Metody numeryczne

Własności estymatorów

Pseudolinearyzacja

P. F. Góra

https://zfs.fais.uj.edu.pl/pawel_gora

2025

Własności wektora estymatorów

Wektor $\mathbf{p}$ obliczamy rozwiązując równanie

$$\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\mathbf{p} = \mathbf{A}^T\mathbf{G}^{-1}\mathbf{y} \tag{1}$$

dla takich wartości pomiarów, jakie faktycznie mamy. Tak obliczony wektor $\mathbf{p}$ jest wektorem estymatorów. Pamiętajmy, że pomiary są obarczone błędami losowymi, a więc także obliczone estymatory są, formalnie, *gaussowskimi liczbami losowymi*. Co można powiedzieć o tych liczbach? $\mathbf{y} = \mathbf{A}\mathbf{p}^* + \boldsymbol{\xi}$, gdzie $\mathbf{p}^*$ jest zbudowany z “prawdziwych”, nieznanych wartości $[a_1, \dots, a_s]^T$ występujących w równaniu

$$\tilde{y}_i = a_1 \cdot f_1(x_i) + a_2 \cdot f_2(x_i) + \cdots + a_s \cdot f_s(x_i)\,. \tag{2}$$

Widać, że

$$\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}(\mathbf{p} - \mathbf{p}^*) = \mathbf{A}^T\mathbf{G}^{-1}\boldsymbol{\xi} \tag{3}$$

wobec czego

$$\langle \mathbf{p} \rangle = \mathbf{p}^* \,, \tag{4}$$

gdyż $\langle \boldsymbol{\xi} \rangle = 0$. Obliczone estymatory są przybliżeniem “prawdziwych” wartości parametrów w sensie równania (4).

Macierz kowariancji estymatorów wynosi

$$\begin{aligned}
\mathbf{C_p} &= \left\langle (\mathbf{p}-\mathbf{p}^*)(\mathbf{p}-\mathbf{p}^*)^T \right\rangle \\
&= \left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1} \mathbf{A}^T\mathbf{G}^{-1} \left\langle \boldsymbol{\xi}\boldsymbol{\xi}^T \right\rangle \mathbf{G}^{-1}\mathbf{A} \left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1} \,,
\end{aligned} \tag{5}$$

gdzie skorzystaliśmy z symetrii macierzy $\mathbf{G}^{-1}$ i macierzy $\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}$. Korzystając z równania

$$\left\langle \boldsymbol{\xi}\boldsymbol{\xi}^T \right\rangle = \mathbf{G} \tag{6}$$

otrzymujemy

$$
\begin{aligned}
\mathbf{C_p} &= \left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{G}^{-1}\overbrace{\underbrace{\mathbf{G}}_{\left\langle\boldsymbol{\xi}\boldsymbol{\xi}^T\right\rangle}\mathbf{G}^{-1}}^{\mathbb{I}}\mathbf{A}\left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1}, \\
&= \left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1}\underbrace{\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1}}_{\mathbb{I}} \\
&= \left(\mathbf{A}^T\mathbf{G}^{-1}\mathbf{A}\right)^{-1}. \qquad (7)
\end{aligned}
$$

Pseudolinearyzacja

Rozważamy nieliniowe zagadnienie najmniejszych kwadratów dla pomiarów niezależnych (nieskorelowanych) i obarczonych statystycznie takimi samymi błędami

$$Q = \text{const} \cdot \frac{1}{2} \sum_{i=1}^{N} (y_i - \tilde{y}_i)^2 = \text{const} \cdot \frac{1}{2} \sum_{i=1}^{N} (y_i - f(x_i; \mathbf{p}))^2 \,, \qquad (8)$$

gdzie $\tilde{y}_i$ oznaczają wartości "idealne", teoretyczne.

Czasami do znalezienia minimum Q stosuje się metodę *pseudolinearyzacji*. Przypuśćmy, że $\mathbf{p}_n$ jest aktualnym przybliżeniem poszukiwanej wartości parametrów $\mathbf{p}$. Stawiamy hipotezę, iż "prawdziwe" wartości parametrów są małą poprawką w stosunku do $\mathbf{p}_n$: $\mathbf{p} \simeq \mathbf{p}_n + \boldsymbol{\delta}\mathbf{p}$ i rozwijamy $\tilde{y}_i = f(x_i; \mathbf{p})$ w szereg Taylora do pierwszego rzędu:

$$\tilde{y}_i = f(x_i; \mathbf{p}_n + \boldsymbol{\delta}\mathbf{p}) \simeq f(x_i; \mathbf{p}_n) + \left[\nabla_{\mathbf{p}} f|_{\mathbf{p}_n}\right]^T \boldsymbol{\delta}\mathbf{p}\,. \tag{9}$$

Podstawiamy to rozwinięcie do (8). Funkcja

$$Q = \frac{1}{2}\sum_{i=1}^{N}\left(y_i - f(x_i; \mathbf{p}_n) - \left[\nabla_{\mathbf{p}} f|_{\mathbf{p}_n}\right]^T \boldsymbol{\delta}\mathbf{p}\right)^2 \tag{10}$$

jest formą kwadratową w poprawkach $\boldsymbol{\delta}\mathbf{p}$. Po znalezieniu znanymi metodami wartości $\boldsymbol{\delta}\mathbf{p}_{\text{min}}$, odpowiadających (jedynemu) minimum (10), podstawiamy $\mathbf{p}_{n+1} = \mathbf{p}_n + \boldsymbol{\delta}\mathbf{p}_{\text{min}}$ i powtarzamy całą procedurę.

Taka procedura dość dobrze działa w wypadku nieliniowej metody najmniejszych kwadratów, choć nie należy jej polecać jako ogólnej metody minimalizacji. Pseudolinearyzacja ma tylko jedno niewątpliwe zastosowanie: Po znalezieniu ostatecznych wartości minimalizujących funkcję (8), za pomocą pseudolinearyzacji wokół tego punktu znajdujemy macierz kowariancji estymatorów, będącą charaktestystyką *liniową*.

Przykład

Przypuśćmy, że do danych dopasowujemy funkcję Gaussa

$$y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \tag{11}$$

zaś aktualnymi przybliżeniami parametrów są $\bar{x}_n$, σ_n^2. Obliczamy

$$f(x;\bar{x},\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \tag{12a}$$

$$\left.\frac{\partial f}{\partial \bar{x}}\right|_{\bar{x}_n,\sigma_n^2} = \frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left(-\frac{(x-\bar{x}_n)^2}{2\sigma_n^2}\right)\cdot\frac{x-\bar{x}_n}{\sigma_n^2}, \tag{12b}$$

$$\left.\frac{\partial f}{\partial(\sigma^2)}\right|_{\bar{x}_n,\sigma_n^2} = \frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left(-\frac{(x-\bar{x}_n)^2}{2\sigma_n^2}\right)\cdot\left[\frac{(x-\bar{x}_n)^2}{4\left(\sigma_n^2\right)^2}-\frac{1}{2\sigma_n^2}\right] \tag{12c}$$

Wyrażenie

$$Q = \frac{1}{2}\sum_{i=1}^{N}\left(y_i - \frac{1}{\sqrt{2\pi\sigma_n^2}}\,\mathrm{exp}\left(-\frac{(x_i-\bar{x}_n)^2}{2\sigma_n^2}\right)\right.$$
$$\left.\cdot\left\{1+\frac{x_i-\bar{x}_n}{\sigma_n^2}\,\delta\bar{x}+\left[\frac{(x_i-\bar{x}_n)^2}{4\left(\sigma_n^2\right)^2}-\frac{1}{2\sigma_n^2}\right]\delta\sigma^2\right\}\right)^2 \tag{13}$$

jest formą kwadratową w zmiennych $\delta\bar{x}$, $\delta\sigma^2$.