THE MULTIVARIATE NORMAL DISTRIBUTION
MARGINAL AND CONDITIONALS DISTRIBUTIONS

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,

$$\underset{\sim}{X} = \left[\begin{array}{c} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{array}\right].$$

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

$$\Sigma = \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array}\right]$$

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

$$\Sigma^{-1} = V = \left[\begin{array}{cc} V_{11} & V_{12} \\ V_{21} & V_{22} \end{array}\right]$$

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

$$\left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array}\right]\left[\begin{array}{cc} V_{11} & V_{12} \\ V_{21} & V_{22} \end{array}\right] = \left[\begin{array}{cc} I_d & 0 \\ 0 & I_{k-d} \end{array}\right]$$

and more specifically the four relations

$$\begin{aligned} \Sigma_{11}V_{11} + \Sigma_{12}V_{21} &= I_d & (2) \\ \Sigma_{11}V_{12} + \Sigma_{12}V_{22} &= 0 & (3) \\ \Sigma_{21}V_{11} + \Sigma_{22}V_{21} &= 0 & (4) \\ \Sigma_{21}V_{12} + \Sigma_{22}V_{22} &= I_{k-d}. & (5) \end{aligned}$$

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 $\underline{x}_2$,

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

for linear terms

$$\underline{x}_2^\mathsf{T} V_{21} \underline{x}_1 = -\underline{x}_2^\mathsf{T} M \underline{m} \qquad \therefore \qquad \underline{m} = -V_{22}^{-1} V_{21} \underline{x}_1 \tag{9}$$

and for constant terms

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

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

$$\underline{x}^\mathsf{T} \Sigma^{-1} \underline{x} = (\underline{x}_2 + V_{22}^{-1} V_{21} \underline{x}_1)^\mathsf{T} V_{22} (\underline{x}_2 + V_{22}^{-1} V_{21} \underline{x}_1) + \underline{x}_1^\mathsf{T} (V_{11} - V_{21}^\mathsf{T} V_{22}^{-1} V_{21}) \underline{x}_1, \tag{11}$$

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

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

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

where

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

giving that

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

and

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

giving that

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

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

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

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

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

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

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}

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

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

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

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