M3S3/S4 STATISTICAL THEORY II

MULTIVARIATE NORMAL DISTRIBUTION: MARGINALS AND CONDITIONALS

Suppose that vector random variable $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)^\mathsf{T}$ has a multivariate normal distribution with pdf given by

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x}\right\} \tag{1}$$

where Σ is the $k \times k$ variance-covariance matrix (we can consider here the case where the expected value $\underset{\sim}{\mu}$ is the $k \times 1$ zero vector; results for the general case are easily available by transformation).

Consider partitioning $\underset{\sim}{X}$ into two components $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ of dimensions d and $k-d$ respectively, that is,

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{bmatrix}.$$

We attempt to deduce

(a) the marginal distribution of $\underset{\sim}{X}_1$, and

(b) the conditional distribution of $\underset{\sim}{X}_2$ given that $\underset{\sim}{X}_1 = \underset{\sim}{x}_1$.

First, write

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where Σ_{11} is $d \times d$, Σ_{22} is $(k-d)\times(k-d)$, $\Sigma_{21} = \Sigma_{12}^\mathsf{T}$, and

$$\Sigma^{-1} = V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

so that $\Sigma V = I_k$ (I_r is the $r \times r$ identity matrix) gives

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \begin{bmatrix} I_d & 0 \\ 0 & I_{k-d} \end{bmatrix}$$

and more specifically the four relations

$$\Sigma_{11}V_{11} + \Sigma_{12}V_{21} = I_d \tag{2}$$
$$\Sigma_{11}V_{12} + \Sigma_{12}V_{22} = 0 \tag{3}$$
$$\Sigma_{21}V_{11} + \Sigma_{22}V_{21} = 0 \tag{4}$$
$$\Sigma_{21}V_{12} + \Sigma_{22}V_{22} = I_{k-d}. \tag{5}$$

From the multivariate normal pdf in equation (1), we can re-express the term in the exponent as

$$\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x} = \underset{\sim}{x}_1^\mathsf{T}V_{11}\underset{\sim}{x}_1 + \underset{\sim}{x}_1^\mathsf{T}V_{12}\underset{\sim}{x}_2 + \underset{\sim}{x}_2^\mathsf{T}V_{21}\underset{\sim}{x}_1 + \underset{\sim}{x}_2^\mathsf{T}V_{22}\underset{\sim}{x}_2. \tag{6}$$

In order to compute the marginal and conditional distributions, we must complete the square in $\underset{\sim}{x}_2$ in this expression. We can write

$$\underset{\sim}{x}^\mathsf{T}\Sigma^{-1}\underset{\sim}{x} = (\underset{\sim}{x}_2 - \underset{\sim}{m})^\mathsf{T}M(\underset{\sim}{x}_2 - \underset{\sim}{m}) + \underset{\sim}{c} \tag{7}$$

and by comparing with equation (6) we can deduce that, for quadratic terms in $\underset{\sim}{x}_2$,

$$\underset{\sim}{x}_2^\mathsf{T}V_{22}\underset{\sim}{x}_2 = \underset{\sim}{x}_2^\mathsf{T}M\underset{\sim}{x}_2 \qquad \therefore \qquad M = V_{22} \tag{8}$$

for linear terms

$$\underset{\sim}{x}_2^\top V_{21}\underset{\sim}{x}_1 = -\underset{\sim}{x}_2^\top M\underset{\sim}{m} \qquad \therefore \qquad \underset{\sim}{m} = -V_{22}^{-1}V_{21}\underset{\sim}{x}_1 \tag{9}$$

and for constant terms

$$\underset{\sim}{x}_1^\top V_{11}\underset{\sim}{x}_1 = \underset{\sim}{c} + \underset{\sim}{m}^\top M\underset{\sim}{m} \qquad \therefore \qquad \underset{\sim}{c} = \underset{\sim}{x}_1^\top (V_{11} - V_{21}^\top V_{22}^{-1}V_{21})\underset{\sim}{x}_1 \tag{10}$$

thus yielding all the terms required for equation (7), that is

$$\underset{\sim}{x}^\top \Sigma^{-1}\underset{\sim}{x} = (\underset{\sim}{x}_2 + V_{22}^{-1}V_{21}\underset{\sim}{x}_1)^\top V_{22}(\underset{\sim}{x}_2 + V_{22}^{-1}V_{21}\underset{\sim}{x}_1) + \underset{\sim}{x}_1^\top (V_{11} - V_{21}^\top V_{22}^{-1}V_{21})\underset{\sim}{x}_1, \tag{11}$$

which, crucially, is a sum of two terms, where the first can be interpreted as a function of $\underset{\sim}{x}_2$, given $\underset{\sim}{x}_1$, and the second is a function of $\underset{\sim}{x}_1$ only.

Hence we have an immediate factorization of the full joint pdf using the chain rule for random variables;

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = f_{\underset{\sim}{X}_2|\underset{\sim}{X}_1}(\underset{\sim}{x}_2|\underset{\sim}{x}_1) f_{\underset{\sim}{X}_1}(\underset{\sim}{x}_1) \tag{12}$$

where

$$f_{\underset{\sim}{X}_2|\underset{\sim}{X}_1}(\underset{\sim}{x}_2|\underset{\sim}{x}_1) \propto \exp\left\{-\frac{1}{2}(\underset{\sim}{x}_2 + V_{22}^{-1}V_{21}\underset{\sim}{x}_1)^\top V_{22}(\underset{\sim}{x}_2 + V_{22}^{-1}V_{21}\underset{\sim}{x}_1)\right\} \tag{13}$$

giving that

$$\underset{\sim}{X}_2|\underset{\sim}{X}_1 = \underset{\sim}{x}_1 \;\sim N\left(-V_{22}^{-1}V_{21}\underset{\sim}{x}_1, V_{22}^{-1}\right) \tag{14}$$

and

$$f_{\underset{\sim}{X}_1}(\underset{\sim}{x}_1) \propto \exp\left\{-\frac{1}{2}\underset{\sim}{x}_1^\top (V_{11} - V_{21}^\top V_{22}^{-1}V_{21})\underset{\sim}{x}_1\right\} \tag{15}$$

giving that

$$\underset{\sim}{X}_1 \sim N\left(0, (V_{11} - V_{21}^\top V_{22}^{-1}V_{21})^{-1}\right). \tag{16}$$

But, from equation (3), $\Sigma_{12} = -\Sigma_{11}V_{12}V_{22}^{-1}$, and then from equation (2), substituting in Σ_{12},

$$\Sigma_{11}V_{11} - \Sigma_{11}V_{12}V_{22}^{-1}V_{21} = I_d \qquad \therefore \qquad \Sigma_{11} = (V_{11} - V_{12}V_{22}^{-1}V_{21})^{-1} = (V_{11} - V_{21}^\top V_{22}^{-1}V_{21})^{-1}.$$

Hence, by inspection of equation (16), we conclude that

$$\boxed{\underset{\sim}{X}_1 \sim N\left(0, \Sigma_{11}\right),} \tag{17}$$

that is, we can extract the Σ_{11} block of Σ to define the marginal variance-covariance matrix of $\underset{\sim}{X}_1$.

Using similar arguments, we can define the conditional distribution from equation (14) more precisely. First, from equation (3), $V_{12} = -\Sigma_{11}^{-1}\Sigma_{12}V_{22}$, and then from equation (5), substituting in V_{12}

$$-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}V_{22} + \Sigma_{22}V_{22} = I_{k-d} \qquad \therefore \qquad V_{22}^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{22} - \Sigma_{12}^\top\Sigma_{11}^{-1}\Sigma_{12}.$$

Finally, from equation (3), taking transposes on both sides, we have that $V_{21}\Sigma_{11} + V_{22}\Sigma_{21} = 0$. Then pre-multiplying by V_{22}^{-1}, and post-multiplying by Σ_{11}^{-1}, we have

$$V_{22}^{-1}V_{21} + \Sigma_{21}\Sigma_{11}^{-1} = 0 \qquad \therefore \qquad V_{22}^{-1}V_{21} = -\Sigma_{21}\Sigma_{11}^{-1},$$

so we have, substituting into equation (14), that

$$\boxed{\underset{\sim}{X}_2|\underset{\sim}{X}_1 = \underset{\sim}{x}_1 \;\sim N\left(\Sigma_{21}\Sigma_{11}^{-1}\underset{\sim}{x}_1, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right).} \tag{18}$$

Thus any marginal, and any conditional distribution of a multivariate normal joint distribution is also multivariate normal, as the choices of $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ are arbitrary.