Método dos Mínimos Quadrados (MMQ)

Domínio Discreto

Nelson Kuhl

IME/USP

8 de setembro de 2020

Reformulando o exemplo

O exemplo da aula anterior pode ser reformulado da seguinte forma, modificando-se também a notação: dadas as alturas y_i medidas nos instantes x_i, $1 \leq i \leq 11$, determinar uma função da forma

$$g(x) = a_0 + a_1 x + a_2 x^2$$

que minimiza o erro quadrático $\sqrt{\sum_{i=1}^{11} [y_i - g(x_i)]^2}$, ou, equivalentemente, determinar a_0, a_1 e a_2 que minimizam o erro quadrático

$$EQ(a_0, a_1, a_2) = \sqrt{\sum_{i=1}^{11} [y_i - a_0 - a_1 x_i - a_2 x_i^2]}$$

Motivados por esta formulação, definiremos agora o problema geral do **método dos mínimos quadrados** (MMQ) para domínios discretos.

MMQ - Domínios discretos

Dada uma tabela

x	x_1	x_2	$\dots$	x_n
$f(x)$	y_1	y_2	$\dots$	y_n

de valores y_i de uma função f em pontos x_i, $1 \leq i \leq n$, determinar

$$g(x) = \sum_{j=0}^{m} a_j g_j(x) \tag{1}$$

que minimiza o erro quadrático $\sqrt{\sum_{i=1}^{n}[y_i - g(x_i)]^2}$, ou, determinar a_j, $0 \leq j \leq m$, que minimizam o erro quadrático

$$EQ(a_0, a_1, \dots, a_m) = \sqrt{\sum_{i=1}^{n}[y_i - \sum_{j=0}^{m} a_j g_j(x_i)]^2} \tag{2}$$

MMQ - Domínios discretos

- A função f pode ser interpretada como a especificação de medidas experimentais. No exemplo do corpo em queda livre, ela é a medida das alturas;

MMQ - Domínios discretos

- A função f pode ser interpretada como a especificação de medidas experimentais. No exemplo do corpo em queda livre, ela é a medida das alturas;
- a função g especifica um modelo que esperamos ajustar-se bem ao experimento. No caso, uma *combinação linear de funções conhecidas* $g_j(x)$, $0 \leq j \leq m$;

MMQ - Domínios discretos

- A função f pode ser interpretada como a especificação de medidas experimentais. No exemplo do corpo em queda livre, ela é a medida das alturas;
- a função g especifica um modelo que esperamos ajustar-se bem ao experimento. No caso, uma *combinação linear de funções conhecidas* $g_j(x)$, $0 \leq j \leq m$;
- no exemplo do corpo em queda livre, g é uma parábola, representada como uma combinação linear de $g_0(x) = 1$, $g_1(x) = x$ e $g_2(x) = x^2$;

MMQ - Domínios discretos

- A função f pode ser interpretada como a especificação de medidas experimentais. No exemplo do corpo em queda livre, ela é a medida das alturas;
- a função g especifica um modelo que esperamos ajustar-se bem ao experimento. No caso, uma *combinação linear de funções conhecidas* $g_j(x)$, $0 \leq j \leq m$;
- no exemplo do corpo em queda livre, g é uma parábola, representada como uma combinação linear de $g_0(x) = 1$, $g_1(x) = x$ e $g_2(x) = x^2$;
- como modelos são imperfeitos e medidas são imprecisas, a função g não se ajustará exatamente aos pontos da tabela. Buscamos então um ajuste segundo um critério de erro adequado. No caso do MMQ, *minimizar o erro quadrático*.

MMQ - Domínios discretos

Trataremos sempre do caso em que o número de medidas é maior do que o número de coeficientes a se determinar, isto é,

$$n > m + 1.$$

É o que ocorre na prática. Assim, o problema (1) satisfazendo (2) é equivalente a resolvermos o sistema linear sobredeterminado

$$\sum_{j=0}^{m} g_j(x_i)a_j = y_i, \quad 1 \leq i \leq n$$

no sentido de mínimos quadrados, e já sabemos como resolvê-lo. A seguir, abordaremos a solução do problema de uma outra forma equivalente, que tem uma interpretação geométrica interessante e permite generalizações usando o mesmo formalismo.

MMQ - Domínios discretos

Se definirmos o **produto interno** entre duas funções u e v, cujos domínios contém os pontos x_i, por

$$\langle u, v \rangle = \sum_{i=1}^{n} u(x_i)v(x_i), \tag{3}$$

então resolver o problema (1)-(2) equivale a determinar os coeficientes da representação de g como combinação linear das funções g_j de forma a **minimizar a distância** $d(f,g)$ entre f e g segundo o produto interno (3), onde

$$d(f,g) := \sqrt{\langle f-g, f-g \rangle} = \sqrt{\sum_{i=1}^{n} [f(x_i) - g(x_i)]^2}$$

MMQ - Domínios discretos

Denote por V o espaço vetorial gerado pelas funções g_j, $0 \leq j \leq m$. Queremos então obter $g \in V$ mais próximo de f segundo a distância definida pelo produto interno (3). Ou seja, queremos a **projeção ortogonal** de f em V segundo este produto interno. Para isto, basta determinar os coeficientes de g de forma que $f - g$ seja ortogonal a cada gerador de V:

$$\langle f - g, g_i \rangle = 0, \quad 0 \leq i \leq m,$$

MMQ - Domínios discretos

Denote por V o espaço vetorial gerado pelas funções g_j, $0 \leq j \leq m$. Queremos então obter $g \in V$ mais próximo de f segundo a distância definida pelo produto interno (3). Ou seja, queremos a **projeção ortogonal** de f em V segundo este produto interno. Para isto, basta determinar os coeficientes de g de forma que $f - g$ seja ortogonal a cada gerador de V:

$$\langle f - g, g_i \rangle = 0, \quad 0 \leq i \leq m,$$

ou

$$\langle f - \sum_{j=0}^{m} a_j g_j, g_i \rangle = 0, \quad 0 \leq i \leq m.$$

MMQ - Domínios discretos

Usando a **linearidade** em relação à primeira componente e a **simetria** do produto interno, podemos reescrever as equações anteriores como o sistema linear de $m+1$ equações e $m+1$ incógnitas

$$\sum_{j=0}^{m} \langle g_i, g_j \rangle a_j = \langle g_i, f \rangle, \quad 0 \leq i \leq m, \tag{4}$$

onde

$$\langle g_i, g_j \rangle = \sum_{k=1}^{n} g_i(x_k) g_j(x_k),\ 0 \leq i, j \leq m, \tag{5}$$

$$\langle g_i, f \rangle = \sum_{k=1}^{n} g_i(x_k) y_k,\ 0 \leq i \leq m, \tag{6}$$

chamado de **Sistema Normal**.

MMQ - Domínios discretos

Matricialmente, podemos escrever o sistema normal como

$$\begin{bmatrix} \langle g_0, g_0 \rangle & \langle g_0, g_1 \rangle & \cdots & \langle g_0, g_m \rangle \\ \langle g_1, g_0 \rangle & \langle g_1, g_1 \rangle & \cdots & \langle g_1, g_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle g_m, g_0 \rangle & \langle g_m, g_1 \rangle & \cdots & \langle g_m, g_m \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \langle g_0, f \rangle \\ \langle g_1, f \rangle \\ \vdots \\ \langle g_m, f \rangle \end{bmatrix}, \tag{7}$$

onde os coeficientes da matriz e as componentes do lado direito são calculados por (5) e (6), respectivamente.

MMQ - Domínios discretos

Matricialmente, podemos escrever o sistema normal como

$$\begin{bmatrix} \langle g_0, g_0 \rangle & \langle g_0, g_1 \rangle & \dots & \langle g_0, g_m \rangle \\ \langle g_1, g_0 \rangle & \langle g_1, g_1 \rangle & \dots & \langle g_1, g_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle g_m, g_0 \rangle & \langle g_m, g_1 \rangle & \dots & \langle g_m, g_m \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \langle g_0, f \rangle \\ \langle g_1, f \rangle \\ \vdots \\ \langle g_m, f \rangle \end{bmatrix}, \qquad (7)$$

onde os coeficientes da matriz e as componentes do lado direito são calculados por (5) e (6), respectivamente.

- Este sistema linear é **idêntico** ao sistema normal que obteríamos da formulação do problema como um sistema linear sobredeterminado;

MMQ - Domínios discretos

Matricialmente, podemos escrever o sistema normal como

$$\begin{bmatrix} \langle g_0, g_0\rangle & \langle g_0, g_1\rangle & \cdots & \langle g_0, g_m\rangle \\ \langle g_1, g_0\rangle & \langle g_1, g_1\rangle & \cdots & \langle g_1, g_m\rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle g_m, g_0\rangle & \langle g_m, g_1\rangle & \cdots & \langle g_m, g_m\rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \langle g_0, f\rangle \\ \langle g_1, f\rangle \\ \vdots \\ \langle g_m, f\rangle \end{bmatrix}, \tag{7}$$

onde os coeficientes da matriz e as componentes do lado direito são calculados por (5) e (6), respectivamente.

- Este sistema linear é **idêntico** ao sistema normal que obteríamos da formulação do problema como um sistema linear sobredeterminado;
- mas esta abordagem nos permitirá formular um problema de mínimos quadrados abstrato em espaços vetoriais gerais munidos de produtos internos gerais, levando-nos a aplicações nas quais o formalismo é o mesmo. Isto será apresentado nas próximas aulas.

Regressão linear

O MMQ é usado com frequência no problema da *regressão linear*, que consiste em ajustar uma tabela

x	x_1	x_2	$\dots$	x_n
$f(x)$	y_1	y_2	$\dots$	y_n

por uma reta, ou seja, por um poliômio de grau menor ou igual a 1

$$g(x) = a_0 + a_1 x.$$

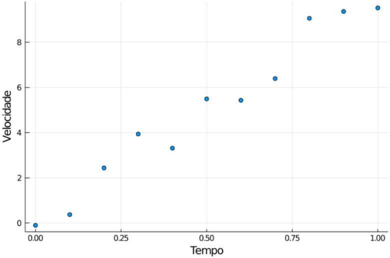
Definido-se $g_0(x) = 1$ e $g_1(x) = x$, concluimos de (5) e (6) que o sistema normal (7) é

$$\begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}. \tag{8}$$

Regressão linear

Exemplo[1] Em um experimento, uma bola é lançada verticalmente, e sua velocidade é medida em função do tempo de acordo com a tabela

Tempo	Velocidade
0.0	−0.10290
0.1	0.37364
0.2	2.43748
0.3	3.93836
0.4	3.31230
0.5	5.49472
0.6	5.43325
0.7	5.43325
0.8	9.06048
0.9	9.36416
1.0	9.52066



[1] ver DeVries, *A First Course in Computational Physics*

Regressão linear

Desprezando-se a resistência do ar, sabemos da cinemática que a velocidade como função do tempo é uma reta

$$v(t) = a_0 + a_1 t.$$

Logo, se formos ajustar os dados pelo MMQ, temos um problema de regressão linear. O sistema linear (8) neste caso é

$$\begin{bmatrix} 11.0 & 5.5 \\ 5.5 & 3.85 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 55.22536 \\ 38.710633 \end{bmatrix},$$

cuja solução é $a_0 = -0.024037$ e $a_1 = 10.089048$. Na próxima figura temos o gráfico dos pontos juntamente com a reta $v(t)$.

Regressão linear