Biost/Stat 533
Spring 2008

# Homework #6

## Problem 1

Let $Y_1, Y_2, \ldots, Y_n$ be random variables with common mean $\theta$ and dispersion matrix $\sigma^2\mathbf{V}$ with $v_{ii} = 1$ and off-diagonal entries $v_{ij} = \rho$. Find the generalized least squares estimate of $\theta$ and show that it is the same as the ordinary least squares estimate.

Solution: The linear model is

$$\boldsymbol{Y} = \mathbf{1}_n\theta + \boldsymbol{\epsilon}$$

where $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^T$, $\mathbf{1}_n \in \mathbb{R}^n$ is the vector whose elements are all 1, and $\boldsymbol{\epsilon} \in \mathbb{R}^n$ satisfies $E[\boldsymbol{\epsilon}] = \mathbf{0}$ and $\text{Var}(\boldsymbol{\epsilon}) = \sigma^2\boldsymbol{V}$. First, we compute $\boldsymbol{V}^{-1}$. To this end, we use the following fact (check this);

$$(\boldsymbol{A} + \boldsymbol{a}\boldsymbol{b}^T)^{-1} = \boldsymbol{A}^{-1} - \frac{(\boldsymbol{A}^{-1}\boldsymbol{a})(\boldsymbol{b}^T\boldsymbol{A}^{-1})}{1 + \boldsymbol{b}^T\boldsymbol{A}^{-1}\boldsymbol{a}}$$

where $\boldsymbol{A}$ is a square matrix and $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors with the same dimension as the number of columns of $\boldsymbol{A}$ and $1 + \boldsymbol{b}^T\boldsymbol{A}^{-1}\boldsymbol{a} \neq 0$. Applying this fact to $\boldsymbol{V} = (1-\rho)\boldsymbol{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^T$, we have (check this)

$$\boldsymbol{V}^{-1} = \frac{1}{1-\rho}\left(\boldsymbol{I}_n - \frac{\rho}{1+(n-1)\rho}\mathbf{1}_n\mathbf{1}_n^T\right)$$

assuming that $1 + (n-1)\rho \neq 0$. The generalized least squares estimate $\theta^*$of $\theta$ is given by

$$\begin{aligned}
\theta^* &= (\mathbf{1}_n^T\boldsymbol{V}^{-1}\mathbf{1}_n)^{-1}\mathbf{1}_n^T\boldsymbol{V}^{-1}\boldsymbol{Y} \\
&= \left(\mathbf{1}_n^T\frac{1}{1-\rho}\left(\boldsymbol{I}_n - \frac{\rho}{1+(n-1)\rho}\mathbf{1}_n\mathbf{1}_n^T\right)\mathbf{1}_n\right)^{-1}\mathbf{1}_n^T\frac{1}{1-\rho}\left(\boldsymbol{I}_n - \frac{\rho}{1+(n-1)\rho}\mathbf{1}_n\mathbf{1}_n^T\right)\boldsymbol{Y} \\
&= (1-\rho)\left(\left(\mathbf{1}_n^T - \frac{n\rho}{1+(n-1)\rho}\mathbf{1}_n^T\right)\mathbf{1}_n\right)^{-1}\frac{1}{1-\rho}\left(\mathbf{1}_n^T - \frac{n\rho}{1+(n-1)\rho}\mathbf{1}_n^T\right)\boldsymbol{Y} \\
&= \left(\left(1 - \frac{n\rho}{1+(n-1)\rho}\right)\mathbf{1}_n^T\mathbf{1}_n\right)^{-1}\left(\left(1 - \frac{n\rho}{1+(n-1)\rho}\right)\mathbf{1}_n^T\right)\boldsymbol{Y} \\
&= \frac{1}{n}\mathbf{1}_n^T\boldsymbol{Y}
\end{aligned}$$

This is the same as the ordinary least squares estimate of $\theta$ obtained in Problem 1 (a) of Homework 4.

Another Solution: Because $\boldsymbol{V}\mathbf{1}_n = (1+(n-1)\rho)\mathbf{1}_n$, it is clear that $\mathcal{R}(\boldsymbol{V}\mathbf{1}_n) = \mathcal{R}(\mathbf{1}_n)$. Then we can apply the Corollay in the page 8 of the Lecture note 12 to conclude the claim.

## Problem 2

Let $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim N_n(\boldsymbol{0}, \sigma^2\boldsymbol{V})$, $\boldsymbol{X}$ is $n \times p$ of rank $p$ and $\boldsymbol{V}$ is a known positive definite $n \times n$ matrix. Let $\boldsymbol{\beta}^*$ be the GLS estimate of $\boldsymbol{\beta}$. Prove that

(a) $Q = ((\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}^*)^T\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}^*)/\sigma^2 \sim \chi^2_{n-p}$.

(b) Find an unbiased estimate of $\sigma^2$.

(c) Define $\boldsymbol{Y}^* = \boldsymbol{X}\boldsymbol{\beta}^*$. If $\boldsymbol{Y}^* = \boldsymbol{P}^*\boldsymbol{Y}$, what is $\boldsymbol{P}^*$? Show that $\boldsymbol{P}^*$ is idempotent but not in general symmetric.

Solution: (a) Because $\boldsymbol{V}$ is positive definite, there exists an invertible matrix $\boldsymbol{K} \in \mathbb{R}^{n\times n}$ such that $\boldsymbol{V} = \boldsymbol{K}\boldsymbol{K}^T$. Thus, let

$$\begin{aligned}\boldsymbol{Z} &\equiv \boldsymbol{K}^{-1}\boldsymbol{Y}\\ \boldsymbol{B} &\equiv \boldsymbol{K}^{-1}\boldsymbol{X}\\ \boldsymbol{\eta} &\equiv \boldsymbol{K}^{-1}\boldsymbol{\epsilon}\end{aligned}$$

and consider the linear model

$$\boldsymbol{Z} = \boldsymbol{B}\boldsymbol{\beta} + \boldsymbol{\eta}.$$

Note that $\boldsymbol{B}$ is of full rank because

$$p = \text{rank}(\boldsymbol{X}) = \text{rank}(\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{X}) \leq \text{rank}(\boldsymbol{K}^{-1}\boldsymbol{X})(= \text{rank}(\boldsymbol{B})) \leq \text{rank}(\boldsymbol{X}) = p$$

and that

$$E[\boldsymbol{\eta}] = 0 \text{ and } Var(\boldsymbol{\eta}) = \sigma^2\boldsymbol{I}.$$

The generalized least squares estimate $\boldsymbol{\beta}^*$ of $\boldsymbol{\beta}$ is the ordinary least squares estimate of $\boldsymbol{\beta}$ in this model and is given by

$$\boldsymbol{\beta}^* = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Z} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}.$$

Applying the Theorem in the page 6 of the Lecture note 8, we have

$$RSS/\sigma^2 \sim \chi^2_{n-p}$$

but

$$\begin{aligned}
RSS &= (\boldsymbol{Z}-\hat{\boldsymbol{Z}})^T(\boldsymbol{Z}-\hat{\boldsymbol{Z}}) \\
&= (\boldsymbol{Z}-\boldsymbol{B}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Z}-\boldsymbol{B}\hat{\boldsymbol{\beta}}) \\
&= (\boldsymbol{K}^{-1}\boldsymbol{Y}-\boldsymbol{K}^{-1}\boldsymbol{X}(\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Z})^T(\boldsymbol{K}^{-1}\boldsymbol{Y}-\boldsymbol{K}^{-1}\boldsymbol{X}(\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Z}) \\
&= (\boldsymbol{Y}-\boldsymbol{X}(\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Z})^T(\boldsymbol{K}^{-1})^T\boldsymbol{K}^{-1}(\boldsymbol{Y}-\boldsymbol{X}(\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Z}) \\
&= (\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}^*)^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}^*).
\end{aligned}$$

Thus the claim that $Q \sim \chi^2_{n-p}$ follows.

(b) From part(a), $(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta^*})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta^*})/(n-p)$ is an unbiased estimate of $\sigma^2$ because

$$E[(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta^*})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta^*})/(n-p)] = \frac{\sigma^2}{n-p}E[Q] = \frac{\sigma^2}{n-p}E[\chi^2_{n-p}] = \sigma^2.$$

(c) As discussed in part(a), we have

$$\boldsymbol{Y}^* = \boldsymbol{X}\boldsymbol{\beta^*} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}$$

so that

$$\boldsymbol{P}^* = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}.$$

$\boldsymbol{P}^*$ is idempotent because

$$\begin{aligned}
\boldsymbol{P}^*\boldsymbol{P}^* &= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1} \\
&= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1} \\
&= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1} \\
&= \boldsymbol{P}^*.
\end{aligned}$$

Also, it can be shown that in general $\boldsymbol{P}^*$ is not symmetric by providing some counterexample (for example, $\boldsymbol{X} = (1,2)^T, \boldsymbol{V} = \text{diag}(1, 1/2)$). Note that computing the transpose is not enough for proof.

## Problem 3

If $\boldsymbol{X}$ is not of full rank, show that any solution $\boldsymbol{\beta}$ of

$$\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}$$

minimizes $(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})$.

Solution: Note that $\boldsymbol{V}$ in this question is assumed to be positive definite. Let $\hat{\boldsymbol{\beta}}$ be an arbitrary solution of $\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta}=\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}$. Then we have $\boldsymbol{X}^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=0$. Using this fact, we have

$$\begin{aligned}
(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}) &= (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}+X(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}))^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}+X(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})) \\
&= (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \\
&\quad +(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\boldsymbol{X}^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \\
&= (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+0+0+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \\
&= (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}).
\end{aligned}$$

Now, note that the first term does not depend on $\boldsymbol{\beta}$ and that the second term is nonnegative because $\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}$ is nonnegative definite (Why?). Since $\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}$ sets the second term to zero, the claim follows.