ISYE 7405

Homework 3

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

1. $X_{p\times n} \sim N_{p\times n}(\mu 1_n^\top, \Delta \otimes \Sigma)$ is a normal random matrix. Let $Y = H_p X H_n$, where $H_p = I_p - 1_p 1_p^\top/p$ and $H_n = I_n - 1_n 1_n^\top/n$. Show that

1. $\sum_{j=1}^n Y_{ij} = \sum_{i=1}^p Y_{ij} = 0$ for all i and j.
2. $Y \sim N_{p\times n}(0, \Delta^Y \otimes \Sigma^Y)$, where

$$\begin{aligned} \sigma_{i,i'}^Y &= \sigma_{i,i'} - \sigma_{i.} - \sigma_{,i'} + \sigma_{..}, i, i' = 1, 2, \cdots, p \\ \Delta_{j,j'}^Y &= \Delta_{j,j'} - \Delta_{j.} - \Delta_{.j'} + \Delta_{..}, j, j' = 1, 2, \cdots, n. \end{aligned}$$

   Here dot "·" indicates an average over the missing index.

3. $\sum_{i'} \sigma_{i,i'}^Y = \sum_{j'} \Delta_{j,j'}^Y = 0$ for all i and j.

Solution. (a) Note that H_p and H_n are two projection matrices orthogonal to 1_p and 1_n, respectively. Therefore, we have $H_p 1_p = 0$ and $H_n 1_n = 0$. We use $e_k = (0, \dots, 1, \dots, 0)^\top$ to denote the vector with only the k-th entry being 1, and the dimension of e_k will be determined later according to matrix multiplication.

Since $\sum_{j=1}^n Y_{ij}$ is the sum of the i-th row of the matrix Y, we have

$$\sum_{j=1}^n Y_{ij} = e_i^\top Y 1_n = e_i^\top H_p X H_n 1_n = 0.$$

Similarly, since $\sum_{i=1}^p Y_{ij}$ is the sum of the j-th column of the matrix Y, we have

$$\sum_{i=1}^p Y_{ij} = 1_p^\top Y e_j = 1_p^\top H_p X H_n = 0.$$

(b) First,

$$Y = H_p X H_n \sim N_{p\times n}(H_p \mu 1_n^\top H_n, (H_n \Delta H_n) \otimes (H_p \Sigma H_p)) = N_{p\times n}(0, (H_n \Delta H_n) \otimes (H_p \Sigma H_p)),$$

and $\Delta^Y = H_n \Delta H_n, \Sigma^Y = H_p \Sigma H_p$. We further notice that

$$H_p e_i = e_i - \frac{1}{p} 1_p 1_p^\top E_i = e_i - \frac{1}{p} e_i.$$

Therefore, the (i,i') entry of Σ^Y is

$$\begin{aligned}
\sigma^Y_{i,i'} &= e_i^\top \Delta^Y e_{i'} = e_i^\top H_p \Sigma H_p e_{i'} = (e_i - p^{-1}1_p)^\top \Sigma (e_{i'} - p^{-1}1_p) \\
&= e_i \Sigma e_{i'} - p^{-1} e_i^\top \Sigma 1_p - p^{-1}\Sigma e_{i'} + p^{-2} 1_p^\top \Sigma 1_p \\
&= \sigma_{i,i'} - \sigma_{i.} - \sigma_{.i'} + \sigma_{..}.
\end{aligned}$$

The result for Δ^Y follows the same argument, except for the difference between p and n and between σ and Δ.
(c) We have

$$\sum_{i'} \sigma^Y_{i,i'} = \sum_{i'} \sigma_{i,i'} - \sigma_{i.} - \sigma_{.i'} + \sigma_{..} = p\sigma_{i.} - p\sigma_{i.} - p\sigma_{..} + p\sigma_{..} = 0,$$

and the same result for Δ^Y.□

2. Consider $X \sim N_{p\times n}(0, \Delta \otimes \Sigma)$, where $\Delta_{ii} = 1$ for all i. $\hat\Sigma = XX^T/n$. Show that $E(\hat\Sigma) = \Sigma$. Now consider $Y_{p\times n} \sim N_{p\times n}(0, I_n \otimes \Sigma)$. Let $\hat\Sigma_0 = YY^T/n$. Show that for all i,j,k,l, we have $Cov(\hat\Sigma_{ij}, \hat\Sigma_{kl}) = Cov((\hat\Sigma_0)_{ij}, (\hat\Sigma_0)_{kl})(1 + 2/n\sum_{g<h}\Delta^2_{gh})$.

Solution.

$$E(\hat\Sigma) = E(XX^T/n) = 1/n\sum_{i=1}^n E(X_iX_i^T) = 1/n\sum_{i=1}^n cov(X_i) = 1/n\sum_{i=1}^n = 1/n\sum_{i=1}^n \Delta_{ii}\Sigma = \Sigma.$$

$$\begin{aligned}
Cov(\hat\Sigma_{ij}, \hat\Sigma_{kl}) &= 1/n^2 Cov(\sum_{r=1}^n X_{ir}X_{jr}, \sum_{s=1}^n X_{ks}X_{ls}) \\
&= 1/n^2 \left(\sum_{r=1}^n Cov(X_{ir}X_{jr}, X_{kr}X_{lr}) + 2\sum_{r<s} Cov(X_{ir}X_{jr}, X_{ks}X_{ls}) \right).
\end{aligned}$$

Notice that X_{ir}'s are normally distributed with mean 0, and by the formula of normal kurtosis :

$$\begin{aligned}
&E((X_1 - E(X_1))(X_2 - E(X_2))(X_3 - E(X_3))(X_4 - E(X_4))) \\
&= cov(X_1,X_2)cov(X_3,X_4) + cov(X_1,X_4)cov(X_2,X_3) + cov(X_1,X_3)cov(X_2,X_4)
\end{aligned}$$

We have:

$$\begin{aligned}
Cov(X_{ir}X_{jr}, X_{ks}X_{ls}) &= E(X_{ir}X_{jr}X_{ks}X_{ls}) - E(X_{rr}X_{jr})E(X_{ks}X_{ls}) \\
&= E(X_{ir}X_{ks})E(X_{ls}X_{jr}) + E(X_{ir}X_{ls})E(X_{is}X_{kr}) \\
&= \Delta^2_{rs}\sigma_{ik}\sigma_{lj} + \Delta^2_{rs}\sigma_{il}\sigma_{kj} \\
&= \Delta^2_{rs}(\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj}) \\
Cov(\hat\Sigma_{ij}, \hat\Sigma_{kl}) &= 1/n^2 \left(\sum_{r=1}^n \Delta^2_{rr}(\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj}) + 2\sum_{r<s}\Delta^2_{rs}(\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj}) \right) \\
&= (\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj})/n(1 + 2/n\sum_{r<s}\Delta^2_{rs})
\end{aligned}$$

Specifically,

$$Cov((\hat{\Sigma}_0)_{ij}, (\hat{\Sigma}_0)_{kl}) = 1/n^2 \left(\sum_{r=1}^{n} 1 \cdot (\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj}) + 2 \sum_{r<s} 0 \cdot (\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj}) \right)$$
$$= (\sigma_{ik}\sigma_{lj} + \sigma_{il}\sigma_{kj})/n.$$

So we have:

$$Cov(\hat{\Sigma}_{ij}, \hat{\Sigma}_{kl}) = Cov((\hat{\Sigma}_0)_{ij}, (\hat{\Sigma}_0)_{kl})(1 + 2/n \sum_{g<h} \Delta_{gh}^2)$$

3. Consider the mapping of $T \to S = T^\top T$, which takes an upper triangular $p \times p$ matrix T into the symmetric matrix S. We mentioned in the class that $J(T \to S) = \left(2^p \prod_{i=1}^{p} t_{ii}^{p-i+1}\right)^{-1}$. Now prove this result.

Solution. From

$$S = T^\top T = \begin{pmatrix} t_{11} & 0 & \cdots & 0 \\ t_{12} & t_{22} & \cdots & 0 \\ \vdots & & & \\ t_{1p} & t_{2p} & \cdots & t_{pp} \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ 0 & t_{22} & \cdots & t_{2p} \\ \vdots & & & \\ 0 & 0 & \cdots & t_{pp} \end{pmatrix},$$

we have

$$\left(\begin{array}{c} s_{11} \\ s_{12} \\ \vdots \\ s_{1p} \\ \hline s_{22} \\ s_{23} \\ \vdots \\ s_{2p} \\ \hline \vdots \\ \hline s_{(p-1)(p-1)} \\ s_{(p-1)p} \\ \hline s_{pp} \end{array}\right) = \left(\begin{array}{c} t_{11}^2 \\ t_{11}t_{12} \\ \vdots \\ t_{11}t_{1p} \\ \hline t_{12}^2 + t_{22}^2 \\ t_{12}t_{13} + t_{22}t_{23} \\ \vdots \\ t_{12}t_{1p} + t_{22}t_{2p} \\ \hline \vdots \\ \hline t_{1(p-1)}^2 + t_{2(p-1)}^2 + \cdots + t_{(p-1)(p-1)}^2 \\ t_{1(p-1)}t_{1p} + t_{2(p-1)}t_{2p} + \cdots + t_{(p-1)(p-1)}t_{(p-1)p} \\ \hline t_{1p}^2 + t_{2p}^2 + \cdots + t_{pp}^2 \end{array}\right).$$

Therefore, the Jacobian matrix of the transformation $T \to S$ is

$$\frac{\partial S}{\partial T} = \frac{\partial(s_{ij} : i \le j)}{\partial(t_{ij} : i \le j)} = \left(\begin{array}{cccccccccc}
2t_{11} \\
* & t_{11} \\
\vdots \\
* & * & \cdots & t_{11} \\ \hline
* & * & \cdots & * & 2t_{22} \\
* & * & \cdots & * & * & t_{22} \\
* & * & \cdots & * & * & \cdots & t_{22} \\ \hline
\vdots \\ \hline
* & * & \cdots & * & * & \cdots & * & \cdots & 2t_{(p-1)(p-1)} \\
* & * & \cdots & * & * & \cdots & * & \cdots & * & t_{(p-1)(p-1)} \\ \hline
* & * & \cdots & * & * & \cdots & * & \cdots & * & * & 2t_{pp}
\end{array}\right).$$

Since it is a lower triangular matrix (the elements denoted by $*$ does not matter in the latter calculation), its determinant is

$$2^p t_{11}^p t_{22}^{p-1} \cdots t_{(p-1)(p-1)}^2 t_{pp} = 2^p \prod_{i=1}^p t_{ii}^{p-i+1}.$$

And the Jacobian $J(T \to S)$ is the inverse of the value above.□

4. Suppose $A_{p\times p}$ is non-singular. Show that the mapping of $S \to \widetilde{S} = ASA^\top$, which takes a symmetric positive definite matrix $S_{p\times p}$ into $\widetilde{S}_{p\times p}$, has Jacobian $J(S \to \widetilde{S}) = (\det(A))^{-(p+1)}$.
Solution. Using the property of the Wishart distribution, if $S \sim W_p(I_p, n)$ has density

$$f_S(S) = c \exp\left\{-\frac{1}{2}\text{tr}(S)\right\} \{\det(S)\}^{(n-p-1)/2},$$

then $\widetilde{S} = ASA^\top \sim W_p(AA^\top, n)$ has density

$$f_{\widetilde{S}}(\widetilde{S}) = c\{\det(AA^\top)\}^{-n/2} \exp\left\{-\frac{1}{2}\text{tr}(A^{-\top}A^{-1}\widetilde{S})\right\} \{\det(\widetilde{S})\}^{(n-p-1)/2}.$$

Since $f_{\widetilde{S}}(\widetilde{S}) = f_S(S) J(S \to \widetilde{S})$, the Jacobian is

$$\begin{aligned}
J(S \to \widetilde{S}) &= \frac{f_{\widetilde{S}}(\widetilde{S})}{f_S(S)} \\
&= \frac{c\{\det(AA^\top)\}^{-n/2} \exp\left\{-\frac{1}{2}\text{tr}(A^{-\top}A^{-1}\widetilde{S})\right\} \{\det(\widetilde{S})\}^{(n-p-1)/2}}{c \exp\left\{-\frac{1}{2}\text{tr}(S)\right\} \{\det(S)\}^{(n-p-1)/2}} \\
&= \frac{c\{\det(AA^\top)\}^{-n/2} \exp\left\{-\frac{1}{2}\text{tr}(A^{-\top}A^{-1}ASA^\top)\right\} \{\det(ASA^\top)\}^{(n-p-1)/2}}{c \exp\left\{-\frac{1}{2}\text{tr}(S)\right\} \{\det(S)\}^{(n-p-1)/2}} \\
&= \{\det(A)\}^{-(p+1)}.□
\end{aligned}$$

5. Let $S \sim W_p(\Sigma, n)$. Calculate the expectation $E(\det(S^m))$, where m is a positive integer.
Solution. According to Bartlett's decomposition of the Wishart matrix S, we have

$$\det(S) = \det(\Sigma^{1/2} T^\top T \Sigma^{1/2}) = \det(\Sigma)\{\det(T)\}^2 = \det(\Sigma) \prod_{i=1}^{p} t_{ii}^2.$$

Therefore, we have

$$\begin{aligned} E(\det(S^m)) &= [\det(\Sigma)]^m E\left\{\prod_{i=1}^{p} t_{ii}^{2m}\right\} = [\det(\Sigma)]^m \prod_{i=1}^{p} E(t_{ii}^{2m}) \\ &= [\det(\Sigma)]^m \prod_{i=1}^{p} E(\chi^2_{n-i+1})^m = [\det(\Sigma)]^m \prod_{i=1}^{p} E\left\{2\text{Gamma}((n-i+1)/2)\right\}^m \\ &= [\det(\Sigma)]^m 2^{mp} \frac{\prod_{i=1}^{p} \Gamma\left(\frac{n-i+1}{2} + m\right)}{\prod_{i=1}^{p} \Gamma\left(\frac{n-i+1}{2}\right)}. \square \end{aligned}$$

6. (Expectation of the inverted Wishart distribution) If $S \sim W_p(\Sigma, n)$ with $n \geq p+2$. Show that $E(S^{-1}) = \Sigma^{-1}/(n-p-1)$. Therefore, $(n-p-1)S^{-1}$ is an unbiased estimate of Σ^{-1}. Hint: Use the fact that for any constant vector a, $a^\top \Sigma^{-1} a / a^\top S^{-1} a \sim \chi^2_{n-p+1}$.
Solution. First, we notice that the expectation of a inverse chi-squared distribution with degrees of freedom ν is

$$E(1/\chi^2_\nu) = E\left\{2\text{Gamma}(\nu/2)\right\}^{-1} = \frac{\Gamma(\nu/2 - 1)}{\Gamma(\nu/2)}/2 = \frac{1}{\nu - 2}.$$

If we take $a = (1, 0, \ldots, 0)$, from $a^\top \Sigma^{-1} a / a^\top S^{-1} a = (\Sigma^{-1})_{11}/(S^{-1})_{11} \sim \chi^2_{n-p+1}$, we know

$$E(S^{-1})_{11} = (\Sigma^{-1})_{11} \times E(\chi^2_{n-p+1})^{-1} = (\Sigma^{-1})_{11}/(n-p-1).$$

Similarly, we also have $E(S^{-1})_{22} = (\Sigma^{-1})_{22}/(n-p-1)$, and the same results hold for other diagonal terms.

If we take $a = (1, 1, 0, \ldots, 0)$, from

$$a^\top \Sigma^{-1} a / a^\top S^{-1} a = \frac{(\Sigma^{-1})_{11} + \Sigma^{-1}_{22} + 2(\Sigma^{-1})_{12}}{(S^{-1})_{11} + (S^{-1})_{22} + 2(S^{-1})_{12}} \sim \chi^2_{n-p+1},$$

we know

$$\begin{aligned} E\left[(S^{-1})_{11} + (S^{-1})_{22} + 2(S^{-1})_{12}\right] &= \left[(\Sigma^{-1})_{11} + \Sigma^{-1}_{22} + 2(\Sigma^{-1})_{12}\right] \times E(\chi^2_{n-p+1})^{-1} \\ &= \left[(\Sigma^{-1})_{11} + \Sigma^{-1}_{22} + 2(\Sigma^{-1})_{12}\right]/(n-p-1). \end{aligned}$$

Since we have proved the diagonal terms, we have

$$E(S^{-1})_{12} = (\Sigma^{-1})_{12}/(n-p-1).$$

Proofs for other off-diagonal terms are similar. $\square$

Remark: you can also use the hint in this way,

$$E\,[\frac{a^T S^{-1} a}{a^T \Sigma^{-1} a}] = E(1/\chi^{n-p+1}) = \frac{1}{n-p-1}.$$

Therefore, for any a, we have

$$a^T\,\left(E\,[S^{-1}] - \frac{1}{n-p-1}\Sigma^{-1}\right)\,a = 0,$$

which implies the symmetric matrix $E\,[S^{-1}] - \frac{1}{n-p-1}\Sigma^{-1}$ is zero, i.e.

$$E\,[S^{-1}] = \frac{1}{n-p-1}\Sigma^{-1}.$$

(In the last step, consider the eigenvalue decomposition of the symmetric matrix.)