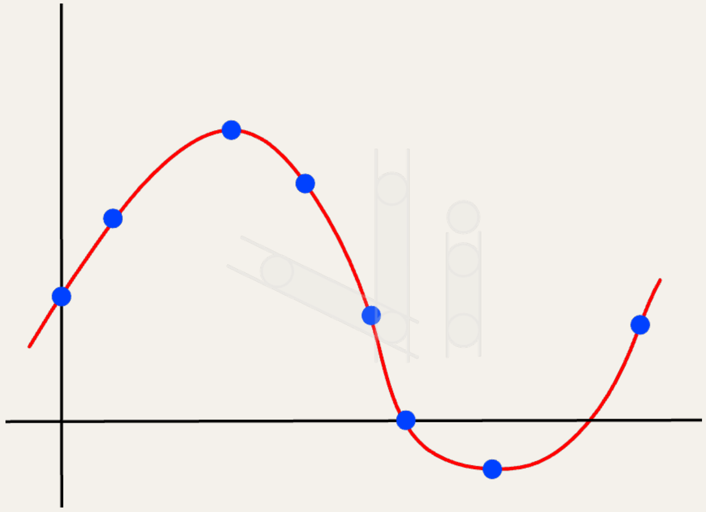
Ajustes de Curvas pelo Método dos Mínimos Quadrados

Prof. Neemias Martins

Centro de Ciências da Natureza

Universidade Federal de São Carlos - Campus Lagoa do Sino

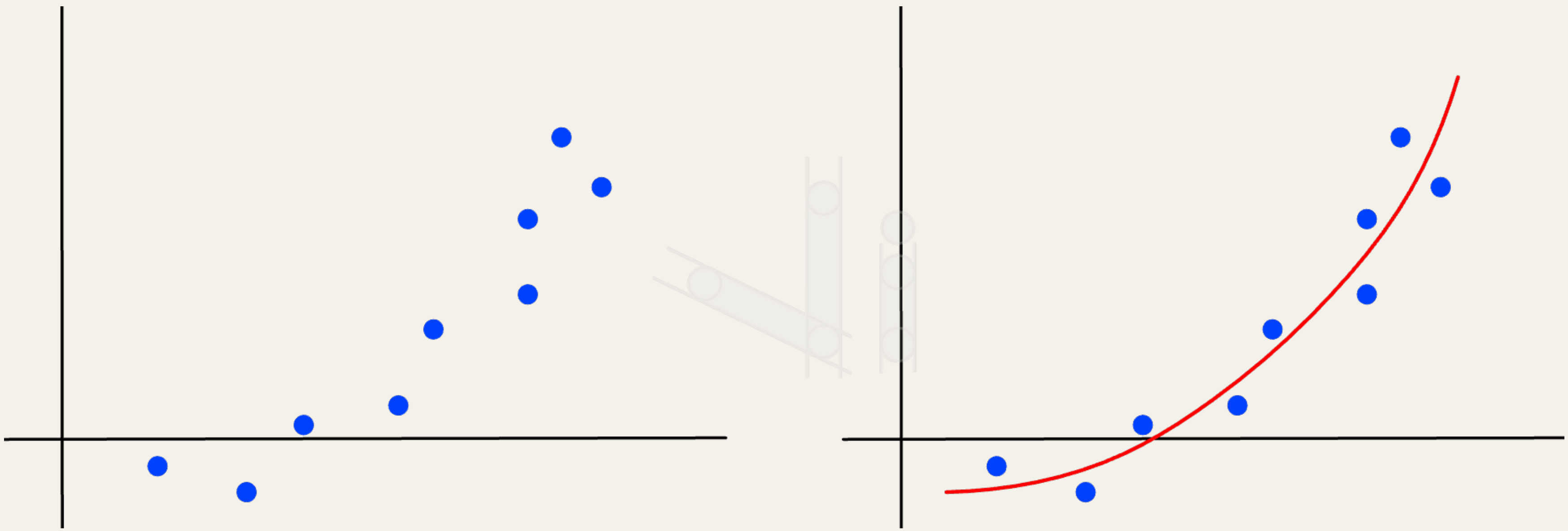
Aula Passada: Interpolação



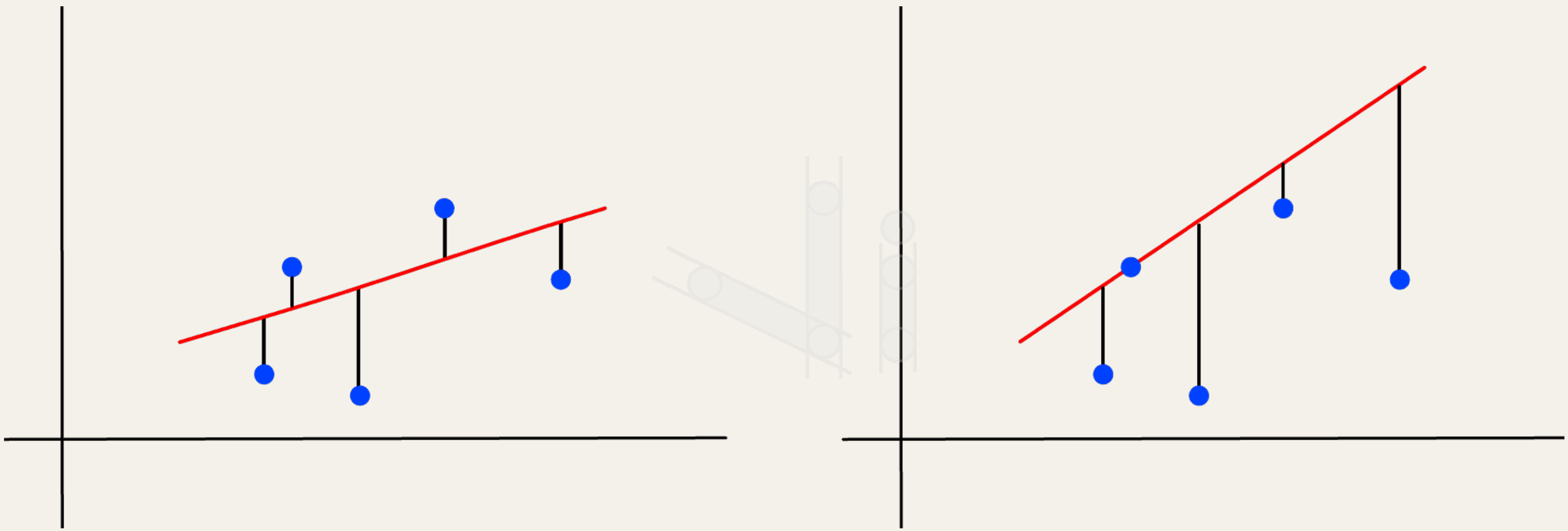
Quando a interpolação **não é recomendada**:

1. Queremos obter um valor aproximado da função em algum ponto fora do intervalo de tabelamento (extrapolação).
2. Os dados são provenientes de experimentos físicos. Neste caso os valores poderão conter erros.

Ajuste de curvas



Ajuste de curvas



Em qual das retas o valor $\sum_{k=1}^{5}((f(x_k)) - \varphi(x_k))^2$ é menor?

O método dos Mínimos Quadrados

Suponha que temos uma tabela de dados, com $x_1, ...x_m \in [a, b]$:

x	x_1	x_2	...	x_m
$f(x)$	$f(x_1)$	$f(x_2)$	...	$f(x_m)$

Escolhidas funções $g_1(x), ..., g_n(x)$ contínuas em $[a, b]$, queremos obter coeficientes $a_1, ..., a_n$ de modo que a função

$$\varphi(x) = a_1 g_1(x) + ... + a_n g_n(x)$$

se aproxime ao máximo de f nos pontos tabelados.

Ou seja, queremos que

$$J(a_1, \ldots, a_n) = \sum_{k=1}^{m} (f(x_k) - \varphi(x_k))^2$$

seja mínimo.

- Note que $J(a_1, \ldots, a_n) = 0$ se e somente se, $f(x_k) = \varphi(x_k)$ em cada x_k.
- Sabemos que o mínimo de $J(a_1, \ldots, a_n)$ deve satisfazer

$$\frac{\partial J}{\partial a_j} = 0, \ \forall j = 1, \ldots, n.$$

$$J(a_1, \ldots, a_n) = \sum_{k=1}^{m} \left(f(x_k) - a_1 g_1(x_k) - a_2 g_2(x_k) \ldots - a_n g_n(x_k)\right)^2,$$

pela regra da cadeia, temos

$$\begin{aligned}
\frac{\partial J}{\partial a_j} &= \sum_{k=1}^{m} 2 \cdot \left(f(x_k) - a_1 g_1(x_k) - a_2 g_2(x_k) - \ldots - a_n g_n(x_k)\right) \cdot \left(-g_j(x_k)\right) \\
&= \sum_{k=1}^{m} 2 \cdot \left(a_1 g_1(x_k) + a_2 g_2(x_k) + \ldots + a_n g_n(x_k) - f(x_k)\right) \cdot g_j(x_k).
\end{aligned}$$

Então,

$$\sum_{k=1}^{m} \left(a_1 g_1(x_k) + a_2 g_2(x_k) + \ldots + a_n g_n(x_k) - f(x_k)\right) \cdot g_j(x_k) = 0, \ \forall j = 1, \ldots, n.$$

Equações normais

Aternativamente, podemos reescrever a expressão no seguinte sistema linear com n equações e incógnitas $a_1, \ldots, a_n$:

$$\begin{cases} \left(\sum_{k=1}^{m} g_1(x_k)g_1(x_k)\right)a_1 + \ldots + \left(\sum_{k=1}^{m} g_n(x_k)g_1(x_k)\right)a_n = \sum_{k=1}^{m} f(x_k)g_1(x_k) \\ \left(\sum_{k=1}^{m} g_1(x_k)g_2(x_k)\right)a_1 + \ldots + \left(\sum_{k=1}^{m} g_n(x_k)g_2(x_k)\right)a_n = \sum_{k=1}^{m} f(x_k)g_2(x_k) \\ \vdots \\ \left(\sum_{k=1}^{m} g_1(x_k)g_n(x_k)\right)a_1 + \ldots + \left(\sum_{k=1}^{m} g_n(x_k)g_n(x_k)\right)a_n = \sum_{k=1}^{m} f(x_k)g_n(x_k) \end{cases}$$

Equações normais

Em termos matriciais, temos $A \cdot a = b$ em que

$$A = \left(a_{ij}\right) \in \mathbb{R}^n \times \mathbb{R}^n$$

$$a = \left(a_j\right) \in \mathbb{R}^n$$

$$b = (b_i) \in \mathbb{R}^n,$$

sendo

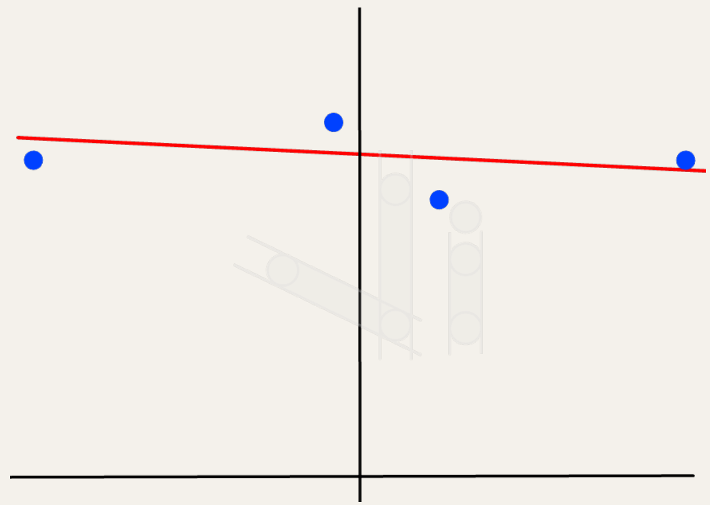
$$a_{ij} = \sum_{k=1}^{m} g_i(x_k) g_j(x_k)$$

$$b_i = \sum_{k=1}^{m} f(x_k) g_i(x_k).$$

Exemplo 1

Usando o Método dos Mínimos Quadrados, queremos encontrar a melhor reta que se ajusta aos seguintes dados tabelados:

x	-1	-0.1	0.2	1
y	1.000	1.099	0.808	1.000



Temos $\varphi(x) = a_1 x + a_2$. Ou seja, $g_1(x) = x$ e $g_2(x) = 1$.

Exemplo 1

Para resolvermos o sistema $A \cdot a = b$,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

Devemos calcular cada um dos termos a seguir

$$a_{ij} = \sum_{k=1}^{4} g_i(x_k) g_j(x_k) \;\text{ e }\; b_i = \sum_{k=1}^{4} y_k g_i(x_k).$$

Exemplo 1

Temos

$$a_{11} = \sum_{k=1}^{4} (x_k)^2 = 2.05$$

$$a_{12} = \sum_{k=1}^{4} x_k = 0.1 = a_{21}$$

$$a_{22} = \sum_{k=1}^{4} 1 = 4$$

$$b_1 = \sum_{k=1}^{4} x_k y_k = 0.0517 \quad b_2 = \sum_{k=1}^{4} y_k = 3.907$$

Exemplo 1

Resolvendo o sistema

$$\begin{pmatrix} 2.05 & 0.1 \\ 0.1 & 4 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.0517 \\ 3.907 \end{pmatrix},$$

i.e

$$\begin{cases} 2.05a_1 + 0.1a_2 = 0.0517 \\ 0.1a_1 + 4a_2 = 3.907 \end{cases},$$

obtemos $a_1 = -0.0224$ e $a_2 = -0.9773$.

Portanto a melhor reta que representa o conjunto de dados apresentados é dada por

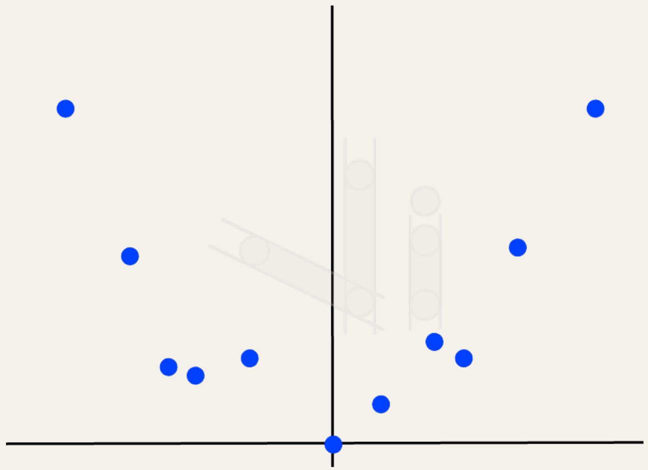
$$\varphi(x) = -0.0224x - 0.9773.$$

Exemplo 2

Um experimento resultou na seguinte tabela de dados:

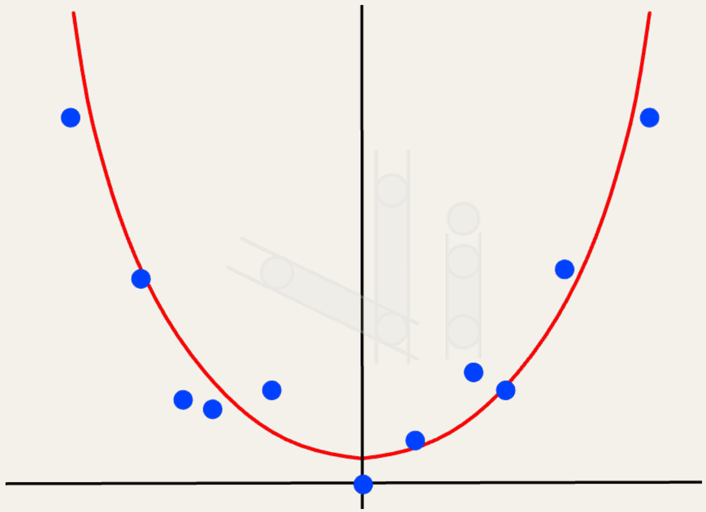
x	-1.00	-0.75	-0.60	-0.50	-0.30	0.00	0.20	0.40	0.50	0.70	1.00
$f(x)$	2.05	1.15	0.45	0.40	0.50	0.00	0.20	0.60	0.51	1.20	2.05

Dispondo todos os 11 pontos da forma $(x, f(x))$ no plano cartesiano, obtemos o seguinte **diagrama de dispersão**



Exemplo 2

O diagrama de dispersão sugere a escolha de uma parábola:



Fazemos $g_1(x) = x^2, g_2(x) = x$ e $g_3(x) = 1$ e assim

$$\begin{aligned}\varphi(x) &= a_1 g_1(x) + a_2 g_2(x) + a_3 g_3(x)\\ &= a_1 x^2 + a_2 x + a_3.\end{aligned}$$

Exemplo 2

Usando o Método dos Mínimos Quadrados, devemos resolver o sistema

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

sendo

$$a_{ij} = \sum_{k=1}^{m} g_i(x_k) g_j(x_k)$$

$$b_i = \sum_{k=1}^{m} f(x_k) g_i(x_k).$$

Exemplo 2

$$a_{11} = \sum_{k=1}^{11} (x_k)^4 = 2.85,\ a_{12} = \sum_{k=1}^{11} (x_k)^3 = -0.25 = a_{21}$$

$$a_{13} = \sum_{k=1}^{11} (x_k)^2 = 4.20 = a_{22} = a_{31}$$

$$a_{23} = \sum_{k=1}^{11} x_k = -0.35 = a_{32},\ a_{33} = \sum_{k=1}^{11} 1 = 11$$

$$b_1 = \sum_{k=1}^{11} (x_k)^2 y_k = 5.87$$

$$b_2 = \sum_{k=1}^{11} x_k y_k = -0.11,\ b_3 = \sum_{k=1}^{11} y_k = 9.11.$$

Exemplo 2

O sistema de equações normais é

$$\begin{pmatrix} 2.85 & -0.25 & 4.20 \\ -0.25 & 4.20 & -0.35 \\ 4.20 & -0.35 & 11 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 5.87 \\ 0.11 \\ 9.11 \end{pmatrix},$$

cuja solução é dada por

$$a_1 = 1.94, \; a_2 = 0.10 \; \text{ e } \; a_3 = 0.09.$$

Portanto, a curva que melhor se ajusta aos dados do experimento é a parábola

$$\varphi(x) = 1.94x^2 + 0.10x + 0.09.$$

Exemplo 2

O mínimo dos quadrados dos desvios é então dado por

$$J(1.94, 0.10, 0.09) = \sum_{k=1}^{11} \left(y_k - 1.94x^2 - 0.10x - 0.09\right)^2 = 0.24$$

Resumo

O método dos mínimos quadrados é usado para encontrar uma função

$$\varphi(x) = a_1 g_1(x) + \ldots + a_n g_{n(x)}$$

que melhor se ajusta a uma tabela de dados

x	x_1	x_2	$\ldots$	x_m
y	y_1	y_2	$\ldots$	y_m

Os coeficientes que minimizam a função

$$J(a_1, \ldots, a_m) = \sum_{k=1}^{m} (f(x_k) - \varphi(x_k))^2,$$

são obtidos resolvendo o sistema linear de equações normais $A \cdot a = b$

Exercícios

1. Usando o método dos mínimos quadrados, encontre a melhor **reta** e a melhor **parábola** que se ajustem aos dados experimentais abaixo:

x	1	2	3	4	5	6	7	8
y	0.5	0.6	0.9	0.8	1.2	1.5	1.7	2.0

2. Compare o mínimo dos quadrados dos desvios dos dois casos e decida qual é das duas curvas é o melhor modelo para o experimento.

Bons estudos!