Basics of Random Vectors

If X is a random variable then X has an expectation $\mathbb{E}(X)$, which is a real number.

Expectation is linear in the sense that if X and Y are random variables on the same probability space and a is a real number then

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) \tag{1}$$

and

$$\mathbb{E}(aX) = a\mathbb{E}(X). \tag{2}$$

The covariance $\operatorname{cov}(X,Y)$ of random variables X and Y defined on the same probability space is defined by

$$\operatorname{cov}(X,Y) = \mathbb{E}\left((X - \mathbb{E}(X))\,(Y - \mathbb{E}(Y))\right),$$

so that

$$\operatorname{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \tag{3}$$

When $X = Y$ this is called the variance of X, written $\operatorname{Var}(X)$, so that

$$\operatorname{Var}(X) = \operatorname{cov}(X,X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2\right) - \left(\mathbb{E}(X)\right)^2.$$

All of the above is in M2504.

Equations (1), (2) and (3) show that

$$\operatorname{cov}(aX,Y) = \mathbb{E}\left(aXY\right) - \mathbb{E}(aX)\mathbb{E}(Y) = a\operatorname{cov}(X,Y) \tag{4}$$

and

$$\begin{aligned} \operatorname{cov}(X_1+X_2,Y) &= \mathbb{E}\left((X_1+X_2)Y\right) - \mathbb{E}(X_1+X_2)\mathbb{E}(Y) \\ &= \mathbb{E}(X_1Y) + \mathbb{E}(X_2Y) - \mathbb{E}(X_1)\mathbb{E}(Y) - \mathbb{E}(X_2)\mathbb{E}(Y) \\ &= \operatorname{cov}(X_1,Y) + \operatorname{cov}(X_2,Y). \end{aligned} \tag{5}$$

In Statistics, we usually have several random variables $Y_1, \ldots, Y_n$, so we gather them into a column vector

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}.$$

Then $\mathbf{Y}$ is called a random vector, and its expectation is defined by

$$\mathbb{E}(\mathbf{Y}) = \begin{pmatrix} \mathbb{E}(Y_1) \\ \mathbb{E}(Y_2) \\ \vdots \\ \mathbb{E}(Y_n) \end{pmatrix}.$$

If A is an $m \times n$ real matrix then $A\mathbf{Y}$ is a random vector of length m. It follows from (1) and (2) that

$$\mathbb{E}(A\mathbf{Y}) = A\mathbb{E}(\mathbf{Y}).$$

The variance-covariance matrix $\mathrm{Cov}(\mathbf{Y})$ of $\mathbf{Y}$ is defined to be the $n \times n$ real matrix (which is necessarily symmetric and positive semi-definite) whose (i, j)-entry is $\mathrm{cov}(Y_i, Y_j)$. In particular, its diagonal entries are $\mathrm{Var}(Y_1)$, …, $\mathrm{Var}(Y_n)$.

Let a_1, …, a_n, b_1, …, b_n be real numbers. Then repeated use of Equations (4) and (5) shows that

$$\mathrm{cov}(a_1Y_1 + \cdots + a_nY_n, b_1Y_1 + \cdots + b_nY_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j \,\mathrm{cov}(Y_i, Y_j). \quad (6)$$

Denote by $\mathbf{a}$ and $\mathbf{b}$ the column vectors whose entries are a_1 ,…, a_n and b_1,…, b_n respectively. Then

$$a_1Y_1 + \cdots + a_nY_n = \mathbf{a} \cdot \mathbf{Y} = \mathbf{a}'\mathbf{Y},$$

where $\cdot$ denotes the standard inner product and $'$ denotes transpose. Then Equation (6) can be rewritten as

$$\mathrm{cov}(\mathbf{a} \cdot \mathbf{Y}, \mathbf{b} \cdot \mathbf{Y}) = \mathbf{a}' \,\mathrm{Cov}(\mathbf{Y})\mathbf{b}. \quad (7)$$

In particular,

$$\mathrm{Var}(\mathbf{a} \cdot \mathbf{Y}) = \mathbf{a}' \,\mathrm{Cov}(\mathbf{Y})\mathbf{a}.$$

Taking $\mathbf{a}'$ and $\mathbf{b}'$ in Equation (7) to be every pair of distinct rows of the matrix A in turn proves that

$$\mathrm{Cov}(A\mathbf{Y}) = A\,\mathrm{Cov}(\mathbf{Y})A'.$$

R. A. Bailey
February 2017