Random Vectors
and the
Variance–Covariance Matrix

**Definition 1.** A *random vector* $\vec{X}$ is a vector $(X_1, X_2, \ldots, X_p)$ of jointly distributed random variables. As is customary in linear algebra, we will write vectors as column matrices whenever convenient.

**Expectation**

**Definition 2.** The expectation $E\vec{X}$ of a random vector $\vec{X} = [X_1, X_2, \ldots, X_p]^T$ is given by

$$E\vec{X} = \begin{bmatrix} E X_1 \\ E X_2 \\ \vdots \\ E X_p \end{bmatrix}.$$  

This is a definition, but it is chosen to merge well with the linear properties of the expectation, so that, for example:

$$E\vec{X} = E\begin{bmatrix} X_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + E\begin{bmatrix} 0 \\ X_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + E\begin{bmatrix} 0 \\ 0 \\ \vdots \\ X_p \end{bmatrix}$$

$$= \begin{bmatrix} E X_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ E X_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ E X_p \end{bmatrix}$$

$$= \begin{bmatrix} E X_1 \\ E X_2 \\ \vdots \\ E X_p \end{bmatrix}.$$
The linearity properties of the expectation can be expressed compactly by stating that for any \( k \times p \)-matrix \( A \) and any \( 1 \times j \)-matrix \( B \),

\[
E(AX) = AE\vec{X} \quad \text{and} \quad E(\vec{X}B) = (E\vec{X})B.
\]

**The Variance–Covariance Matrix**

**Definition 3.** The **variance–covariance matrix** (or simply the **covariance matrix**) of a random vector \( \vec{X} \) is given by:

\[
\text{Cov}(\vec{X}) = E\left[ (\vec{X} - E\vec{X})(\vec{X} - E\vec{X})^T \right].
\]

**Proposition 4.**

\[
\text{Cov}(\vec{X}) = E[\vec{X}\vec{X}^T] - E\vec{X}(E\vec{X})^T.
\]

**Proposition 5.**

\[
\text{Cov}(\vec{X}) =
\begin{bmatrix}
\text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_p) \\
\text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_p) \\
\vdots & \vdots & \ddots & \vdots \\
\text{Cov}(X_p, X_1) & \text{Cov}(X_p, X_2) & \cdots & \text{Var}(X_p)
\end{bmatrix}.
\]

Thus, \( \text{Cov}(\vec{X}) \) is a symmetric matrix, since \( \text{Cov}(X,Y) = \text{Cov}(Y,X) \).

**Exercise 1.** Prove Propositions 4 and 5.

**Linear combinations of random variables**

Consider random variables \( X_1, \ldots, X_p \). We want to find the expectation and variance of a new random variable \( L(X_1, \ldots, X_p) \) obtained as a linear combination of \( X_1, \ldots, X_p \); that is,

\[
L(X_1, \ldots, X_p) = \sum_{i=1}^{p} a_i X_i.
\]

Using vector–matrix notation we can write this in a compact way:

\[
L(\vec{X}) = \vec{a}^T \vec{X},
\]
where $\vec{a}^T = [a_1, \ldots, a_p]$. Then we get:
\[
E[L(\vec{X})] = E[\vec{a}^T \vec{X}] = \vec{a}^T \overline{E\vec{X}}
\]
and
\[
\text{Var}[L(\vec{X})] = E[\vec{a}^T \vec{X} \vec{X}^T \vec{a}] - \overline{E(\vec{a}^T \vec{X})}[\overline{E(\vec{a}^T \vec{X})}]^T
\]
\[
= \vec{a}^T \overline{E[\vec{X} \vec{X}^T]} \vec{a} - \vec{a}^T \overline{E\vec{X}(E\vec{X})^T} \vec{a}
\]
\[
= \vec{a}^T \left( \overline{E[\vec{X} \vec{X}^T]} - \overline{E\vec{X}(E\vec{X})^T} \right) \vec{a}
\]
\[
= \vec{a}^T \overline{\text{Cov}(\vec{X})} \vec{a}
\]
Thus, knowing $E\vec{X}$ and $\text{Cov}(\vec{X})$, we can easily find the expectation and variance of any linear combination of $X_1, \ldots, X_p$.

**Corollary 6.** If $\Sigma$ is the covariance matrix of a random vector, then for any constant vector $\vec{a}$ we have
\[
\vec{a}^T \Sigma \vec{a} \geq 0.
\]
That is, $\Sigma$ satisfies the property of being a positive semi-definite matrix.

**Proof.** $\vec{a}^T \Sigma \vec{a}$ is the variance of a random variable. \qed

This suggests the question: Given a symmetric, positive semi-definite matrix, is it the covariance matrix of some random vector? The answer is yes.

**Exercise 2.** Consider a random vector $\vec{X}$ with covariance matrix $\Sigma$. Then, for any $k$ dimensional constant vector $\vec{c}$ and any $p \times k$-matrix $A$, the $k$-dimensional random vector $\vec{c} + A^T \vec{X}$ has mean $\vec{c} + A^T \overline{E\vec{X}}$ and has covariance matrix
\[
\text{Cov}(\vec{c} + A^T \vec{X}) = A^T \Sigma A.
\]

**Exercise 3.** If $X_1, X_2, \ldots, X_p$ are i.i.d. (independent identically distributed), then $\text{Cov}([X_1, X_2, \ldots, X_p]^T)$ is the $p \times p$ identity matrix, multiplied by a non-negative constant.

**Theorem 7** (Classical result in Linear Algebra). If $\Sigma$ is a symmetric, positive semi-definite matrix, there exists a matrix $\Sigma^{1/2}$ (not unique) such that
\[
(\Sigma^{1/2})^T \Sigma^{1/2} = \Sigma.
\]

**Exercise 4.** Given a symmetric, positive semi-definite matrix $\Sigma$, find a random vector with covariance matrix $\Sigma$. 

3
The Multivariate Normal Distribution

A $p$-dimensional random vector $\vec{X}$ has the multivariate normal distribution if it has the density function

$$f(\vec{X}) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp \left( -\frac{1}{2} (\vec{X} - \vec{\mu})^T \Sigma^{-1} (\vec{X} - \vec{\mu}) \right),$$

where $\vec{\mu}$ is a constant vector of dimension $p$ and $\Sigma$ is a $p \times p$ positive semi-definite which is invertible (called, in this case, positive definite). Then, $E\vec{X} = \vec{\mu}$ and $\text{Cov}(\vec{X}) = \Sigma$.

The standard multivariate normal distribution is obtained when $\vec{\mu} = 0$ and $\Sigma = I_p$, the $p \times p$ identity matrix:

$$f(\vec{X}) = (2\pi)^{-p/2} \exp \left( -\frac{1}{2} \vec{X}^T \vec{X} \right).$$

This corresponds to the case where $X_1, \ldots, X_p$ are i.i.d. standard normal.

**Exercise 5.** Let $X_1$ and $X_2$ be random variables with standard deviation $\sigma_1$ and $\sigma_2$, respectively, and with correlation $\rho$. Find the variance–covariance matrix of the random vector $[X_1, X_2]^T$.

**Exercise 6** (The bivariate normal distribution). Consider a 2-dimensional random vector $\vec{X}$ distributed according to the multivariate normal distribution (in this case called, for obvious reasons, the bivariate normal distribution). Starting with the formula for the density in matrix notation, derive the formula for the density of $\vec{X}$ depending only on $\mu_1, \mu_2$ (the means of $X_1$ and $X_2$), $\sigma_1, \sigma_2$ (the standard deviations of $X_1$ and $X_2$), and the correlation coefficient $\rho$, and write it out without using matrix notation.

**Exercise 7.** Consider a bivariate normal random vector $\vec{X} = [X_1, X_2]^T$, where $E\vec{X} = [5, -4]^T$, the standard deviations are $\text{StDev}(X_1) = 2$ and $\text{StDev}(X_2) = 3$, and the correlation coefficient of $X_1$ and $X_2$ is $-4/5$. Use $R$ (or any other software package) to generate 100 independent draws of $\vec{X}$, and plot them as points on the plane.

*Hint: To find $\Sigma^{1/2}$, find the eigenvalue decomposition of $\Sigma$ as:*

$$\Sigma = PDP^T,$$
where $D$ is diagonal. Construct $D^{1/2}$ by taking the square root of each diagonal entry, and define

$$\Sigma^{1/2} = PD^{1/2}P^T.$$ 

In R, you can find the eigenvalue decomposition of $\Sigma$ using:

```r
ed <- eigen(sigma)
D <- diag(ed$values)
P <- ed$vectors
```