Recitation Note: The Multivariate Gaussian Distribution

Tyler Maunu
maunut@mit.edu

September 3, 2019

This note will review the multivariate Gaussian distribution. It will begin by considering the simpler bivariate case, and then discuss the full multivariate case. We will also use 2×2 matrix formulas to show that the formulas for their pdfs do in fact coincide in the 2-dimensional case.

The Multivariate Gaussian Distribution

The pdf of the bivariate Gaussian distribution is

$$\begin{aligned} &f(x, y, \mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho) \\ &\quad = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right)\right]. \end{aligned} \tag{1}$$

The pdf for the general d-dimensional multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is given by

$$f(\boldsymbol{x}; \mu, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{d/2} \det(\Sigma)^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]. \tag{2}$$

Questions

1. Show that the marginal distributions of X and Y are $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$.

2. What is $\mathrm{Cov}(X, Y)$?

3. Show that X and Y can equivalently be defined in the following way. Let Z_1 and Z_2 be independent $N(0, 1)$ random variables. Then, $X = \sigma_x Z_1 + \mu_x$ and $Y = \sigma_y[\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_y$.

4. Show that the multivariate case in (2) reduces to the pdf in (1) when $d = 2$.

Solutions

1.

Notice that

$$\frac{1}{(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}-\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right)=\left(\frac{(x-a(y))^2}{(1-\rho^2)\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right), \tag{3}$$

where $a(y)=\mu_x+\rho\frac{\sigma_x}{\sigma_y}(y-\mu_y)$. Therefore, the marginal distribution of y is

$$\begin{aligned}
&f(y;\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho) \qquad (4)\\
&\quad=\int\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}-\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right)\right]dx\\
&\quad=\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{1}{2}\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\int\frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}\frac{(x-a(y))^2}{(1-\rho^2)\sigma_x^2}\right)dx\\
&\quad=\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{1}{2}\frac{(y-\mu_y)^2}{\sigma_y^2}\right].
\end{aligned}$$

This means that Y must be Gaussian with mean μ_y and variance σ_y^2. The same argument yields that X is $N(\mu_x,\sigma_x^2)$.

2. We have

$$\mathrm{Cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}X\mathbb{E}Y \tag{5}$$

Since the marginal distributions are $N(\mu_x,\sigma_x^2)$ and $N(\mu_y,\sigma_y^2)$, we know that $\mathbb{E}X=\mu_x$ and $\mathbb{E}Y=\mu_y$.

Using the answer to 1., the random variable $X|Y$ has pdf

$$f(X=x|Y=y)\propto\exp\left(-\frac{1}{2}\frac{(x-a(y))^2}{(1-\rho^2)\sigma_x^2}\right), \tag{6}$$

where again $a(y)=\mu_x+\rho\frac{\sigma_x}{\sigma_y}(y-\mu_y)$. Therefore, $\mathbb{E}(X|Y=y)=a(y)$. The law of total expectation yields

$$\begin{aligned}
\mathbb{E}(XY)=\mathbb{E}[\mathbb{E}(XY|Y)]&=\mathbb{E}\left[Y\left(\mu_x+\rho\frac{\sigma_x}{\sigma_y}(Y-\mu_y)\right)\right] \qquad (7)\\
&=\mathbb{E}\left[Y\mu_x+\rho\frac{\sigma_x}{\sigma_y}(Y^2-Y\mu_y)\right]\\
&=\mu_y\mu_x+\rho\frac{\sigma_x}{\sigma_y}\left(\sigma_y^2+\mu_y^2-\mu_y^2\right)\\
&=\mu_y\mu_x+\rho\sigma_x\sigma_y.
\end{aligned}$$

Therefore, we conclude that $\mathrm{Cov}(X,Y)=\rho\sigma_x\sigma_y$.

3. To show this, we will show that X and Y have the proper marginal distributions, and that $\mathrm{Cov}(X,Y)=\rho\sigma_x\sigma_y$. The conclusion then follows.

$$\mathbb{E}X=\mathbb{E}(\sigma_xZ_1+\mu_x)=\mu_x. \tag{8}$$

$$\mathbb{E}Y = \mathbb{E}(\sigma_y \rho Z_1 + \sigma_y \sqrt{1-\rho^2} Z_2 + \mu_y) = \mu_y. \tag{9}$$

$$\mathrm{Var}(X) = \mathrm{Var}(\sigma_x Z_1 + \mu_x) = \sigma_x^2 \mathrm{Var}(Z_1) = \sigma_x^2. \tag{10}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(\sigma_y \rho Z_1 + \sigma_y \sqrt{1-\rho^2} Z_2 + \mu_y) = \sigma_y^2 \rho^2 \mathrm{Var}(Z_1) + \sigma_y^2 (1-\rho^2) \mathrm{Var}(Z_2) + \sigma_y^2. \tag{11}$$

$$\begin{aligned} \mathrm{Cov}(X, Y) &= \mathrm{Cov}(\sigma_x Z_1 + \mu_x, \sigma_y \rho Z_1 + \sigma_y \sqrt{1-\rho^2} Z_2 + \mu_y) \\ &= \mathrm{Cov}(\sigma_x Z_1, \sigma_y \rho Z_1 + \sigma_y \sqrt{1-\rho^2} Z_2) \\ &= \mathrm{Cov}(\sigma_x Z_1, \sigma_y \rho Z_1) + \mathrm{Cov}(\sigma_x Z_1, \sigma_y \sqrt{1-\rho^2} Z_2) \\ &= \sigma_x \sigma_y \rho \mathrm{Cov}(Z_1, Z_1) = \sigma_x \sigma_y \rho. \end{aligned} \tag{12}$$

As we see, this completely specifies the parameters of the bivariate Gaussian distribution, and therefore (X, Y) has the distribution specified in (1).

4. First, recall that for a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

we have

$$\det(A) = a_{11}a_{22} - a_{21}a_{22}, \qquad A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{13}$$

Let

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}, \tag{14}$$

which is the covariance matrix in the bivariate case. Its inverse is

$$\mathbf{\Sigma}^{-1} = \frac{1}{(1-\rho^2)\sigma_x^2\sigma_y^2} \begin{pmatrix} \sigma_y^2 & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix} \tag{15}$$

where $(1-\rho^2)\sigma_x^2\sigma_y^2 = \det(\mathbf{\Sigma})$. Let $\boldsymbol{X} = (X, Y)^\top$ be the two-dimensional multivariate Gaussian random vector, and $\boldsymbol{x} = (x, y)^\top$. Then,

$$\begin{aligned} (\boldsymbol{x} - \boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) &= \frac{1}{(1-\rho^2)\sigma_x^2\sigma_y^2} (\boldsymbol{x} - \boldsymbol{\mu})^\top \begin{pmatrix} \sigma_y^2(x-\mu_x) - \rho\sigma_x\sigma_y(y-\mu_y) \\ -\rho\sigma_x\sigma_y(x-\mu_x) + \sigma_x^2(y-\mu_y) \end{pmatrix} \\ &= \frac{1}{(1-\rho^2)\sigma_x^2\sigma_y^2} \Big[\sigma_y^2(x-\mu_x)^2 - \rho\sigma_x\sigma_y(y-\mu_y)(x-\mu_x) \\ &\quad - \rho\sigma_x\sigma_y(x-\mu_x)(y-\mu_y) + \sigma_x^2(y-\mu_y)^2 \Big] \\ &= \left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right) \right]. \end{aligned} \tag{16}$$