Métodos dos quadrados mínimos no caso geral discreto linear

MS211 – Cálculo Numérico

Giuseppe Romanazzi

26 Novembro 2020

Conteúdo

Problemas dos quadrados mínimos

Analisaremos nesta aula os problemas de quadrados mínimos (PQM) no caso geral discreto linear. Resolveremos dois exercícios. Verificaremos que para resolver sistemas lineares sobredeterminados (com mais equações que incógnitas $m > n$) podemos usar estes métodos de quadrados mínimos, que levará a usar as equações normais.

Problemas dos quadrados mínimos no caso geral discreto linear

Suponhamos de ter muitas observações (x_k, y_k), $k = 1, \ldots, m$ (com m grande) do problema analisado, que sabemos que satisfaz uma lei linear $\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_n g_n(x)$ onde as n funções base $g_j(x)$ são conhecidas e vale $n < m$.
Então o problema de encontrar os "melhores" α_i parâmetros (ou coeficientes) do problema analisado tais que

$$\varphi(x_k) = \sum_{j=1}^{n} \alpha_j g_j(x_k) \approx y_k$$

é dito problema dos quadrados mínimos no caso discreto linear e é resolvido com um método chamado método dos quadrados mínimos discreto que generaliza o método da reta dos quadrados mínimos.

Método dos quadrados mínimos no caso discreto linear

Iremos minimizar as distâncias $d_k = y_k - \varphi(x_k)$ onde $\varphi(x) = \sum_{j=1}^{n} \alpha_j g_j(x)$, que é equivalente a minimizar os quadrados delas, e para obter isso para todas as distâncias uma vez só minimizamos a soma destas distancias ao quadrado

$$F(\alpha_1, \ldots, \alpha_n) = \sum_{k=1}^{m} d_k^2 = \sum_{k=1}^{m} (y_k - \varphi(x_k))^2,$$

então vamos encontrar os n coeficientes $\tilde{\alpha}_j$ tais que

$$F(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n) \leq F(\alpha_1, \ldots, \alpha_n)$$

onde

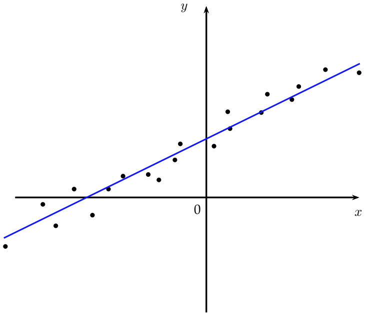
$$F(\alpha_1, \ldots, \alpha_n) = \sum_{k=1}^{m} (y_k - \varphi(x_k))^2 = \sum_{k=1}^{m} \left(y_k - \sum_{j=1}^{n} \alpha_j g_j(x_k) \right)^2$$

Como escolher as funções base $g_j(x)$?

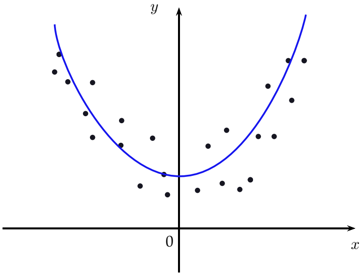
Dois casos são possíveis:

1. a função $\varphi(x)$ que descreve o problema é conhecida e se desconhecem os seus parâmetros α_j. Este é o caso onde as funções bases $g_j(x)$ são dadas ou já se conhecem
2. Outro caso é quando se desconhecem também as funções bases. Neste caso observando a distribuição dos dados (x_k, y_k) disponíveis com $k = 1, \ldots, m$ poderemos achar as funções bases $g_j(x)$

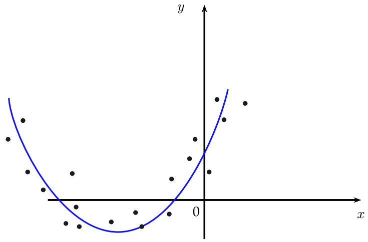
Exemplos: observando a distribuição dos dados, achamos as funções bases $g_j(x)$



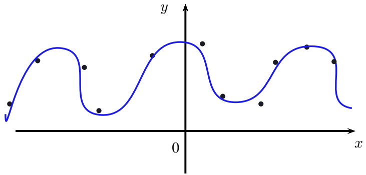
$g_1(x) = 1, g_2(x) = x$
$\varphi(x) = \alpha_1 + \alpha_2 x =$
$\alpha_1 g_1(x) + \alpha_2 g_2(x)$



$g_1(x) = x^2$, $g_2(x) = 1$
$\varphi(x) = \alpha_1 x^2 + \alpha_2 =$
$\alpha_1 g_1(x) + \alpha_2 g_2(x)$



$g_1(x) = 1, g_2(x) = x, \ g_3(x) = x^2$
$\varphi(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2$



$g_1(x) = 1,$
$g_2(x) = \sin(x), g_3(x) = \cos(x)$
$\varphi(x) = \alpha_1 + \alpha_2 \sin(x) + \alpha_3 \cos(x)$

Método dos quadrados mínimos no caso geral discreto linear

Temos de minimizar

$$F(\alpha_1, \ldots, \alpha_n) = \sum_{k=1}^{m} (y_k - \varphi(x_k))^2 = \sum_{k=1}^{m} (y_k - \sum_{j=1}^{n} \alpha_j g_j(x_k))^2$$

Os $\{\alpha_j\}_{j=1,\ldots,n}$ que minimizam F serão aqueles para que :

$$\frac{\partial F}{\partial \alpha_i}(\alpha_1, \ldots, \alpha_n) = 0 \qquad \text{para cada } i = 1, \ldots, n$$

Note que conseguimos chegar a um sistema de n equações $(\dfrac{\partial F}{\partial \alpha_i} = 0)$ nas n incógnitas α_j. Este é um sistema linear, como é descrito na slide sucessiva.

Determinamos os coeficientes α_j

Para cada $i = 1, \dots, n$

$$\frac{\partial F}{\partial \alpha_i}(\alpha_1, \dots, \alpha_n) = 2\sum_{k=1}^{m}(y_k - \sum_{j=1}^{n}\alpha_j g_j(x_k)) \cdot g_i(x_k) = 0$$

Multiplicando por $\frac{1}{2}$ obtemos para cada $i = 1, \dots, n$

$$\sum_{j=1}^{n}\sum_{k=1}^{m} g_i(x_k)g_j(x_k)\alpha_j = \sum_{k=1}^{m} y_k g_i(x_k)$$

Este resulta ser a linha i do sistema linear

$$A\alpha = b$$

onde $A = (a_{ij})_{i,j=1,\dots,n}$, $b = (b_1, \dots, b_n)^t$ e $\alpha = (\alpha_1, \dots, \alpha_n)^t$ com

$$a_{ij} = \bar{g}_i^t \bar{g}_j = \sum_{k=1}^{m} g_i(x_k)g_j(x_k), \quad b_i = \bar{y}^t \bar{g}_i, \quad i, j = 1, \dots, n$$

onde usamos os vetores $\bar{g}_i = (g_i(x_1), \dots, g_i(x_m))^t \in \mathbb{R}^m$ e $\bar{y} = (y_1, \dots, y_m)^t \in \mathbb{R}^m$.

Com as muitas observações do problema (x_k, y_k) com $k = 1, \ldots, m$, com m grande $(m > n)$, encontramos as (melhores, no sentido dos quadrados mínimos) α_i do modelo $\varphi(x) = \sum_{i=1}^{n} \alpha_i g_i(x)$ resolvendo o sistema linear de dimensão n (pequena) $A\alpha = b$.
O problema da reta dos quadrados mínimos é um caso particular do problema dos quadrados mínimos no caso discreto linear: porque usando $n = 2$ e $g_1(x) = 1$, $g_2(x) = x$ o problema resulta no escolher os melhores α_1 e α_2 que minimizam a soma dos quadrados das distâncias $F(\alpha_1, \alpha_2) = \sum_{k=1}^{m} (y_k - (\alpha_1 + \alpha_2 x))^2$ o sistema resultante $A\alpha = b$ tem

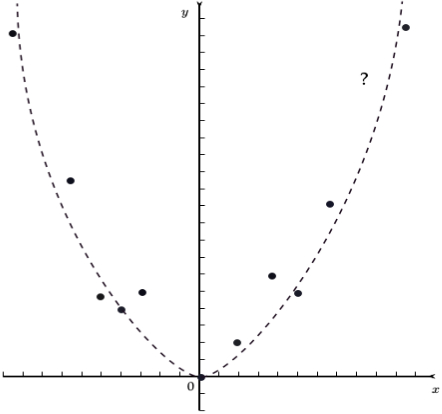
$$A = \begin{pmatrix} \bar{g}_1^t \bar{g}_1 & \bar{g}_1^t \bar{g}_2 \\ \bar{g}_2^t \bar{g}_1 & \bar{g}_2^t \bar{g}_2 \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{k=1}^{m} 1 & \displaystyle\sum_{k=1}^{m} 1 \cdot x_k \\ \displaystyle\sum_{k=1}^{m} 1 \cdot x_k & \displaystyle\sum_{k=1}^{m} x_k^2 \end{pmatrix} = \begin{pmatrix} m & \displaystyle\sum_{k=1}^{m} x_k \\ \displaystyle\sum_{k=1}^{m} x_k & \displaystyle\sum_{k=1}^{m} x_k^2 \end{pmatrix}$$

$$b = \begin{pmatrix} \bar{g}_1^t y \\ \bar{g}_2^t y \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{k=1}^{m} 1 \cdot y_k \\ \displaystyle\sum_{k=1}^{m} x_k y_k \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{k=1}^{m} y_k \\ \displaystyle\sum_{k=1}^{m} x_k y_k \end{pmatrix}$$

Exemplo

Determinamos a melhor α para que $\varphi(x) = \alpha x^2$ passe perto dos seguintes m pontos (x_k, y_k)

x_k	-1.0	-0.75	-0.6	-0.5	-0.3	0	0.2	0.4	0.5	0.7	1
y_k	2.05	1.153	0.45	0.4	0.5	0	0.2	0.6	0.512	1.2	2.05



Observamos que temos $m = 11$ pontos (x_k, y_k), e somente uma incógnita $\alpha_1 = \alpha$. Portanto se $n = 1$ e o sistema do método dos quadrados mínimos $A\alpha = b$ é de dimensão 1, é uma equação linear, sendo que $\varphi(x) = \alpha x^2$ então $g_1(x) = x^2$ assim podemos rescrever $\varphi(x)$ como $\varphi(x) = \alpha_1 g_1(x^2)$, e do método dos quadrados mínimos obtemos que α tem de satisfazer a equação $\bar{g}_1^t \bar{g}_1 \alpha = \bar{g}_1^t y$ ou seja

$$\sum_{k=1}^{11} g_1(x_k)^2 \alpha = \sum_{k=1}^{m} g_1(x_k) y_k.$$

Sendo $g_1(x_k) = x_k^2$ temos que $\displaystyle\sum_{k=1}^{11} g_1(x_k)^2 = \sum_{k=1}^{11} x_k^4 = 2.8464$

$\displaystyle\sum_{k=1}^{11} g_1(x_k) y_k = \sum_{k=1}^{11} x_k^2 y_k = 5.8765$ e portanto

$$\alpha = \frac{5.8765}{2.8464} = 2.0642.$$

Segundo Exemplo

Suponhamos agora de querer melhorar a função modelo $\varphi(x)$ que descreve o problema observado e descrito dos pontos (x_k, y_k). Por exemplo podemos achar um modelo com mais um parâmetro $\varphi(x) = \alpha_1 + \alpha_2 x^2$. Notamos neste caso que $\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x)$, com $g_1(x) = 1$ e $g_2(x) = x^2$. Assim temos $n = 2$ (número das incógnitas) e $m = 11$ (números dos dados). Aplicando o PQM no caso discreto temos de resolver um sistema linear de dimensão 2

$$\begin{pmatrix} \bar{g}_1^t \bar{g}_1 & \bar{g}_1^t \bar{g}_2 \\ \bar{g}_2^t \bar{g}_1 & \bar{g}_2^t \bar{g}_2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \bar{g}_1^t y \\ \bar{g}_2^t y \end{pmatrix}$$

onde $\bar{g}_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = (1, \ldots, 1)^t \in \mathbb{R}^{11}$,
$\bar{g}_2 = (g_2(x_1), \ldots, g_2(x_{11}))^t = (x_1^2, \ldots, x_{11}^2) \in \mathbb{R}^{11}$,
$y = (y_1, \ldots, y_{11})^t \in \mathbb{R}^{11}$.

$$\begin{pmatrix} \bar{g}_1^t\bar{g}_1 & \bar{g}_1^t\bar{g}_2 \\ \bar{g}_2^t\bar{g}_1 & \bar{g}_2^t\bar{g}_2 \end{pmatrix} = \begin{pmatrix} 11 & \displaystyle\sum_{k=1}^{11} x_k^2 \\ \displaystyle\sum_{k=1}^{11} x_k^2 & \displaystyle\sum_{k=1}^{11} x_k^4 \end{pmatrix} = \begin{pmatrix} 11 & 4.2025 \\ 4.2025 & 2.8464 \end{pmatrix}$$

$$\begin{pmatrix} \bar{g}_1^t y \\ \bar{g}_2^t y \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{k=1}^{11} y_k \\ \displaystyle\sum_{k=1}^{11} x_k^2 y_k \end{pmatrix} = \begin{pmatrix} 9.1150 \\ 5.8756 \end{pmatrix}$$

Então o sistema para resolver é

$$\begin{pmatrix} 11 & 4.2025 \\ 4.2025 & 2.8464 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 9.1150 \\ 5.8756 \end{pmatrix}$$

que dá como solução

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0.0915 \\ 1.9287 \end{pmatrix}$$

A melhor parábola do tipo $\varphi(x) = a + bx^2$ é $\varphi(x) = 0.0915 + 1.9287x^2$.

Resolução de sistemas lineares sobredeterminados

Suponhamos de ter um problema onde queremos achar n incógnitas $x_1, \dots, x_n$ que satisfazem o sistema linear de $m > n$ equações $Ax = b$. Sabemos, ver aula 7 (Sistemas Lineares) slide 17, que este sistema admite uma solução única se $posto(A) = posto(A|b) = n$ e este acontece muito raramente! Porque tem de ter no máximo n equações linearmente independentes.

Por exemplo $\left\{ \begin{array}{lcl} 2x_1 + 4x_2 & = & 1 \\ x_1 - x_2 & = & -1 \\ 2x_1 + x_2 & = & 2 \\ x_1 + 3x_2 & = & 4 \end{array} \right.$ tem 3 equações linearmente independentes: $posto(A|b) = 3 > n = 2$, $posto(A) = 2$ então não admite solução. Mas será que podemos achar a melhor aproximação de x_1, x_2 que satisfaz o sistema? A resposta é sim, porque achar as melhores x_1, x_2 pode ser visto como um problema dos quadrados mínimos no caso discreto linear.

Agora as incógnitas são x_1, x_2 , então $n = 2$ e o numero de informações que conhecemos são $m = 4$. Queremos minimizar a soma dos quadrados das distâncias
$F(x_1, x_2) = (1 - (2x_1 + 4x_2))^2 + (-1 - (x_1 + x_2))^2 + (2 - (2x_1 + x_2))^2 + (4 - (x_1 + 3x_2))^2) = \|b - Ax\|^2$. Onde $\| \cdot \|$ é a norma euclideana, então $\|z\|^2 := \sum_{k=1}^{n} z_k^2$.
No caso geral de um sistema $Ax = b$ de mais equações que incógnitas que não admite única solução podemos achar a melhor solução possível minimizando

$$F(x_1, \ldots, x_n) = \|b - Ax\|^2.$$

Note que no caso que o sistema admita única solução x (que acontece quando $posto(A|b) = posto(A) = n$) então a mesma pode ser achada resolvendo o sistema quadrado com as n equações linearmente independentes ou resolvendo o problema dos quadrados mínimos que admitirá o mínimo valor 0. Isso porque $F(x) = 0 \iff Ax = b$.

Minimizar $F(x_1, \dots, x_n) = \|b - Ax\|^2$ equivale a resolver a equação normal $A^t Ax = A^t b$

Observamos que minimizar

$$F(x_1, \dots, x_n) = \|b - Ax\|^2 = \sum_{k=1}^{m} (b_k - \sum_{j=1}^{n} a_{kj} x_j)^2$$

é equivalente a minimizar a soma das distâncias ao quadrado
$\displaystyle\sum_{k=1}^{m} (b_k - \sum_{j=1}^{n} x_j g_j(z_k))^2$ onde $g_j(z_k) = a_{kj}$ para cada $k = 1, \dots, m$ e $j = 1, \dots, n$.
O resultante sistema linear é

$$\begin{pmatrix} \sum_{k=1}^{m} g_1(z_k)^2 & \sum_{k=1}^{m} g_2(z_k)g_1(z_k) & \dots & \sum_{k=1}^{m} g_n(z_k)g_1(z_k) \\ \sum_{k=1}^{m} g_2(z_k)g_1(z_k) & \sum_{k=1}^{m} g_2(z_k)^2 & \dots & \sum_{k=1}^{m} g_n(z_k)g_2(z_k) \\ \vdots & \vdots & \dots & \vdots \\ \sum_{k=1}^{m} g_n(z_k)g_1(z_k) & \sum_{k=1}^{m} g_n(z_k)g_2(z_k) & \dots & \sum_{k=1}^{m} g_n(z_k)^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} g_1(z_k)b_k \\ \sum_{k=1}^{m} g_2(z_k)b_k \\ \vdots \\ \sum_{k=1}^{m} g_n(z_k)b_k \end{pmatrix}$$

Note que, sendo $g_j(z_k) = a_{kj}$, onde a_{kj} são os coeficientes da matriz $A \in \mathbb{R}^{m \times n}$ do problema sobredeterminado original $Ax = b$, temos que o sistema para resolver pode ser também escrito como

$$A^t A x = A^t b$$

que é muito mais simples para lembrar. Este último sistema chama-se equação normal, que então permite de encontrar a x que minimiza a soma ao quadrado das distâncias $F(x) = \|b - Ax\|^2$ que pode ser usado para achar uma boa x para que $Ax \approx b$.
Note que o sistema $A^t A x = A^t b$ é um sistema linear simétrico de dimensão n.

Resolução de $Ax = b$ sobredeterminado com $m > n$ através a equação normal

$$\begin{cases} 2x_1 + 4x_2 & = & 1 \\ x_1 - x_2 & = & -1 \\ 2x_1 + x_2 & = & 2 \\ x_1 + 3x_2 & = & 4 \end{cases}$$

Sabemos já que não existe a solução deste sistema (porque $posto(A|b) = 3 > posto(A) = 2$). Usando o método dos quadrados mínimos discretos ou analogamente a equação normal podemos achar a melhor x que minimiza a distancia $F(x) = \|b - Ax\|^2$ ousseja temos de resolver $A^t Ax = A^t b$ com

$$A = \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} \text{ e } b = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 4 \end{pmatrix}$$

e então $A^tA = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 4 & -1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 12 \\ 12 & 27 \end{pmatrix}$ e
$A^tb = \begin{pmatrix} 9 \\ 19 \end{pmatrix}$ e assim a solução será $x = \begin{pmatrix} 0.1190 \\ 0.6508 \end{pmatrix}$.