Homework 4
Due Date: November 18 2008

These are all based on problems in Mardia, Kent and Bibby (with minor changes of notation and content): 3.2.6, 3.3.1, 3.4.13, 5.3.3, 5.3.5.

1. Suppose $\begin{pmatrix} X_1 & X_2 & X_3 \end{pmatrix}$ has mean $\begin{pmatrix} \mu_1 & \mu_2 & \mu_3 \end{pmatrix}$ and covariance matrix $\Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & 0 \\ \rho^2 & 0 & 1 \end{pmatrix}$.

   Show that the conditional distribution of $\begin{pmatrix} X_1 & X_2 \end{pmatrix}$ given X_3 has mean $\begin{pmatrix} \mu_1 + \rho^2(X_3 - \mu_3) & \mu_2 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} 1-\rho^4 & \rho \\ \rho & 1 \end{pmatrix}$.

2. If $X \sim MVN_p(\mu, \Sigma)$ and $Q\Sigma Q^T$ $(q \times q)$ is nonsingular, then given $QX = q$, show that the conditional distribution of X is normal with mean $\mu + \Sigma Q^T(Q\Sigma Q^T)^{-1}(q - Q\mu)$ and (singular) covariance matrix $\Sigma - \Sigma Q^T(Q\Sigma Q^T)^{-1}Q\Sigma$.

3. If $M \sim W_p(\Sigma, m)$, show that $E(M^{-1}) = \frac{\Sigma^{-1}}{m-p-1}$.

4. (a) Consider the hypothesis $H_0: \ \mu = k\mu_0$ with Σ known (in other words, the hypothesis is that μ is some unknown multiple k of a given vector μ_0). Show that under H_0, the MLE of k is $\hat{k} = \mu_0^T\Sigma^{-1}\bar{X}/\mu_0^T\Sigma^{-1}\mu_0$ where $\bar{X}$ is the mean of independent $X_1, ..., X_n$, each $MVN_p(\mu, \Sigma)$. Also show that the LRT statistic $\lambda = \frac{L_0}{L_1}$ satisfies

   $$-2\log\lambda \;=\; n\bar{X}^T\Sigma^{-1}\{\Sigma - (\mu_0^T\Sigma^{-1}\mu_0)^{-1}\mu_0\mu_0^T\}\Sigma^{-1}\bar{X}.$$

   Deduce that the *exact* distribution of $-2\log\lambda$ is χ^2_{p-1} when H_0 is true.

   (b) Now consider the hypothesis $H_0: \ \mu = k\mu_0$ with Σ unknown. In this case the MLE of k under H_0 is $\hat{k} = \mu_0^T S_0^{-1}\bar{X}/\mu_0^T S_0^{-1}\mu_0$ (you can assume this without proof). With $d = \bar{X} - \hat{k}\mu_0$, show that

   $$\begin{aligned} -2\log\lambda &= n\log(1 + d^T S_0^{-1} d), \\ d^T S_0^{-1} d &= \bar{X}^T S_0^{-1}\{S_0 - (\mu_0^T S_0^{-1}\mu_0)^{-1}\mu_0\mu_0^T\}S_0^{-1}\bar{X}. \end{aligned}$$

   Hence show that the exact distribution of $(n-1)d^T S_0^{-1} d$ is $T^2_{p-1}(n-1)$.

5. Assume we have a sample of size n from $X_i = \begin{pmatrix} X_i^{(1)} & X_i^{(2)} & \dots & X_i^{(p)} \end{pmatrix} \sim MVN_p(\mu, \Sigma)$.

   (a) Show that the LRT for the hypothesis that $X_i^{(1)}$ is uncorrelated with $\begin{pmatrix} X_i^{(2)} & \dots & X_i^{(p)} \end{pmatrix}$ is given by $\lambda = \frac{L_0}{L_1} = (1 - R^2)^{n/2}$ where $R^2 = s_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$; here $S = \begin{pmatrix} s_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$.

   (b) Consider the null hypothesis that Σ has entries σ^2 on the diagonal and $\rho\sigma^2$ off the diagonal. Let S be the sample covariance matrix, v the average of the diagonal entries of S and vr the average of the off-diagonal entries. Show that the LRT is

   $$\lambda \;=\; \left\{\frac{|S|}{v^p(1-r)^{p-1}(1+(p-1)r)}\right\}^{n/2}.$$

Some Hints

Hints on Question 4(b)

The question as stated in Mardia, Kent and Bibby refers to two pieces of theory we have not done in class. Therefore, I am repeating those here.

First, I show the derivation of the formula for $\hat{k}$ (MKB, page 106): this is actually a bit more involved than just substituting S_0 for Σ in the formula derived in (a). Based on (a) and substituting $\hat{\Sigma}$, we have $\hat{k} = \bar{X}^T \hat{\Sigma}^{-1} \mu_0 / \mu_0^T \hat{\Sigma}^{-1} \mu_0$. Meanwhile, the equation for $\hat{\Sigma}$ assuming $\hat{\mu}$ known is

$$\hat{\Sigma} = S_0 + (\bar{X} - \hat{\mu})(\bar{X} - \hat{\mu})^T \tag{1}$$

(this comes down to minimizing $\log|\Sigma| + \operatorname{tr}\{S_0 + (\bar{X} - \hat{\mu})(\bar{X} - \hat{\mu})^T\}$ and the solution to that comes from the same argument given in class for the MLE of Σ when μ is unconstrained). The problem is to show these two equations are satisfied simultaneously when S_0 is substituted for $\hat{\Sigma}$ in the formula for $\hat{k}$. But if we simultaneously premultiply (1) by $\hat{\Sigma}^{-1}$ and postmultiply by S_0^{-1}, then

$$S_0^{-1} = \hat{\Sigma}^{-1} + \hat{\Sigma}^{-1}(\bar{X} - \hat{\mu})(\bar{X} - \hat{\mu})^T S_0^{-1} \tag{2}$$

Now premultiply (2) by μ_0^T: we have $\mu_0^T \hat{\Sigma}^{-1}(\bar{X} - \hat{\mu}) = 0$ from the formula for $\hat{\mu} = \hat{k}\mu_0$. Therefore $\mu_0^T S_0^{-1} = \mu_0^T \hat{\Sigma}^{-1}$, and it then follows that the two alternative forms of $\hat{k}$ are the same.

Second, there is a piece of theory for testing a constrained hypothesis about μ in the case that Σ is unknown (MKB pp. 132–133). Suppose we want to test the null $H_0: \ R\mu = r$ where R is a $q \times p$ constraint matrix and r is given. In this case

$$-2\log\lambda \quad = \quad n\log(1 + d^T S_0^{-1} d)$$

where $d = S_0 R^T (R S_0 R^T)^{-1}(R\bar{X} - r)$. Then

$$(n-1) d^T S_0^{-1} d \quad = \quad (n-1)(R\bar{X} - r)^T (R S_0 R^T)^{-1}(R\bar{X} - r) \tag{3}$$

where $R\bar{X} \sim MVN_q(r, n^{-1} R\Sigma R^T)$ independently of $nRS_0R^T = (n-1)RSR^T \sim W_q(R\Sigma R^T, n-1)$ using Prop. 4 of the notes. Therefore, by definition of Hotelling's T^2, (3) has distribution $T^2_{n-1}(q)$.

Hints on Question 5

(a) If $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ then

$$|A| = |A_{11}| \cdot |A_{22} - A_{21}A_{11}^{-1}A_{12}| = |A_{22}| \cdot |A_{11} - A_{12}A_{22}^{-1}A_{21}|.$$

(b) Let I be the $p \times p$ identity matrix, J the $p \times p$ matrix of ones. If $E = (1-\rho)I + \rho J$ is a matrix with 1 on the diagonal and ρ in all off-diagonal entries, then $|E| = (1-\rho)^{p-1}(1 - \rho + p\rho)$ and $E^{-1} = \frac{1}{1-\rho}\left(I - \frac{\rho}{1-\rho+p\rho}J\right)$.