CSIR-UGC NET-Dec 2015-Problem(109)

Chirag Mehta

AI20BTECH11006

Multivariate gaussian

Multivariate Gaussian expression, definition

The multivariate normal distribution of a n dimensional vector $\boldsymbol{X} = (X_1, X_2, ..., X_n)^\top$ can be written as

$$f_{\boldsymbol{X}}(x_1, x_2, ..., x_n) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}} \tag{1}$$

Mean vector $\boldsymbol{\mu}$ is defined as

$$\boldsymbol{\mu} = E[\boldsymbol{X}] = (E[X_1], E[X_2], ..., E[X_n])^\top \tag{2}$$

Covariance matrix $\boldsymbol{\Sigma}$ is defined as

$$\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j) \tag{3}$$

multivariate gaussian contd.

equivalence

A random vector $\boldsymbol{X} = (X_1, X_2, ..., X_n)^\top$ has a multivariate gaussian distribution if it satifies

- For every linear combination $Y = a_1X_1 + a_2X_2 + ... + a_nX_n$ of its components in normally distributed. That is for any vector $\boldsymbol{a} \in \mathbb{R}^n$, the random variable $Y = a^\top X$ has a univariate normal distribution

marginal probability

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \tag{4}$$

dirac delta function

dirac delta function

An important property of dirac delta function that will be used at multiple ocassions in this solution is

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a) \tag{5}$$

Question

CSIR-UGC NET-Dec 2015-Problem(109)

Suppose $\begin{pmatrix} X \\ Y \end{pmatrix}$ is a random vector such that the marginal distribution of X and the marginal distribution of Y are the same and each is normally distributed with mean 0 and variance 1. Then, which of the following conditions imply independence of X and Y?

1. $\text{Cov}(X, Y) = 0$
2. $aX + bY$ is normally distributed with mean 0 and variance $a^2 + b^2$ for all real a and b
3. $\Pr(X \leq 0, Y \leq 0) = \frac{1}{4}$
4. $E\left[e^{itX+isY}\right] = E\left[e^{itX}\right] E\left[e^{isY}\right]$ for all real s and t

Solution

Given $X \sim N(0,1)\,, Y \sim N(0,1)$

1.

$$Cov(X,Y) = 0 \tag{6}$$
$$E[XY] - E[X]E[Y] = 0 \tag{7}$$
$$E[XY] = 0 \tag{8}$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x,y)\,dx\,dy = 0 \tag{9}$$

This doesn't imply independence. Counter example given below

Solution Contd.

1. Lets consider a case where X and Y are dependent based on the following relation, Y being independent of K

$$X = KY \tag{10}$$

PMF for K is given as

$$p_K(k) = \begin{cases} \frac{1}{2} & k = 1 \\ \frac{1}{2} & k = -1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Solution Contd.

1. A simulation is given below, Y is gaussian, then X also follows gaussian

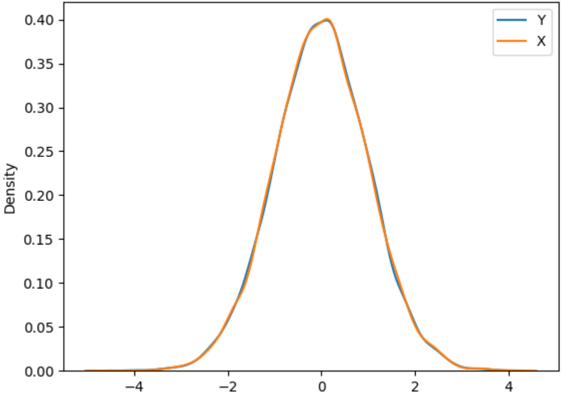

Figure: X and Y, if Y is normal

1. Theoretically it can be proved in the following manner, Since K and Y are independent

$$f_X(x) = \Pr(K = 1) f_Y(x) + \Pr(K = -1) f_Y(-x) \tag{12}$$
$$= \frac{1}{2}\left(f_Y(x) + f_Y(-x)\right) \tag{13}$$
$$= f_Y(x) \tag{14}$$

Therefore, X follows identical but not independent distribution as Y, An alternative proof is given below as a proof for marginal probability

Solution Contd.

1. Now consider that X is normally distributed, we will establish Y is also normally distributed. The joint probability distribution is therefore

$$
\begin{aligned}
f_{XY}(x,y) &= f_{X|Y}(x|y)f_X(x) \\
&= f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y)) \qquad (15)
\end{aligned}
$$

The marginal probability distribution function for X is given as

$$\int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dy \qquad (16)$$

Using (5), we get

$$\int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dy = f_X(x) \qquad (17)$$

We know that $X \sim N(0,1)$, $f_X(x)$ represents gaussian probability distribution function.

Solution Contd.

1. Futher, using symmetry of (10), we can establish that marginal distribution of Y is gaussian. Here is a proof anyways

$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dx \tag{18}$$

Using (5), we get

$$f_Y(y) = \frac{1}{2}\left(f_X(y)+f_X(-y)\right) = f_X(y) \tag{19}$$

Since Y has identical probability distribution function, $Y \sim N(0,1)$

Solution Contd.

1. The covariance is given as

$$Cov(X, Y) = E[XY] - E[X]E[Y] = E[XY] \tag{20}$$

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x, y)\, dy\, dx \tag{21}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)\frac{1}{2}(\delta(x + y) + \delta(x - y))\, dy\, dx \tag{22}$$

$$= \int_{-\infty}^{\infty} xf_X(x) \int_{-\infty}^{\infty} y\frac{1}{2}(\delta(x + y) + \delta(x - y))\, dy\, dx \tag{23}$$

Using (5)

$$E[XY] = \int_{-\infty}^{\infty} xf_X(x)\frac{1}{2}(x - x)\, dx = 0 \tag{24}$$

Solution Contd.

2. Defining the following matrices/vectors

vector/matrix	expression
$\mathbf{Z}$	$\begin{pmatrix} X & Y \end{pmatrix}^\top$
$\mathbf{C}$	$\begin{pmatrix} a & b \end{pmatrix}^\top$
$\boldsymbol{\mu}$	$\begin{pmatrix} 0 & 0 \end{pmatrix}^\top$
$\boldsymbol{\Sigma}$	$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

Table: vectors/matrices and their expressions

Solution Contd.

2. Given

$$\boldsymbol{C}^\top \boldsymbol{Z} \sim N\left(0, a^2 + b^2\right) \tag{25}$$

Since this is true for all a and b, it is equivalent to X and Y being jointly gaussian

$$\boldsymbol{Z} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{26}$$

For correlated random variables X and Y in bivariate normal distribution, we have

$$\sigma_Z^2 = \sum_{i,j} \Sigma_{ij} \tag{27}$$

$$a^2 + b^2 = a^2 + b^2 + 2\rho ab \tag{28}$$

$$\therefore \rho = 0 \tag{29}$$

Solution Contd.

2. The joint distribution is given as

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{z}-\boldsymbol{\mu}\right)^{\top}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{z}-\boldsymbol{\mu}\right)\right)}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}} \tag{30}$$

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\begin{pmatrix}x & y\end{pmatrix} I_2 \begin{pmatrix}x & y\end{pmatrix}^{\top}\right)}{\sqrt{(2\pi)^2}} \tag{31}$$

Where I_2 is the identity matrix of order 2

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}x & y\end{pmatrix}^{\top}\right)}{\sqrt{(2\pi)^2}} \tag{32}$$

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right)}{\sqrt{(2\pi)^2}} = f_X(x)f_Y(y) \tag{33}$$

$\therefore$ **Option(2) is correct**

Solution Contd.

2. A simulation for bivariate gaussian is given below

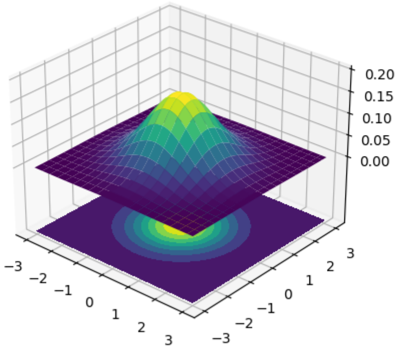

Figure: bivariate gaussian with 0 mean vector and identity covariance matrix

Solution Contd.

3. $$\Pr(X \leq 0, Y \leq 0) = \frac{1}{4} \tag{34}$$

This doesn't imply independence, it can be true even for dependent X and Y, the counter example is (15), the joint probability function is symmetric across all 4 quadrants

$$\therefore \Pr(X \leq 0, Y \leq 0) = \frac{1}{4} \tag{35}$$

Alternatively, here is proof

$$\Pr(X \leq 0) = F_X(0) = \frac{1}{2} \tag{36}$$

Using (10)

$$\Pr(Y \leq 0 | X \leq 0) = \frac{1}{2} \tag{37}$$

Solution Contd.

3. Using (36) and (37)

$$\Pr\left(X \le 0, Y \le 0\right) = \frac{1}{4} \tag{38}$$

4.

$$E\left[e^{itX+isY}\right] = E\left[e^{itX}\right]E\left[e^{isY}\right] \tag{39}$$

$$E\left[e^{itX+isY}\right] = \varphi_X(t)\varphi_Y(s) \tag{40}$$

Solution Contd.

4. The inverse is given as

$$f_{XY}(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-itX-isY} E\left[e^{itX+isY}\right] ds\, dt \tag{41}$$

Using (40)

$$f_{XY}(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-itX-isY} \varphi_X(t)\varphi_Y(s)\, ds\, dt \tag{42}$$

$$f_{XY}(x,y) = f_X(x) f_Y(y) \tag{43}$$

$\therefore$ **Option(4) is correct**

Reading material

1. Sum of 2 gaussian random variables need not be gaussian
   https://github.com/planetmath/62_Statistics/blob/master/pdf/62E15-SumsOfNormalRandomVariablesNeedNotBeNormal.pdf
2. Multivariate gaussian
   - https://bit.ly/2RComEB
   - https://en.wikipedia.org/wiki/Multivariate_normal_distribution
3. Characteristic function of multivariate gaussian
   https://bit.ly/3gbWZeF