Multivariate Distributions

A random vector T is a vector whose elements have a joint distribution. i.e., if $(\Omega, \mathcal{A}, P)$ is a probability space, $T_{p\times 1} : \Omega \to \mathcal{R}^p$ is such that $T^{-1}(B) \in \mathcal{A}$, and hence $Pr(T \in B) = P(T^{-1}(B))$.

Thus, $\mathbf{X} = (X_1, \ldots, X_p)'$ is a random vector if X_i's are random variables with a joint distribution. If the joint density exists, we have $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{R}^p$ such that

$$\int_{\mathcal{R}^p} f(\mathbf{x})\, d\mathbf{x} = 1 \text{ and } P(\mathbf{X} \in A) = \int_A f(\mathbf{x})\, d\mathbf{x}, \qquad A \subset \mathcal{R}^p.$$

Example. $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right)$, $-1 < \rho < 1$, if

$$f(x_1, x_2) = \frac{1}{2\pi} \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right\}}.$$

Check that $E(X_i) = \mu_i$, $Var(X_i) = \sigma_i^2$, $i = 1, 2$ and $Cov(X_1, X_2) = \rho\sigma_1\sigma_2$.

Example. $\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim$ Uniform on unit ball if

$$f(x_1, x_2, x_3) = \begin{cases} \frac{3}{4\pi} & \text{if } x_1^2 + x_2^2 + x_3^2 \leq 1; \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{X} = (X_1, \ldots, X_p)'$ be a random vector and assume $\mu_i = E(X_i)$ exists for all i. Then define $E(\mathbf{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$ as the mean vector of $\mathbf{X}$. A random matrix $Z_{p\times q} = ((z_{ij}))$ is a matrix whose elements are jointly distributed random variables. If $G(Z)$ is a matrix valued function of Z, then $E(G(Z)) = ((E(G_{ij}(Z))))$.
If $G(Z) = AZB$, where A and B are constant matrices, $E(G(Z)) = AE(Z)B$.
If (Z, T) has a joint distribution, and A, B, C, D are constant matrices, $E(AZB + CTD) = AE(Z)B + CE(T)D$.
If Z is symmetric and positive semi-definite (nnd) with probability 1, $E(Z)$ is also symmetric and positive semi-definite. i.e., show $a'E(Z)a \geq 0$ for all a. Note that $a'E(Z)a = E(a'Za) \geq 0$, since for all a, $a'Za \geq 0$ wp 1.

Suppose $Z_{p\times p}$ is p.s.d. with wp 1. Then its spectral decomposition gives $Z = \Gamma D_\lambda \Gamma'$, where Γ is orthogonal and D_λ is diagonal. Let $\lambda_i(Z) = i$th diagonal element of D_λ, $\lambda_1(Z) \geq \lambda_2(Z) \geq \ldots \geq \lambda_p(Z) \geq 0$ wp 1. What about $E(Z)$? Is $\lambda_i(E(Z)) = E(\lambda_i(Z))$? No. However, $E(Z)$ is p.s.d., so $\lambda_i(E(Z)) \geq 0$.

Suppose $X_{p\times 1}$ has mean μ and also $E[(X_i-\mu_i)(X_j-\mu_j)] = Cov(X_i, X_j) = \sigma_{ij}$ exists for all i, j. i.e., $\sigma_{ii} < \infty$ for all i. Then the covariance matrix (or the variance-covariance matrix or the dispersion matrix) of X is defined as

$$Cov(X) = \Sigma = E\left[(X-\mu)(X-\mu)'\right] = ((E[(X_i-\mu_i)(X_j-\mu_j)])) = ((\sigma_{ij})).$$

Σ is symmetric, $\sigma_{ii} = Var(X_i) \geq 0$ and Σ is p.s.d.

Theorem. $\Sigma_{p\times p}$ is a covariance matrix (of some X) iff Σ is symmetric p.s.d.

Proof. (i) If $\Sigma = Cov(X)$ for some X and $E(X) = \mu$, then for any $\alpha \in \mathcal{R}^p$,

$$\begin{aligned} \alpha'\Sigma\alpha &= \alpha' Cov(X)\alpha = \alpha' E\left[(X-\mu)(X-\mu)'\right]\alpha \\ &= E\left[\alpha'(X-\mu)(X-\mu)'\alpha\right] = E\left[\{\alpha'(X-\mu)\}^2\right] \\ &= E\left[(\alpha'X - \alpha'\mu)^2\right] = Var(\alpha'X) \geq 0, \end{aligned}$$

so Σ is p.s.d. It is actually p.d. unless there exists $\alpha \neq 0$ such that $Var(\alpha'X) = 0$ (i.e., $\alpha'X = c$ w.p.1)

(ii) Now suppose Σ is any symmetric p.s.d matrix of rank $r \leq p$. Then $\Sigma = CC'$, $C_{p\times r}$ of rank r. Let $Y_1, \ldots, Y_r$ be i.i.d with $E(Y_i) = 0$, $Var(Y_i) = 1$. Let $Y = (Y_1, \ldots, Y_r)'$. Then $E(Y) = 0$, $Cov(Y) = I_r$. Let $X = CY$. Then $E(X) = 0$ and
$Cov(X) = E(XX') = E(CYY'C') = CE(YY')C' = CC' = \Sigma$.

For $a \neq 0$, $a'Cov(X)a = 0$ iff $Cov(X)a = 0$, or $Cov(X)$ has a zero eigen value.

If $X_{p\times 1}$ and $Y_{q\times 1}$ are jointly distributed with finite second moments for their elements, and with $E(X) = \mu$, $E(Y) = \nu$, then
$Cov(X_{p\times 1}, Y_{q\times 1}) = ((Cov(X_i, Y_j)))_{p\times q} = ((E(X_i-\mu_i)(Y_j-\nu_j))) = ((E(X_iY_j)-\mu_i\nu_j)) = E(XY') - \mu\nu' = E[(X-E(X))(Y-E(Y))']$.
$Cov(X) = Cov(X, X) = E[(X-E(X))(X-E(X))'] = E(XX') - E(X)(E(X))'$.
$Cov(AX, BY) = A\,Cov(X, Y)B'$,
$Cov(AX) = Cov(AX, AX) = A\,Cov(X, X)A' = A\,Cov(X)A'$.

Consider $X_{p\times 1} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ and $Y_{q\times 1} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$. Then

$$\begin{aligned} Cov(X,Y) &= \begin{pmatrix} Cov(X_1,Y_1) & Cov(X_1,Y_2) \\ Cov(X_2,Y_1) & Cov(X_2,Y_2) \end{pmatrix} \\ &\neq Cov(Y,X) = \begin{pmatrix} Cov(Y_1,X_1) & Cov(Y_1,X_2) \\ Cov(Y_2,X_1) & Cov(Y_2,X_2) \end{pmatrix} \end{aligned}$$

in general. Further, note,

$$\begin{aligned} Cov(X+Y) &= Cov(X+Y,X+Y) \\ &= Cov(X,X)+Cov(X,Y)+Cov(Y,X)+Cov(Y,Y) \\ &= Cov(X)+Cov(Y)+Cov(X,Y)+Cov(X,Y)' \\ &\neq Cov(X)+Cov(Y)+2Cov(X,Y), \end{aligned}$$

in general. If X and Y are independent, we do have, $Cov(X,Y) = ((Cov(X_i,Y_j))) = 0$ since $Cov(X_i,Y_j) = 0$ for all i and j.

Quadratic Forms.

$X'AX$ is called a quadratic form of X. Note that $E(X'AX) = E[tr(X'AX)] = E[tr(AXX')] = tr[E(AXX')] = tr[AE(XX')] = tr[A(\Sigma+\mu\mu')] = tr(A\Sigma)+tr(A\mu\mu') = tr(A\Sigma)+\mu'A\mu$, since $Cov(X) = \Sigma = E((X-\mu)(X-\mu)') = E(XX'-X\mu'-\mu X'+\mu\mu') = E(XX')-\mu\mu'$.