Tratamento Estatístico de Dados em Física Experimental

Prof. Zwinglio Guimarães

2º semestre de 2016

Tópico 7 - Ajuste de parâmetros de funções

(Máxima Verossimilhança e Mínimos Quadrados)

Método da máxima verossimilhança

- Para estimar os valores dos parâmetros das funções (densidade) de probabilidade que regem os dados obtido em um experimento, é razoável supor que o conjunto de dados obtidos seja um conjunto com grande probabilidade de ocorrer.
- O método da máxima verossimilhança consiste em se estimar os valores desses parâmetros como sendo os que maximizam o produto das funções (densidade) de probabilidades de se obter cada um dos dados obtidos no experimento.

A função verossimilhança

- A função verossimilhança, $\mathcal{L}(\{x_i\}|\vec{a})$, de se obter um conjunto de dados $\{x_i\} = \{x_1, x_2, \dots, x_N\}$, sendo que a medida de cada dado x_i é descrita pela função (densidade) de probabilidade $f(x_i|\vec{a})$ é:

$$\mathcal{L}(\{x_i\}|\vec{a}) = f(x_1|\vec{a}).\,f(x_2|\vec{a}) \dots f(x_N|\vec{a})$$

Onde $f(x_i|\vec{a})$ é uma função densidade de probabilidade que depende de um conjunto de parâmetros $\vec{a}$ (por exemplo, para uma gaussiana os parâmetros podem ser o valor médio verdadeiro e o desvio-padrão).

Aspectos práticos do método da máxima verossimilhança

O método da máxima verossimilhança consiste nas seguintes etapas:

1. Escrever a função verossimilhança $\mathcal{L}(\{x_i\}|\vec{a})$
2. Determinar o vetor de parâmetros $\vec{a} = \vec{\tilde{a}}$ que maximizam $\mathcal{L}$ (na prática, o $\ln(\mathcal{L})$, porque isso simplifica manipular algebricamente as derivadas)
3. Estimar as incertezas dos parâmetros estimados $\vec{\tilde{a}}$ por propagação de incertezas.

Após isso, é conveniente avaliar se as expressões obtidas para os estimadores não são tendenciosas.

Um exemplo do uso do método da máxima verossimilhança

A função verossimilhança de 2 medições independentes de uma mesma grandeza $\{x_i\} = \{x_1, x_2\}$, com função densidade de probabilidade gaussiana de valor verdadeiro x_0 e desvios padrões conhecidos (σ_1 e σ_2) é:

$$\mathcal{L}(\{x_1, x_2\}/x_0) = \frac{1}{2\pi\ \sigma_1\ \sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x_1 - x_0}{\sigma_1}\right)^2 + \left(\frac{x_2 - x_0}{\sigma_2}\right)^2\right]}$$

E a estimativa $\tilde{x}$ de x_0 por máxima verossimilhança é:

$$\tilde{x} = \frac{x_1\left(\frac{1}{\sigma_1}\right)^2 + x_2\left(\frac{1}{\sigma_2}\right)^2}{\left(\frac{1}{\sigma_1}\right)^2 + \left(\frac{1}{\sigma_2}\right)^2} \quad \text{com incerteza} \quad \sigma_{\tilde{x}} = \sqrt{\frac{1}{\left(\frac{1}{\sigma_1}\right)^2 + \left(\frac{1}{\sigma_2}\right)^2}}$$

A estimativa (tendenciosa) do desvio padrão por máxima verossimilhança

A função verossimilhança de um conjunto de medições repetitivas de uma mesma grandeza depende do valor médio verdadeiro e do desvio padrão verdadeiro das medições. Nessas condições, as estimativas da média e da variância por máxima verossimilhança, são:

$$\tilde{x} = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \text{e} \qquad \tilde{\sigma}^2 = \frac{\sum_{i=1}^{N} (x_i - \tilde{x})^2}{N}$$

No entanto, <u>a estimativa da variância por máxima verossimilhança é tendenciosa</u>, pois:

$$\langle \tilde{\sigma}^2 \rangle = \frac{N-1}{N} \sigma_0^2 .$$

O método dos mínimos quadrados

- O método dos mínimos quadrados consiste em determinar os parâmetros como os que minimizam a seguinte somatória:

$$Q(\vec{a}) = \sum_{i=1}^{N} \left(\frac{y_i - G(x_i, \vec{a})}{\sigma_i} \right)^2$$

onde a função modelo, $G(x_i, \vec{a})$, descreve a relação entre o valor esperado do $i - ésimo$ dado experimental com os parâmetros a serem estimados. Por exemplo, no caso do ajuste de uma reta: $G(x, a_1, a_2) = a_1 + a_2 x$

Ajuste de funções lineares nos parâmetros pelo MMQ

No caso de funções lineares nos parâmetros, $\frac{\partial}{\partial a_i}\left(\frac{\partial G}{\partial a_j}\right) = 0$, a função modelo pode ser escrita como:

$$G(x, \vec{a}) = a_1 g_1(x) + a_2 g_2(x) + \cdots + a_M g_M(x)$$

e os parâmetros que minimizam a variável $Q(\vec{a})$ correspondem às soluções do seguinte sistema linear:

$$\begin{aligned}
\sum_{i=1}^{N} \frac{y_i g_1(x_i)}{\sigma_i^2} &= \tilde{a}_1 \sum_{i=1}^{N} \frac{g_1(x_i) g_1(x_i)}{\sigma_i^2} + \tilde{a}_2 \sum_{i=1}^{N} \frac{g_1(x_i) g_2(x_i)}{\sigma_i^2} + \cdots \\
\sum_{i=1}^{N} \frac{y_i g_2(x_i)}{\sigma_i^2} &= \tilde{a}_1 \sum_{i=1}^{N} \frac{g_2(x_i) g_1(x_i)}{\sigma_i^2} + \tilde{a}_2 \sum_{i=1}^{N} \frac{g_2(x_i) g_2(x_i)}{\sigma_i^2} + \cdots \\
&\vdots
\end{aligned}$$

Ajuste de funções lineares nos parâmetros pelo MMQ - II

- O sistema de equações do MMQ pode ser escrito como:

$$D = \mathbf{M}\tilde{A}$$

onde $\tilde{A}_l = \tilde{a}_l$, e:

$$D_l = \sum_{i=1}^{N} \frac{y_i g_l(x_i)}{\sigma_i^2} \qquad \text{e} \qquad M_{l,c} = \sum_{i=1}^{N} \frac{g_l(x_i) g_c(x_i)}{\sigma_i^2}$$

- A solução é dada por: $\tilde{A} = (\mathbf{M}^{-1})D$

- A matriz de covariância de $\tilde{A}$ é dada por: $\mathrm{V}_{\tilde{A}} = (\mathbf{M}^{-1})$

Obs: No Octave, a inversa da matriz M é obtida por: **inv(M)**

Covariâncias e correlações (revisão)

- Interpretação da matriz de covariâncias, $\boldsymbol{V_A} = \boldsymbol{M}^{-1}$:

$$\mathbf{V}_A = \begin{bmatrix} \sigma_{a_1}^2 & cov(a_1, a_2) \\ cov(a_1, a_2) & \sigma_{a_2}^2 \end{bmatrix}$$

- As correlações correspondentes, $\rho_{a_1,a_2} = \frac{cov(a_1,a_2)}{\sigma_{a_1}\sigma_{a_2}}$,

podem ser fornecidas em uma matriz de correlações:

$$\mathbf{C}_A = \begin{bmatrix} 1 & \rho(a_1, a_2) \\ \rho(a_1, a_2) & 1 \end{bmatrix}$$

Exemplo de aplicação do MMQ em um ajuste pouco usual

Exemplo baseado em exercício do livro *"Método dos Mínimos Quadrados com Formalismo Matricial"* do prof. Otaviano Helene

Exemplo numérico de ajuste de uma função linear nos parâmetros pelo MMQ

$$G(x,\vec{a}) = a_1 g_1(x) + a_2 g_2(x) + \cdots + a_M g_M(x)$$

- Considere o volume de combustível em trajetos com diferentes composições de trechos urbano e rodoviário

  a) **42,4** litros em **120,3** km na cidade e **451,6** km na estrada
  b) **28,0** litros em **195,1** km na cidade e **115,3** km na estrada
  c) **34,3** litros em **10,2** km na cidade e **523,5** km na estrada
  d) **36,5** litros em **320,9** km na cidade e **54,2** km na estrada
  e) **29,5** litros em **110,6** km na cidade e **277,4** km na estrada

$$y = \begin{bmatrix} 42{,}4 \\ 28{,}0 \\ 34{,}3 \\ 36{,}5 \\ 29{,}5 \end{bmatrix} \qquad g_1 = \begin{bmatrix} 120{,}3 \\ 195{,}1 \\ 10{,}2 \\ 320{,}9 \\ 110{,}6 \end{bmatrix} \qquad g_2 = \begin{bmatrix} 451{,}6 \\ 115{,}3 \\ 523{,}5 \\ 54{,}2 \\ 277{,}4 \end{bmatrix}$$

Ajustes lineares pelo MMQ (revisão)

- Escrevendo a função modelo como:

$$G(x,\vec{a}) = a_1 g_1(x) + a_2 g_2(x) + \cdots + a_M g_M(x)$$

- O sistema linear de equações do MMQ pode ser escrito de forma matricial: $\overline{D} = \mathbf{M}\overline{\tilde{A}}$

$$D_l = \sum_{i=1}^{N} \frac{y_i g_l(x_i)}{\sigma_i^2} \qquad M_{l,c} = \sum_{i=1}^{N} \frac{g_l(x_i) g_c(x_i)}{\sigma_i^2}$$

cuja solução é: $\overline{\tilde{A}} = \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \\ \vdots \\ \tilde{a}_M \end{bmatrix} = (\mathbf{M}^{-1})\overline{D}$ com $\mathrm{V}_{\tilde{A}} = \mathbf{M}^{-1}$

Obs: No Octave, a inversa da matriz M pode ser obtida por: **inv(M)**

Resultados do exemplo numérico:

$$y = \begin{bmatrix} 42{,}4 \\ 28{,}0 \\ 34{,}3 \\ 36{,}5 \\ 29{,}5 \end{bmatrix} \qquad g_1 = \begin{bmatrix} 120{,}3 \\ 195{,}1 \\ 10{,}2 \\ 320{,}9 \\ 110{,}6 \end{bmatrix} \qquad g_2 = \begin{bmatrix} 451{,}6 \\ 115{,}3 \\ 523{,}5 \\ 54{,}2 \\ 277{,}4 \end{bmatrix}$$

Considerando que a incerteza dos valores de y sejam $\sigma_i = 0{,}5\ l$

$$\tilde{A} = \begin{bmatrix} 0{,}1041\ (13)\ \ l/km \\ 0{,}0674\ \ (7)\ \ l/km \end{bmatrix}$$

$$cov(\tilde{a}_1, \tilde{a}_2) = -4{,}13.10^{-7}\ \ l^2/km^2 \qquad \rho_{\tilde{a}_1,\tilde{a}_2} = -0{,}42$$

E a qualidade do ajuste pode ser avaliada pelo teste de χ^2

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - G(x_i, \tilde{a})}{\sigma_i} \right)^2 = 4{,}23$$