Random vectors and properties

Mean and variance of a random vector

Let $Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$ be a random vector with $EY = \begin{pmatrix} EY_1 \\ EY_2 \\ \vdots \\ EY_n \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \mu$. The variance covariance matrix of $Y$ denoted with $\Sigma = \text{var}(Y)$ is defined as follows:

$$\text{var}(Y) = E(Y - \mu)(Y - \mu)^t$$

$$= E\begin{pmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \\ \vdots \\ Y_n - \mu_n \end{pmatrix}(Y_1 - \mu_1, Y_2 - \mu_2, \ldots, Y_n - \mu_n)$$

$$= E\begin{pmatrix} (Y_1 - \mu_1)^2 & (Y_1 - \mu_1)(Y_2 - \mu_2) & \ldots & (Y_1 - \mu_1)(Y_n - \mu_n) \\ (Y_2 - \mu_2)(Y_1 - \mu_1) & (Y_2 - \mu_2)^2 & \ldots & (Y_2 - \mu_2)(Y_n - \mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ (Y_n - \mu_n)(Y_1 - \mu_1) & (Y_n - \mu_n)(Y_2 - \mu_2) & \ldots & (Y_n - \mu_n)^2 \end{pmatrix}$$

Take now expectation for each element of the matrix above. What do we get?

So $\Sigma$ is the variance covariance matrix of the vector $Y$. It is symmetric and positive definite.

Suppose $Y_1, \ldots, Y_n$ are independent identically distributed (i.i.d.) random variables. This means that $E[Y_i] = \mu, i = 1, \ldots, n, \text{var}[Y_i] = \sigma^2, i = 1, \ldots, n$ and $\text{cov}[Y_i, Y_j] = 0, i \neq j$. Find the expression of the mean vector of $Y$ and the variance covariance matrix of $Y$ using the vector $1 = (1, 1, \ldots, 1)'$ and the identity matrix $I$ for this special case.
Two important results are given below: The mean and variance of a linear combination of the elements of a random vector and the mean and variance of a set of linear combinations of the elements of a random vector. In the first case we will examine \( \mathbf{a}' \mathbf{Y} = a_1 Y_1 + \ldots + a_n Y_n \) while in the second case we will examine a set of \( p \) of these combinations.

1. Expected value and variance of a linear combination of \( \mathbf{Y} \). Let \( \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \) be a vector of constants and let \( q = \mathbf{a}' \mathbf{Y} \). Then \( E(q) = E(\mathbf{a}' \mathbf{Y}) = \mathbf{a}' E(\mathbf{Y}) = \mathbf{a}' \mu \). The variance of \( q \) can be found as follows:

\[
\text{var}(q) = E(q - \mu_q)^2 = E(\mathbf{a}' \mathbf{Y} - \mathbf{a}' \mu)^2 \\
= E(\mathbf{a}' \mathbf{Y} - \mathbf{a}' \mu)(\mathbf{a}' \mathbf{Y} - \mathbf{a}' \mu) \\
= \mathbf{a}' E(\mathbf{Y} - \mu)(\mathbf{Y} - \mu)' \mathbf{a} \\
= \mathbf{a}' \Sigma \mathbf{a}.
\]

Note: \( q \) is a scalar and therefore its variance should be a scalar and not a matrix. We can verify that \( \text{var}(q) = \mathbf{a}' \Sigma \mathbf{a} \) is \( 1 \times 1 \).

We can also express the variance of a linear combination using summation notation as follows:

\[
\text{var} \left( \sum_{i=1}^{n} a_i Y_i \right) = \text{cov} \left( \sum_{i=1}^{n} a_i Y_i, \sum_{j=1}^{n} a_j Y_j \right) \\
= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \text{cov}(Y_i, Y_j) \\
= \sum_{i=1}^{n} a_i^2 \text{var}(Y_i) + \sum_{i=1}^{n} \sum_{j \neq i} a_i a_j \text{cov}(Y_i, Y_j) \\
= \sum_{i=1}^{n} a_i^2 \text{var}(Y_i) + 2 \sum_{i=1}^{n-1} \sum_{j > i} a_i a_j \text{cov}(Y_i, Y_j)
\]

Example;

Let \( \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} \), \( \mu = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 4 \end{pmatrix} \), and \( \Sigma = \begin{pmatrix} 3 & 2 & 3 & 3 \\ 2 & 5 & 5 & 4 \\ 3 & 5 & 9 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix} \). Find the mean and variance of \( q = 4Y_1 - 2Y_2 + Y_3 + 3Y_4 \).
2. Let \( A \) be a \( p \times n \) matrix of constants. We will examine now \( Q = AY \). Unlike \( q \) (see result (1) above), \( Q \) is a \( p \times 1 \) vector and therefore its variance should be a \( p \times p \) matrix. Let’s find the expected value of \( Q \) first. \( E(Q) = E(AY) = A\mu \). For the variance of \( Q \) use the definition of the variance covariance matrix of a random vector.

\[
\text{var}(Q) =
\]

3. **Expectation of a quadratic expression**

Let \( Y \) be a random vector \( n \times 1 \) and let \( A \) be an \( n \times n \) matrix of constants. Consider the quadratic expression \( Y'AY \). We want to find the expected value of this quadratic expression. Such expressions appear in linear models and the result here provides a method for finding such expectations. For example, suppose \( n = 2 \). Find the expected value of

\[
\begin{pmatrix}
  Y_1 & Y_2
\end{pmatrix}
\begin{pmatrix}
  2 & 4 \\
  5 & 3
\end{pmatrix}
\begin{pmatrix}
  Y_1 \\
  Y_2
\end{pmatrix} = 2Y_1^2 + 3Y_2^2 + 9Y_1Y_2
\]

To find the expected value of this quadratic expression we will use properties of the trace of a square matrix. We can do this because \( Y'AY \) is a scalar. We will also need this result:

\[
E[YY'] = \Sigma + \mu\mu'.
\]

(Show this result using the definition of the variance covariance matrix).

\[
E[Y'AY] = E[tr(Y'AY)]
\]
4. Other results:

a. Covariance between two linear combinations:
\[
\text{cov}(\mathbf{a}'\mathbf{Y}, \mathbf{b}'\mathbf{Y}) = \mathbf{a}\Sigma\mathbf{b}'. 
\]
This is a scalar.

Example;

Let \( \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}, \mu = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 4 \end{pmatrix}, \text{ and } \Sigma = \begin{pmatrix} 3 & 2 & 3 & 3 \\ 2 & 5 & 5 & 4 \\ 3 & 5 & 9 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}. \) Find the covariance between

\( q_1 = 4Y_1 - 2Y_2 + Y_3 + 3Y_4 \) and \( q_2 = Y_1 + 3Y_2 - 5Y_3 - 4Y_4. \)

b. \( \text{cov}(\mathbf{A}\mathbf{Y}, \mathbf{B}\mathbf{Y}) = \mathbf{A}\Sigma\mathbf{B}'. \) This is a matrix.