STAT 510 Homework 1

Yifan Zhu

January 17, 2017

1.

$$
\begin{aligned}
\boldsymbol{AB} &= \begin{bmatrix} 1 & 5 \\ 4 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & -2 & 4 \\ 5 & -1 & 2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 & -1 & 2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 3 & -2 & 4 \\ 12 & 12 & -8 & 16 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 25 & -5 & 10 & 15 \\ -5 & 1 & -2 & -3 \\ 5 & -1 & 2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 28 & -2 & 8 & 19 \\ 7 & 13 & -10 & 13 \\ 5 & -1 & 2 & 3 \end{bmatrix}
\end{aligned}
$$

2. (a) $\boldsymbol{W} = \begin{bmatrix} 1 & 5 \\ 3 & -1 \end{bmatrix}$, thus $\boldsymbol{W}^{-1} = \begin{bmatrix} -0.0625 & 0.3125 \\ 0.1875 & -0.0625 \end{bmatrix}$. Hence

$$
\boldsymbol{G} = \begin{bmatrix} (\boldsymbol{W}^{-1})^T & \boldsymbol{0} \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{W}^{-1} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} -0.0625 & 0.3125 \\ 0.1875 & -0.0625 \\ 0 & 0 \end{bmatrix}
$$

(b) From R we have

$$
\boldsymbol{G} = ginv(\boldsymbol{A}) = \begin{bmatrix} 0.01414514 & 0.04797048 \\ 0.18880689 & -0.05535055 \\ 0.02091021 & 0.11439114 \end{bmatrix}
$$

3. From the definition, for $u \sim t_m(\delta)$, there exist independent $y \sim N(\delta, 1)$ and $w \sim \chi^2_m$ such that $u = \frac{y}{\sqrt{w/m}}$.

Hence $u^2 = \frac{y^2}{w/m} = \frac{y/1}{w/m}$. From the definition of non-central chi-squared distribution, we know that $y^2 \sim \chi^2_1(\delta^2/2)$. Also, y^2 and w are independent. Then from the definition of F distribution, we have $u^2 \sim F_{1,m}(\delta^2/2)$.

4. Let $\boldsymbol{z} = \begin{bmatrix} x \\ y \end{bmatrix}$, then $\boldsymbol{z} \sim N(\boldsymbol{\mu}, \boldsymbol{I}_{2\times 2})$, where $\boldsymbol{\mu} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Then we have $\boldsymbol{z}^T\boldsymbol{z} \sim \chi^2_2(\boldsymbol{\mu}^T\boldsymbol{\mu}/2 = 5/2)$. Thus

$$
\begin{aligned}
&P(\{\text{the string will need to be longer than 6 units}\}) \\
=&P(\boldsymbol{z}^T\boldsymbol{z} > 36) = 1 - pchisq(36, df = 2, ncp = 5) = 0.00014096
\end{aligned}
$$

5. (a) Let $\boldsymbol{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, then $\boldsymbol{z} \sim N(\boldsymbol{0}, \boldsymbol{I}_{2\times 2})$. Thus $z_1 - z_2 = \begin{bmatrix} 1 & -1 \end{bmatrix} \boldsymbol{z} \sim N(0,2)$ and hence $(z_1 - z_2)/\sqrt{2} \sim N(0,1)$. From the definition of χ^2 distribution, we know that

$$(z_1 - z_2)^2/2 = \left((z_1 - z_2)/\sqrt{2}\right)^2 \sim \chi_1^2$$

(b) $\begin{bmatrix} z_1 + z_2 \\ z_1 - z_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N(\boldsymbol{0}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix})$. Hence $z_1 + z_2$and $z_1 - z_2$ are independent, and $(z_1 + z_2)/\sqrt{2} \perp (z_1 - z_2)^2/2$. From (a) we know $(z_1 - z_2)^2/2 \sim \chi_1^2$. And $(z_1 + z_2)/\sqrt{2} \sim N(0,1)$. Then from the definition of t distribution,

$$(z_1 + z_2)/|z_1 - z_2| = \frac{(z_1 + z_2)/\sqrt{2}}{\sqrt{(z_1 - z_2)^2/2}} \sim t_1$$

6. (a) $\bar{y} = \frac{1}{n}\boldsymbol{1}^T\boldsymbol{y}$, thus $\boldsymbol{y} - \bar{y}\boldsymbol{1} = \boldsymbol{y} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{y} = (\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y}$. Hence

$$\begin{aligned}
s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \\
&= \frac{1}{n-1}((\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y})^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y} \\
&= \frac{1}{n-1}\boldsymbol{y}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y} \\
&= \frac{1}{n-1}\boldsymbol{y}(\boldsymbol{I} - \frac{2}{n}\boldsymbol{1}\boldsymbol{1}^T + \frac{1}{n^2}\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y} \\
&= \frac{1}{n-1}\boldsymbol{y}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y} \\
&= \boldsymbol{Y}^T\frac{1}{n-1}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{B}\boldsymbol{y}
\end{aligned}$$

Here $\boldsymbol{B} = \frac{1}{n-1}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)$

(b) $(n-1)s^2/\sigma^2 = \boldsymbol{y}^T\frac{1}{\sigma^2}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{y}$. Let $\boldsymbol{A} = \frac{1}{\sigma^2}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)$, then $\boldsymbol{A\Sigma} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T$. From the proof of (a) we know that $\boldsymbol{A\Sigma} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T$ is idempotent. $rank(\boldsymbol{A}) = rank(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T) = tr(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T) = tr(\boldsymbol{I}) - \frac{1}{n}tr(\boldsymbol{1}\boldsymbol{1}^T) = n - \frac{1}{n}n = n - 1$. Then we know

$$(n-1)s^2/\sigma^2 \sim \chi_{n-1}^2(\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}/2)$$

We also have $\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu} = \mu^2\boldsymbol{1}^T\boldsymbol{A}\boldsymbol{1} = \frac{\mu^2}{\sigma^2}\boldsymbol{1}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{1} = \frac{\mu^2}{\sigma^2}(\boldsymbol{1}^T\boldsymbol{1} - \frac{1}{n}\boldsymbol{1}^T\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{1}) = \frac{\mu^2}{\sigma^2}(n - \frac{1}{n}n^2) = 0$, hence

$$(n-1)s^2/\sigma^2 \sim \chi_{n-1}^2$$

7. (a) If $\boldsymbol{A} = \boldsymbol{0}$ then $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{0}$.
If $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{0}$, then $(\boldsymbol{A}^T\boldsymbol{A})_{ii}$ in the diagonal line should also be 0. And for every i, we have $(\boldsymbol{A}^T\boldsymbol{A})_{ii} = \sum_{j=1}^{n} a_{ji}^2 = 0 \Rightarrow a_{ji} = 0$ for all $j = 1, 2, \cdots, n$, where a_{ji} is the (j,i) component of $\boldsymbol{A}$. Hence for every element in $\boldsymbol{A}$, we have $a_{ij} = 0, i, j = 1, 2, \cdots, n$.Thus $\boldsymbol{A} = \boldsymbol{0}$.

(b) $\boldsymbol{XA} = \boldsymbol{XB} \Rightarrow \boldsymbol{X}^T\boldsymbol{XA} = \boldsymbol{X}^T\boldsymbol{XB}$.
$\boldsymbol{X}^T\boldsymbol{XA} = \boldsymbol{X}^T\boldsymbol{XB} \Rightarrow \boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \Rightarrow (\boldsymbol{A} - \boldsymbol{B})^T\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \Rightarrow (\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}))^T(\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B})) = \boldsymbol{0}$. From (a) we know that $\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \Rightarrow \boldsymbol{XA} = \boldsymbol{XB}$.

(c) For any generalized inverse $(\boldsymbol{X}^T\boldsymbol{X})^-$, we have $\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}^T\boldsymbol{X}$. Let $(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{A}$ and $\boldsymbol{I} = \boldsymbol{B}$, from (b) we have

$$\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}$$

(d) $\boldsymbol{G}$ is a generalized inverse of $\boldsymbol{A}$, then $\boldsymbol{AGA} = \boldsymbol{A}$. Take transpose on both sides we have $\boldsymbol{A}^T\boldsymbol{G}^T\boldsymbol{A}^T = \boldsymbol{A}^T$. $\boldsymbol{A}$ is symmetric then $\boldsymbol{A}^T = \boldsymbol{A}$, then $\boldsymbol{AG}^T\boldsymbol{A} = \boldsymbol{A}$. Hence $\boldsymbol{G}^T$ is also a generalized inverse of $\boldsymbol{A}$.

(e) Denote $\boldsymbol{G} = (\boldsymbol{X}^T\boldsymbol{X})^-$. As $\boldsymbol{X}^T\boldsymbol{X}$ is symmetric, then $\boldsymbol{G}^T$ is also a generalized inverse of $\boldsymbol{X}^T\boldsymbol{X}$. Hence we have

$$\boldsymbol{XG}^T\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}$$

Take transpose on both sides and we have

$$\boldsymbol{X}^T\boldsymbol{XGX}^T = \boldsymbol{X} \Rightarrow \boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = \boldsymbol{X}^T$$

(f) From (c):

$$\boldsymbol{P_X P_X} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = (\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X})(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = \boldsymbol{P_X}$$

From (e):

$$\boldsymbol{P_X P_X} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-(\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T) = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T = \boldsymbol{P_X}$$

(g) From (e) we have $\boldsymbol{X}^T = \boldsymbol{X}^T\boldsymbol{XG}_2\boldsymbol{X}^T$. Hence we can plug this into $\boldsymbol{XG}_1\boldsymbol{X}^T$ and we have

$$\boldsymbol{XG}_1\boldsymbol{X}^T = \boldsymbol{XG}_1\boldsymbol{X}^T\boldsymbol{XG}_2\boldsymbol{X}^T = (\boldsymbol{XG}_1\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{G}_2\boldsymbol{X}^T = \boldsymbol{XG}_2\boldsymbol{X}^T$$

$\boldsymbol{XG}_1\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{X}$ is from (c).

(h) Denote $(\boldsymbol{X}^T\boldsymbol{X})^- = \boldsymbol{G} = \boldsymbol{G}_1$ and $\boldsymbol{G}^T = \boldsymbol{G}_2$. From (d) we know that $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are both generalized inverse of $\boldsymbol{X}^T\boldsymbol{X}$, and from (g) we have $\boldsymbol{XGX}^T = \boldsymbol{XG}_1\boldsymbol{X}^T = \boldsymbol{XG}_2\boldsymbol{X}^T = \boldsymbol{XG}^T\boldsymbol{X}^T$. Hence

$$\boldsymbol{P}_{\boldsymbol{X}}^T = \boldsymbol{XG}^T\boldsymbol{X}^T = \boldsymbol{XGX}^T = \boldsymbol{P_X}$$