Métodos Estatísticos em Física Experimental

Prof. Zwinglio Guimarães

1º semestre de 2015

Aulas 9 e 10

Método da máxima verossimilhança

- Para estimar os valores verdadeiros 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 ocorrer (dado as funções densidade de probabilidade dos dados).
- O método da máxima verossimilhança consiste em estimar os valores verdadeiros desses parâmetros como sendo aqueles que maximizam essa probabilidade.

A função verossimilhança

- A função verossimilhança, $\mathcal{L}(\{x_i\}|\vec{a})$, referente à probabilidade de se obter o conjunto de dados $\{x_i\} = \{x_1, x_2, \dots, x_N\}$, dado que as funções densidade de probabilidade de cada dado dependam dos parâmetros $\vec{a} = (a_1, a_2, \dots, a_M)$ é dada por:

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

Onde $f(x_i|\vec{a})$ é a f.d.p. que rege a medida do dado x_i, a qual depende dos parâmetros $\vec{a}$ (por exemplo, para uma f.d.p. 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. Calcular o vetor de valores do parâmetros $\vec{a} = \tilde{\vec{a}}$ que maximizam $\mathcal{L}$ (na prática, se maximiza o logaritmo de $\mathcal{L}$, porque isso transforma os produtos em somas, simplificando manipular algebricamente as derivadas)
3. As incertezas nos parâmetros estimados $\tilde{\vec{a}}$ são obtidas por propagação de incertezas.

Um exemplo do uso do método da máxima verossimilhança

A função verossimilhança de 2 medições independentes de uma mesma grandeza $\{x_i\} = \{x_1, x_2\}$, que tenham f.d.p. gaussiana 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:

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

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 do desvio padrão por máxima verossimilhança, são:

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

No entanto, <u>a estimativa do desvio padrão por máxima verossimilhança é tendenciosa</u>, pois:

$$\langle\tilde{\sigma}\rangle = \sqrt{\frac{N}{N-1}}\,\sigma_0.$$

O método dos mínimos quadrados

- O método dos mínimos quadrados consiste em determinar os parâmetros como àqueles 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, $\frac{\partial}{\partial a_i}\left(\frac{dG}{da_j}\right) = 0$, 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 linear de equações do MMQ pode ser escrito de forma matricial:

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

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: $\overline{\tilde{A}} = (\mathbf{M}^{-1})\overline{D}$

Sendo que a matriz de covariância de $\overline{\tilde{A}}$ é: $\mathrm{V}_{\tilde{A}} = (\mathbf{M}^{-1})$

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:

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