DCC008 - Cálculo Numérico

Método dos Mínimos Quadrados

Bernardo Martins Rocha

Departamento de Ciência da Computação
Universidade Federal de Juiz de Fora
bernardomartinsrocha@ice.ufjf.br

Conteúdo

Introdução

Existem duas classes de métodos para a aproximação de dados, e a distinção entre elas está em considerarmos, ou não, a existência de erro nos dados.

No primeiro caso, consideramos que não existem erros nos dados e podemos exigir que a curva passe pelos pontos dados. Como vimos, esse problema é resolvido com **interpolação**.

Nesse caso aproximamos uma função $f(x)$ por uma função polinomial $p(x)$ que passa exatamente pelos pontos dados $(x_0, y_0), \ldots, (x_n, y_n)$. Assim dado $p(x)$ é possível estimar o valor de $f(\bar{x})$ para um ponto $\bar{x}$ diferente dos pontos x_i, para $i = 0, 1, \ldots, n$.

A outra classe de métodos leva em consideração possíveis erros introduzidos na obtenção dos dados (limitações do instrumento, condições experimentais, etc).

Introdução

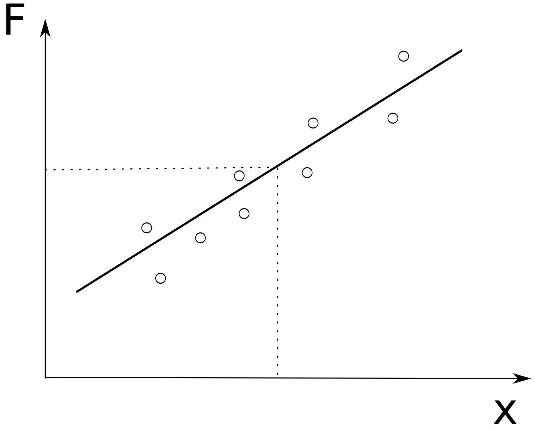
No segundo caso o **método dos mínimos quadrados** tem sido amplamente utilizado.

Para ilustrar a idéia, 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.

$$\text{Lei de Hooke} \quad \Rightarrow \quad F = kx$$

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

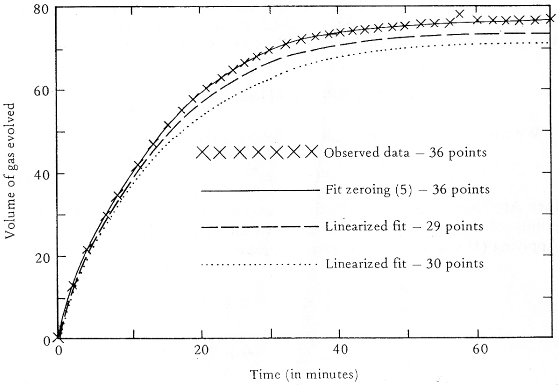
Introdução



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

Introdução

Outro exemplo: $f(x) = ae^{bx}$



O método dos mínimos quadrados é uma das técnicas mais usadas em problemas práticos porque em geral buscamos aproximações para dados que são medidas obtidas experimentalmente e possuem um certo grau de incerteza.

Introdução

Interpolação e Mínimos Quadrados

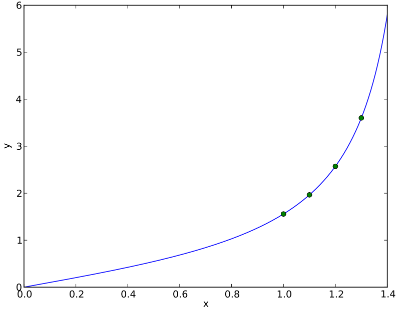


Figura: Interpolação

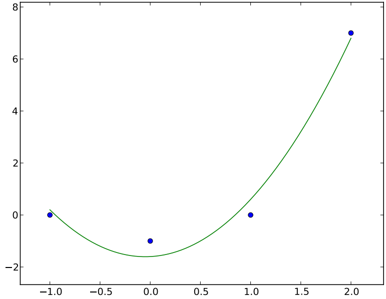


Figura: Mínimos Quadrados

Método dos mínimos 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,

$$g(x) = c_0 + c_1 x$$

ou que é uma parábola

$$g(x) = c_0 + c_1 x + c_2 x^2$$

ou que tenha alguma outra forma específica.

Método dos mínimos quadrados

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

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

onde $\phi_0(x), \phi_1(x), \ldots, \phi_n(x)$ são funções pré-estabelecidas.

MMQ linear: a função $g(x)$ que aproxima $f(x)$ é linear nos seus parâmetros $c_0, c_1, \ldots, c_n$.

Exemplos:

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_2(x) = x^2, \quad \ldots$$
$$\phi_0(x) = \sin(\pi x), \quad \phi_1(x) = \sin(2\pi x), \quad \ldots$$

Veremos mais adiante como trabalhar no **caso não-linear**, como por exemplo quando $g(x) = c_0\ e^{c_1 x}$.

Método dos mínimos quadrados

Para cada conjunto de coeficientes c_i, $i = 0, 1, \ldots, n$, o desvio ou resíduo da aproximação dada pela eq. (1) no ponto x_k é dado por

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

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

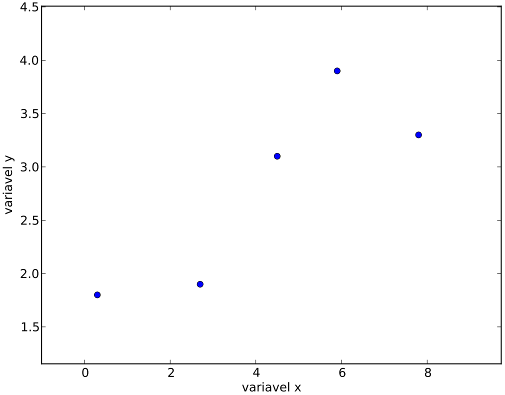
Vamos ver um exemplo para começar a discussão.

Método dos mínimos quadrados

Suponha que temos 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

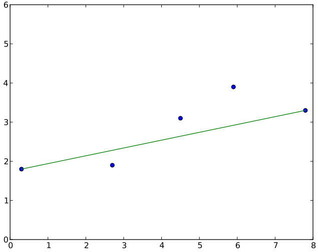
Vamos analisar os dados.

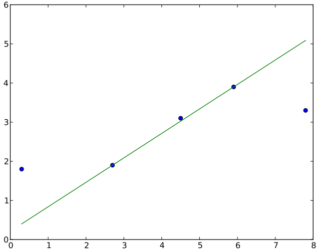


Método dos mínimos quadrados

Podemos considerar que existe uma relação aproximadamente linear entre as variáveis. Logo, podemos desprezar alguns pontos da tabela, escolher 2 pontos e usar polinômios interpoladores lineares para aproximação. Ou seja $g(x) = c_0 + c_1 x$.

- Para $(0.3, 1.8)$ e $(7.8, 3.3)$ temos $g_1(x) = 1.74 + 0.2x$.
- Para $(2.7, 1.9)$ e $(5.9, 3.9)$ temos $g_2(x) = 0.2125 + 0.625x$.



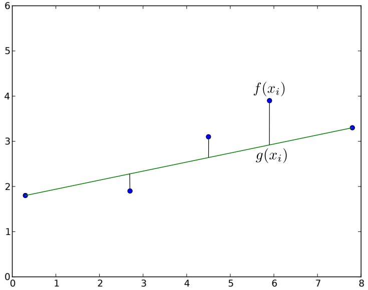


Qual aproximação é melhor? $g_1(x)$ ou $g_2(x)$?
Como verificar a qualidade da aproximação?

Método dos mínimos quadrados

Para obter as aproximações g_1 e g_2 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 de $f(x_i)$ e $g(x_i)$ ao quadrado, isto é



$$E^2 = \sum_{i=1}^{m} r(x_i)^2 = \sum_{i=1}^{m} [f(x_i) - g(x_i)]^2$$
$$= \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i]^2$$

Método dos mínimos quadrados

Para $g_1(x) = 1.74 + 0.2x$

i	x_i	$f(x_i)$	$g(x_i)$	r_i^2
1	0.3	1.8	1.8	0
2	2.7	1.9	2.28	0.144
3	4.5	3.1	2.64	0.2116
4	5.9	3.9	2.92	0.9604
5	7.8	3.3	3.3	0

$$E^2 = 1.316$$

Para $g_2(x) = 0.2125 + 0.625x$

i	x_i	$f(x_i)$	$g(x_i)$	r_i^2
1	0.3	1.8	0.4	1.96
2	2.7	1.9	1.9	0
3	4.5	3.1	3.025	0.0056
4	5.9	3.9	3.9	0
5	7.8	3.3	5.0875	3.1952

$$E^2 = 3.0166$$

Logo, $g_1(x)$ é mais adequado de acordo com esse critério de erros ao quadrado.

Como escolher c_0 e c_1 de forma que E^2 seja mínimo?

Método dos mínimos quadrados

Até então falamos em aproximação de dados discretos, isto é, aproximar $f(x)$ sabendo os valores de f em um conj. de pontos x_i.

Podemos trabalhar com o caso de aproximar $f(x)$, uma função conhecida para todo x, por uma função mais simples como $g(x)$. No **caso contínuo** do método dos mínimos quadrados temos que minimizar

$$\int_a^b r(x)^2 \, dx = \int_a^b (f(x) - g(x))^2 \, dx$$

Vamos introduzir formalmente a definição de produto escalar entre funções para que possamos medir o resíduo:

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

Método dos mínimos quadrados

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), ou,
- a integral da função resíduo ao quadrado (caso contínuo).

Desta forma, introduzindo a notação do produto escalar temos

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

Como vimos $< r, r >$ é função dos parâmetros $c_0, c_1, \ldots, c_n$, isto é, $< r, r >=< r, r > (c_0, c_1, \ldots, c_n)$.

Antes de estudar o método de forma geral no caso discreto, vamos ver um exemplo simples: **regressão linear**.

Método dos mínimos quadrados

Regressão linear

Dada uma tabela de dados $(x_i, f(x_i))$, $i = 1, \ldots, m$, queremos encontrar a reta que melhor se ajusta aos dados no critério dos mínimos quadrados. Como o ajuste será feito por uma reta temos que $\phi_0(x) = 1$ e $\phi_2(x) = x$, ou seja

$$f(x) \approx g(x) = c_0 + c_1 x$$

o resíduo é dado por

$$r = f(x) - c_0 - c_1 x$$

Pelo método dos mínimos quadrados devemos encontrar c_0 e c_1 que **minimizem** a função

$$
\begin{aligned}
< r, r > &= < f(x) - c_0 - c_1 x, f(x) - c_0 - c_1 x > \qquad (2)\\
&= \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i]^2 \qquad (3)
\end{aligned}
$$

Método dos mínimos quadrados

Regressão linear

Do Cálculo sabemos que a condição necessária de ponto crítico é que as derivadas nele sejam nulas, isto é

$$\frac{\partial}{\partial c_0} < r, r >= 0, \quad \frac{\partial}{\partial c_1} < r, r >= 0$$

Derivando a Eq. (2) com relação a c_0 temos

$$\begin{aligned}
\frac{\partial}{\partial c_0} < r, r > &= \frac{\partial}{\partial c_0} \left\{ \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i]^2 \right\} \\
&= -2 \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i]
\end{aligned}$$

De forma similar, derivando a Eq. (3) com relação a c_1 temos

$$\frac{\partial}{\partial c_1} < r, r > = \frac{\partial}{\partial c_1} \left\{ \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i]^2 \right\}$$
$$= -2 \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i](x_i)$$

Condição de ponto crítico:

$$\frac{\partial}{\partial c_0} < r, r > = 0 \quad \Rightarrow \quad \sum_{i=1}^{m} [f(x_i) - c_0 - c_1 x_i] = 0$$
$$\frac{\partial}{\partial c_1} < r, r > = 0 \quad \Rightarrow \quad \sum_{i=1}^{m} (x_i)[f(x_i) - c_0 - c_1 x_i] = 0$$

Manipulando as expressões temos

$$\sum_{i=1}^{m} f(x_i) = \sum_{i=1}^{m} c_0 + c_1 x_i = c_0 m + c_1 \sum_{i=1}^{m} x_i$$

$$\sum_{i=1}^{m} x_i f(x_i) = \sum_{i=1}^{m} x_i (c_0 + c_1 x_i) = c_0 \sum_{i=1}^{m} x_i + c_1 \sum_{i=1}^{m} x_i^2$$

chegamos a

$$c_0 m + c_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} f(x_i)$$

$$c_0 \sum_{i=1}^{m} x_i + c_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i \; f(x_i)$$

Método dos mínimos quadrados

Regressão linear

$$c_0 m + c_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} f(x_i)$$
$$c_0 \sum_{i=1}^{m} x_i + c_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i \; f(x_i)$$

pode ser escrito na forma:

$$\begin{bmatrix} m & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} f(x_i) \\ \sum_{i=1}^{m} x_i \; f(x_i) \end{bmatrix} \tag{4}$$

Temos que, resolvendo o sistema de equações lineares encontramos c_0 e c_1, e assim determinamos $g(x) = c_0 + c_1 x$.

Vejamos um exemplo numérico para ilustrar o procedimento. Considere a seguinte tabela de dados:

x_i	0	0.25	0.5	0.75	1
$f(x_i)$	1	1.2840	1.6487	2.1170	2.7183

Precisamos montar o sistema de equações lineares dado por (4) para encontrar c_0 e c_1 e assim determinar $g(x)$.

$$\sum_{i=1}^{5} x_i = 2.5 \qquad \sum_{i=1}^{5} x_i^2 = 1.875$$

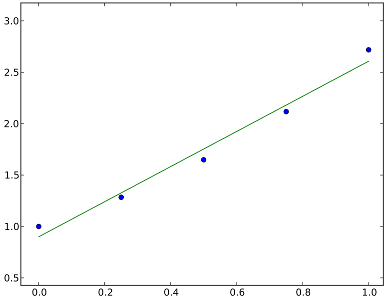
$$\sum_{i=1}^{5} f(x_i) = 8.768 \qquad \sum_{i=1}^{5} x_i f(x_i) = 5.4514$$

Logo temos o seguinte sistema

$$5c_0 + 2.5c_1 = 8.768$$
$$2.5c_0 + 1.875c_1 = 5.4514$$

Resolvendo encontramos

$$c_0 = 0.89968$$
$$c_1 = 1.70784 \quad \Rightarrow \quad g(x) = 0.89968 + 1.70784x$$



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

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

onde $\phi_j(x)$ são funções conhecidas. Assim para encontrar os parâmetros $c_0, c_1, \ldots, c_n$ é preciso minimizar a função:

$$\begin{aligned}< r, r > &=< f - c_0\phi_0 - \ldots - cn\phi_n, f - c_0\phi_0 - \ldots - cn\phi_n > \\ &= \sum_{i=1}^{m}\left[f(x_i) - \sum_{j=0}^{n} c_j\phi_j(x_i)\right]^2\end{aligned}$$

Derivando com relação a cada um dos parâmetros c_k e igualando a zero temos

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

Método dos mínimos quadrados

Caso Geral

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

Lembrando que

$$
< f, g >= \sum_{i=1}^{m} f(x_i)g(x_i)
$$

Método dos mínimos quadrados

Caso Geral

Usando a notação de produto escalar podemos escrever

$$\Rightarrow \quad \sum_{j=0}^{n} c_j \underbrace{\left(\sum_{i=1}^{m} \phi_j(x_i)\phi_k(x_i) \right)}_{<\phi_j,\phi_k>=<\phi_k,\phi_j>} = \underbrace{\sum_{i=1}^{m} f(x_i)\phi_k(x_i)}_{<f,\phi_k>=<\phi_k,f>}$$

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

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

<u>Exemplo para $n = 2$:</u>

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

Método dos mínimos quadrados

Caso Geral

Escrevendo de forma matricial temos

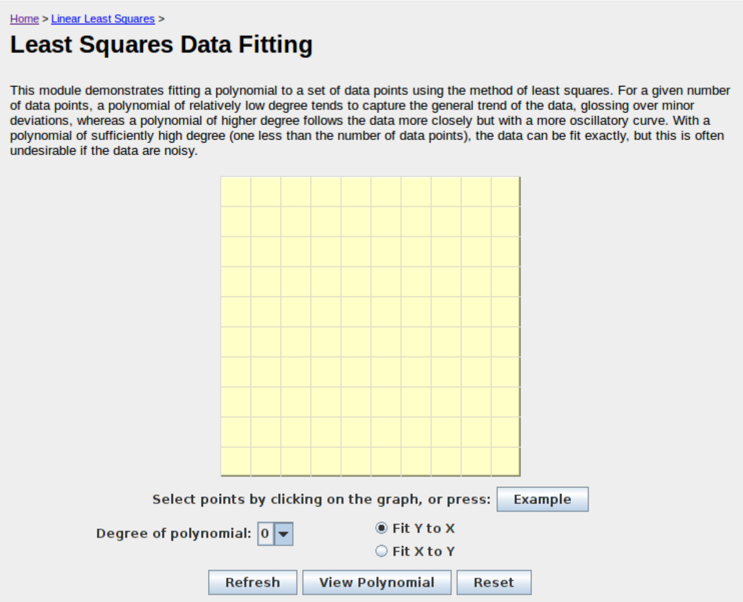
$$\begin{bmatrix} <\phi_0,\phi_0> & <\phi_0,\phi_1> & <\phi_0,\phi_2> \\ <\phi_1,\phi_0> & <\phi_1,\phi_1> & <\phi_1,\phi_2> \\ <\phi_2,\phi_0> & <\phi_2,\phi_1> & <\phi_2,\phi_2> \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} <\phi_0,f> \\ <\phi_1,f> \\ <\phi_2,f> \end{bmatrix}$$

De forma geral o sistema é dado por

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

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

Método dos mínimos quadrados



Exemplo

Dada a seguinte tabela

x	-1	0	1	2
y	0	-1	0	7

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

Solução do Exemplo

Queremos encontrar $g(x) = c_0 + c_1 x + c_2 x^2$ onde temos que

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_2(x) = x^2$$

Sendo assim, vamos primeiro montar os vetores que correspondem a avaliação de cada $\phi_j(x)$ nos pontos x_i dados na tabela.

Método dos mínimos quadrados

Solução do Exemplo

$$\phi_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \phi_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \phi_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 4 \end{bmatrix}, \quad f = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 7 \end{bmatrix}$$

Calculando os produtos escalares temos

$$< \phi_0, \phi_0 > = 4$$
$$< \phi_0, \phi_1 > = 2$$
$$< \phi_0, \phi_2 > = 6$$
$$< \phi_1, \phi_1 > = 6$$
$$< \phi_1, \phi_2 > = 8$$
$$< \phi_2, \phi_2 > = 18$$

$$< f, \phi_0 > = 6$$
$$< f, \phi_1 > = 14$$
$$< f, \phi_2 > = 28$$

Método dos mínimos quadrados

Solução do Exemplo

Assim temos o seguinte sistema de equações lineares (equações normais)

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

Resolvendo o sistema encontramos a seguinte solução

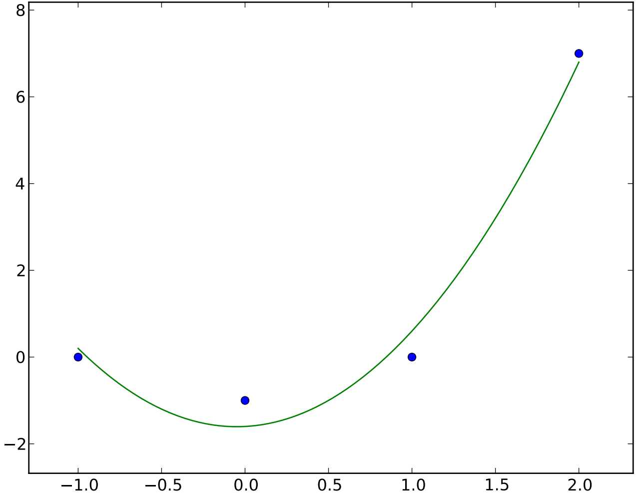
$$c_0 = -\tfrac{8}{5}, \quad c_1 = \tfrac{1}{5}, \quad c_2 = 2$$

logo a aproximação pelo MMQ é dada por

$$g(x) = -\tfrac{8}{5} + \tfrac{1}{5}x + 2x^2$$

Método dos mínimos quadrados

Solução do Exemplo

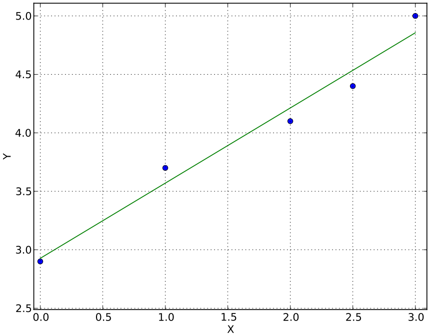


Exemplo - Sala de Aula

x	0.0	1.0	2.0	2.5	3.0
y	2.9	3.7	4.1	4.4	5.0

Solução do Exemplo

$g(x) = 2.92672414 + 0.64310345x$



Conteúdo

- Aula passada
  - Caso discreto
  - Regressão linear
  - Caso geral
- Aula de hoje
  - Caso não-linear
  - Caso contínuo
  - Polinômio ortogonais

Mínimos Quadrados

Dados m pontos $(x_1, y_1), \ldots, (x_m, y_m)$ deseja-se encontrar a função

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

que melhor se ajusta aos dados no sentido do método dos mínimos quadrados, considerando que as funções $\phi_j(x)$, $j = 0, 1, \ldots, n$ são conhecidas.

Nesta situação estamos trabalhando com o MMQ no caso linear, pois a função $g(x)$ é uma função *linear* dos parâmetros $c_0, c_1, \ldots, c_n$.

Veremos a seguir como lidar com o caso em que $g(x)$ é uma função não-linear dos parâmetros $c_0, c_1, \ldots, c_n$.

Mínimos Quadrados

Caso Não-Linear

Vamos considerar agora que estamos interessados em ajustar uma curva $g(x)$ que é uma função não-linear dos parâmetros.

Alguns exemplos:

$$g(x) = c_0 \; e^{c_1 x}$$
$$g(x) = c_0 \; x^{c_1}$$
$$g(x) = \frac{1}{c_0 + c_1 x}$$
$$g(x) = c_0 \; c_1^x$$

Uma das formas de tratar essa situação é através da **transformação** ou **linearização** do modelo não-linear em um modelo linear.

Vamos ilustrar o procedimento através de um exemplo.

Exemplo

Seja $f(x) = 20e^{-15x}$. Considere a seguinte tabela de dados

x_i	0.0	0.1	0.2	0.3	0.4
y_i	20.5	4.60	1.00	0.15	0.05

Solução do Exemplo

Queremos fazer um ajuste da seguinte forma

$$f(x) \approx g(x) = c_0 \; e^{c_1 x}$$

Temos que fazer uma transformação e linearizar o modelo

$$\begin{aligned} \ln f &\approx \ln\left(c_0 \; e^{c_1 x}\right) \\ &\approx \ln\left(c_0\right) + \ln\left(e^{c_1 x}\right) \\ &\approx \ln\left(c_0\right) + c_1 x \end{aligned}$$

Exemplo

Assim temos

$$F(x) = a_0\phi_0(x) + a_1\phi_1(x)$$

onde

$$\phi_0(x) = 1$$
$$\phi_1(x) = x$$

com a seguinte mudança

$$a_0 = \ln(c_0)$$
$$a_1 = c_1$$
$$F = \ln f(x)$$

Exemplo

Construimos os vetores

$$\phi_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \phi_1 = \begin{bmatrix} 0.0 \\ 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \end{bmatrix}, F = \begin{bmatrix} \ln f(x_1) \\ \ln f(x_2) \\ \ln f(x_3) \\ \ln f(x_4) \\ \ln f(x_5) \end{bmatrix} = \begin{bmatrix} 3.02 \\ 1.52 \\ 0 \\ -1.89 \\ -2.99 \end{bmatrix}$$

para montar o sistema precisamos calcular

$$< \phi_0, \phi_0 >, < \phi_0, \phi_1 >, < \phi_1, \phi_1 >$$
$$< \phi_0, F >, < \phi_1, F >$$

Exemplo

assim podemos montar o sistema

$$\begin{bmatrix} 5 & 1.0 \\ 1.0 & 0.3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0.35 \\ -1.335 \end{bmatrix}$$

cuja solução é dada por $a_0 = 2.88$ e $a_1 = -14.05$.

E assim, como

$$a_0 = \ln c_0 \quad \Rightarrow \quad e^{a_0} = c_0 \quad \Rightarrow \quad c_0 = e^{2.88} \quad \Rightarrow \quad \boxed{c_0 = 17.81}$$
$$a_1 = c_1 \quad \Rightarrow \quad \boxed{c_1 = -14.05}$$

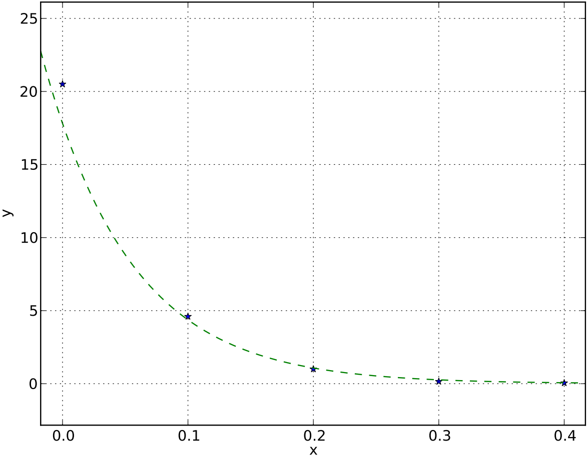
logo a função que melhor aproxima os dados do exemplo é dada por

$$\boxed{g(x) = c_0 e^{c_1 x} = 17.81\ e^{-14.05x}}$$

Mínimos Quadrados

Caso Não-Linear

Exemplo



Mínimos Quadrados

Caso Não-Linear

- Exemplo (1)

$$g(x) = \frac{1}{c_0 + c_1 x} \quad \Rightarrow \quad F(x) = \frac{1}{f(x)} \quad \text{e} \quad G(x) = c_0 + c_1 x$$

- Exemplo (2)

$$g(x) = c_0\ x^{c_1} \quad \Rightarrow \quad F(x) = \ln f \quad \text{e} \quad \begin{aligned} G(x) &= \ln c_0 + \ln x^{c_1} \\ G(x) &= \ln c_0 + c_1 \ln x \\ G(x) &= a_0 + a_1 \ln x \end{aligned}$$

com $\phi_0 = 1$, $\phi_1 = \ln x$, $a_0 = \ln c_0$ e $a_1 = c_1$.

- Exemplo (3)

$$g(x) = c_0\ c_1^x \quad \Rightarrow \quad F(x) = \ln f \quad \text{e} \quad \begin{aligned} G(x) &= \ln c_0 + \ln c_1^x \\ G(x) &= \ln c_0 + x \ln c_1 \\ G(x) &= a_0 + a_1 x \end{aligned}$$

com $\phi_0 = 1$, $\phi_1 = x$, $a_0 = \ln c_0$ e $a_1 = \ln c_1$.

Em situações que só conhecemos a função através dos dados experimentais tabelados, surge a questão: *qual família de funções melhor aproxima os dados?*

Qual função $g(x)$ devemos usar: $g(x) = \frac{1}{c_0 + c_1 x}$ ou $g(x) = c_0 \; c_1^x$?

A idéia é aplicar o chamado **teste de alinhamento**:

1. Transformar $y = f(x)$ em

$$F_1(x) = a_0 \phi_0(x) + a_1 \phi_1(x)$$
$$F_2(x) = b_0 \bar{\phi_0}(x) + b_1 \bar{\phi_1}(x)$$

2. Plotar x vs $F_1(x)$ e x vs $F_2(x)$
3. Escolher aquela que o gráfico estiver *"mais linear"*

Exemplo

Considere a seguinte tabela

x_i	-8	-6	-4	-2	0	2	4
$f(x_i)$	30	10	9	6	5	4	4

Qual é a melhor opção ?

a) $g(x) = \frac{1}{c_0 + c_1 x}$

b) $g(x) = c_0 \; c_1^x$

Solução do Exemplo

Vamos linearizar as funções

a)

$$f(x) \approx g(x) = \frac{1}{c_0 + c_1 x}$$

obtemos

$$\frac{1}{f(x)} \approx c_0 + c_1 x$$
$$F_1(x) = \frac{1}{f(x)} = a_0 + a_1 x$$

onde $a_0 = c_0$, $a_1 = c_1$, $\phi_0 = 1$ e $\phi_1 = x$.

Solução do Exemplo

b)
$$f(x) \approx g(x) = c_0 \; c_1^x$$

obtemos

$$\ln f \approx \ln c_0 + x \ln c_1$$
$$F_2(x) = a_0 + a_1 x$$

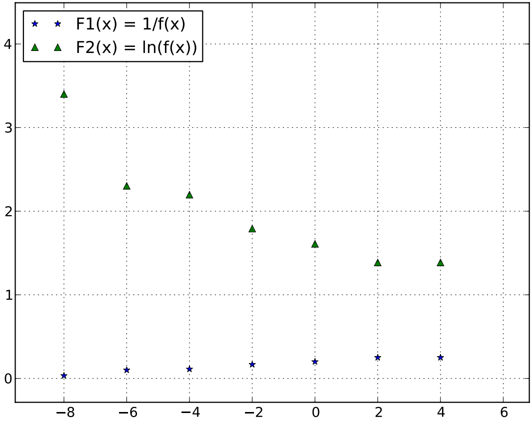
onde $a_0 = \ln c_0$, $a_1 = \ln c_1$, $\phi_0 = 1$ e $\phi_1 = x$.

Temos que

$$F_1(x) = \frac{1}{f(x)}, \qquad F_2(x) = \ln f(x)$$

Vamos fazer os gráficos de x contra $F_1(x)$ e x contra $F_2(x)$ para o teste de alinhamento.

Solução do Exemplo



Logo vemos que $F_1(x)$ é mais adequada para o ajuste de curva.

$\square$

Mínimos Quadrados

Caso Contínuo

Até então trabalhamos em como *aproximar* uma função $f(x)$ por uma função $g(x)$ em um conjunto de pontos x_i onde conhecemos o valor de $f(x_i)$. **(caso discreto)**

Vamos considerar agora o **caso contínuo**, onde queremos aproximar $f(x)$ que é uma função conhecida para todo x em um intervalo $[a,b]$, por uma função $g(x)$ na forma

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

onde $\phi_0(x), \ldots, \phi_n(x)$ são funções contínuas e conhecidas.

Para apresentar o método vamos considerar o seguinte caso

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

Mínimos Quadrados
Caso Contínuo

Lembrando que o produto escalar entre duas funções contínuas no intervalo $[a, b]$ é definido por

$$< f, g >= \int_a^b f(x)g(x)\ dx$$

podemos então medir a distância de $f(x)$ a $g(x)$ fazendo

$$< f - g, f - g >= \int_a^b [f(x) - g(x)]^2\ dx$$

no caso discreto tinhamos

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

Mínimos Quadrados

Caso Contínuo

Considerando que $r = f - g$, temos

$$
\begin{aligned}
< r, r > &= \int_a^b [f(x) - g(x)]^2 \, dx = \int_a^b [f - c_0\phi_0 - c_1\phi_1]^2 \, dx \\
&= \int_a^b [f^2 - 2c_0 f\phi_0 - 2c_1 f\phi_1 + c_0^2\phi_0^2 + 2c_0c_1\phi_0\phi_1 + c_1^2\phi_1^2] \, dx \\
&= \int_a^b f^2 \, dx - 2c_0 \int_a^b f\phi_0 \, dx - 2c_1 \int_a^b f\phi_1 \, dx \\
&+ c_0^2 \int_a^b \phi_0^2 \, dx + 2c_0c_1 \int_a^b \phi_0\phi_1 \, dx + c_1^2 \int_a^b \phi_1^2 \, dx
\end{aligned}
$$

No método dos mínimos quadrados procuramos o ponto de mínimo da função $< r, r > = < r, r > (c_0, c_1)$. Precisamos calcular as derivadas parciais de $< r, r >$ com relação a c_0 e c_1.

Calculando as derivadas parciais e igualando a zero temos

$$\frac{\partial < r, r >}{\partial c_0} = -2\int_a^b f\phi_0 \, dx + 2c_0 \int_a^b \phi_0\phi_0 \, dx + 2c_1 \int_a^b \phi_0\phi_1 \, dx = 0$$
$$\frac{\partial < r, r >}{\partial c_1} = -2\int_a^b f\phi_1 \, dx + 2c_1 \int_a^b \phi_1\phi_1 \, dx + 2c_0 \int_a^b \phi_0\phi_1 \, dx = 0$$

organizando os termos, temos

$$c_0 \int_a^b \phi_0\phi_0 \, dx + c_1 \int_a^b \phi_0\phi_1 \, dx = \int_a^b f\phi_0 \, dx$$
$$c_1 \int_a^b \phi_1\phi_1 \, dx + c_0 \int_a^b \phi_0\phi_1 \, dx = \int_a^b f\phi_1 \, dx$$

usando a notação $< \phi_i, \phi_j >= \int_a^b \phi_i(x)\phi_j(x) \, dx$, podemos escrever

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

que é um sistema de equações lineares (*equações normais*), o qual pode ser escrito de forma matricial como

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

Obs: o procedimento é similar ao caso discreto, a diferença é na forma de se calcular os produtos escalares, pois no caso contínuo temos que calcular as integrais

$$< \phi_0(x), \phi_1(x) >= \int_a^b \phi_0(x)\phi_1(x)\ dx$$
$$< f(x), \phi_0(x) >= \int_a^b f(x)\phi_0(x)\ dx,\ \ldots$$

Exemplo

Seja $f(x) = x^4 - 5x$, $x \in [-1, 1]$. Vamos aproximar $f(x)$ por um polinômio de segundo grau usando o MMQ.

Solução do Exemplo

$f(x)$ é contínua no intervalo $[-1, 1]$.
Queremos

$$f(x) \approx g(x) = c_0 + c_1 x + c_2 x^2$$

onde usamos as funções $\phi_0(x) = 1$, $\phi_1(x) = x$, $\phi_2(x) = x^2$.
Precisamos resolver o sistema de equações normais

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

Exemplo

Substituindo $\phi_0(x) = 1, \phi_1(x) = x, \phi_2(x) = x^2$, temos

$$\begin{bmatrix} <1,1> & <1,x> & <1,x^2> \\ <x,1> & <x,x> & <x,x^2> \\ <x^2,1> & <x^2,x> & <x^2,x^2> \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} <f,\phi_0> \\ <f,\phi_1> \\ <f,x^2> \end{bmatrix}$$

usando o produto escalar usual em $[-1,1]$ definido por

$$<f,g> = \int_a^b f(x)g(x)\ dx \quad \Rightarrow \quad \int_{-1}^{1} f(x)g(x)\ dx$$

temos

$$<1,1> = \int_{-1}^{1} dx = x]_{-1}^{1} = 2$$

$$<1,x> = \int_{-1}^{1} x\ dx = \frac{x^2}{2}]_{-1}^{1} = 0 = <x,1>$$

$$<1,x^2> = \int_{-1}^{1} x^2\ dx = \frac{x^3}{3}]_{-1}^{1} = \frac{2}{3} = <x^2,1>$$

Mínimos Quadrados

Caso Contínuo

Exemplo

$$< x, x > = \int_{-1}^{1} x^2 \ dx = \tfrac{x^3}{3}]_{-1}^{1} = \tfrac{2}{3}$$

$$< x, x^2 > = \int_{-1}^{1} x^3 \ dx = \tfrac{x^4}{4}]_{-1}^{1} = 0 = < x^2, x >$$

$$< x^2, x^2 > = \int_{-1}^{1} x^4 \ dx = \tfrac{x^5}{5}]_{-1}^{1} = \tfrac{2}{5}$$

e ainda

$$< f, 1 > = \int_{-1}^{1} x^4 - 5x \ dx = \left. \tfrac{x^5}{2} - \tfrac{5x^2}{5} \right]_{-1}^{1} = \left(\tfrac{1}{5} - \tfrac{5}{2}\right) - \left(-\tfrac{1}{5} - \tfrac{5}{2}\right) = \tfrac{2}{5}$$

$$< f, x > = \int_{-1}^{1} x^5 - 5x^2 \ dx = \left. \tfrac{x^6}{6} - \tfrac{5x^3}{3} \right]_{-1}^{1} = -\tfrac{10}{3}$$

$$< f, x^2 > = \int_{-1}^{1} x^6 - 5x^3 \ dx = \left. \tfrac{x^7}{7} - \tfrac{5x^4}{4} \right]_{-1}^{1} = \tfrac{2}{7}$$

Exemplo

Assim o sistema de equações normais é dado por

$$\begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ -\frac{10}{3} \\ \frac{2}{7} \end{bmatrix}$$

resolvendo com o método de Cholesky, por exemplo, obtemos a seguinte solução

$$c_0 = -\frac{3}{35}, \quad c_1 = -5, \quad c_2 = \frac{6}{7}$$

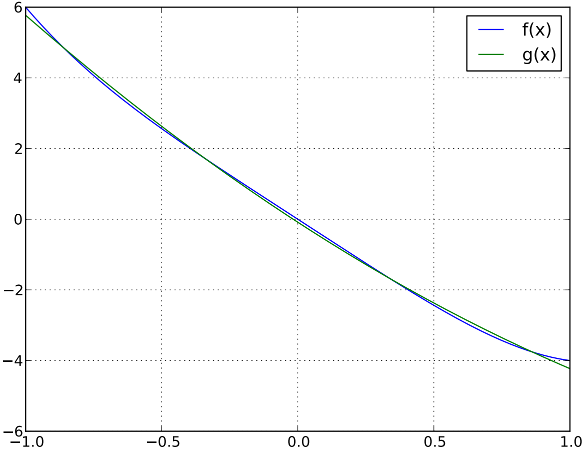
e assim

$$\boxed{g(x) = -\frac{3}{35} - 5x + \frac{6}{7}x^2}$$

Mínimos Quadrados

Caso Contínuo

Exemplo



Quando o produto escalar, no caso discreto ou no contínuo, é igual a zero, dizemos que os vetores (caso discreto) ou funções (caso contínuo) são ortogonais.

Exemplos

Discreto

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$< \mathbf{x}, \mathbf{y} >= \sum_{i=1}^{2} x_i y_i = x_1 y_1 + x_2 y_2 = 1(0) + 0(1) = 0$$

Contínuo

$$f(x) = \sin(x), \quad g(x) = \cos(x)$$

$$< f, g >= \int_{-\pi}^{\pi} \sin(x) \cos(x) \; dx = 0$$

Mínimos Quadrados
Ortogonalidade

Podemos também usar o produto escalar com peso, dado por uma função $w(x)$. Nesse caso o produto escalar é dado por:

- caso contínuo

$$< f, g >= \int_a^b w(x)f(x)g(x)\ dx$$

- caso discreto

$$< f, g >= \sum_{i=1}^{m} w(x_i)f(x_i)g(x_i)$$

e temos também a mesma noção de vetores ou funções ortogonais.

Note que até agora trabalhamos com o seguinte conjunto de funções

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \dots, \quad \phi_n(x) = x^n$$

o qual constitui uma base para o espaço vetorial dos polinômios de grau menor ou igual a n. Observe que se trabalharmos com um conjunto de funções ϕ_j que são ortogonais entre si, isto é, que satisfazem

$$< \phi_i, \phi_j >= 0, \quad \forall i \neq j \tag{5}$$

então o sistema normal tem a seguinte forma (diagonal)

$$\begin{bmatrix} < \phi_0, \phi_0 > & 0 & \dots & 0 \\ 0 & < \phi_1, \phi_1 > & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & < \phi_n, \phi_n > \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} < f, \phi_0 > \\ < f, \phi_1 > \\ \vdots \\ < f, \phi_n > \end{bmatrix}$$

Neste caso a solução é facilmente calculada tomando

$$
\begin{aligned}
&< \phi_0, \phi_0 > c_0 =< f, \phi_0 > \quad \Rightarrow \quad c_0 = \frac{< f, \phi_0 >}{< \phi_0, \phi_0 >} \\
&< \phi_1, \phi_1 > c_1 =< f, \phi_1 > \quad \Rightarrow \quad c_1 = \frac{< f, \phi_1 >}{< \phi_1, \phi_1 >} \\
&\qquad\qquad\vdots \\
&< \phi_n, \phi_n > c_n =< f, \phi_n > \quad \Rightarrow \quad c_n = \frac{< f, \phi_n >}{< \phi_n, \phi_n >}
\end{aligned}
$$

isto é, para $i = 0, 1, \ldots, n$

$$
\boxed{c_i = \frac{< f, \phi_i >}{< \phi_i, \phi_i >}} \tag{6}
$$

Existem várias *famílias* de polinômios ortogonais, e estes podem ser construídos com *diferentes produtos escalares*, dependendo do intervalo, da função peso, ou de ser contínua ou discreta a aplicação.

A seguir veremos 2 exemplos de famílias de polinômios ortogonais:

- Polinômios de Legendre
- Polinômios de Chebyshev

Os polinômios de Legendre são ortogonais no intervalo $[-1, 1]$ com a função peso $w(x) = 1$.

Os primeiros polinômios são dados por

$$
\begin{aligned}
P_0(x) &= 1\\
P_1(x) &= x\\
P_2(x) &= \tfrac{1}{2}(3x^2 - 1)\\
P_3(x) &= \tfrac{1}{2}(5x^3 - 3x)\\
P_4(x) &= \tfrac{1}{8}(35x^4 - 30x^2 + 3)\\
&\dots
\end{aligned}
$$

De forma geral eles podem ser escritos como

$$P_k(x) = \frac{1}{(2^k)k!}\frac{d^k}{dx^k}[(x^2 - 1)]^k, \quad k = 1, \dots$$

Mínimos Quadrados

Polinômios de Legendre

Vamos verificar para alguns casos, que de fato os polinômios de Legendre são ortogonais.

$$< P_1, P_0 > = \int_{-1}^{1} x\ dx = \left[\frac{x^2}{2}\right]_{-1}^{1} = 0$$

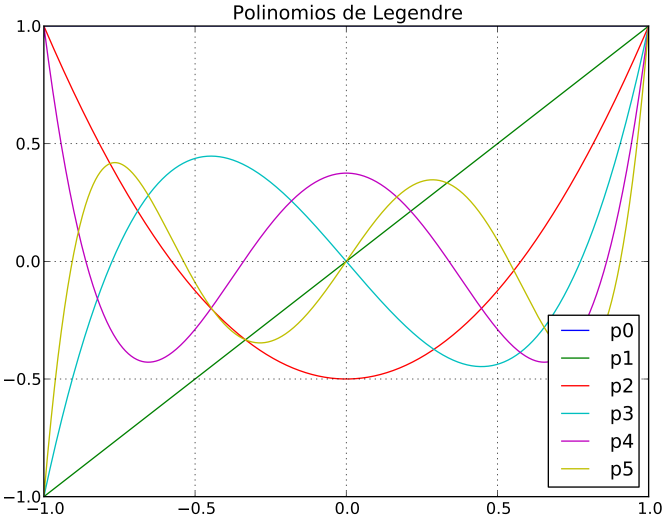
$$< P_1, P_1 > = \int_{-1}^{1} x^2\ dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{2}{3}$$

$$\begin{aligned} < P_1, P_2 > &= \int_{-1}^{1} x\tfrac{1}{2}(3x^2 - 1)\ dx = \tfrac{1}{2}\int_{-1}^{1} 3x^3 - x\ dx \\ &= \tfrac{1}{2}\left[\tfrac{3}{4}x^4 - \tfrac{x^2}{2}\right]_{-1}^{1} = \tfrac{1}{2}\left[\left(\tfrac{3}{4} - \tfrac{1}{2}\right) - \left(\tfrac{3}{4} - \tfrac{1}{2}\right)\right] = 0 \end{aligned}$$

$$< P_1, P_3 > = \int_{-1}^{1} x\tfrac{1}{2}\left(5x^3 - 3x\right)\ dx = \ldots = 0$$

Mínimos Quadrados

Polinômios de Legendre



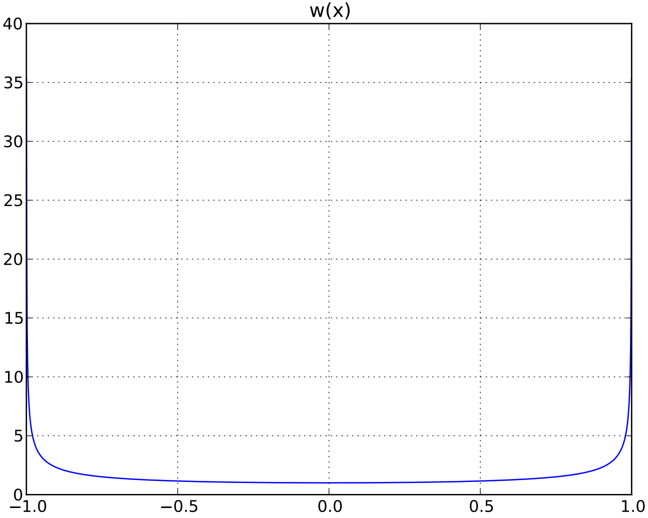
Os polinômios de Chebyshev são ortogonais no intervalo $[-1, 1]$ com a seguinte função peso: $w(x) = \frac{1}{\sqrt{1-x^2}}$.

Os primeiros polinômios são dados por

$$
\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= 2x^2 - 1 \\
P_3(x) &= 4x^3 - 3x \\
P_4(x) &= 8x^4 - 8x^2 + 1 \\
&\dots
\end{aligned}
$$

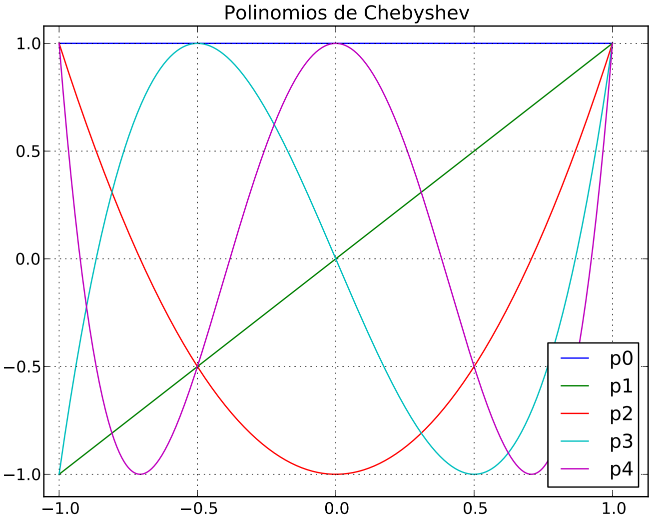
Mínimos Quadrados

Polinômios de Chebyshev



Mínimos Quadrados

Polinômios de Chebyshev



Exemplo

Use o MMQ com os polinômios de Legendre para encontrar a função $g(x) = c_0\phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x)$ considerando que $\phi_0(x) = 1$, $\phi_1(x) = x$ e $\phi_2(x) = \frac{1}{2}(3x^2 - 1)$, que melhor se ajusta à $f(x) = e^x$ no intervalo $[-1, 1]$.

Solução do Exemplo

Como estamos trabalhando com polinômios ortogonais, podemos encontrar os parâmetros c_0, c_1 e c_2 usando

$$c_i = \frac{< f, \phi_i >}{< \phi_i, \phi_i >}$$

Para isso temos que calcular

- $< f, \phi_0 >$, $< f, \phi_1 >$, $< f, \phi_2 >$
- $< \phi_0, \phi_0 >$, $< \phi_1, \phi_1 >$, $< \phi_2, \phi_2 >$

Solução do Exemplo

Calculando

$$< \phi_0, \phi_0 > = \int_{-1}^{1} dx = x]_{-1}^{1} = 1 - (-1) = 2$$

$$< \phi_1, \phi_1 > = \int_{-1}^{1} x^2 \ dx = \left.\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3} - \left(-\frac{1}{3}\right) = \frac{2}{3}$$

$$\begin{aligned} < \phi_2, \phi_2 > &= \int_{-1}^{1} \tfrac{1}{2}(3x^2 - 1)\tfrac{1}{2}(3x^2 - 1) \ dx \\ &= \tfrac{1}{4}\int_{-1}^{1} (9x^4 - 6x^2 + 1) \ dx = \tfrac{1}{4}\left[\tfrac{9x^5}{5} - \tfrac{6x^3}{3} + x\right]_{-1}^{1} \\ &= \tfrac{1}{4}\left[\left(\tfrac{9}{5} - 2 + 1\right) - \left(-\tfrac{9}{5} + 2 - 1\right)\right] \\ &= \ldots = \frac{2}{5} \end{aligned}$$

Solução do Exemplo

Calculando

$$< f, \phi_0 > = \int_{-1}^{1} e^x \, dx = e^x]_{-1}^{1} = e^1 - e^{-1} = 2.35$$

$$< f, \phi_1 > = \int_{-1}^{1} x \, e^x \, dx \quad \Rightarrow \quad \text{IPP}: \boxed{\int udv = uv| - \int v \, du}$$

$$= xe^x \Big|_{-1}^{1} - \int_{-1}^{1} e^x \, dx = xe^x \Big|_{-1}^{1} - e^x \Big|_{-1}^{1}$$

$$= e^x(x-1) \Big|_{-1}^{1} = \left[e^1(1-1) - e^{-1}(-2)\right]$$

$$= 2e^{-1} = 0.736$$

Solução do Exemplo

Calculando

$$
\begin{aligned}
< f, \phi_2 > &= \int_{-1}^{1} e^x \tfrac{1}{2}(3x^2 - 1)\, dx \quad \text{IPP} \quad \Rightarrow \quad u = (3x^2-1),\ dv = e^x \\
&= \frac{1}{2}\left[(3x^2-1)e^x \Big|_{-1}^{1} - 6 \int_{-1}^{1} e^x x \, dx \right] \\
&= \frac{1}{2}\left[(2e^1 - 2e^{-1}) - 6\ (0.736) \right] = \frac{1}{2}\left[2\ (2.35) - 4.414 \right] \\
&= 0.143
\end{aligned}
$$

Solução do Exemplo

Assim temos que os coeficientes são

$$
\begin{aligned}
c_0 &= \frac{<f,\phi_0>}{<\phi_0,\phi_0>} = \frac{2.35}{2} = 1.175 \\
c_1 &= \frac{<f,\phi_1>}{<\phi_1,\phi_1>} = \frac{0.736}{2/3} = 1.104 \\
c_2 &= \frac{<f,\phi_2>}{<\phi_2,\phi_2>} = \frac{0.143}{2/5} = 0.358
\end{aligned}
$$

e portanto

$$
\begin{aligned}
g(x) &= c_0\phi_0 + c_1\phi_1 + c_2\phi_2 \\
&= c_0 + c_1 x + c_2[0.5\ (3x^2 - 1)] \\
&= 1.175 + 1.104\ x + 0.358\ [0.5(3x^2 - 1)]
\end{aligned}
$$

é a função que melhor aproxima $f(x) = e^x$ em $[-1, 1]$ segundo o critério dos mínimos quadrados.

Mínimos Quadrados

Exemplo com polinômios de Legendre

Solução do Exemplo