Ajuste de curvas por quadrados mínimos lineares

Felipe Leonardo de Aguiar e Wanderley Innocêncio Moreira Júnior
Engenharia de Minas – 1°. Período
Professor: Rodney Josué Biezuner
Disciplina: Geometria Analítica e Álgebra Linear

1. Introdução

Utilizamos este método quando temos uma distribuição de pontos e queremos ajustar a melhor curva a este conjunto de dados.

Inicialmente, vamos analisar o caso em que a curva de ajuste é uma função linear:

$$y_i = a + bx_i$$

Para que esta seja a reta que melhor se ajusta aos dados, devemos minimizar a soma das diferenças entre os valores de f(x) tabelados y_i e os valores da curva de ajuste $a+bx_i$ em cada ponto. Mas esta diferença pode ser tanto positiva quanto negativa, o que pode ocasionar em uma soma nula das diferenças mesmo com os valores muito distantes da reta.

Uma forma de evitar o cancelamento é minimizar o quadrado da diferença. Poderíamos ter escolhido minimizar o módulo da diferença, mas isto acarretaria em uma complicação nos cálculos, devido à necessidade de se obter as primeiras derivadas. Supondo que sejam p pontos tabelados, definimos a função:

$$S(a,b) = \sum_{i=1}^{p} \left(y_i - (a + bx_i)\right)^2$$

Nossa problema agora é encontrar valores de *a* e de *b* que minimizam *S(a,b).*

Usando notação matricial, com os resíduos definidos por

$$r_i = y_i - (a + bx_i)$$

e definindo as matrizes

$$X = \begin{pmatrix} a \\ b \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}, R = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_p \end{pmatrix}$$

segue que $y_i = a + bx_i$ para todo i variando de 1 até p é o mesmo que $AX = Y$. Assim, como queremos minimizar

$$S(a,b) = \sum_{i=1}^{p} {r_i}^2$$

em notação matricial temos que

$$\sum_{i=1}^{p} r_i^{\,2} = R^T R$$

onde

$$R = Y - AX$$

Denotando $M = S(a,b)$, temos

$$M = (Y - AX)^T (Y - AX) = Y^T Y - X^T A^T Y - Y^T AX + X^T A^T AX$$

Queremos obter os parâmetro *a* e *b* ou, em notação matricial, o vetor *X* de modo a minimizar *M*. Para isso, o gradiente de M (ou seja, a derivada primeira da função de duas variáveis *M*) deve ser nulo:

$$\nabla M = -A^T Y - Y^T A + 2A^T AX = 0 \Rightarrow A^T AX = A^T Y$$

Assim, para encontrarmos *a* e *b* que faça com que a soma do quadrado das diferenças entre y_i e $a+bx_i$ seja mínima basta resolvermos o sistema linear

$$A^T AX = A^T Y$$

Como a matriz A^TA é simétrica definida positiva, o sistema linear admite solução única e esta solução será o ponto crítico que será o ponto de mínimo.

Efetuando os cálculos de A^TA e de A^TY temos:

$$A^T A = \begin{pmatrix} \sum_{i=1}^{p} 1 & \sum_{i=1}^{p} x_i \\ \sum_{i=1}^{p} x_i & \sum_{i=1}^{p} x_i^2 \end{pmatrix}, A^T Y = \begin{pmatrix} \sum_{i=1}^{p} y_i \\ \sum_{i=1}^{p} x_i y_i \end{pmatrix}$$

2. Ajuste de Curvas por Polinômios e outras Funções

Podemos generalizar este resultado para ajustarmos qualquer polinômio da forma

$$y = a_0 + a_1 x + \ldots + a_n x^n$$

aos pontos (x_i, y_i). Basta fazermos:

$$r_i = y_i - \left(a_0 + a_1 x_i + \ldots + a_n x_i^n\right)$$

$$X = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}, R = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{pmatrix}, A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_p & x_p^2 & \cdots & x_p^n \end{pmatrix}$$

Então, para encontrarmos os pontos a_0, a_1, ..., a_n, temos que resolver o mesmo sistema $A^TAX=A^TY$. Efetuando os cálculos de A^TA e de A^TY, temos:

$$\begin{pmatrix} \sum_{i=1}^{p} x_i^0 & \sum_{i=1}^{p} x_i^1 & \cdots & \sum_{i=1}^{p} x_i^n \\ \sum_{i=1}^{p} x_i^1 & \sum_{i=1}^{p} x_i^2 & \cdots & \sum_{i=1}^{p} x_i^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{p} x_i^n & \sum_{i=1}^{p} x_i^{n+1} & \cdots & \sum_{i=1}^{p} x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{p} x_i^0 y_i \\ \sum_{i=1}^{p} x_i^1 y_i \\ \vdots \\ \sum_{i=1}^{p} x_i^n y_i \end{pmatrix}$$

Este procedimento pode ser generalizado para qualquer curva de ajuste da forma:

$$y = a_0 g_0(x) + \ldots + a_n g_n(x)$$

desde que as funções $g_j(x)$ avaliadas nos pontos resultem em vetores linearmente independentes, que é uma condição necessária para que a matriz $A^t A$ seja invertível.

3. Um exemplo de aplicação do método de quadrados mínimos

Podemos usar o método dos quadrados mínimos para ajustar uma curva aos dados da população brasileira entre os anos de 1872 e 1996; com isso podemos prever qual será a população em um ano posterior.

Considere a tabela abaixo da população brasileira (em milhões):

ano	1872	1890	1900	1920	1940	1950	1960	1970	1980	1991	1996
população	9.9	14.3	17.4	30.6	41.2	51.9	70.2	93.1	119.0	146.2	157.1

Vamos ajustar uma curva da forma de um polinômio de segundo grau $y=a+bx+cx^2$ onde y denota a população e x o ano.

De posse da tabela podemos construir um sistema AX=Y onde:

$$A = \begin{pmatrix} 1 & 1872 & 1872^2 \\ 1 & 1890 & 1890^2 \\ 1 & 1900 & 1900^2 \\ 1 & 1920 & 1920^2 \\ 1 & 1940 & 1940^2 \\ 1 & 1950 & 1950^2 \\ 1 & 1960 & 1960^2 \\ 1 & 1970 & 1970^2 \\ 1 & 1980 & 1980^2 \\ 1 & 1991 & 1991^2 \\ 1 & 1996 & 1996^2 \end{pmatrix}, \quad X = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad Y = \begin{pmatrix} 9.9 \\ 14.3 \\ 17.4 \\ 30.6 \\ 41.2 \\ 51.9 \\ 70.2 \\ 93.1 \\ 119.0 \\ 146.2 \\ 157.1 \end{pmatrix}$$

logo, a solução X é obtida resolvendo o sistema $A^T AX = A^T Y$ cuja solução é

$$X = \begin{pmatrix} 46044.8 \\ -48.7163 \\ 0.0128889 \end{pmatrix}$$

Daí podemos calcular a população em 2000 calculando y(2000)=167.9, cujo valor podemos comparar com os dados oficiais do IBGE que informa que a população brasileira em 2000 era de 169.8 milhões de habitantes.

Ajustando um polinômio de segundo grau pelo método dos quadrados mínimos conseguimos uma previsão para a população em 2000 que difere da real em 1.1%. A partir desses dados podemos prever a população brasileira no ano de 2010 calculando y(2010)=197.6.

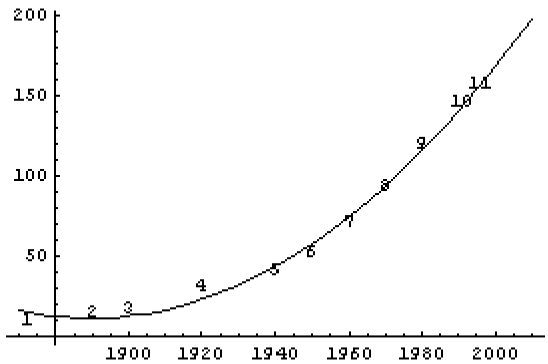

Gráfico do polinômio de grau 2 obtida por quadrados mínimos, juntamente com os pontos da tabela.

Vamos agora tentar ajustar aos mesmos pontos uma função do tipo $y(x) = a\exp(bx)$. Neste caso a curva de ajuste não é linear nos parâmetros a,b e para aplicar o procedimento descrito anteriormente (quadrados mínimos lineares) devemos linearizar a curva de ajuste. Para este exemplo, é aplicando-se o logaritmo na expressão acima, obtemos $\ln(y) = \ln a + bx$. Portanto calculamos $\ln(y)$, e com isso ajustamos uma reta aos pontos $(x, \ln(y))$ pelo método dos quadrados mínimos. Note que os parâmetros a e b encontrados não são os que minimizam a função

$$S(a,b) = \sum_{i=1}^{p} \left(y_i - a\exp(bx_i) \right)^2$$

mas os que minimizam a função

$$G(a,b) = \sum_{i=1}^{p} \left[\ln y_i - \left(\ln a + bx_i \right) \right]^2$$

ano	1872	1890	1900	1920	1940	1950	1960	1970	1980	1991	1996
população	9.9	14.3	17.4	30.6	41.2	51.9	70.2	93.1	119.0	146.2	157.1
ln(população)	2.29253	2.66026	2.85647	3.421	3.71844	3.94932	4.25135	4.53367	4.77912	4.98498	5.05688

Agora, vamos construir a matriz A e os vetores X e Y:

$$A = \begin{pmatrix} 1 & 1872 \\ 1 & 1890 \\ 1 & 1900 \\ 1 & 1920 \\ 1 & 1940 \\ 1 & 1950 \\ 1 & 1960 \\ 1 & 1970 \\ 1 & 1980 \\ 1 & 1991 \\ 1 & 1996 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Y = \begin{pmatrix} 2.29253 \\ 2.66026 \\ 2.85647 \\ 3.421 \\ 3.71844 \\ 3.94932 \\ 4.25135 \\ 4.53367 \\ 4.77912 \\ 4.98498 \\ 5.05688 \end{pmatrix}$$

resolvendo o sistema $A^T AX = A^T Y$, temos como solução o vetor:

$$X = \begin{pmatrix} -40.4266 \\ 0.0227992 \end{pmatrix}$$

logo a solução é ln(y)=-40.4266+0.0227992x. Aplicando a função exponencial temos:

$$y=\exp(-40.4266)\exp(0.0227992x)=2.77287\ 10^{-18}\exp(0.0227992x),$$

Calculando o valor dessa função no ponto 2000 temos y(2000)=176.2, cujo erro percentual comparado com os dados do IBGE é de 3.7%. A partir desses dados podemos prever a população brasileira no ano de 2010 calculando y(2010)=221.35.

Podemos então traçar o gráfico da nova função

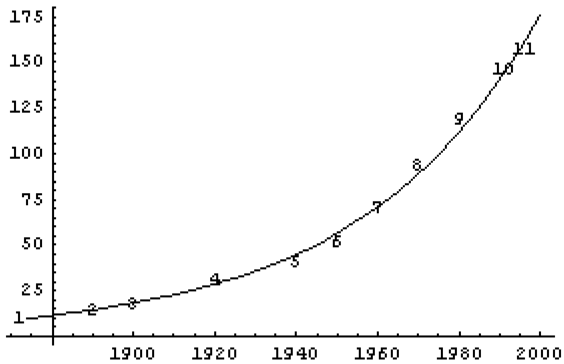

Gráfico da função exponencial obtida por quadrados mínimos, juntamente com os pontos da tabela.

Note que quando ajustamos por uma função exponencial, conseguimos ajustar melhor os pontos iniciais, já no caso de uma parábola, conseguimos ajustar melhor os pontos finais da tabela.

Em ambos os casos, podemos calcular o resíduo, dado pela norma 2 do vetor formado pela diferença entre o valor real e o valor calculado pela função obtida para cada ponto. Para a parábola temos que o resíduo é 14.2. Já para a exponencial temos que o resíduo é 11.45.

Note que apesar do resíduo da exponencial ter sido menor do que o resíduo da parábola, a previsão feita pela parábola para o ano de 2000 foi melhor do que a da exponencial, portanto a maneira como a curva se ajusta aos pontos tabelados não diz nada a respeito da maneira como essa curva fará previsões para outros pontos.