partial c_i} = -2 \frac{\partial c^\top}{\partial c_i} A^\top b + \frac{\partial c^\top}{\partial c_i} A^\top A c + c^\top A^\top \frac{\partial c}{\partial c_i} \quad (9)$$

Já que,  $\frac{\partial c}{\partial c_i} = e_i$  (vetor de base canônica), pode-se reescrever (9) como

$$\frac{\partial h}{\partial c_i} = -2e_i^\top A^\top b + e_i^\top A^\top A c + c^\top A^\top A e_i.$$

Observe que  $e_i^\top A^\top$  é a  $i$ -ésima linha da matriz transposta e que

$$e_i^\top A^\top A c + c^\top A^\top A e_i = 2e_i^\top A^\top A c \quad (\text{pela propriedade de produto interno})$$

usando a seguinte notação

$$e_i^\top A^\top = (A^\top)_{i*}$$

e o fato

$$\frac{\partial h}{\partial c_i} = 0 \quad (\text{uma vez que deseja-se os valores críticos de } h),$$

tem-se  $n$  equações

$$(A^\top)_{i*} A c = (A^\top)_{i*} b \quad \text{para} \quad i = 1, \dots, n. \quad (10)$$

Reescrevendo na forma matricial,

$$A^\top A c = A^\top b. \quad (11)$$

Nota-se que o mínimo de  $h$  ocorre em alguma solução do sistema (11). A partir disto, surge a seguinte questão, isso é verdade sempre? Em outras palavras, cada solução do  $A^\top A c = A^\top b$  é uma solução de mínimos quadrados? Para responder, deve-se mostrar que a função  $h$  atinge seu valor mínimo em cada solução do sistema (11).

Tome  $z$  uma solução qualquer da equação normal, ou seja,  $A^\top A z = A^\top b$ , e escreva  $h$  aplicada a tal ponto,

$$h(z) = b^\top b - 2z^\top A^\top b + z^\top A^\top A z = b^\top b - 2z^\top A^\top b + z^\top A^\top b,$$

isto é,

$$h(z) = b^\top b - z^\top A^\top b.$$

Agora considere  $y \in \mathbb{R}^{n \times 1}$ , qualquer, e seja  $u = y - z$ , assim  $y = z + u$ . Note que,

$$h(y) = h(z) + v^\top v, \quad \text{sendo} \quad v = Au. \quad (12)$$

De fato,

$$\begin{aligned} h(y) &= b^\top b - 2(z + u)^\top A^\top b + (z + u)^\top A^\top A (z + u) \\ &= b^\top b - 2[z^\top A^\top b + u^\top A^\top b] + z^\top A^\top A z + z^\top A^\top A u + u^\top A^\top A z + u^\top A^\top A u \\ &= b^\top b - 2z^\top A^\top b + z^\top A^\top A z - 2u^\top A^\top b + 2u^\top A^\top A z + u^\top A^\top A u, \end{aligned} \quad (13)$$



uma vez que  $z$  é solução da equação normal  $A^\top Az = A^\top b$  a expressão

$$-2u^\top A^\top b + 2u^\top A^\top Az = -2u^\top A^\top b + 2u^\top A^\top b = 0$$

e (13) pode ser simplificada, ou seja,

$$h(y) = b^\top b - 2z^\top A^\top b + z^\top A^\top Az + u^\top A^\top Au,$$

que é justamente a equação (12).

Usando a equação (12), e uma vez que  $v^\top v = \|v\|^2 \geq 0$ , conclui-se que

$$h(z) \leq h(y) \quad \text{para todo} \quad y \in \mathbb{R}^{n \times 1},$$

assim,  $h$  atinge seu valor mínimo para cada solução do sistema normal.

Logo quando se tem um problema de mínimos quadrados (caso discreto, ou discretização do caso contínuo) a busca de solução implica resolver o sistema normal  $A^\top Ac = A^\top b$ , sendo  $c$  o vetor coluna dos coeficientes da melhor aproximação. Vejamos uma ilustração abaixo.

**Exemplo 4.1** 1. *Resolva o sistema*

$$\begin{cases} x_1 + x_2 & = & 3 \\ -2x_1 + 3x_2 & = & 1 \\ 2x_1 - x_2 & = & 2 \end{cases}$$

A matriz aumentada  $[A|b]$  é,

$$\begin{bmatrix} 1 & 1 & 3 \\ -2 & 3 & 1 \\ 2 & -1 & 2 \end{bmatrix}$$

e escalonando tem-se

$$\begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 7/5 \\ 0 & 0 & 1/5 \end{bmatrix}$$

logo o sistema não tem solução. Mas pode-se achar a solução por mínimos quadrados. Ou seja,

2. *Encontre a solução do sistema acima,  $Ac = b$ , por mínimos quadrados.*

*Para isto, resolve-se a equação normal  $A^\top Ac = A^\top b$ . Resolvendo o produto  $A^\top A$ , encontramos*

$$A^\top A = \begin{bmatrix} 9 & -7 \\ -7 & 11 \end{bmatrix}$$

*Podemos perceber que a matriz  $A^\top A$  é não singular, logo a solução será  $c = (A^\top A)^{-1} A^\top b$ . Ou seja,*

$$c = \left[ \frac{83}{50}, \frac{71}{50} \right]^\top.$$

Mostrou-se que os espaços das funções aproximadoras é sempre constituído pelas combinações lineares de uma *base de aproximação*  $\phi_1, \phi_2, \dots, \phi_n$ ; isto é, função aproximadora tem a forma

$$s(x) = \sum_{j=1}^n c_j \phi_j(x), \quad (14)$$

onde  $\phi$  é um elemento da base do espaço aproximante e uma maneira de encontrar os coeficientes da melhor aproximação. Observa-se, também que além da determinação dos coeficientes  $c_1, c_2, \dots, c_n$  pode-se ter como objetivo a determinação de uma base. O método trabalha bem com várias bases, por exemplo, monômios  $x^n$ , funções trigonométricas  $\sin(\pi nx)$ , funções radiais (exemplo: gaussianas) e outras. A escolha da base vai depender do problema em questão.

O conjunto dos polinômios de grau menor ou igual a  $n-1$  é um espaço vetorial de dimensão  $n$ . Se  $\phi_1, \phi_2, \dots, \phi_n$  são elementos da base deste espaço escreve-se  $p(x)$  com se segue

$$p(x) = \sum_{j=1}^n c_j \phi_j(x).$$

Uma escolha da base para a base deste espaço é

$$\phi_1 = 1, \phi_2 = x, \phi_3 = x^2 \dots, \phi_n = x^{n-1}.$$

Neste trabalho, limita-se somente a esta escolha.

#### 4.0.1 Uma pequena observação:

O sistema normal  $A^T A c = A^T b$  pode ser escrito como

$$\begin{bmatrix} \sum_{i=1}^m \phi_1(x_i)\phi_1(x_i) & \sum_{i=1}^m \phi_1(x_i)\phi_2(x_i) & \dots & \sum_{i=1}^m \phi_1(x_i)\phi_n(x_i) \\ \sum_{i=1}^m \phi_2(x_i)\phi_1(x_i) & \sum_{i=1}^m \phi_2(x_i)\phi_2(x_i) & \dots & \sum_{i=1}^m \phi_2(x_i)\phi_n(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m \phi_n(x_i)\phi_1(x_i) & \sum_{i=1}^m \phi_n(x_i)\phi_2(x_i) & \dots & \sum_{i=1}^m \phi_n(x_i)\phi_n(x_i) \end{bmatrix}_{n \times n} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} \sum_{i=1}^m \phi_1(x_i)f(x_i) \\ \sum_{i=1}^m \phi_2(x_i)f(x_i) \\ \vdots \\ \sum_{i=1}^m \phi_n(x_i)f(x_i) \end{bmatrix}_{n \times 1}.$$

Ou seja,

$$\begin{bmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \dots & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \dots & \langle \phi_2, \phi_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \dots & \langle \phi_n, \phi_n \rangle \end{bmatrix}_{n \times n} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} \langle f, \phi_1 \rangle \\ \langle f, \phi_2 \rangle \\ \vdots \\ \langle f, \phi_n \rangle \end{bmatrix}_{n \times 1}.$$

Observe que este último é exatamente o sistema obtido na teoria da projeção.

Por exemplo, no espaço dos polinômio de grau 1, nota-se que o sistema normal é

$$\begin{bmatrix} \sum_{i=1}^m \phi_1(x_i)\phi_1(x_i) & \sum_{i=1}^m \phi_1(x_i)\phi_2(x_i) \\ \sum_{i=1}^m \phi_2(x_i)\phi_1(x_i) & \sum_{i=1}^m \phi_2(x_i)\phi_2(x_i) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m \phi_1(x_i)f(x_i) \\ \sum_{i=1}^m \phi_2(x_i)f(x_i) \end{bmatrix}.$$

Com a escolha da base  $\phi_1 = 1$  e  $\phi_2 = x$ , obtem-se

$$\begin{cases} mc_1 + c_2 \sum_{i=1}^m x_i & = \sum_{i=1}^m f(x_i) \\ c_1 \sum_{i=1}^m x_i + c_2 \sum_{i=1}^m x_i^2 & = \sum_{i=1}^m x_i f(x_i) \end{cases} .$$

E deve-se encontrar os coeficientes  $c_1$  e  $c_2$ .

**Ilustração:** Suponha que deseja-se obter uma função que melhor aproxime os dados tabelados, veja a Tabela 1: A Figura 2 é obtida através dos dados da Tabela 1.

Tabela 1: Valores de uma função.

$x_i$	$y_i$	$x_i$	$y_i$
1	0,3	6	6,7
2	1,1	7	8,2
3	2,1	8	10,5
4	3,0	9	12,1
5	4,3	10	15,6

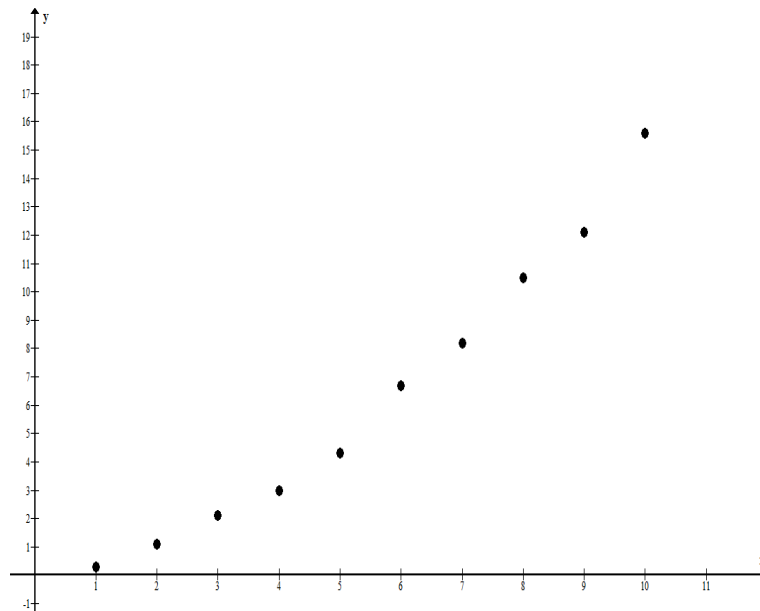


Figura 2: Gráfico com valores da Tabela 1

Observa-se uma relação entre  $x$  e  $y$ , isto leva a seguinte conclusão, o melhor ajuste deve ser linear. Lembre-se, não se deseja uma função de aproximação que se ajuste perfeitamente aos dados, mas sim a que melhor descreve simultaneamente os mesmos. Assim, da observação acima, considera-se o polinômio de grau 1 e deve-se encontrar as constantes  $c_1$  e  $c_2$ .

**Exemplo 4.2** Considerando os dados apresentados na Tabela 1, pode-se aumentar a tabela e somar as colunas, conforme mostrado abaixo, na terceira e na quarta coluna da Tabela 2.

Na quinta coluna da Tabela 2 temos a reta que aproxima esses dados e implicam em

Tabela 2: Valores de uma função.

$x_i$	$y_i$	$x_i^2$	$x_i y_i$	$P(x_i) = 1,665x_i - 2,767$
1	0,3	1	0,3	-1,102
2	1,1	4	2,2	0,563
3	2,1	9	6,3	2,228
4	3,0	16	12,0	3,893
5	4,3	25	21,5	5,558
6	6,7	36	40,2	7,223
7	8,2	49	57,4	8,888
8	10,5	64	84,0	10,553
9	12,1	81	108,9	12,218
10	15,6	100	156,0	13,883
55	63,9	385	488,8	$E_2 = \sum_{i=1}^{10} (y_i - P(x_i))^2 \approx 8,35$

$$c_1 = \frac{385(63,9) - 55(488,8)}{10(385) - (55)^2} = -2,767 \quad e \quad c_2 = \frac{10(488,8) - 55(63,9)}{10(385) - (55)^2} = 1,665$$

de modo que  $P(x) = 1,665x - 2,767$ .

O gráfico dessa reta e os pontos dados são mostrados na Figura 3, abaixo.

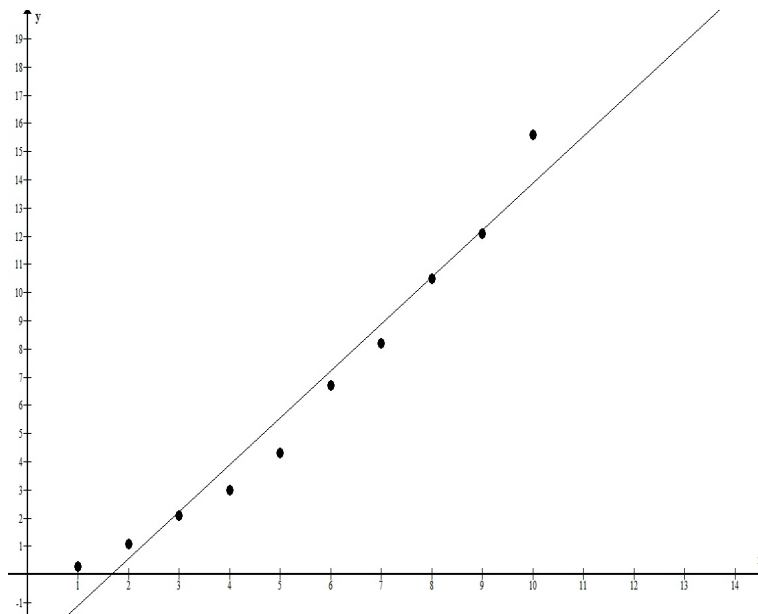


Figura 3: Gráfico com a reta que aproxima os dados

**Exemplo 4.3** Veja os valores fictícios da Tabela 3. Ajuste estes dados com o polinômio de grau dois dos mínimos quadrados discreto.

Tabela 3: Valores pontuais de um Polinômio

$i$	$x_i$	$y_i$
1	-1,00	1,0000
2	-0,50	0,5100
3	-0,20	0,1000
4	0,50	0,5400
5	1,00	2,1000

Neste caso,  $P(x) = c_1 + c_2x + c_3x^2$ . Fazendo

$$y = \begin{bmatrix} 1,00 \\ 0,51 \\ 0,10 \\ 0,54 \\ 2,10 \end{bmatrix}, \quad u_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad u_1 = \begin{bmatrix} -1,0 \\ -0,5 \\ -0,2 \\ 0,5 \\ 1,0 \end{bmatrix} \quad e \quad u_2 = \begin{bmatrix} 1,00 \\ 0,25 \\ 0,04 \\ 0,25 \\ 1,00 \end{bmatrix}$$

Deve-se, então, resolver o sistema linear:

$$\begin{pmatrix} \langle u_0, u_0 \rangle & \langle u_1, u_0 \rangle & \langle u_2, u_0 \rangle \\ \langle u_0, u_1 \rangle & \langle u_1, u_1 \rangle & \langle u_2, u_1 \rangle \\ \langle u_0, u_2 \rangle & \langle u_1, u_2 \rangle & \langle u_2, u_2 \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \langle y, u_0 \rangle \\ \langle y, u_1 \rangle \\ \langle y, u_2 \rangle \end{pmatrix}$$

Assim obtem-se:

$$\begin{pmatrix} 5 & -0,2 & 2,54 \\ -0,2 & 2,54 & -0,008 \\ 2,54 & -0,008 & 2,1266 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 4,25 \\ 1,095 \\ 3,3665 \end{pmatrix}$$

cuja solução é:  $c_1 \approx 0,1599$ ,  $c_2 \approx 0,4481$  e  $c_3 \approx 1,3937$ .

Desse modo, o polinômio de mínimos quadrados de grau dois que ajusta os dados precedentes é  $P(x) = 0,1599 + 0,4480x + 1,3937x^2$ , cujo gráfico está na Figura 4.

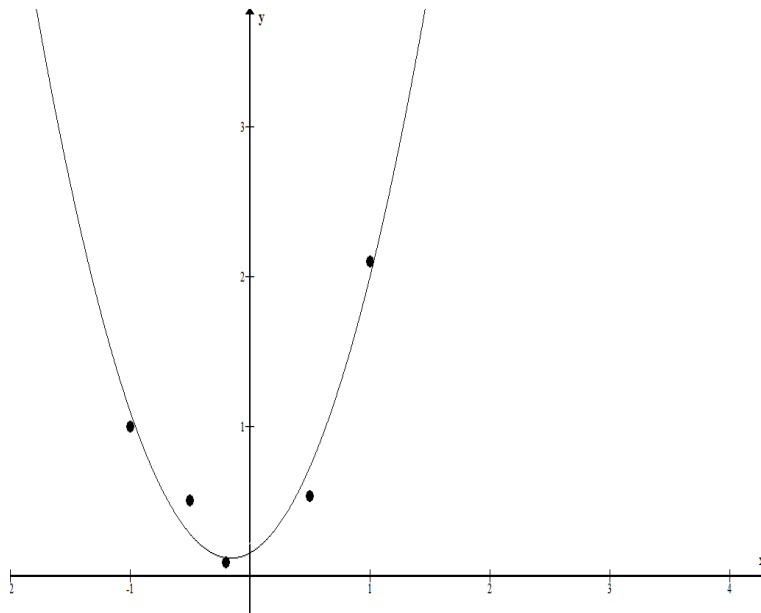


Figura 4: Gráfico com valores da Tabela 3

### 3.2 MÍNIMOS QUADRADOS no Ensino Médio – aplicações em sala de aula

Nesta seção propomos uma aplicação didática no Ensino Médio do método dos Mínimos Quadrados. Observamos que a matemática, do ensino médio, em vários de seus conteúdos, é apresentada sem o rigor das demonstrações. Muitas dessas noções são desenvolvidas de forma intuitiva pelo aluno e o processo de compreensão e aplicação da ideia matemática envolvida não fica prejudicado. Não queremos apresentar ao aluno do ensino médio o rigor matemático do método dos Mínimos Quadrados, portanto apresentamos de uma maneira intuitiva, para que o mesmo possa testar e aplicar.

Sob essa percepção, a apresentação, aos alunos do ensino médio, do método para aproximação de funções, pode oferecer um ganho no desenvolvimento matemático. Podemos apresentar aos mesmos algumas aproximações, utilizando o caso discreto, como: ajuste de pontos por retas; aproximação polinomial, até no máximo do 2º grau.

Entre as várias razões nas quais podemos basear tal afirmação, destacamos:

- O estudo dos Mínimos Quadrados apresenta uma oportunidade para o aluno revisar e reforçar a sua compreensão de vários conceitos aprendidos durante o Ensino Médio, entre os quais citamos o estudo das funções do 1º grau e 2º grau, representação gráfica de funções, propriedades gráficas das funções elementares, domínio e imagem, concavidade, entre outros.
- A aproximação de funções pelo Método dos Mínimos Quadrados pode ser visualizada geometricamente, sem recorrer a recursos inadequados a essa faixa etária. Para isso, pode-se utilizar softwares, como o GeoGebra ou o Graph. Estes são aplicativos livres usados para desenhar gráficos matemáticos em um sistema de coordenadas. No Graph, especificamente, podemos visualizar uma função facilmente. Também, com ele, é possível fazer alguns cálculos matemáticos sobre as funções.

- Enfim, a apresentação de conceitos mais avançados a alunos do ensino médio pode estimular a curiosidade daqueles que têm maior afinidade com a Matemática e estão inclinados a seguir seus estudos na área de exatas.

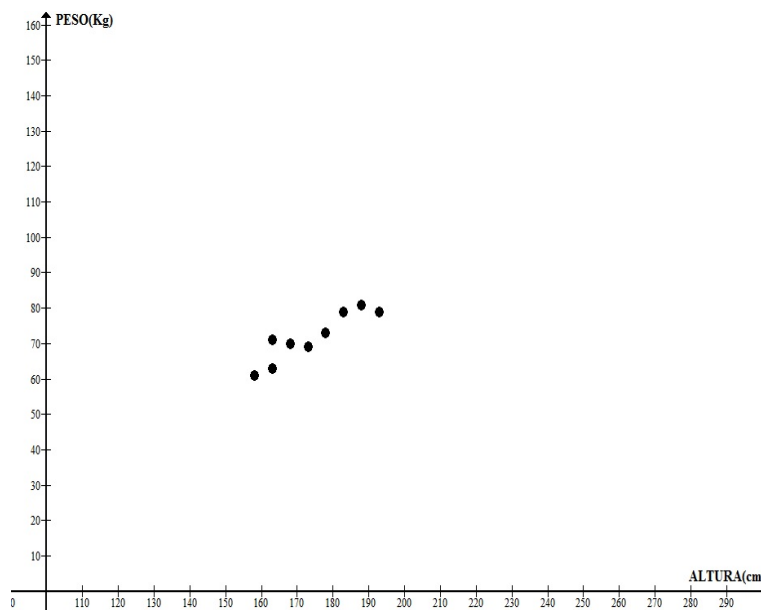
Analisando o primeiro objetivo entre os destacados acima, recomenda-se, antes do início desse estudo, dedicar um momento para uma efetiva retomada de todos os conceitos citados, que foram ensinados ao longo das três séries do ensino Médio e que tenham relação com os exemplos que serão apresentados.

A partir deste ponto, e com foco nos objetivos didáticos listados acima, apresentaremos a seguir alguns exemplos que podem ser trabalhados em sala de aula, com a exploração geométrica no Graph da função e seus ajustes.

**Exemplo 4.4** *A tabela abaixo mostra as alturas(cm) e pesos(kg) de uma amostra de nove jovens entre as idades de 17 e 20 anos, extraída ao acaso entre alunos de uma escola:*

Altura	183	173	168	188	158	163	193	163	178
Peso	79	69	70	81	61	63	79	71	73

a) *Faça o diagrama de dispersão dos dados no “Graph” e observe que relação existe entre altura e o peso. Vemos abaixo a dispersão feita no “Graph”.*



b) *Ajuste uma reta que descreve o comportamento do peso em função da altura, isto é, peso =  $f$ (altura). Para isto vamos expandir a tabela de dados, a última linha é formada pela soma de cada coluna.*

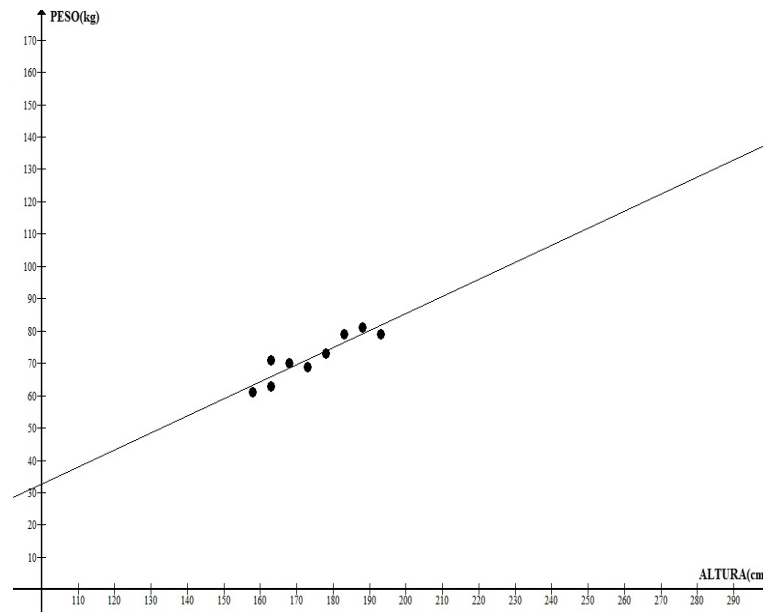
$x$	$y$	$xy$	$x^2$
183	79	14457	33489
173	69	11937	29929
168	70	11760	28224
188	81	15228	35344
158	61	9638	24964
163	63	10269	26569
193	79	15247	37249
163	71	11573	26569
178	73	12994	31684
1567	646	113103	274021

Sabemos que o polinômio de grau 1 tem a forma  $P(x) = c_2(x) + c_1$ , onde  $P$  é o peso em kg e  $x$  é a altura em cm.

Encontrando os coeficientes  $c_2$  e  $c_1$ , temos;

$$c_2 = \frac{1.567(646) - 9(113.103)}{(1.567)^2 - 9(274.021)} \approx 0,5276 \quad e \quad c_1 = \frac{113.103(1.567) - 274.021(646)}{(1.567)^2 - 9(274.021)} \approx -20,078$$

de modo que  $P(x) = 0,5276x - 20,078$ . A reta apresentada é feita no "Graph".



c) Estime o peso de um aluno dessa escola com 175 cm de altura; estime a altura de um aluno com 68 kg.

Para resolver este item basta substituir as variáveis corretas:

$$\text{Altura}(x) = 175 \Rightarrow P(175) = 0,5276(175) - 20,078 \Rightarrow P(x) \approx 72,25 \text{ kg.}$$

$$\text{Peso}(y) = 80 \Rightarrow 80 = 0,5276x - 20,078 \Rightarrow x \approx 189,69 \text{ cm.}$$



**Exemplo 4.5** A queda na taxa de fecundidade e o envelhecimento populacional são comuns em países que atingiram um patamar mediano de desenvolvimento, e isso vem acontecendo com o Brasil. Esses avanços sociais trazem consigo desafios no que diz respeito aos gastos com o sistema previdenciário. “É fácil entender por quê. Se as pessoas vivem mais, elas receberão aposentadoria durante um período de tempo maior. Por mais nobre que seja uma despesa destinada a assegurar a velhice digna, a questão é: como financiá-la?” [3].

Cresce a cada ano o número de aposentados e o número de pessoas na ativa, que contribuem com o INSS não avança na mesma proporção. O economista Marcelo Caetano, do Instituto de Pesquisa Econômica Aplicada (Ipea), estimou, com base na atual taxa de fecundidade das brasileiras, de 1,9 filho por mulher, que, se o ritmo se mantiver estável nos próximos anos, já em 2032 haverá mais gente recebendo aposentadoria do que contribuintes sustentando o INSS. Segundo ele, se não houver ajuste no sistema, o rombo nas contas da Previdência assumirá proporções explosivas. Veja o gráfico abaixo, retirado da revista Veja, edição 2011 de 2008:

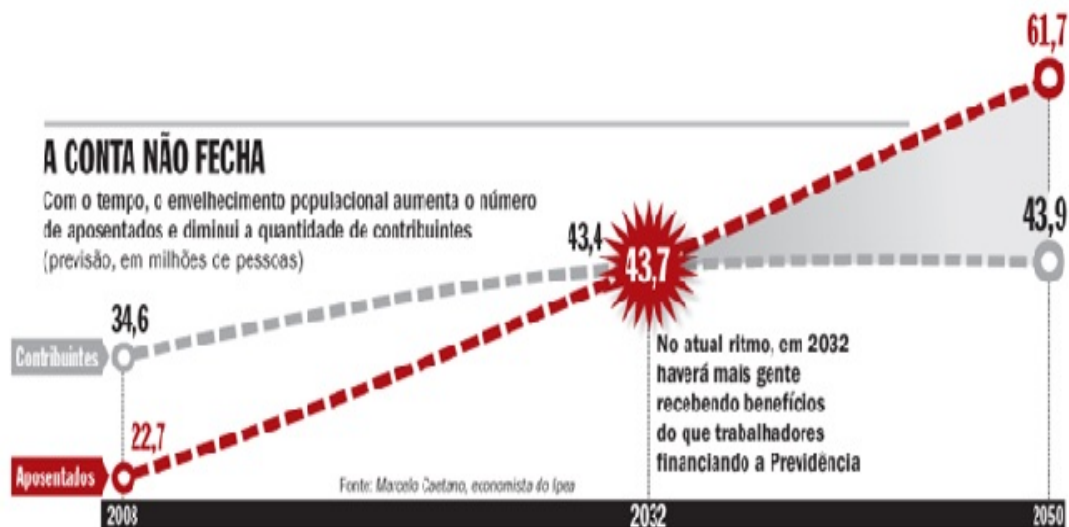


Figura 5: Número de contribuintes e aposentados brasileiros

Diante de tal situação, podemos investigar o seguinte problema: Considerando as previsões, no ano de 2040 qual será o número de aposentados?

Podemos considerar como hipótese, a partir da análise do gráfico acima, que o crescimento do número de aposentados no decorrer do tempo é linear. Consideremos, então, o polinômio  $P(x) = c_2x + c_1$ .

Vamos colocar os dados numa tabela e somar as colunas, conforme mostrado, na segunda, terceira, quarta e quinta colunas da Tabela 4. Apresentamos os cálculos na tabela seguinte:

Tabela 4: Dados do problema “A Conta não fecha”

ANO	TEMPO( $x_i$ )	APOSENTADOS x ( $10^6$ )( $y_i$ )	$x_i^2$	$x_i y_i$	$P(x_i)$
2008	8	22,7	64	181,60	22,37
2032	32	43,7	1.024	1.398,40	44,69
2050	50	61,7	2.500	3.085	61,43
$\Sigma$	90	128,10	3.588,00	4.665,00	

As equações abaixo nos dão  $P(x)$  que implica em:

$$c_1 = \frac{3.588(128,10) - 90(4.665)}{3(3.588) - (90)^2} \approx 14,93 \quad e \quad c_2 = \frac{3(4.665) - 90(128,1)}{3(3.588) - (90)^2} \approx 0,93$$

de modo que  $P(x) = 0,93x + 14,93$ .

Respondendo a pergunta, em 2040 teremos ( $x$  é dado pela seguinte convenção:  $x = 0$ , corresponde ao ano de 2000 e assim por diante).

$$P(x_i) = 0,93 \cdot (40) + 14,93 = 52,13 \text{ milhões de aposentados.}$$

Abaixo o gráfico do número de aposentados em função do tempo, feito no software “Graph”:

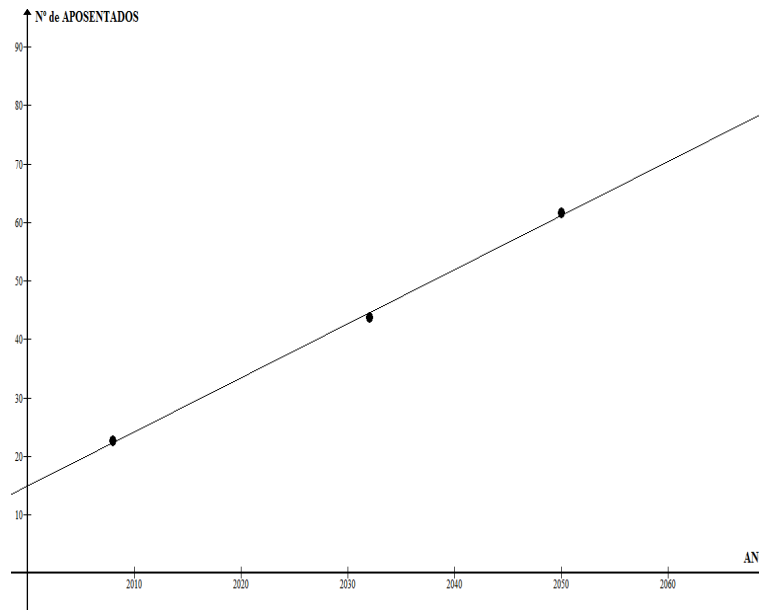
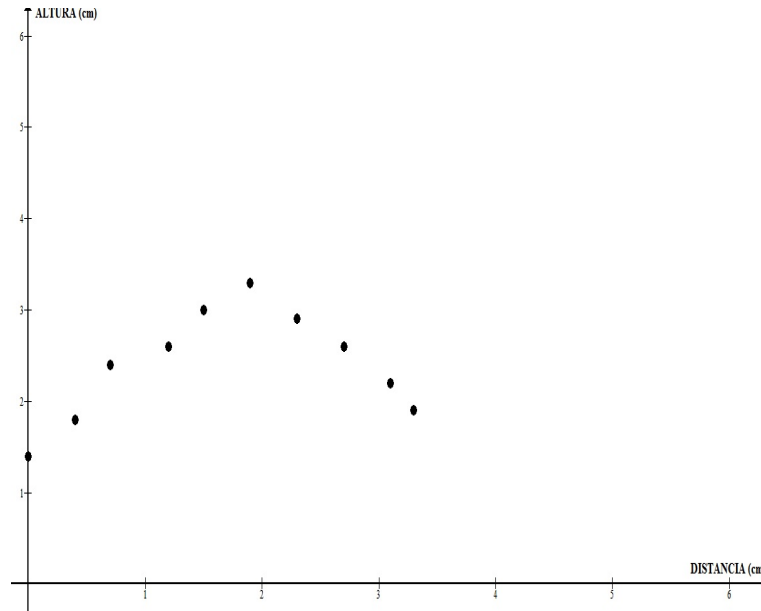


Figura 6: Número de aposentados brasileiros-Reta Ajustada

**Exemplo 4.6** Uma bola de ping pong é atirada para cima por uma criança. Sua trajetória é fotografada por uma máquina que registra a sua altura em relação a distância percorrida na horizontal. Os dados foram registrados na tabela abaixo:

<i>Altura(y)</i>	1,4	1,8	2,4	2,6	3	3,3	2,9	2,6	2,2	1,9	<i>metros</i>
<i>Distância(x)</i>	0	0,4	0,7	1,2	1,5	1,9	2,3	2,7	3,1	3,3	<i>metros</i>

a) Faça o diagrama de dispersão dos dados:  
Abaixo temos o diagrama feito no “Graph”.



b) Estime a que distância da criança a bolinha de ping pong vai atingir o solo.

A dispersão dos dados sugere um polinômio do 2º grau, neste caso queremos

$$f(x) \approx P(x) = c_1 + c_2x + c_3x^2.$$

Faremos a aproximação e determinaremos onde esta parábola cruza o eixo horizontal. Propomos que os alunos do ensino médio resolvam esta atividade usando a representação do sistema com somatório, pois os mesmo não vêem produto interno nesse segmento. Escrevendo as equações:

$$\begin{cases} c_1 \sum_{i=0}^m x_i^0 + c_2 \sum_{i=1}^m x_i^1 + c_3 \sum_{i=1}^m x_i^2 = \sum_{i=0}^m y_i x_i^0 \\ c_1 \sum_{i=0}^m x_i^1 + c_2 \sum_{i=1}^m x_i^2 + c_3 \sum_{i=1}^m x_i^3 = \sum_{i=0}^m y_i x_i^1 \\ c_1 \sum_{i=0}^m x_i^2 + c_2 \sum_{i=1}^m x_i^3 + c_3 \sum_{i=1}^m x_i^4 = \sum_{i=0}^m y_i x_i^2 \end{cases}$$

Devemos, então, resolver o sistema linear:

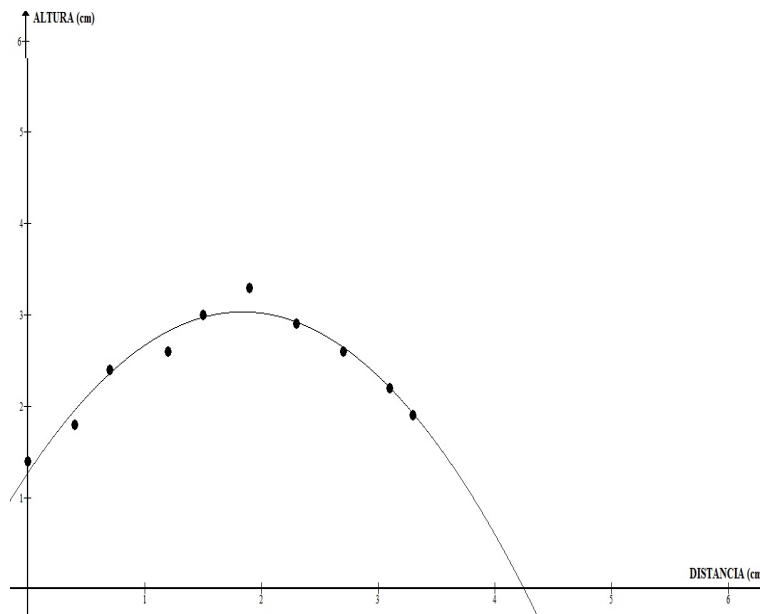
$$\begin{pmatrix} 10 & 17,10 & 41,03 \\ 17,10 & 41,03 & 109,947 \\ 41,03 & 109,947 & 312,5063 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 24,10 \\ 43,07 \\ 99,999 \end{pmatrix}$$

cuja solução é:  $c_1 \approx 1,2707$ ,  $c_2 \approx 1,9173$  e  $c_3 \approx -0,5214$ .

Portanto, a parábola que melhor aproxima a função tabelada é

$$f(x) \approx P(x) = 1,2707 + 1,9173x - 0,5214x^2$$

Abaixo temos o parábola feita no “Graph”



Para atingir o solo, a altura da bolinha tem que ser zero, logo faremos  $P(x) = 0$ . Sugerimos aos alunos que resolvam a equação do 2º grau acima com o uso da calculadora.

$$x' \approx -0,5734 \text{ e } x'' \approx 4,2500.$$

Logo a bolinha vai atingir o solo a uma distância de 4,25 metros de onde partiu, aproximadamente.

## Comentário e agradecimento final

Esperamos que os alunos tenham muito entusiasmo, curiosidade e empenho. Claro que certa dificuldade pode ocorrer, já que o assunto é novo para eles. Propomos que a atividade seja feita com calculadora e/ou softwares matemáticos. Desejamos que os alunos gostem da atividade e apliquem em outras aproximações.

Há tantos a agradecer, por tanto que dedicaram a mim, não somente por terem ensinado, mas por terem me feito aprender.

- À CAPES pela oportunidade de fazer este curso, aos professores da UFJS-Campus Alto Paraopeba pelo apoio e dedicação ao ensino, à orientadora, pela paciência e apoio permanente.

- A Minha família, Adriana, Ana Clara, Laura e Davi que nos momentos de minha ausência dedicados ao estudo, sempre fizeram entender que o futuro, é feito a partir da constante dedicação no presente.

- Aos meus colegas de turma, que fortaleceram os laços da igualdade, num ambiente fraterno e respeitoso. Jamais lhes esquecerei.

- Por fim, à aquele, que me permitiu tudo isso, ao longo de toda a minha vida, e, não somente nestes anos como mestrando, à você meu DEUS, obrigado, reconheço cada vez mais

em todos os meus momentos, que você é o maior mestre, que uma pessoa pode conhecer e reconhecer.

## Referências

- [1] AGUIAR, R., *Apostila: Álgebra Linear*. Joinville CCT-UDESC.
- [2] BARROSO, L. C., BARROSO, M., FILHO, F., CARVALHO, M. E MAIA, M. , *Cálculo Numérico: com aplicações* São Paulo: Harbra,1987.
- [3] BORSATO CINTIA, *Desse jeito, quebra*. Revista Veja, São Paulo, Edição 2071, 31 de Julho de 2008. p.94-103.
- [4] BURDEN, R. E FAIRES, J. , *Análise Numérica* São Paulo: Cengage Learning, 2008.
- [5] FRANCO, N. , *Cálculo Numérico* São Paulo: Pearson Prentice Hall, 2006.
- [6] GALVÃO, L. C., NUNES, L. F., *Apostila: Notas de aula: Cálculo numérico*. Paraná Universidade Tecnológica Federal do Paraná.
- [7] GRAPH, versão 4.3, Disponível em < [graph.softonic.com.br](http://graph.softonic.com.br) >. Acesso em 24/01/2014.
- [8] HAMMERLIN, G., HOFFMANN, K., SCHUMAKER L., *Numerical Mathematics*,New York: Springer-Verlag, 1991.
- [9] MEYER, C. , *Matrix Analysis and Applied Linear Algebra* Sian.
- [10] RUGGIERO, M. E LOPES, V. , *Cálculo Numérico: Aspectos teóricos e Computacionais* São Paulo: Makron Books,1996.
- [11] SPERANDIO, D., MENDES, J., MONKEN, L. , *Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos* São Paulo: Pearson Prentice Hall, 2003.