L E C T U R E 10

Theorem 10.1. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be mutually independent with $\mathbf{X}_j$ distributed as $N_m(\mu_j, \Sigma)$ (Note that each X_j has the same covariance matrix Σ.) Then

$$\mathbf{V}_1 = c_1\mathbf{X}_1 + c_2\mathbf{X}_2 + \cdots + c_n\mathbf{X}_n$$

is distributed as $N_m(\sum_{j=1}^n c_j\mu_j, (\sum_{j=1}^n c_j^2)\Sigma)$. Moreover, $\mathbf{V}_1$ and $\mathbf{V}_2 = b_1\mathbf{X}_l + b_2\mathbf{X}_2 + \cdots + b_n\mathbf{X}_n$ are jointly multivariate normal with covariance matrix

$$\begin{bmatrix} (\sum\limits_{j=1}^n c_j^2)\Sigma & (\mathbf{b}'c)\Sigma \\ (\mathbf{b}'c)\Sigma & (\sum\limits_{j=1}^n b_j^2)\Sigma \end{bmatrix}$$

Consequently, $\mathbf{V}_1$ and $\mathbf{V}_2$ are independent if $\mathbf{b}'c = \sum\limits_{j=1}^n c_j b_j = 0$.

Proof. The nm component vector

$$[X_{11}, \ldots, X_{1m}, X_{12}, \ldots, X_{2m}, \ldots, X_{nm}] = [\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_n'] = \underbrace{\mathbf{X}'}_{(1\times nm)}$$

is multivariate normal. In particular, $\underbrace{\mathbf{X}}_{(nm\times 1)}$ is distributed as $N_{nm}(\mu, \Sigma_{\mathbf{X}})$, where

$$\underbrace{\mu}_{(nm\times 1)} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \quad \text{and} \quad \underbrace{\Sigma_{\mathbf{X}}}_{(nm\times nm)} = \begin{bmatrix} \Sigma & 0 & \ldots & 0 \\ 0 & \Sigma & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \Sigma \end{bmatrix}$$

The choice

$$\underbrace{\mathbf{A}}_{(2m\times nm)} = \begin{bmatrix} c_1\mathbf{I} & c_2\mathbf{I} & \ldots & c_n\mathbf{I} \\ b_1\mathbf{I} & b_2\mathbf{I} & \ldots & b_n\mathbf{I} \end{bmatrix}$$

where $\mathbf{I}$ is the $m \times m$ identity matrix, gives

$$\mathbf{AX} = \begin{bmatrix} \sum\limits_{j=1}^n c_j\mathbf{X}_j \\ \sum\limits_{j=1}^n b_j\mathbf{X}_j \end{bmatrix} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix}$$

and $\mathbf{AX}$ is normal $N_{2m}(\mathbf{A}\mu, \mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}')$. Straightforward block multiplication shows that $\mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}'$ has the first block diagonal term

$$[c_1\Sigma, c_2\Sigma, \ldots, c_n\Sigma][c_1\mathbf{I}, c_2\mathbf{I}, \ldots, c_n\mathbf{I}]' = \left(\sum_{j=1}^n c_j^2\right)\Sigma.$$

The off-diagonal term is

$$[c_1\Sigma, c_2\Sigma, \dots, c_n\Sigma][b_1\mathbf{I}, b_2\mathbf{I}, \dots, b_n\mathbf{I}]' = \left(\sum_{j=1}^{n} c_j b_j\right)\Sigma.$$

This term is the covariance matrix for $\mathbf{V}_1, \mathbf{V}_2$. Consequently, when $\sum_{j=1}^{n} c_j b_j = b'c = 0$, so that $\sum_{j=1}^{n} c_j b_j)\Sigma = 0$, $\mathbf{V}_1$ and $\mathbf{V}_2$ are independent.

Therefore, the property of zero correlation is equivalent to requiring the coefficient vectors $\mathbf{b}$ and $\mathbf{c}$ to be perpendicular.

Example 10.1. (Linear combinations of random vectors). Let $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and $\mathbf{X}_4$ be independent and identically distributed 3×1 random vectors with

$$\mu = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

We first consider a linear combination $\mathbf{a}'\mathbf{X}_1$ of the three components of $\mathbf{X}_1$. This is a random variable with mean

$$\mathbf{a}'\mu = 3a_1 - a_2 + a_3$$

and variance

$$\mathbf{a}'\Sigma\mathbf{a} = 3a_1^2 + a_2^2 + 2a_3^2 - 2a_1a_2 + 2a_1a_3.$$

That is, a linear combination $\mathbf{a}'\mathbf{X}_1$ of the components of a random vector is a single random variable consisting of a sum of terms that are each a constant times a variable. This is very different from a linear combination of random vectors, say,

$$c_1\,\mathbf{X}_1 + c_2\,\mathbf{X}_2 + c_3\,\mathbf{X}_3 + c_4\,\mathbf{X}_4$$

which is itself a random vector. Here each term in the sum is a constant times a random vector.

Now consider two linear combinations of random vectors

$$\frac{1}{2}\mathbf{X}_1 + \frac{1}{2}\mathbf{X}_2 + \frac{1}{2}\mathbf{X}_3 + \frac{1}{2}\mathbf{X}_4$$

and

$$\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3 - 3\mathbf{X}_4.$$

Find the mean vector and covariance matrix for each linear combination of vectors and also the covariance between them.

By Theorem 10.1 with $c_1 = c_2 = c_3 = c_4 = 1/2$, the first linear combination has mean vector

$$(c_1 + c_2 + c_3 + c_4)\mu = 2\,\mu = \begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix}$$

and covariance matrix

$$(c_1^2 + c_2^2 + c_3^2 + c_4^2)\Sigma = 1 \times \Sigma = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

For the second linear combination of random vectors, we apply Result 4.8 with $b_1 = b_2 = b_3 = 1$ and $b_4 = -3$ to get mean vector

$$(b_1 + b_2 + b_3 + b_4)\mu = 0\,\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and covariance matrix

$$(b_1^2 + b_2^2 + b_3^2 + b_4^2)\Sigma = 12 \times \Sigma = \begin{bmatrix} 36 & -12 & 12 \\ -12 & 12 & 0 \\ 12 & 0 & 24 \end{bmatrix}$$

Finally, the covariance matrix for the two linear combinations of random vectors is

$$(c_1 b_1 + c_2 b_2 + c_3 b_3 + c_4 b_4)\Sigma = 0\Sigma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Every component of the first linear combination of random vectors has zero covariance with every component of the second linear combination of random vectors.

If, in addition, each $\mathbf{X}$ has a trivariate normal distribution, then the two linear combinations have a joint six-variate normal distribution, and the two linear combinations of vectors are independent.

§10.1. BIVARIATE NORMAL DISTRIBUTION

Up to this point all of the random variables have been of one dimensions. A very important two–dimensional probability law which is a generalization of the one–dimensional normal probability law is called *Bivariate normal distribution*.

The random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ are said to have a Bivariate normal distribution with parameters $(a_1, a_2, \sigma_1, \sigma_2, \rho)$ if their joint density function is given by

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ \times \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1-a_1}{\sigma_1}\right)^2 - 2\rho \cdot \frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-a_2}{\sigma_2}\right)^2 \right] \right\}. \tag{104}$$

We see that Bivariate normal distribution is determined by 5 parameters. These are a_1, a_2, σ_1, σ_2 and ρ such that $a_1 \in (-\infty, +\infty)$, $a_2 \in (-\infty, +\infty)$, $\sigma_1 > 0$, $\sigma_2 > 0$ and $\rho \in (-1, +1)$.

By (51) the density function of η_1 is defined by the following formula

$$f_{\eta_1}(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_2 = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{(x_1-a_1)^2}{2\sigma_1^2} \right).$$

Similarly, for random variable η_2, we obtain:

$$f_{\eta_2}(x_2) = \int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_1 = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{(x_2-a_2)^2}{2\sigma_2^2} \right).$$

Therefore, $\eta_1(\omega)$ and $\eta_2(\omega)$ are both normal random variables with respective parameters $\mathcal{N}(a_1, \sigma_1)$ and $\mathcal{N}(a_2, \sigma_2)$.

Thus, the marginal distributions of $\eta_1(\omega)$ and $\eta_2(\omega)$ are both normal, even though the joint distributions for $\rho = 0$, $\rho = 0.3$, $\rho = 0.6$ and $\rho = 0.9$ are quite different from each other.

It is not difficult to verify that

$$\mathrm{Cov}\,(\eta_1, \eta_2) = E\,[(\eta_1 - a_1) \cdot (\eta_2 - a_2)] = \rho\,\sigma_1\sigma_2,$$

i. e. covariance between $\eta_1(\omega)$ and $\eta_2(\omega)$ is equal to $\rho\,\sigma_1\,\sigma_2$ and therefore

$$\rho(\eta_1, \eta_2) = \rho.$$

Thus, we have proved that

$$\eta_1 \quad \text{and} \quad \eta_2 \quad \text{are independent} \quad \text{if and only if the correlation is equal to } 0.$$

Indeed, substituting $\rho = 0$ in (104) we obtain

$$f(x_1, x_2) = f_{\eta_1}(x_1) \cdot f_{\eta_2}(x_2).$$

Definition 10.1. If $\eta_1(\omega)$ and $\eta_2(\omega)$ have a joint density function $f(x_1, x_2)$, then the conditional density function of $\eta_1(\omega)$, given that $\eta_2(\omega) = x_2$ is defined for all values of x_2 such that $f_{\eta_2}(x_2) \neq 0$, by

$$f_{\eta_1/\eta_2}(x_1|x_2) = \frac{f(x_1, x_2)}{f_{\eta_2}(x_2)}.$$

To motivate this definition, multiply the left–hand side by dx_1 and the right–hand side by $(dx_1 \cdot dx_2)/(dx_2)$ to obtain

$$f_{\eta_1/\eta_2}(x_1|x_2)dx_1 = \frac{f(x_1, x_2)\, dx_1\, dx_2}{f_{\eta_2}(x_2)\, dx_2} \approx \frac{P\{\omega\colon\ x_1 \leq \eta_1(\omega) \leq x_1 + dx_1, x_2 \leq \eta_2(\omega) \leq x_2 + dx_2\}}{P\{\omega\colon\ x_2 \leq \eta_2(\omega) \leq x_2 + dx_2\}} =$$
$$= P\{\omega\colon\ x_1 \leq \eta_1(\omega) \leq x_1 + dx_1 | x_2 \leq \eta_2(\omega) \leq x_2 + dx_2\}.$$

In other words, for small values of dx_1 and dx_2, $f_{\eta_1/\eta_2}(x_1|x_2)$ represents the conditional probability that $\eta_1(\omega)$ is between x_1 and $x_1 + dx_1$, given that $\eta_2(\omega)$ is between x_2 and $x_2 + dx_2$.

One can show that the conditional density of η_1, given that $\eta_2 = x_2$, is the normal density with parameters

$$\mathcal{N}\left(a_1 + \rho\frac{\sigma_1}{\sigma_2}\,(x_2 - a_2), \sigma_1^2\,(1 - \rho^2)\right).$$

Similarly, the conditional density of η_2, given that $\eta_1 = x_1$, is the normal density with parameters

$$\mathcal{N}\left(a_2 + \rho\frac{\sigma_2}{\sigma_1}\,(x_1 - a_1), \sigma_2^2\,(1 - \rho^2)\right).$$