Suppose that vector random variable $\mathbf{X} = (X_1, X_2, \dots, X_k)^{\top}$ has a multivariate normal distribution with pdf given by

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{\left|\Sigma\right|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{x}^{\top}\Sigma^{-1}\mathbf{x}\right\}$$
(1)

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

Consider partitioning **X** into two components \mathbf{X}_1 and \mathbf{X}_2 of dimensions k_1 and $k_2 = k - k_1$ respectively, that is, $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2]^{\top}$. We attempt to deduce

(a) the **marginal** distribution of X_1 , and

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

First, write

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

where Σ_{11} is $k_1 \times k_1$, Σ_{22} is $k_2 \times k_2$, $\Sigma_{21} = \Sigma_{12}^{\top}$, and

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

so that $\Sigma \mathbf{V} = \mathbf{I}_k$ (\mathbf{I}_r is the $r \times r$ identity matrix) gives

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k_2} \end{bmatrix}$$

where 0 represents the zero matrix of appropriate dimension. More specifically,

$$\Sigma_{11}\mathbf{V}_{11} + \Sigma_{12}\mathbf{V}_{21} = \mathbf{I}_{k_1} \tag{2}$$

$$\Sigma_{11}\mathbf{V}_{12} + \Sigma_{12}\mathbf{V}_{22} = \mathbf{0}$$

$$\Sigma_{21}\mathbf{V}_{11} + \Sigma_{22}\mathbf{V}_{21} = \mathbf{0}$$

$$\Sigma_{21}\mathbf{V}_{12} + \Sigma_{22}\mathbf{V}_{21} = \mathbf{0}$$
(4)

$$\sum_{i=1}^{21} \mathbf{v}_{11} + \sum_{i=2}^{22} \mathbf{v}_{21} = \mathbf{0} \tag{4}$$

$$\Sigma_{21} \mathbf{V}_{12} + \Sigma_{22} \mathbf{V}_{22} = \mathbf{I}_{k_2}.$$
 (5)

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

$$\mathbf{x}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{x} = \mathbf{x}_1^{\top} \mathbf{V}_{11} \mathbf{x}_1 + \mathbf{x}_1^{\top} \mathbf{V}_{12} \mathbf{x}_2 + \mathbf{x}_2^{\top} \mathbf{V}_{21} \mathbf{x}_1 + \mathbf{x}_2^{\top} \mathbf{V}_{22} \mathbf{x}_2.$$
(6)

In order to compute the marginal and conditional distributions, we must complete the square in x_2 in this expression. We can write

$$\mathbf{x}^{\top} \Sigma^{-1} \mathbf{x} = (\mathbf{x}_2 - \mathbf{m})^{\top} \mathbf{M} (\mathbf{x}_2 - \mathbf{m}) + \mathbf{c}$$
(7)

and by comparing with equation (6) we can deduce that, for quadratic terms in x_2 ,

$$\mathbf{x}_{2}^{\top}\mathbf{V}_{22}\mathbf{x}_{2} = \mathbf{x}_{2}^{\top}\mathbf{M}\mathbf{x}_{2} \qquad \therefore \qquad \mathbf{M} = \mathbf{V}_{22}$$

$$\tag{8}$$

for linear terms

$$\mathbf{x}_{2}^{\top}\mathbf{V}_{21}\mathbf{x}_{1} = -\mathbf{x}_{2}^{\top}\mathbf{M}\mathbf{m} \qquad \therefore \qquad \mathbf{m} = -\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_{1} \tag{9}$$

and for constant terms

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

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

$$\mathbf{x}^{\top} \Sigma^{-1} \mathbf{x} = (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1)^{\top} \mathbf{V}_{22} (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1) + \mathbf{x}_1^{\top} (\mathbf{V}_{11} - \mathbf{V}_{21}^{\top} \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{x}_1,$$
(11)

which, crucially, is a sum of two terms, where the first can be interpreted as a function of x_2 , given x_1 , and the second is a function of x_1 only.

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

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

where

$$f_{\mathbf{X}_{2}|\mathbf{X}_{1}}(\mathbf{x}_{2}|\mathbf{x}_{1}) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}_{2} + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_{1})^{\top}\mathbf{V}_{22}(\mathbf{x}_{2} + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_{1})\right\}$$
(13)

giving that

$$\mathbf{X}_{2}|\mathbf{X}_{1} = \mathbf{x}_{1} \sim \mathcal{N}_{k_{2}}\left(-\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_{1}, \mathbf{V}_{22}^{-1}\right)$$
(14)

and

$$f_{\mathbf{X}_1}(\mathbf{x}_1) \propto \exp\left\{-\frac{1}{2}\mathbf{x}_1^{\top}(\mathbf{V}_{11} - \mathbf{V}_{21}^{\top}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1\right\}$$
(15)

giving that

$$\mathbf{X}_{1} \sim \mathcal{N}_{k_{1}} \left(\mathbf{0}, (\mathbf{V}_{11} - \mathbf{V}_{21}^{\top} \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1} \right).$$
(16)

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

$$\Sigma_{11}\mathbf{V}_{11} - \Sigma_{11}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21} = \mathbf{I}_d \qquad \therefore \qquad \Sigma_{11} = (\mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1} = (\mathbf{V}_{11} - \mathbf{V}_{21}^{\top}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1}.$$

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

$$\mathbf{X}_{1} \sim \mathcal{N}_{k_{1}}\left(\mathbf{0}, \Sigma_{11}\right), \tag{17}$$

that is, we can extract the Σ_{11} block of Σ to define the marginal sigma matrix of \mathbf{X}_1 .

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

$$-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{V}_{22} + \Sigma_{22}\mathbf{V}_{22} = \mathbf{I}_{k-d} \qquad \therefore \qquad \mathbf{V}_{22}^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{22} - \Sigma_{12}^{\top}\Sigma_{11}^{-1}\Sigma_{12}.$$

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

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

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

$$\mathbf{X}_{2}|\mathbf{X}_{1}=\mathbf{x}_{1} \sim \mathcal{N}_{k_{2}}\left(\Sigma_{21}\Sigma_{11}^{-1}\mathbf{x}_{1}, \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right).$$
(18)

Thus any marginal, and any conditional distribution of a multivariate normal joint distribution is also multivariate normal, as the choices of X_1 and X_2 are arbitrary.