coefficient to be

$$R=\frac{(\sigma_{12}\Sigma_{22}^{-1}\sigma_{21})^{\frac{1}{2}}}{\sigma_{11}^{\frac{1}{2}}}\geqslant \mathrm{Corr}(X_1,\boldsymbol{\alpha}'\mathbf{X}_2),\quad 0\leqslant R\leqslant 1. \tag{15.8}$$

In the case where

$$\mathbf{X}=\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix},\quad \mathbf{X}_1: k\times 1,\quad \mathbf{X}: (n-k)\times 1,\quad \boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix},\quad k\geqslant 1,$$

we could define the r.v.'s $Z_1=\boldsymbol{\alpha}_1'\mathbf{X}_1$ and $Z_2=\boldsymbol{\alpha}_2'\mathbf{X}_2$ whose correlation coefficient is

$$\mathrm{Corr}(Z_1,Z_2)=\frac{\boldsymbol{\alpha}_1'\boldsymbol{\Sigma}_{12}\boldsymbol{\alpha}_2}{(\boldsymbol{\alpha}_1'\boldsymbol{\Sigma}_{11}\boldsymbol{\alpha}_1)^{\frac{1}{2}}(\boldsymbol{\alpha}_2'\boldsymbol{\Sigma}_{22}\boldsymbol{\alpha}_2)^{\frac{1}{2}}}. \tag{15.9}$$

From the above inequality it follows that for $\boldsymbol{\alpha}_2=\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\alpha}_1$

$$C_{12}=\frac{(\boldsymbol{\alpha}_1'\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\alpha}_1)^{\frac{1}{2}}}{(\boldsymbol{\alpha}_1'\boldsymbol{\Sigma}_{11}\boldsymbol{\alpha}_1)^{\frac{1}{2}}}\geqslant \mathrm{Corr}(Z_1,Z_2). \tag{15.10}$$

$\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ has at most k non-zero eigenvalues which measure the association between $\mathbf{X}_1$ and $\mathbf{X}_2$ and are called *canonical correlations.*

Let

$$\mathbf{X}=\begin{pmatrix}X_1\\ X_2\\ \mathbf{X}_3\end{pmatrix}\quad \text{where } \mathbf{X}_3:(n-2)\times 1$$

$$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_{11} & \sigma_{12} & \boldsymbol{\sigma}_{13}\\ \sigma_{21} & \sigma_{22} & \boldsymbol{\sigma}_{23}\\ \boldsymbol{\sigma}_{31} & \boldsymbol{\sigma}_{32} & \boldsymbol{\Sigma}_{33}\end{pmatrix}.$$

Another form of correlation of interest in this context is the correlation between X_1 and X_2 given that the effect of $\mathbf{X}_3$ is taken away. For this we form the r.v.'s

$$Y_1=X_1-\mathbf{b}_1'\mathbf{X}_3\quad\text{and}\quad Y_2=X_2-\mathbf{b}_2'\mathbf{X}_3\quad\text{and}\quad \mathrm{Corr}(Y_1,Y_2)$$

is maximised by $\mathbf{b}_1=\boldsymbol{\Sigma}_{33}^{-1}\boldsymbol{\sigma}_{31}$ and $\mathbf{b}_2=\boldsymbol{\Sigma}_{33}^{-1}\boldsymbol{\sigma}_{32}$ as seen above. Hence we define the *partial correlation coefficient* between X_1 and X_2 given $\mathbf{X}_3$ to be

$$\rho_{12.3}=\frac{\sigma_{12}-\boldsymbol{\sigma}_{13}\boldsymbol{\Sigma}_{33}^{-1}\boldsymbol{\sigma}_{32}}{[\sigma_{11}-\boldsymbol{\sigma}_{13}\boldsymbol{\Sigma}_{33}^{-1}\boldsymbol{\sigma}_{31}]^{\frac{1}{2}}[\sigma_{22}-\boldsymbol{\sigma}_{23}\boldsymbol{\Sigma}_{33}^{-1}\boldsymbol{\sigma}_{32}]^{\frac{1}{2}}}\geqslant \mathrm{Corr}(Y_1,Y_2). \tag{15.11}$$

15.2 The multivariate normal distribution

The univariate normal density function discussed above was of the form

$$f(x;\mu,\sigma^2)=(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}. \tag{15.12}$$

The density function of $\mathbf{X}\equiv(X_1,X_2,\ldots,X_n)'$ when the X_is are IID normally distributed r.v.'s was shown to be of the form

$$f(\mathbf{x};\mu,\sigma^2)=\prod_{i=1}^{n} f(x_i;\mu,\sigma^2)$$

$$=(2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}. \tag{15.13}$$

Similarly, the density function of $\mathbf{X}$ when the X_is are only independent, i.e. $X_i\sim N(\mu_i,\sigma_i^2)$, $i=1,2,\ldots,n$, takes the form

$$f(\mathbf{x};\mu_1,\ldots,\mu_n,\sigma_1^2,\ldots,\sigma_n^2)=\prod_{i=1}^{n} f(x_i;\mu_i,\sigma_i^2)$$

$$=(2\pi)^{-n/2}(\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right\}. \tag{15.14}$$

Comparing the above three density functions we can discern a pattern developing which is very suggestive for the density function of an arbitrary normal vector $\mathbf{X}$ with $E(\mathbf{X})=\boldsymbol{\mu}$ and $\mathrm{Cov}(\mathbf{X})=\boldsymbol{\Sigma}$, which takes the form

$$f(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})=(2\pi)^{-n/2}(\det\boldsymbol{\Sigma})^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\}, \tag{15.15}$$

and we write $\mathbf{X}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$. If the X_is are IID r.v.'s $\boldsymbol{\Sigma}=\sigma^2\mathbf{I}_n$ and $(\det\boldsymbol{\Sigma})=(\sigma^2)^n$. On the other hand, if the X_is are independent but not identically distributed

$$\boldsymbol{\Sigma}=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_n^2)\quad\text{and}\quad(\det\boldsymbol{\Sigma})=\prod_{i=1}^{n}(\sigma_i^2)=(\sigma_1^2\cdots\sigma_n^2).$$

In the case of $n=2$

$$\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\quad \boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},\quad \text{where } \rho=\frac{\sigma_{12}}{\sigma_1\sigma_2},$$

$$\Rightarrow (\det\boldsymbol{\Sigma})=\sigma_1^2\sigma_2^2(1-\rho^2)>0\quad\text{for } -1<\rho<1$$

and

$$\Sigma^{-1}=\frac{1}{(1-\rho^2)}\begin{pmatrix}\frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2}\\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{pmatrix}. \tag{15.16}$$

Thus the bivariate normal density function is

$$f(x_1,x_2;\,\boldsymbol{\mu},\Sigma)=\frac{[\sigma_1^2\sigma_2^2(1-\rho^2)]^{-\frac{1}{2}}}{2\pi}\times\exp\left\{-\frac{(1-\rho^2)^{-1}}{2}\times\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \tag{15.17}$$

(see Chapter 6). The standard bivariate density function can be obtained by defining the new r.v.'s

$$Z_i=\left(\frac{X_i-\mu_i}{\sigma_i}\right),\quad i=1,2,$$

whose density function is

$$f(z_1,z_2;\,\rho)=\frac{(1-\rho^2)^{-\frac{1}{2}}}{2\pi}\exp\left\{-\frac{(1-\rho^2)^{-1}}{2}(z_1^2-2\rho z_1z_2+z_2^2)\right\}. \tag{15.18}$$

(1) ***Properties***

(*N1*) *Let* $\mathbf{X}\sim N(\boldsymbol{\mu},\Sigma)$ *then* $\mathbf{Y}=(\mathbf{AX}+\mathbf{b})\sim N(\mathbf{A}\boldsymbol{\mu}+\mathbf{b},\,\mathbf{A}\Sigma\mathbf{A}')$ *for* $\mathbf{A}$: $m\times n$ *and* $\mathbf{b}$: $m\times 1$ *constant matrices, e.g. if* $\mathbf{Y}=c\mathbf{X}$, $c\neq 0$, $\mathbf{Y}\sim N(c\boldsymbol{\mu},c^2\Sigma)$.

This property shows that if $\mathbf{X}$ is normally distributed then any *linear function* of $\mathbf{X}$ is also normally distributed.

(*N2*) *Let* $\mathbf{X}_t\sim N(\boldsymbol{\mu}_t,\Sigma_t)$, $t=1,2,\ldots,T$, *be independently distributed random vectors, then for any arbitrary constant matrices* $\mathbf{A}_t$, $t=1,2,\ldots,T$,

$$\left(\sum_{t=1}^{T}\mathbf{A}_t\mathbf{X}_t\right)\sim N\left(\sum_{t=1}^{T}\mathbf{A}_t\boldsymbol{\mu}_t,\ \sum_{t=1}^{T}(\mathbf{A}_t\Sigma\mathbf{A}_t')\right).$$

The converse also holds. If the $\mathbf{X}_t$s are IID then $\boldsymbol{\mu}_t=\boldsymbol{\mu}$, $\Sigma_t=\Sigma$, $t=1,2,\ldots,T$, and

$$\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_t\right)\sim N\left(\boldsymbol{\mu},\frac{1}{T}\Sigma\right).$$

(*N3*) *Let* $\mathbf{X}\sim N(\boldsymbol{\mu},\Sigma)$ *then the* X_i*s are* **independent** *if and only if* $\sigma_{ij}=0$, $i\neq j$, $i,j=1,2,\ldots,n$, *i.e.* $\Sigma=\text{diag}(\sigma_{11},\ldots,\sigma_{nn})$. *In general, zero covariance does not imply independence but in the case of normality the two are equivalent.*

(*N4*) *If* $\mathbf{X}\sim N(\boldsymbol{\mu},\Sigma)$ *then the* **marginal distribution** *of any* $k\times 1$ *subset* $\mathbf{X}_1$ *where*

$$\mathbf{X}\equiv\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix},\quad \boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix},\quad \Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix},$$

$\mathbf{X}_1\sim N(\boldsymbol{\mu}_1,\Sigma_{11})$. *This follows from property N1 for* $\mathbf{A}=(\mathbf{I}_k\colon 0)$, $(k\times n)$, $\mathbf{b}=\mathbf{0}$. *Similarly,* $\mathbf{X}_2\sim N(\boldsymbol{\mu}_2,\Sigma_{22})$.

These can be verified directly using

$$f(\mathbf{x}_1;\boldsymbol{\theta}_1)=\int f(\mathbf{x};\boldsymbol{\theta})\,\mathrm{d}\mathbf{x}_2\quad\text{and}\quad f(\mathbf{x}_2;\boldsymbol{\theta}_2)=\int f(\mathbf{x};\boldsymbol{\theta})\,\mathrm{d}\mathbf{x}_1,$$

although the manipulations involved are rather too cumbersome. Taking $k=1$, this property implies that each component of $\mathbf{X}\sim N(\boldsymbol{\mu},\Sigma)$ is also normally distributed; the converse, however, is not true.

(*N5*) *For the same partition of* $\mathbf{X}$ *considered in N4 the* **conditional distribution** *of* $\mathbf{X}_1$ *given* $\mathbf{X}_2$ *takes the form*

$$(\mathbf{X}_1/\mathbf{X}_2),\quad N(\boldsymbol{\mu}_1+\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2),\quad \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}).$$

This follows from property N1 for

$$\mathbf{A}=\begin{pmatrix}\mathbf{I}_k & -\Sigma_{12}\Sigma_{22}^{-1}\\ 0 & \mathbf{I}_{n-k}\end{pmatrix}\quad\text{and}\quad\mathbf{b}=\mathbf{0} \tag{15.19}$$

since

$$\mathbf{AX}=\begin{pmatrix}\mathbf{X}_1-\Sigma_{12}\Sigma_{22}^{-1}\mathbf{X}_2\\ \mathbf{X}_2\end{pmatrix},\quad \text{Cov}(\mathbf{AX})=\begin{pmatrix}\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0\\ 0 & \Sigma_{22}\end{pmatrix}. \tag{15.20}$$

From this we can deduce that if $\Sigma_{12}=0$ then $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent since $(\mathbf{X}_1/\mathbf{X}_2)\sim N(\boldsymbol{\mu}_1,\Sigma_{11})$. Moreover, for any Σ_{12}, $(\mathbf{X}_1-\Sigma_{12}\Sigma_{22}^{-1}\mathbf{X}_2)$ and $\mathbf{X}_2$ are independent given that their covariance is zero. Similarly, $(\mathbf{X}_2/\mathbf{X}_1)\sim N(\boldsymbol{\mu}_2+\Sigma_{21}\Sigma_{11}^{-1}(\mathbf{X}_1-\boldsymbol{\mu}_1),\ \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$. In the case $n=2$

$$(X_1/X_2)\sim N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(X_2-\mu_2),\ \sigma_1^2(1-\rho^2)\right). \tag{15.21}$$

These results can be verified using the formula

$$f(\mathbf{x}_1/\mathbf{x}_2;\,\boldsymbol{\phi})=\frac{f(\mathbf{x};\,\boldsymbol{\theta})}{f(\mathbf{x}_2;\,\boldsymbol{\theta}_2)}. \tag{15.22}$$