Cálculo Numérico / Métodos Numéricos

Sistemas lineares

Minimos Quadrados - Caso discreto

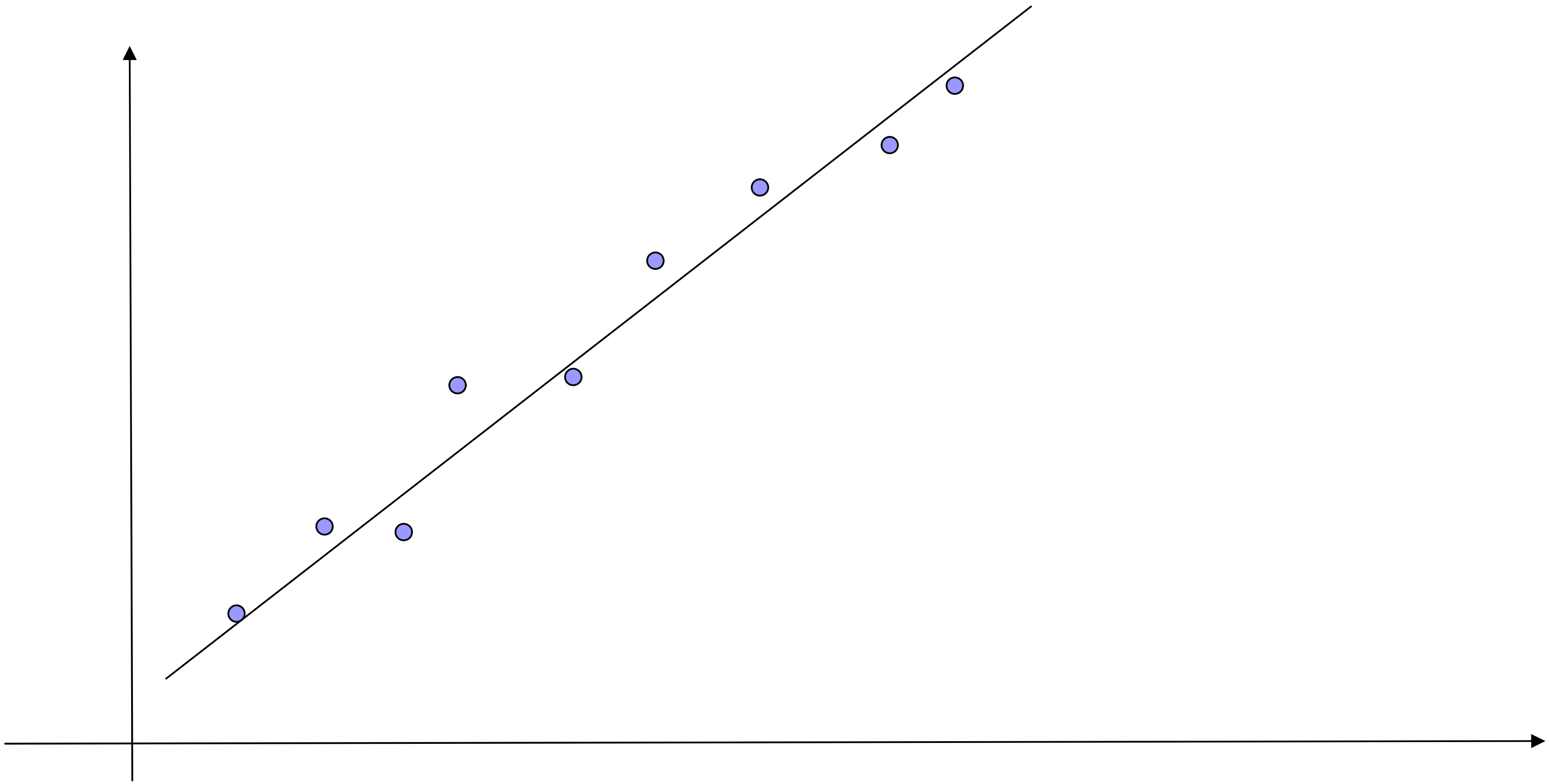
6 May 2008 . 17:21

Caso discreto

- Vamos inicialmente considerar o caso em que sabemos a função a aproximar em apenas alguns pontos:

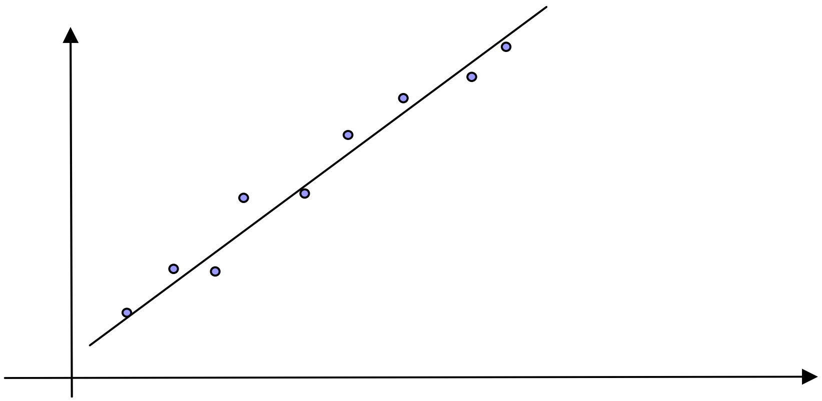
x	x_1	x_2	x_3	...	x_m
$f(x)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	...	$f(x_m)$

Exemplo gráfico



Vemos que os pontos parecem uma reta.
A pergunta é: qual a melhor reta que os aproximaria ?

Reta (regressão linear)



$$f(x) \approx g(x) = a_1 g_1(x) + a_2 g_2(x)$$

$$f(x) \approx g(x) = a_1 x + a_2$$

$g_1(x) = x$

$g_2(x) = 1$

tarefa: escolher a_1 e a_2 de modo que o erro seja mínimo!

Mínimos quadrados

- Como vimos, usamos os "mínimos quadrados".

$$Min \sum_{i=1}^{m} e(x_i)^2$$

$$=$$

$$Min \sum_{i=1}^{m} (f(x_i) - g(x_i))^2$$

E queremos o mínimo em referência aos parâmetros a_1 e a_2

Do cálculo diferencial, se a função $e(a_1,a_2) = \sum_{i=1}^{m} e(x_i)^2$ tem mínimo, então:

$$\frac{\partial E}{\partial a_1} = 0 \qquad \text{e} \qquad \frac{\partial E}{\partial a_2} = 0$$

Derivando em relação a a_1

$$\frac{\partial E}{\partial a_1} = \frac{\partial}{\partial a_1}\left(\sum_{i=1}^{m} e(x_i)^2\right)$$

$$= \frac{\partial}{\partial a_1}\left(\sum_{i=1}^{m}(a_1 x_i + a_2 - f(x_i))^2\right)$$

$$= \sum_{i=1}^{m} 2(a_1 x_i + a_2 - f(x_i))x_i$$

$$= 2\left[a_1 \sum_{i=1}^{m} x_i^2 + a_2 \sum_{i=1}^{m} x_i\right] - 2\left[\sum_{i=1}^{m} x_i f(x_i)\right] \quad = 0$$

Assim:

i) $$\left(\sum_{i=1}^{m} x_i^2\right) a_1 + \left(\sum_{i=1}^{m} x_i\right) a_2 = \sum_{i=1}^{m} x_i f(x_i)$$

Derivando em relação a a_2

$$\frac{\partial E}{\partial a_2} = \frac{\partial}{\partial a_2}\left(\sum_{i=1}^{m} e(x_i)^2\right)$$

$$= \frac{\partial}{\partial a_2}\left(\sum_{i=1}^{m}(a_1 x_i + a_2 - f(x_i))^2\right)$$

$$= \sum_{i=1}^{m} 2(a_1 x_i + a_2 - f(x_i))$$

$$= 2\left[a_1 \sum_{i=1}^{m} x_i + m \cdot a_2 - \sum_{i=1}^{m} f(x_i)\right] \quad = 0$$

Assim:

ii) $$\left(\sum_{i=1}^{m} x_i\right) a_1 + m \cdot a_2 = \sum_{i=1}^{m} f(x_i)$$

Sistema

- Portanto, os parâmetros que minimizam $E(a_1,a_2)$ obrigatoriamente respeitam o sistema abaixo:

$$\begin{cases} \left(\sum_{i=1}^{m} x_i^2\right) a_1 + \left(\sum_{i=1}^{m} x_i\right) a_2 = \sum_{i=1}^{m} x_i f(x_i) \\ \left(\sum_{i=1}^{m} x_i\right) a_1 + m \cdot a_2 = \sum_{i=1}^{m} f(x_i) \end{cases}$$

Sistema de equações normais.

Note que a matriz A é simétrica e definida positiva (podemos aplicar Cholesky)

pode-se provar que o ponto obtido realmente minimiza a função $E(a_1,a_2)$

Exemplo

- Obter a reta que melhor ajusta os dados:

x	0	1	2	3	4
f(x)	0.98	-3.01	-6.99	-11.01	-15

Solução:

Como vimos, a reta $g(x) = a_1 x + a_2$ que melhor se ajusta é aquela cujos parâmetros resolve o sistema:

$$\begin{cases} \left(\sum_{i=1}^{m} x_i^2\right) a_1 + \left(\sum_{i=1}^{m} x_i\right) a_2 = \sum_{i=1}^{m} x_i f(x_i) \\ \left(\sum_{i=1}^{m} x_i\right) a_1 + m \cdot a_2 = \sum_{i=1}^{m} f(x_i) \end{cases}$$

Exemplo (solução)

Sistema:

$$\begin{bmatrix} \sum_{k=1}^{5} x_i^2 & \sum_{k=1}^{5} x_i \\ \sum_{k=1}^{5} x_i & 5 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} \sum_{k=1}^{5} f(x_i)x_i \\ \sum_{k=1}^{5} f(x_i) \end{bmatrix}$$

	xi	f(xi)	xi^2	f(xi)xi	
	0.00	0.98	0.00	0.00	
	1.00	-3.01	1.00	-3.01	
	2.00	-6.99	4.00	-13.98	
	3.00	-11.01	9.00	-33.03	
	4.00	-15.00	16.00	-60.00	
soma:	**10.00**	**-35.03**	**30.00**	**-110.02**	

$$\begin{bmatrix} 30 & 10 \\ 10 & 5 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} -110.02 \\ -35.03 \end{bmatrix} \longrightarrow$$

$a_1 = -3.9960$

$a_2 = 0.9860$

Exemplo (solução)

Logo:

$f(x) \approx g(x) = -3.9960\,x + 0.9860$

Erro:

$e(x_1)^2 = (f(0)-g(0))^2 = 0.0000$

$e(x_2)^2 = (f(1)-g(1))^2 = 0.0000$

$e(x_3)^2 = (f(2)-g(2))^2 = 0.0003$

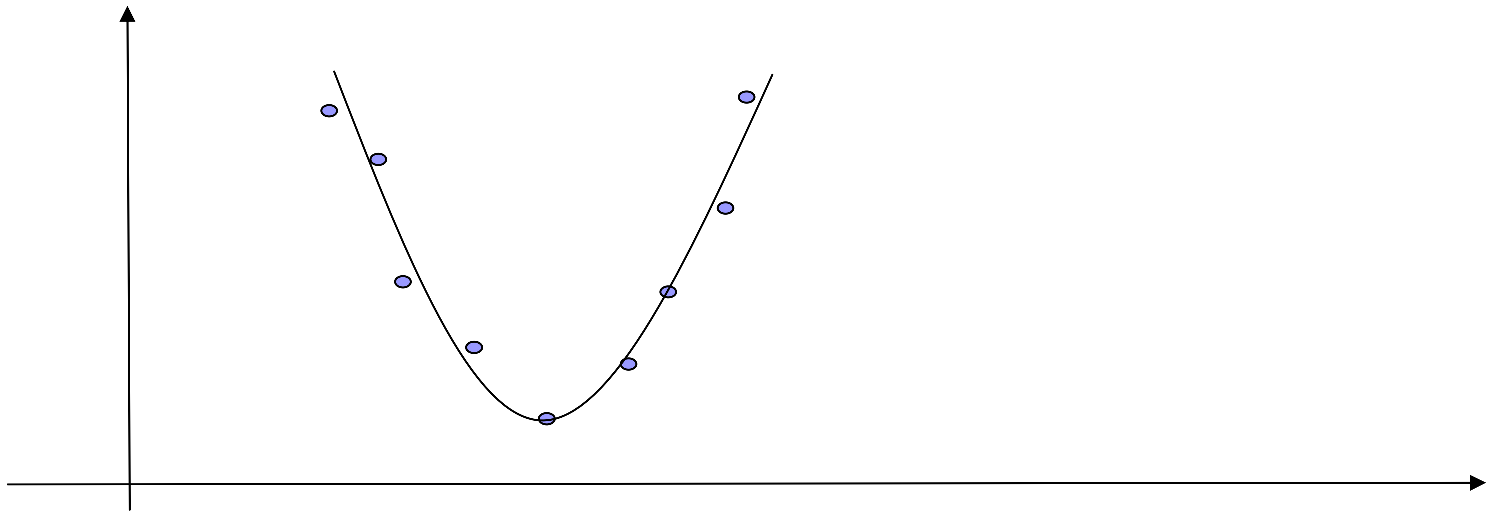
$e(x_4)^2 = (f(3)-g(3))^2 = 0.0001$

$e(x_5)^2 = (f(4)-g(4))^2 = 0.0000$

$\sum_{i=1}^{5} e(x_i)^2 = 0.0004$

Outras funções

- Obviamente, nem toda função que desejaremos aproximar será uma reta.
- Por exemplo:



Nesse caso: $g(x) = a_1 g_1(x) + a_2 g_2(x) + a_3 g_3(x)$

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

Caso geral:

$g(x) = a_1g_1(x) + a_2g_2(x) + \ldots + a_ng_n(x)$

Procedendo de maneira análoga, temos que derivar a função de erro parcialmente em relação a cada um dos n parâmetros e igualar a zero (condição necessária para que seja um ponto de mínimo):

$$\frac{\partial E}{\partial a_1} = 0 \iff (\sum_{i=1}^{m} g_1(x_i)g_1(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_1(x_i))a_2 \ldots + \\ + \ldots + (\sum_{i=1}^{m} g_n(x_i)g_1(x_i))a_n = \sum_{i=1}^{m} f(x_i)g_1(x_i)$$

$$\frac{\partial E}{\partial a_1} = 0 \iff (\sum_{i=1}^{m} g_1(x_i)g_2(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_2(x_i))a_2 \ldots + \\ + \ldots + (\sum_{i=1}^{m} g_n(x_i)g_2(x_i))a_n = \sum_{i=1}^{m} f(x_i)g_2(x_i)$$

...

$$\frac{\partial E}{\partial a_1} = 0 \iff (\sum_{i=1}^{m} g_1(x_i)g_n(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_n(x_i))a_2 \ldots + \\ + \ldots + (\sum_{i=1}^{m} g_n(x_i)g_n(x_i))a_n = (\sum_{i=1}^{m} f(x_i)g_n(x_i)$$

Sistema

- E obtemos o sistema:

$$
\begin{cases}
(\sum_{i=1}^{m} g_1(x_i)g_1(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_1(x_i))a_2 ... + (\sum_{i=1}^{m} g_n(x_i)g_1(x_i))a_n = \sum_{i=1}^{m} f(x_i)g_1(x_i) \\
(\sum_{i=1}^{m} g_1(x_i)g_2(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_2(x_i))a_2 ... + (\sum_{i=1}^{m} g_n(x_i)g_2(x_i))a_n = \sum_{i=1}^{m} f(x_i)g_2(x_i) \\
... \\
(\sum_{i=1}^{m} g_1(x_i)g_n(x_i))a_1 + (\sum_{i=1}^{m} g_2(x_i)g_n(x_i))a_2 ... + (\sum_{i=1}^{m} g_n(x_i)g_n(x_i))a_n = \sum_{i=1}^{m} f(x_i)g_n(x_i)
\end{cases}
$$

Exemplo

- Considere a função f(x) definida conforme a tabela:

x	-2	-1	0	1	2	3
f(x)	19.01	3.99	-1.00	4.01	18.99	45.00

ao traçarmos o gráfico, vemos que os pontos se assemelham a uma parábola. Encontre, pois, o polinômio de grau dois que melhor se ajusta aos pontos.

$g(x) = a_1x^2 + a_2x + a_3$

isto é: $g_1 = x^2$, $g_2 = x$ e $g_3 = 1$

Exemplo (solução)

- Temos o sistema de equações normais:

$$\begin{bmatrix} \sum_{i=1}^{6} x_i^4 & \sum_{i=1}^{6} x_i^3 & \sum_{i=1}^{6} x_i^2 \\ \sum_{i=1}^{6} x_i^3 & \sum_{i=1}^{6} x_i^2 & \sum_{i=1}^{6} x_i \\ \sum_{i=1}^{6} x_i^2 & \sum_{i=1}^{6} x_i & 6 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{6} f(x_i)x_i^2 \\ \sum_{i=1}^{6} f(x_i)x_i \\ \sum_{i=1}^{6} f(x_i) \end{bmatrix}$$

$\boxed{\sum_{i=1}^{6} g_3(x_i)\, g_2(x_i)}$

$\boxed{\sum_{i=1}^{6} g_1(x_i)\, g_1(x_i)}$

e assim por diante...

Numericamente

	xi	xi^2	xi^3	xi^4	f(xi)	xif(xi)	xi^2f(xi)
	-2.00	4.00	-8.00	16.00	19.01	-38.02	76.04
	-1.00	1.00	-1.00	1.00	3.99	-3.99	3.99
	0.00	0.00	0.00	0.00	-1	0	0
	1.00	1.00	1.00	1.00	4.01	4.01	4.01
	2.00	4.00	8.00	16.00	18.99	37.98	75.96
	3.00	9.00	27.00	81.00	45.00	135	405
soma:	**3.00**	**19.00**	**27.00**	**115.00**	**90.00**	**134.98**	**565.00**

$$\begin{bmatrix} 115 & 27 & 19 \\ 27 & 19 & 3 \\ 19 & 3 & 6 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} 565 \\ 134.98 \\ 90.00 \end{bmatrix}$$

$a_1 = 5.0893$

$a_2 = 0.0515$

$a_3 = -1.1403$

Função g(x)

- $g(x) = 5.0893x^2 + 0.0515x - 1.1403$

- Nenhuma outra função quadrática apresentará um menor erro quadrático para aqueles pontos.