

MODELOWANIE REGRESJI PORZĄDKOWEJ PRZY UŻYCIU PROCESU GAUSSOWSKIEGO

$\{f(x_i)\}_{i=1}^n$ - realizacja procesu gaussowskiego o średniej 0 i macierzy kowariancji Σ zadanej wzorem:

$$\Sigma = (K(x_i, x_j))_{i,j=1\dots n} = \left(e^{-\frac{\kappa}{2} \sum_{\xi=1}^d (x_i^\xi - x_j^\xi)^2} \right)_{i,j=1\dots n},$$

gdzie $\kappa > 0$, a x_i^ξ to ξ -ty element x_i .

Wtedy f ma rozkład łączny o gęstości:

$$\mathbb{P}(f) = \frac{1}{Z_f} e^{-\frac{1}{2} f^T \Sigma^{-1} f},$$

gdzie $Z_f = (2\Pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}$, a $f = [f(x_1), \dots, f(x_n)]^T$.

Wtedy:

$$\mathbb{P}(D|f) = \prod_{i=1}^n \mathbb{P}(y_i|f(x_i)),$$

gdzie $D = \{y_1, \dots, y_r\}$.

Intuicyjnie:

$$\mathbb{P}_{ideal}(y_i|f(x_i)) = \mathbb{I}\{f(x_i) \in (b_{y_{i-1}}, b_{y_i}]\},$$

gdzie $b_0 = -\infty, b_r = +\infty$.

Można wygodniej sparametryzować b_i jako: $b_i \in \mathbb{R}$, $b_i = \sum_{t=2}^j \Delta_t + b_1$, gdzie $\Delta_t > 0$ oraz $j = 2, \dots, r-1$. Bardzo rzadko mamy jednak do czynienia z idealną sytuacją, dlatego będziemy budować model zakładając dodatkowy szum δ o rozkładzie $\mathcal{N}(0, \sigma^2)$. Wtedy prawdopodobieństwo zmienia się następująco:

$$\mathbb{P}(y_i|f(x_i)) = \Phi(z_1^i) - \Phi(z_2^i),$$

gdzie $z_1^i := \frac{b_{y_i} - f(x_i)}{\sigma}$ oraz $z_2^i := \frac{b_{y_{i-1}} - f(x_i)}{\sigma}$.

Dowód

$$\begin{aligned} \mathbb{P}(y_i|f(x_i)) &= \int \mathbb{P}_{ideal}(y_i|f(x_i) + \delta_i) d\delta_i = \int \mathbb{P}(\delta_i) \mathbb{I}\{f(x_i) + \delta_i \in (b_{y_{i-1}}, b_{y_i}]\} d\delta_i = \\ &= \int \frac{1}{2\Pi\sigma} e^{-\frac{u^2}{2\sigma^2}} \mathbb{I}\{u \in (b_{y_{i-1}} - f(x_i), b_{y_i} - f(x_i)]\} du = \int_{b_{y_i} - f(x_i)}^{b_{y_{i-1}} - f(x_i)} \frac{1}{2\Pi\sigma} e^{-\frac{u^2}{2\sigma^2}} du = \\ &= \int_{\frac{b_{y_i} - f(x_i)}{\sigma}}^{\frac{b_{y_{i-1}} - f(x_i)}{\sigma}} \frac{1}{2\Pi} e^{-\frac{u^2}{2}} du = \Phi\left(\frac{b_{y_i} - f(x_i)}{\sigma}\right) - \Phi\left(\frac{b_{y_{i-1}} - f(x_i)}{\sigma}\right) \end{aligned}$$

■

Wprowadźmy następującą funkcję straty:

$$l(y_i, f(x_i)) := -\ln \mathbb{P}(y_i|f(x_i))$$

Jej pochodne to:

$$\begin{aligned}\frac{\partial l(y_i, f(x_i))}{\partial f(x_i)} &= \frac{1}{\sigma} \frac{\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)} \\ \frac{\partial^2 l(y_i, f(x_i))}{\partial^2 f(x_i)} &= \frac{1}{\sigma^2} \left(\frac{\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)} \right)^2 + \frac{1}{\sigma^2} \frac{z_1^i \frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - z_2^i \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)}\end{aligned}$$

Dowód

$$\begin{aligned}\frac{\partial l(y_i, f(x_i))}{\partial f(x_i)} &= -\ln[\Phi(z_1^i) - \Phi(z_2^i)] = -\frac{1}{\Phi(z_1^i) - \Phi(z_2^i)} \cdot \Phi'(z_1^i) \cdot \left(-\frac{1}{\sigma}\right) - \Phi'(z_2^i) \cdot \left(-\frac{1}{\sigma}\right) = \\ &= \frac{1}{\sigma} \frac{\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 l(y_i, f(x_i))}{\partial^2 f(x_i)} &= \frac{\partial}{\partial f(x_i)} \left(\frac{\partial l(y_i, f(x_i))}{\partial f(x_i)} \right) = \frac{\partial}{\partial f(x_i)} \left(\frac{1}{\sigma} \frac{\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)} \right) = \\ &= \frac{1}{\sigma} \frac{1}{[\Phi(z_1^i) - \Phi(z_2^i)]^2} \{ [\Phi(z_1^i) - \Phi(z_2^i)] \cdot \\ &\cdot \left[\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} \left(-\frac{1}{\sigma} \cdot z_1^i \right) \left(-\frac{1}{\sigma} \right) - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}} \left(-\frac{1}{\sigma} \cdot z_2^i \right) \left(-\frac{1}{\sigma} \right) \right] - \\ &- \left(\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}} \right) \cdot \left(-\frac{1}{\sigma} \right) \cdot \left(\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}} \right) \} = \\ &= \frac{1}{\sigma^2} \left(\frac{\frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)} \right)^2 + \frac{1}{\sigma^2} \frac{z_1^i \frac{1}{2\Pi} e^{-\frac{z_1^i{}^2}{2}} - z_2^i \frac{1}{2\Pi} e^{-\frac{z_2^i{}^2}{2}}}{\Phi(z_1^i) - \Phi(z_2^i)}\end{aligned}$$

■