Gaussiana bivariada: distribuições condicionais

A gaussiana bivariada é definida como

$$P(x,y|\mu_x,\mu_y,\sigma_x,\sigma_y,\sigma_{xy}) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{Z^2}{2(1-\rho^2)}\right)$$

onde

$$Z^2 = \frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}$$

e ρ é o coeficiente de correlação:

$$\rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$$

Estamos interessados em calcular $P(x|y)$, isto é, a pdf de x para um valor fixo de y. Podemos reescrever a expressão acima como

$$P(x,y|\mu_x,\mu_y,\sigma_x,\sigma_y,\sigma_{xy}) \propto f(x,y) = \exp\left[-\frac{1}{2}\left(ax^2-2bx+c\right)\right]$$

É fácil ver que

$$a = \frac{1}{\sigma_x^2(1-\rho^2)}$$

e

$$b = \frac{1}{(1-\rho^2)}\left[\frac{\mu_x}{\sigma_x^2} + \frac{\rho(y-\mu_y)}{\sigma_x\sigma_y}\right].$$

O valor de c não importa pois ele não aparece nos resultados, como veremos.

Vamos agora calcular $P(x|y)$:

$$P(x|y) = \frac{\exp\left[-\frac{1}{2}\left(ax^2-2bx+c\right)\right]}{\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(ax^2-2bx+c\right)\right]dx}$$

$$f(x,y) = \exp\left[-\frac{1}{2}\left(ax^2-2bx+c\right)\right] = \exp\left[-\frac{a}{2}\left(x^2-2\frac{b}{a}x+\frac{c}{a}\right)\right] =$$

$$= \exp\left[-\frac{a}{2}\left(x^2-2\frac{b}{a}x+\frac{c}{a}+(\frac{b^2}{a^2}-\frac{c}{a})-(\frac{b^2}{a^2}-\frac{c}{a})\right)\right] =$$

$$= \exp\left[-\frac{a}{2}\left(x^2-2\frac{b}{a}x+(\frac{c}{a}+\frac{b^2}{a^2}-\frac{c}{a})-(\frac{b^2}{a^2}-\frac{c}{a})\right)\right] =$$

$$= \exp\left[-\frac{a}{2}\left(x^2-2\frac{b}{a}x+\frac{b^2}{a^2}-(\frac{b^2}{a^2}-\frac{c}{a})\right)\right] =$$

$$= \exp\left[-\frac{a}{2}\left((x^2-2\frac{b}{a}x+\frac{b^2}{a^2})+\frac{a}{2}(\frac{b^2}{a^2}-\frac{c}{a})\right)\right] =$$

$$= \exp\left[-\frac{a}{2}\left(x^2-2\frac{b}{a}x+\frac{b^2}{a^2}\right)\right]\exp\left[\frac{a}{2}\left(\frac{b^2}{a^2}-\frac{c}{a}\right)\right] =$$

$$= c_1 \exp\left[-\frac{a}{2}\left(x^2 - 2\frac{b}{a}x + \frac{b^2}{a^2}\right)\right] =$$

$$= c_1 \exp\left[-\frac{a}{2}\left(x - \frac{b}{a}\right)^2\right]$$

$$\int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} c_1 \exp\left[-\frac{a}{2}\left(x - \frac{b}{a}\right)^2\right] dx = c_1 \frac{\sqrt{2\pi}}{\sqrt{a}}$$

$$P(x|y) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx} = \frac{\sqrt{a}}{\sqrt{2\pi}} \exp\left[-\frac{a}{2}\left(x - \frac{b}{a}\right)^2\right]$$

Logo, $P(x|y)$ é uma distribuição gaussiana de precisão $a = 1/\sigma^2$ e média b/a. Como

$$a = \frac{1}{(1-\rho^2)\sigma_x^2}$$

e

$$b = \frac{1}{(1-\rho^2)}\left[\frac{\mu_x}{\sigma_x^2} + \frac{\rho(y-\mu_y)}{\sigma_x\sigma_y}\right],$$

temos que

$$P(x|y) = N(\mu, \sigma)$$

onde

$$\mu = \mu_x - \frac{\rho\sigma_x(y-\mu_y)}{\sigma_y}$$

e

$$\sigma = \frac{\sigma_x}{(1-\rho^2)}.$$