Método dos Mínimos Quadrados

Prof. Jonathan Esteban Arroyo Silva

Departamento de Ciência da Computação
Universidade Federal de São João del-Rei
silva.jea@ufsj.edu.br

Sumário

Introdução

- Diferentemente das condições para utilizar interpolação, ao levar em conta a existência de erros nos dados obtidos (limitações dos instrumentos de medição, condições experimentais, etc), surgem novos desafios a serem enfrentados
- Neste tipo de problemas, o método dos mínimos quadrados é amplamente utilizado

Introdução

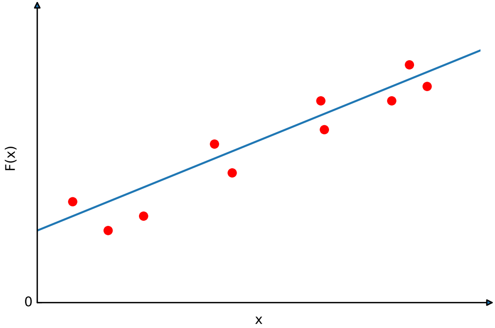
- Para ilustrar a ideia, considere agora o problema de determinar a constante de uma mola. A Lei de Hooke nos diz que o deslocamento de uma mola é proporcional à força nela aplicada. A questão é como encontrar essa constante de proporcionalidade a partir de dados experimentais

$$F = kx$$

- Suponha que possamos realizar vários experimentos medindo forças aplicadas à mola e os seus respectivos deslocamentos

Introdução

O problema consiste em encontrar uma reta que melhor aproxime esses dados. A inclinação da reta irá nos fornecer a constante k da mola



Método dos Mínimo Quadrados

- Vamos chamar de $f(x)$ a função que queremos aproximar por uma outra função $g(x)$
- Nem sempre temos uma expressão analítica para $f(x)$
- No método dos mínimos quadrados (MMQ) partimos da hipótese de que temos algumas informações sobre a forma de $g(x)$
- Poderíamos saber, *através da observação dos dados*, por exemplo, que $g(x)$ é uma reta, uma parábola ou alguma outra forma específica

Método dos Mínimo Quadrados

- De forma geral, no caso linear, vamos considerar que a aproximação será por uma função do tipo:

$$g(x) = c_0\phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x) \tag{1}$$

sendo $\phi_0(x), \phi_1(x), \ldots, \phi_n(x)$ **funções pré estabelecidas**

- Neste caso, a linearidade se dá em relação aos parâmetros $c_0, c_1, \ldots, c_n$
- Exemplo de algumas famílias de funções utilizadas:

$$\phi_0(x) = 1;\;\; \phi_1(x) = x;\;\; \phi_2(x) = x^2; \ldots$$
$$\phi_0(x) = \text{sen}(\pi x);\;\; \phi_1(x) = \text{sen}(2\pi x); \ldots$$

- Para cada conjunto de coeficientes $c_j, i = 0, 1, \ldots, n$, o **desvio** ou **resíduo da aproximação** no ponto x_k é dado por:

$$
\begin{aligned}
r(x_k) &= f(x_k) - g(x_k) \\
&= f(x_k) - [c_0\phi_0(x_k) + c_1\phi_1(x_k) + \ldots + c_n\phi_n(x_k)] \\
&= r(x_k; c_0, c_1, \ldots, c_n)
\end{aligned}
$$

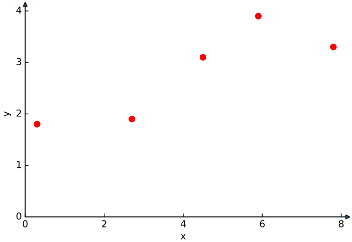
sendo assim, precisamos de estabelecer critérios de aproximação para encaminhar o problema da determinação dos parâmetros $c_0, c_1, \ldots, c_n$ que nos levarão à **melhor aproximação**.

Exemplo

e seu respectivo gráfico:

Tendo a seguinte tabela de dados:

x_i	0.3	2.7	4.5	5.9	7.8
y_i	1.8	1.9	3.1	3.9	3.3

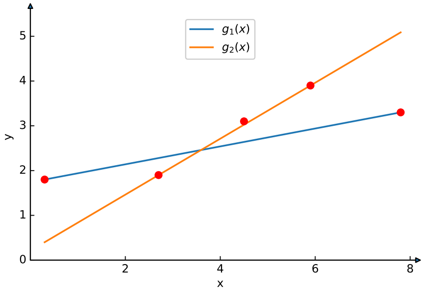


Podemos considerar que existe uma relação aproximadamente linear entre as variáveis.

Exemplo

Assim, desprezando alguns pontos da tabela e escolhemos 2 deles, podemos utilizar polinômios interpoladores lineares para obter uma solução aproximada:

- Escolhendo os pontos (0.3,1.8) e (7.8,3.3) teremos $g_1(x) = 1.74 + 0.2x$
- Escolhendo os pontos (2.7,1.9) e (5.9,3.9) teremos $g_2(x) = 0.2125 + 0.625x$

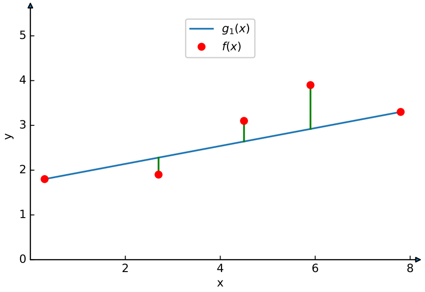


Questionamentos

- Qual aproximação é melhor? $g_1(x)$ ou $g_2(x)$?
- Como verificar a qualidade da aproximação?

Exemplo

- Para obter as aproximações $g_1(x)$ e $g_2(x)$ desprezamos vários dados da tabela para fazer a interpolação, o que não é muito conveniente de se fazer
- Um modo de se verificar a qualidade da aproximação é calculando a soma de todas as distâncias verticais entre $f(x_i)$ e $g(x_i)$ (representadas pelas linhas verdes no gráfico a seguir) ao quadrado:



Exemplo

Utilizando a seguinte fórmula:

$$E^2 = \sum_{i=1}^{m} r^2(x_i) = \sum_{i=1}^{m} [f(x_i) - g(x_i)]^2 \qquad (2)$$

e as seguintes tabelas, é possível calcular a qualidade da aproximação de $g_1(x)$ e $g_2(x)$:

x_i	$f(x_i)$	$g_1(x_i)$	$g_2(x_i)$	$r_1^2(x_i)$	$r_2^2(x_i)$
0.3	1.8	1.8	0.4	0	1.96
2.7	1.9	2.78	1.9	0.144	0
4.5	3.1	2.64	3.025	0.2116	0.0056
5.9	3.9	2.94	3.9	0.9604	0
7.8	3.3	3.3	2.275	0	1.051

Questionamentos

- Obtendo $E_1^2 = 1.316$ e $E_2^2 = 3.0166$, podemos afirmar que $g_1(x)$ se aproxima melhor que $g_2(x)$
- Como escolher c_0 e c_1 de forma que E^2 seja mínimo?

Definição - I

Antes de partir para águas mais profundas, é importante definir o produto escalar entre funções:

$$< f, g >= \begin{cases} \sum_{i=1}^{m} f(x_i)g(x_i), & \text{caso discreto} \\ \int_a^b f(x)g(x)dx, & \text{caso contínuo} \end{cases}$$

Definição - II

- O método dos mínimos quadrados consiste na procura de parâmetros $(c_0, c_1, \ldots, c_n)$ que minimizem:
  - a soma dos quadrados dos resíduos (caso discreto)
  - a integral da função resíduo ao quadrado (caso contínuo)
- Utilizando a notação de produto escalar, podemos escrever:

$$< r, r >= \sum_{i=1}^{m} [r(x_i)]^2 = \sum_{i=1}^{m} [f(x_i) - g(x_i)]^2$$

MMQ - Caso Discreto Linear

De forma geral, queremos aproximar $f(x)$ por

$$g(x) = c_0\phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x)$$

sendo ϕ_j funções conhecidas. Assim, para encontrar os melhores parâmetros $c_0, c_1, \ldots, c_n$ é preciso minimizar a função com relação aos parâmetros c_j:

$$< r, r >= \sum_{i=1}^{m} \left[f(x_i) - \sum_{j=0}^{n} c_j\phi_j(x_i) \right]^2 \tag{3}$$

Derivando a Eq.3 com relação a cada um dos parâmetros c_k e igualando a zero temos:

$$-2\sum_{i=1}^{m}\left[f(x_i)-\sum_{j=0}^{n}c_j\phi_j(x_i)\right]\phi_k(x_i)=0$$

$$\sum_{i=1}^{m}f(x_i)\phi_k(x_i)-\sum_{i=1}^{m}\left[\sum_{j=0}^{n}c_j\phi_j(x_i)\phi_k(x_i)\right]=0$$

$$\sum_{i=1}^{m}\left[\sum_{j=0}^{n}c_j\phi_j(x_i)\phi_k(x_i)\right]=\sum_{i=1}^{m}f(x_i)\phi_k(x_i)$$

$$\sum_{j=0}^{n}c_j\left[\sum_{i=1}^{m}\phi_j(x_i)\phi_k(x_i)\right]=\sum_{i=1}^{m}f(x_i)\phi_k(x_i)$$

MMQ - Caso Discreto Linear

Através da notação de produto escalar podemos reescrever a última igualdade como:

$$\sum_{j=0}^{n} c_j < \phi_j, \phi_k >=< f, \phi_k >, \text{para } k = 0, 1, \ldots, n$$

que é um sistema de equações lineares $(n+1) \times (n+1)$.

Caso - n=2

Expandindo a expressão anterior para o caso $n = 2$ temos:

$$< \phi_0, \phi_0 > c_0 + < \phi_0, \phi_1 > c_1 + < \phi_0, \phi_2 > c_2 = < f, \phi_0 >$$
$$< \phi_1, \phi_0 > c_0 + < \phi_1, \phi_1 > c_1 + < \phi_1, \phi_2 > c_2 = < f, \phi_1 >$$
$$< \phi_2, \phi_0 > c_0 + < \phi_2, \phi_1 > c_1 + < \phi_2, \phi_2 > c_2 = < f, \phi_2 >$$

ou de forma matricial:

$$\begin{bmatrix} < \phi_0, \phi_0 > & < \phi_0, \phi_1 > & < \phi_0, \phi_2 > \\ < \phi_1, \phi_0 > & < \phi_1, \phi_1 > & < \phi_1, \phi_2 > \\ < \phi_2, \phi_0 > & < \phi_2, \phi_1 > & < \phi_2, \phi_2 > \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} < f, \phi_0 > \\ < f, \phi_1 > \\ < f, \phi_2 > \end{bmatrix}$$

Caso geral

Para caso geral, teremos:

$$\begin{bmatrix} <\phi_0,\phi_0> & <\phi_0,\phi_1> & \cdots & <\phi_0,\phi_n> \\ <\phi_1,\phi_0> & <\phi_1,\phi_1> & \cdots & <\phi_1,\phi_n> \\ \vdots & \vdots & \ddots & \vdots \\ <\phi_n,\phi_0> & <\phi_n,\phi_1> & \cdots & <\phi_n,\phi_n> \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} <f,\phi_0> \\ <f,\phi_1> \\ \vdots \\ <f,\phi_n> \end{bmatrix}$$

o qual é chamado de **sistema normal** ou **equações normais**.
Observe que a matriz é simétrica, pois $<f,g>=<g,f>$.

Exemplo - I

Dada a seguinte tabela:

x_i	0	0.25	0.5	0.75	1.0
$f(x_i)$	1	1.284	1.6487	2.117	2.7183

aproxime $f(x)$ por um polinômio linear usando o MMQ.

Solução

Como queremos encontrar $g(x) = c_0\phi_0(x) + c_1\phi_1(x)$ sendo

$$\phi_0(x) = 1, \text{ e } \phi_1(x) = x$$

vamos primeiro montar os vetores que correspondem a avaliação de cada $\phi_j(x)$ nos pontos x_i dados na tabela, tendo o seguinte sistema linear:

$$\begin{bmatrix} <\phi_0,\phi_0> & <\phi_0,\phi_1> \\ <\phi_1,\phi_0> & <\phi_1,\phi_1> \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} <f,\phi_0> \\ <f,\phi_1> \end{bmatrix}$$

Solução

Como cada $\phi_j(x)$ deve ser calculado em todos os pontos x_i, teremos

$$\phi_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \text{ e } \phi_1 = \begin{bmatrix} 0 \\ 0.25 \\ 0.5 \\ 0.75 \\ 1.0 \end{bmatrix}$$

E calculando os produtos escalares temos:

$$< \phi_0, \phi_0 >= 5; \;\; < \phi_1, \phi_0 >= 2.5; \;\; < \phi_1, \phi_1 >= 1.875$$
$$< f, \phi_0 >= 8.768; \;\; < f, \phi_1 >= 5.4514$$

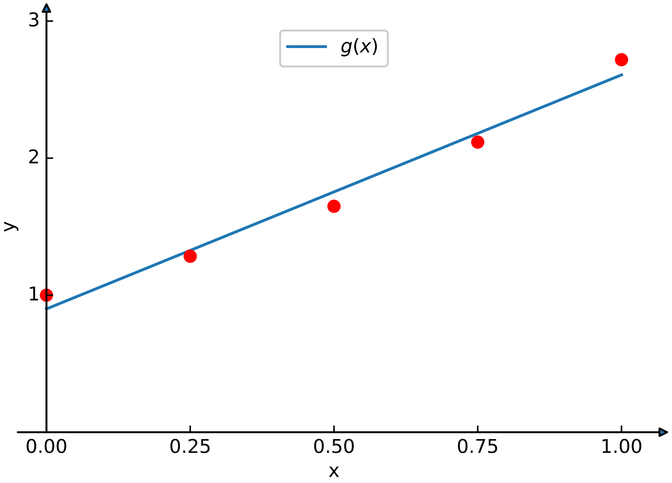
Solução

Resultando no seguinte sistema:

$$\begin{bmatrix} 5 & 2.5 \\ 2.5 & 1.875 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 8.768 \\ 5.4514 \end{bmatrix}$$

cuja solução é o vetor $\mathbf{c} = \begin{bmatrix} 0.89968 \\ 1.70784 \end{bmatrix}$, resultando na função:

$$g(x) = 0.89968 + 1.70784x$$



$g(x)$
3
2
1
y
0.00
0.25
0.50
0.75
1.00
x

Exemplo - II

Dada a seguinte tabela:

x_i	-1	0	1	2
$f(x_i)$	0	-1	0	7

aproxime $f(x)$ por um polinômio quadrático usando o MMQ.

Solução

Como queremos encontrar $g(x) = c_0\phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$ sendo

$$\phi_0(x) = 1, \;\; \phi_1(x) = x, \;\text{e}\; \phi_2(x) = x^2$$

vamos primeiro montar os vetores que correspondem a avaliação de cada $\phi_j(x)$ nos pontos x_i dados na tabela, tendo o seguinte sistema linear:

$$\begin{bmatrix} <\phi_0,\phi_0> & <\phi_0,\phi_1> & <\phi_0,\phi_2> \\ <\phi_1,\phi_0> & <\phi_1,\phi_1> & <\phi_1,\phi_2> \\ <\phi_2,\phi_0> & <\phi_2,\phi_1> & <\phi_2,\phi_2> \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} <f,\phi_0> \\ <f,\phi_1> \\ <f,\phi_2> \end{bmatrix}$$

Solução

Como cada $\phi_j(x)$ deve ser calculado em todos os pontos x_i, teremos

$$\phi_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ \phi_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \text{ e } \phi_2(x) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 4 \end{bmatrix}$$

E calculando os produtos escalares temos:

$$< \phi_0, \phi_0 >= 4; \quad < \phi_1, \phi_0 >= 2; \quad < \phi_2, \phi_0 >= 6;$$
$$< \phi_1, \phi_1 >= 6; \quad < \phi_2, \phi_1 >= 8; \quad < \phi_2, \phi_2 >= 18;$$
$$< f, \phi_0 >= 6; \quad < f, \phi_1 >= 14; \quad < f, \phi_2 >= 28$$

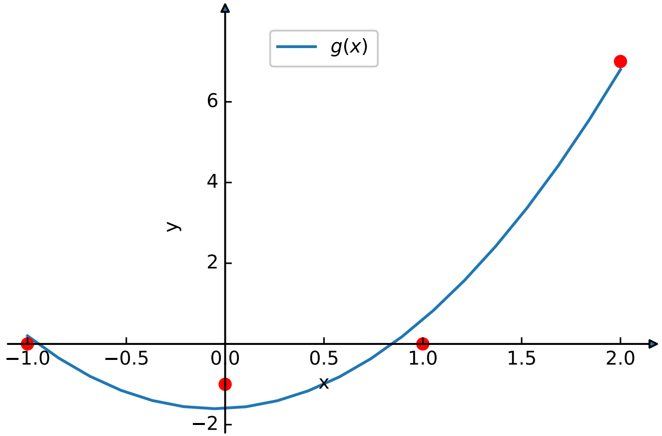
Solução

Resultando no seguinte sistema:

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 28 \end{bmatrix}$$

cuja solução é o vetor $\mathbf{c} = \begin{bmatrix} -8/5 \\ 1/5 \\ 2 \end{bmatrix}$, resultando na função:

$$g(x) = -8/5 + x/5 + 2x^2$$



$g(x)$
6
4
2
−2
y
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x

Conclusão

Foram abordados os seguintes assuntos:

- Introdução à aproximação de curvas pelo método dos mínimos quadrados
- MMQ - caso linear

OBRIGADO PELA SUA PARTICIPAÇÃO

Dragon Ball ©