STAT 830

The Multivariate Normal Distribution

In this section I present the basics of the multivariate normal distribution as an example to illustrate our distribution theory ideas.

Definition: A random variable $Z \in R^1$ has a standard normal distribution (we write $Z \sim N(0,1)$) if and only if Z has the density

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

Note: To see that this is a density let

$$I = \int_{-\infty}^{\infty} \exp(-u^2/2) du.$$

Then

$$\begin{aligned} I^2 &= \left\{ \int_{-\infty}^{\infty} \exp(-u^2/2) du. \right\}^2 \\ &= \left\{ \int_{-\infty}^{\infty} \exp(-u^2/2) du \right\} \left\{ \int_{-\infty}^{\infty} \exp(-v^2/2) dv \right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-(u^2+v^2)/2\} du dv \end{aligned}$$

Now do this integral in polar co-ordinates by the substitution $u = r\cos\theta$ and $v = r\sin\theta$ for $0 < r < \infty$ and $-\pi < \theta \le \theta$. The Jacobian is r and we get

$$\begin{aligned} I^2 &= \int_0^{\infty} \int_{-\pi}^{\pi} r \exp(-r^2/2) d\theta dr \\ &= 2\pi \int_0^{\infty} r \exp(-r^2/2) dr \\ &= -2\pi \exp(-r^2/2)\Big|_{r=0}^{\infty} \\ &= 2\pi. \end{aligned}$$

Thus

$$I = \sqrt{2\pi}.$$

Definition: A random vector $Z \in R^p$ has a standard multivariate normal distribution, written $Z \sim MVN(0, I)$ if and only if $Z = (Z_1, \ldots, Z_p)^t$ with the Z_i independent and each $Z_i \sim N(0, 1)$.

In this case according to our theorem **??**

$$\begin{aligned} f_Z(z_1, \ldots, z_p) &= \prod \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} \\ &= (2\pi)^{-p/2} \exp\{-z^t z/2\}\,; \end{aligned}$$

here, superscript t denotes matrix transpose.

Definition: $X \in R^p$ has a multivariate normal distribution if it has the same distribution as $AZ + \mu$ for some $\mu \in R^p$, some $p \times p$ matrix of constants A and $Z \sim MVN(0, I)$.

Remark: If the matrix A is singular then X does not have a density. This is the case for example for the residual vector in a linear regression problem.

Remark: If the matrix A is invertible we can derive the multivariate normal density by change of variables:

$$X = AZ + \mu \Leftrightarrow Z = A^{-1}(X - \mu)$$

$$\frac{\partial X}{\partial Z} = A \qquad \frac{\partial Z}{\partial X} = A^{-1}\,.$$

So

$$\begin{aligned} f_X(x) &= f_Z(A^{-1}(x - \mu))|\det(A^{-1})| \\ &= \frac{\exp\{-(x-\mu)^t (A^{-1})^t A^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}|\det A|}\,. \end{aligned}$$

Now define $\Sigma = AA^t$ and notice that

$$\Sigma^{-1} = (A^t)^{-1}A^{-1} = (A^{-1})^t A^{-1}$$

and

$$\det \Sigma = \det A \det A^t = (\det A)^2\,.$$

Thus f_X is

$$\frac{\exp\{-(x-\mu)^t \Sigma^{-1}(x-\mu)/2\}}{(2\pi)^{p/2}(\det \Sigma)^{1/2}}\,;$$

the $MVN(\mu, \Sigma)$ density. Note that this density is the same for all A such that $AA^t = \Sigma$. This justifies the usual notation $MVN(\mu, \Sigma)$.

Here is a question: for which μ, Σ is this a density? The answer is that this is a density for any μ but if $x \in R^p$ then

$$\begin{aligned} x^t \Sigma x &= x^t A A^t x \\ &= (A^t x)^t (A^t x) \\ &= \sum_1^p y_i^2 \geq 0 \end{aligned}$$

where $y = A^t x$. The inequality is strict except for $y = 0$ which is equivalent to $x = 0$. Thus Σ is a positive definite symmetric matrix.

Conversely, if Σ is a positive definite symmetric matrix then there is a square invertible matrix A such that $AA^t = \Sigma$ so that there is a $MVN(\mu, \Sigma)$ distribution. (This square root matrix A can be found via the Cholesky decomposition, e.g.)

When A is singular X will not have a density because $\exists a$ such that $P(a^t X = a^t \mu) = 1$; in this case X is confined to a hyperplane. A hyperplane has p dimensional volume 0 so no density can exist.

It is still true that the distribution of X depends only on $\Sigma = AA^t$: if $AA^t = BB^t$ then $AZ + \mu$ and $BZ + \mu$ have the same distribution. This can be proved using the characterization properties of moment generating functions.

I now make a list of three basic properties of the MVN distribution.

1. All margins of a multivariate normal distribution are multivariate normal. That is, if
$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$
$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$
and
$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$
then $X \sim MVN(\mu, \Sigma) \Rightarrow X_1 \sim MVN(\mu_1, \Sigma_{11})$.

2. All conditionals are normal: the conditional distribution of X_1 given $X_2 = x_2$ is $MVN(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$

3. If $X \sim MVN_p(\mu, \Sigma)$ then $MX + \nu \sim MVN(M\mu + \nu, M\Sigma M^t)$. We say that an affine transformation of a multivariate normal vector is normal.