Problemas e métodos dos quadrados mínimos - Reta dos quadrados mínimos

MS211 – Cálculo Numérico

Giuseppe Romanazzi

24 Novembro 2020

Conteúdo

Problemas dos quadrados mínimos

Os problemas dos quadrados mínimos (PQM) são referidos a problemas onde as informações disponíveis são em número maior respeito as incógnitas do problema.
Estes incógnitas como sabemos são normalmente algum parâmetro incógnito necessário para descrever através o modelo o problema analisado.
Estudaremos dois diferentes PQM : discretos(aulas 20-21) e contínuos (aula 22). Estas duas classes de problemas podem ser subdividas em duas subclasses: problemas lineares e não lineares (ver aula 22).
Vamos resolver no fim desta aula um caso discreto particular chamado problema da reta dos quadrados mínimos.

Problemas dos quadrados mínimos discretos lineares

Os problemas discretos lineares consistem no encontrar os parâmetros α_i da função $\varphi(x)$ que descreve o problema e definida com uma combinação linear de n funções bases $g_i(x)$:

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_n g_n(x)$$

As funções $g_i(x)$ são conhecidas e dependem do problema analisado. Estes parâmetros são encontrados utilizando muitas informações (observações) discretas y_i num número m de pontos x_i (espaciais, temporais, etc.) x_i.
Falaremos do problema dos quadrados mínimos (PQM) quando o problema discreto linear tem um número m de observações $(x_i, y_i)_{i=1,\ldots,m}$ maior do número n dos parâmetros incógnitos $(\alpha_i)_{i=1,\ldots,n}$, ou seja quando

$$m > n$$

Caso $m = n$ é equivalente a resolver um sistema linear

Note que se for $m = n$ com $g_j(x)$ funções lineares (em x) o problema será equivalente a resolver um sistema linear onde a $i-$sima equação linear é

$$\sum_{j=1}^{n} g_j(x_i)\alpha_j = y_i;$$

e onde as incógnitas são as α_i e a matriz $A \in \mathbb{R}^{n\times n}$ de n linhas e colunas dos coeficientes é

$$A = (a_{ij}) = (g_j(x_i))_{i,j=1,\ldots,n};$$

o vetor b de termos independentes com coeficientes $b_i = y_i$, com $i = 1, \ldots, n$. Então é como se temos de resolver o sistema linear de dimensão $m = n$

$$A\alpha = b$$

onde o vetor das incógnitas é $\alpha = (\alpha_1, \ldots, \alpha_n)^t \in \mathbb{R}^n$.

Exemplos

Existem muitos fenômenos, eventos, onde é necessários resolver um problema dos quadrados mínimos linear discreto.
Cada vez que temos fenômenos descritos de modelos lineares e temos muitas informações ou observações do fenômeno teremos de resolver um PQM linear discreto.
Por exemplo considere um circuito elétrico onde a (diferença de) voltagem V medida em Volts é dada da seguinte soma de produtos de varias resistência R_i medida em Ohm para funções que dependem da correntes elétricas I medida em Ampere: $V = R_1 I + 2R_2 \sin(I + 1) - R_3$
Se queremos achar as resistência R_1, R_2, R_3 do circuito usando os vários valores medidos (I_i, V_i), com $i = 1, \ldots, m$, e $m > 3$, por exemplo $m = 100$, de voltagens associadas a diferentes intensidades de correntes, (ou vice-versa de correntes associadas a diferentes voltagens) estaremos tratando um problema dos quadrados mínimos linear e discreto de $n = 3$ incógnitas e com m um número grande de dados observados ou medidos.

Sabe se que o cumprimento de uma planta que no dia t_0 da plantação tem um cumprimento L_0 e que se fornece calor e água suficiente tem um crescimento quadrático no tempo de taxa constante $r > 0$, ou seja o cumprimento $L(t)$ no tempo $t \geq t_0$ satisfaz

$$L(t) = L_0 + r \cdot t^2,$$

após t dias da plantação; Quer-se encontrar a taxa r de uma planta que foi plantada no dia 1 de Março ($t_0 = 61$ dias do inicio do ano) sabendo que o seu cumprimento inicial foi $L_0 = 10cm$ e nos dias seguintes do ano teve as seguintes medidas

$t(dias)$	$L(cm)$
70	25
80	55
90	95
100	130

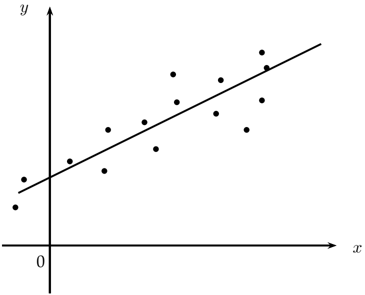
Qual'e' a taxa r? o problema tem como incógnita somente $\alpha = r$, então $n = 1$; A função $\varphi(t) = \sum_{i=1}^{n} \alpha_i g_i(t)$ que descreve o problema é $\varphi(t) = g_1(t) + \alpha g_2(t) = 10 + \alpha t^2$ então temos o numero de incógnitas é $n = 1$, e o número de observações $(x_i, y_i) = (t_i, L_i)$ são $m = 4$. Sendo $m > n$, podemos encontrar r resolvendo um problema dos quadrados mínimos.

Se L_0 não foi conhecido o problema tem duas incógnitas $(n = 2)$ L_0 e α e com $g_1(t) = 1$ e $g_2(t) = t^2$. Se conhessemos L_0 em vez o problema reduz-se no achar a única incógnita .

Vamos tratar a seguir um caso particular dos problemas dos quadrados mínimos chamado reta dos quadrados mínimos.

Reta dos quadrados mínimos, Regressão linear

Queremos determinar a melhor reta $y = ax + b$ que passa perto dos m pontos (x_i, y_i) com $i = 1, \ldots, m$.



Este problema equivale a determinar os melhores parâmetros a, b da reta que passa perto dos dadps pontos

Reta dos quadrados mínimos, Regressão linear

Há varias estratégias para fazer isso, uma destas é de achar a reta dos mínimos quadrados que será aquela para que a soma das distancias ao quadrado dos pontos da reta procurada seja a menor possível. Esta reta será portanto chamada **reta dos quadrados mínimos**. Claramente, se os pontos fossem distribuídos já sobre uma reta então esta será a reta dos quadrados mínimos.

A reta procurada tem equação $y = a + bx$, o nosso problema consiste então no achar a e b ($n = 2$) usando os m pontos (x_i, y_i).
Queremos minimizar as somas das distância $y_i - (a + bx_i)$ ao quadrado $[y_i - (a + bx_i)]^2$, portanto vamos encontrar $\tilde{a}, \tilde{b}$ tais que posto

$$F(a, b) := \sum_{i=1}^{m} [y_i - (a + bx_i)]^2$$

$F(\tilde{a}, \tilde{b})$ toma o menor valor de F ao variar dos a, b:

$$F(\tilde{a}, \tilde{b}) = \min_{a,b} F(a, b)$$

Achar a reta dos quadrados mínimos

Sendo que a dupla $(\tilde{a}, \tilde{b})$ tem de minimizar F então é um ponto critico de F e portanto tem de satisfazer $\dfrac{\partial F}{\partial a}(\tilde{a}, \tilde{b}) = 0, \quad \dfrac{\partial F}{\partial b}(\tilde{a}, \tilde{b}) = 0,$ temos portanto duas equações e duas incognitas. Desenvolvendo as derivadas

$$\frac{\partial F}{\partial a}(\tilde{a}, \tilde{b}) = 2\sum_{i=1}^{m}(y_i - (\tilde{a} + \tilde{b}x_i)) \cdot 1 = 0 \iff m\tilde{a} + (\sum_{i=1}^{m} x_i)\tilde{b} = \sum_{i=1}^{m} y_i$$

$$\frac{\partial F}{\partial b} = 2\sum_{i=1}^{m}(y_i - (\tilde{a} + \tilde{b}x_i)) \cdot x_i = 0 \iff \sum_{i=1}^{m} x_i\tilde{a} + (\sum_{i=1}^{m} x_i^2)\tilde{b} = \sum_{i=1}^{m} y_i x_i$$

Ou seja $\tilde{a}, \tilde{b}$ são solução do sistema linear (sistema linear da reta dos quadrados mínimos)

$$\begin{pmatrix} m & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \end{pmatrix}$$

Usamos uma notação vetorial, para ver os coeficientes do sistema numa maneira mais compacta. Definimos os vetores em $\mathbb{R}^m$: $e = (1, \ldots, 1)^t$, $x = (x_1, \ldots, x_m)^t$, $y = (y_1, \ldots, y_m)^t$ então teremos o sistema

$$\begin{pmatrix} m & e^t x \\ e^t x & x^t x \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e^t y \\ x^t y \end{pmatrix}$$

Resolvemos ele usando o método de Cramer. O determinante $|A|$ da matriz A é $|A| = m x^t x - (e^t x)^2$ e os coeficiente serão então

$$a = \frac{\begin{vmatrix} e^t y & e^t x \\ y^t x & x^t x \end{vmatrix}}{|A|} \quad b = \frac{\begin{vmatrix} m & e^t y \\ e^t x & y^t x \end{vmatrix}}{|A|}$$

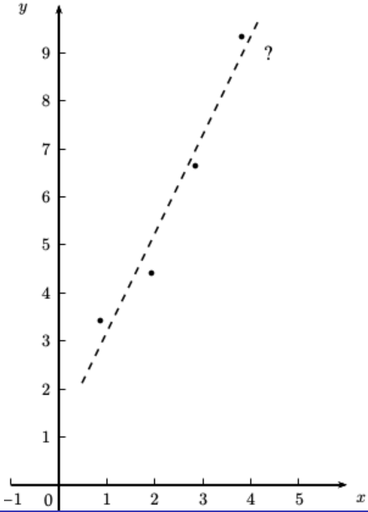
ou seja, sendo $\|x\| = \sqrt{\sum_{i=1}^{m} x_i^2} = \sqrt{x^t x}$ temos

$$a = \frac{e^t y \|x\|^2 - (e^t x)(y^t x)}{|A|}, \quad b = \frac{m y^t x - (e^t x)(e^t y)}{|A|}.$$

Exemplo

Encontrar a reta (dos quadrados mínimos) que passa mais próxima aos $m = 4$ dados conhecidos (x_i, y_i) da tabela

x	y
1	3.5
2	4.5
3	6.8
4	9.2



A reta terá equação $y = a + bx$ com

$a = \dfrac{e^t y\|x\|^2 - (e^t x)(y^t x)}{m\|x\|^2 - (e^t x)^2} \quad b = \dfrac{m y^t x - (e^t x)(e^t y)}{m\|x\|^2 - (e^t x)^2}$ Observamos que

$$\|x\|^2 = x^t x = \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} i^2 = \frac{m(m+1)(2m+1)}{6} = \frac{4 \cdot 5 \cdot 9}{6} = 30$$

$y^t x = \sum_{i=1}^{4} y_i x_i = 3.5 + 9 + 20.4 + 38.8 = 69.7$

$e^t x = \displaystyle\sum_{i=1}^{4} x_i = 10, \quad e^t y = \sum_{i=1}^{4} y_i = 24$ e o determinante da matriz é

então $|A| = m\|x\|^2 - (e^t x)^2 = 4 \cdot 30 - 10^2 = 120 - 100 = 20.$ Então temos

$$a = \frac{24 \cdot 30 - 10 \cdot 69.7}{20} = \frac{23}{20} = 1.15$$

é a interseção da reta dos quadrados mínimos com o eixo das y.

$$b = \frac{4 \cdot 69.7 - 10 \cdot 24}{20} = \frac{278.8 - 240}{20} = \frac{38.8}{20} = 1.94 \approx 2$$

é a inclinação da reta

A reta dos quadrados mínimos tem equação $y = 1.15 + 1.94x$