ISYE 7405

Homework 4

Jacob Aguirre *Email: aguirre@gatech.edu*
Instructor: Dr. Shihao Yang

1. $X_{n\times p} \sim N_{p\times n}(\mu 1_n^T, I_n \otimes \Sigma)$ is a normal random matrix. Let a be a fixed p-dimensional vector and $y = a^T X = (y_1, y_2, \ldots, y_n)$. Let $t(a)$ be the corresponding one-sample t statistic

$$t(a) = \frac{\sqrt{n}\bar{y}}{\sqrt{\sum_i (y_i - \bar{y})^2/(n-1)}}.$$

Show that Hotelling's T^2 statistic is equal to $\max_{a\neq 0} t^2(a)$.
Solution. Let $J_n = 1/n 1_n 1_n^T$.

$$\begin{aligned} t^2(a) &= \frac{n(n-1)a^T X 1_n/n 1_n/n^T X^T a}{a^T X(I - J_n)X^T a} \\ &= \frac{(n-1)a^T X J_n^T X^T a}{a^T X(I - J_n)X^T a} \end{aligned}$$

Therefore,

$$1 - \frac{1}{t^2(a)/(n-1)+1} = \frac{(n-1)a^T X(I-J_n)X^T a}{a^T X X^T a},$$

$$\hat{a} = argmax\, t^2(a) = argmax \frac{(n-1)a^T X J_n X^T a}{a^T X X^T a}$$

Let $A = XJ_nX^T$, $B = XX^T$, then $\hat{a} = B^{-1/2}\xi_1$, where ξ_1 is the first eigenvector of $B^{-1/2}AB^{-1/2}$. Notice that $rank(B^{-1/2}AB^{-1/2}) = 1$, first eigenvalue $\lambda_1 = tr(B^{-1/2}AB^{-1/2}) = 1/n 1_n^T X(X^TX)^{-1}X^T 1_n$, $\xi_1 = (X^TX)^{-1/2}X^T 1_n$.

$$\begin{aligned} t^2(\hat{a}) &= \frac{(n-1)(1_n^T X(X^TX)^{-1}X^T 1_n)^2}{n(1_n^T X(X^TX)^{-1}X^T(I-J-n)X(X^TX)^{-1}X^T 1_n)} \\ &= \frac{(n-1)1_n^T X(X^TX)^{-1}X^T 1_n}{n(1 - 1/n 1_n^T X(X^TX)^{-1}X^T 1_n)}. \end{aligned}$$

By Woodbury formula,

$$\begin{aligned} T^2 &= n(n-1)1_n^T/nX^T(X(I-J_n)X^T)^{-1}X1_n/n \\ &= (n-1)/n 1_n^T X^T((XX^T)^{-1} - (XX^T)^{-1}(-1 + 1_n/\sqrt{n}X(X^TX)^{-1}X^T 1_n/\sqrt{n})^{-1}(XX^T)^{-1})X1_n \\ &= (n-1)/n(1_n^T X(X^TX)^{-1}X^T 1_n - \frac{1/n(1_n^T X(X^TX)^{-1}X^T 1_n)^2}{1/n 1_n^T X(X^TX)^{-1}X^T 1_n - 1}) \\ &= \frac{(n-1)1_n^T X(X^TX)^{-1}X^T 1_n}{n(1 - 1/n 1_n^T X(X^TX)^{-1}X^T 1_n)} = t^2(\hat{a}). \end{aligned}$$

□

2. Show that Hotelling's two-sample T^2 test is distributed as

$$T^2 \sim \frac{(n-2)p}{n-p-1} F_{p,n-p-1}(\delta^2), \text{ where } \delta^2 = \frac{n_1 n_2}{n}(\mu_2-\mu_1)^\top \Sigma^{-1}(\mu_2-\mu_1).$$

Solution. The T^2 statistic can be represented as

$$T^2 = \frac{(\bar{Y}-\bar{X})^\top \left\{\Sigma(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X})}{(\bar{Y}-\bar{X})^\top \left\{\Sigma(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X}) \Big/ (\bar{Y}-\bar{X})^\top \left\{\left(\frac{S_X^2+S_Y^2}{n-2}\right)(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X})} \equiv \frac{B}{A}.$$

First, $\bar{X} \sim N_p(\mu_1, \Sigma/n_1), \bar{Y} \sim N_p(\mu_2, \Sigma/n_2)$, and $\bar{X} \perp\!\!\!\perp \bar{Y}$. Then,

$$\bar{Y}-\bar{X} \sim N_p\left\{\mu_2-\mu_1, \Sigma(n_1^{-1}+n_2^{-1})\right\}.$$

By the definition of non-central chi-squared distribution, we have

$$B = (\bar{Y}-\bar{X})^\top \left\{\Sigma(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X}) \sim \chi_p^2(\delta^2),$$

where

$$\delta^2 = (\mu_2-\mu_1)^\top \left\{\Sigma(n_1^{-1}+n_2^{-1})\right\}^{-1}(\mu_2-\mu_1) = \frac{n_1 n_2}{n}(\mu_2-\mu_1)^\top \Sigma^{-1}(\mu_2-\mu_1).$$

Second,

$$A = \frac{(\bar{Y}-\bar{X})^\top \left\{\Sigma(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X})}{(\bar{Y}-\bar{X})^\top \left\{\left(\frac{S_X^2+S_Y^2}{n-2}\right)(n_1^{-1}+n_2^{-1})\right\}^{-1}(\bar{Y}-\bar{X})} = (n-2)^{-1}\frac{(\bar{Y}-\bar{X})^\top \Sigma^{-1}(\bar{Y}-\bar{X})}{(\bar{Y}-\bar{X})^\top (S_X^2+S_Y^2)^{-1}(\bar{Y}-\bar{X})}.$$

Since $(\bar{X},\bar{Y}) \perp\!\!\!\perp (S_X^2, S_Y^2)$ and $S_X^2+S_Y^2 \sim W_p(\Sigma, n-2)$, we have

$$A|(\bar{X},\bar{Y}) \sim (n-2)^{-1}\chi^2_{(n-2)-(p-1)} \sim (n-2)^{-1}\chi^2_{n-p-1},$$

which implies that $A \perp\!\!\!\perp B$, and marginally, $A \sim (n-2)^{-1}\chi^2_{n-p-1}$. Therefore, T^2 statistic can be represented as

$$T^2 = \frac{(n-2)p}{n-p-1} \times \frac{\chi_p^2(\delta^2)/p}{\chi^2_{n-p-1}/(n-p-1)} \sim \frac{(n-2)p}{n-p-1} \times F_{p,n-p-1}(\delta^2),$$

since $\chi_p^2(\delta^2) \perp\!\!\!\perp \chi^2_{n-p-1}$ above. □