Biostatistics 533

Classical Theory of Linear Models

Spring 2007

Midterm

Name: KEY

Problems do not have equal value and some problems will take more time than others. Spend your time wisely. This test has six pages including this title page.

Problem	1	2	3	4	5	6	Total
Possible Points	20	5	10	20	10	15	80
Score							

All problems pertain to the linear model

$$\mathbf{Y}_{n\times 1} = \mathbf{X}_{n\times p}\boldsymbol{\beta}_{p\times 1} + \boldsymbol{\varepsilon}_{n\times 1}$$

with $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$.

1. (20 points) Let $\mathbf{P}$ be the projection operator onto $\mathcal{R}(\mathbf{X})$. For least-squares estimation, recall that $\hat{\boldsymbol{\varepsilon}} = (\mathbf{I} - \mathbf{P})\mathbf{Y}$. Derive

(a) $E(\hat{\boldsymbol{\varepsilon}})$

$E(\hat{\boldsymbol{\varepsilon}}) = E((\mathbf{I} - \mathbf{P})\mathbf{Y}) = (\mathbf{I} - \mathbf{P})E(\mathbf{Y}) = (\mathbf{I} - \mathbf{P})\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta} - \mathbf{P}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta} = 0$

(using $\mathbf{P}\mathbf{X} = \mathbf{X}$)

(b) $\text{cov}(\hat{\boldsymbol{\varepsilon}})$

$\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \text{cov}((\mathbf{I} - \mathbf{P})\mathbf{Y}) = (\mathbf{I} - \mathbf{P})\text{cov}(\mathbf{Y})(\mathbf{I} - \mathbf{P})' = (\mathbf{I} - \mathbf{P})\sigma^2\mathbf{I}(\mathbf{I} - \mathbf{P})' = \sigma^2(\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P})' = \sigma^2(\mathbf{I} - \mathbf{P})$ since $\mathbf{I} - \mathbf{P}$ is symmetric and idempotent.

(c) $\text{cov}(\hat{\boldsymbol{\varepsilon}}, \mathbf{P}\mathbf{Y})$

$\text{cov}(\hat{\boldsymbol{\varepsilon}}, \mathbf{P}\mathbf{Y}) = \text{cov}((\mathbf{I} - \mathbf{P})\mathbf{Y}, \mathbf{P}\mathbf{Y}) = (\mathbf{I} - \mathbf{P})\text{cov}(\mathbf{Y})\mathbf{P}' = \sigma^2\mathbf{I}(\mathbf{I} - \mathbf{P})\mathbf{P} = \sigma^2(\mathbf{P} - \mathbf{P}^2) = \sigma^2(\mathbf{P} - \mathbf{P}) = \mathbf{0}$ since $\mathbf{P}$ is symmetric and idempotent.

(d) $E[RSS]$

$E[RSS] = E(\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}) = E(\mathbf{Y}'(\mathbf{I} - \mathbf{P})'(\mathbf{I} - \mathbf{P})\mathbf{Y}) =$

$E(\mathbf{Y}'(\mathbf{I} - \mathbf{P})\mathbf{Y}) = \text{tr}((\mathbf{I} - \mathbf{P})\sigma^2\mathbf{I}) + (\mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \mathbf{P})(\mathbf{X}\boldsymbol{\beta})$

We've used the fact that $\mathbf{I} - \mathbf{P}$ is symmetric and idempotent, and we've used our result for the expectation of a quadratic form. The second term is 0 because $(\mathbf{I} - \mathbf{P})\mathbf{X}$ is 0. So continue:

$E[RSS] = \sigma^2\text{tr}((\mathbf{I} - \mathbf{P})) = \sigma^2(\text{tr}(\mathbf{I}) - \text{tr}(\mathbf{P})) = \sigma^2(n - \text{rank}(\mathbf{P}))$

2. (5 points) Suppose $\hat{\boldsymbol{\beta}}_1 \neq \hat{\boldsymbol{\beta}}_2$ are two different least-squares estimates of $\boldsymbol{\beta}$. Show there are infinitely many least-squares estimates of $\boldsymbol{\beta}$.

We have $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}_1$ and $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}_2$ since $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are both least-squares estimates. Let $\hat{\boldsymbol{\beta}}_p = p\hat{\boldsymbol{\beta}}_1 + (1-p)\hat{\boldsymbol{\beta}}_2$ for any number $p \in (0,1)$. Then $\mathbf{X}\hat{\boldsymbol{\beta}}_p = \mathbf{X}p\hat{\boldsymbol{\beta}}_1 + \mathbf{X}(1-p)\hat{\boldsymbol{\beta}}_2 = p\hat{\mathbf{Y}} + (1-p)\hat{\mathbf{Y}} = \hat{\mathbf{Y}}$ so $\hat{\boldsymbol{\beta}}_p$ is also a solution for any $p \in (0,1)$. This gives infinitely many solutions.

3. (10 points) Suppose rank($\mathbf{X}$) $< p$. Show $\boldsymbol{\beta}$ is not estimable. That is, show there is no matrix $\mathbf{C}$ such that $\mathbf{CY}$ is an unbiased estimate of $\boldsymbol{\beta}$. (Equivalently, show that if $\boldsymbol{\beta}$ is estimable then $\mathbf{X}$ has full rank.)

Solution 1: Suppose there exists a matrix $\mathbf{C}$ such that $E(\mathbf{CY}) = \boldsymbol{\beta}$. $E(\mathbf{CY}) = \mathbf{C}E(\mathbf{Y}) = \mathbf{CX}\boldsymbol{\beta}$. So $\mathbf{CX}\boldsymbol{\beta} = \boldsymbol{\beta}$ $\underline{\text{for all } \boldsymbol{\beta}}$. Therefore $\mathbf{CX} = \mathbf{I}_{p\times p}$. We have rank($\mathbf{I}$) $= p$ but also rank($\mathbf{I}$) $\leq$ rank($\mathbf{X}$) $< p$. Contradiction. Such a $\mathbf{C}$ cannot exist.

If you are not comfortable with proof by contradiction, you might prefer this way of saying things:

Solution 2: If $\boldsymbol{\beta}$ is estimable then in particular each β_i is estimable. That is, $\mathbf{e}_i'\boldsymbol{\beta}$ is estimable for all p-vectors $\mathbf{e}_i$ that are all 0's except for a 1 in the i^{th} position. That means each $\mathbf{e}_i$ is in the row space of $\mathbf{X} \Rightarrow$ rank($\mathbf{X}$) must be at least p. but $\mathbf{X}$ is $n \times p \Rightarrow \mathbf{X}$ has rank at most p. Therefore the rank of $\mathbf{X}$ is p – $\mathbf{X}$ has full rank.

4. (20 points)

(a) What does BLUE stand for?

Best Linear Unbiased Estimate

(b) What does BLUE mean?

An estimator is BLUE for a parameter θ if it is linear, unbiased, and has minimum variance among all linear unbiased estimators.

(c) We proved in class that for the least squares estimator $\hat{\boldsymbol{\theta}}$ of the mean vector of $\mathbf{Y}$, $\mathbf{c}'\hat{\boldsymbol{\theta}}$ is the BLUE of $\mathbf{c}'\boldsymbol{\theta}$ for any $\mathbf{c}$. Using this fact in the case rank$(\mathbf{X}) = p$, prove that $\mathbf{d}'\hat{\boldsymbol{\beta}}$ is the BLUE of $\mathbf{d}'\boldsymbol{\beta}$.

In the full rank case, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ so $\mathbf{d}'\hat{\boldsymbol{\beta}} = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{c}'\mathbf{Y}$ where $\mathbf{c}' = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Therefore $\mathbf{d}'\hat{\boldsymbol{\beta}}$ is linear.
Also, $E[\mathbf{d}'\hat{\boldsymbol{\beta}}] = E[\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}] = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{Y}] = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{d}'\boldsymbol{\beta}$, so $\mathbf{d}'\hat{\boldsymbol{\beta}}$ is unbiased.

$$
\begin{aligned}
\mathbf{c}'\boldsymbol{\theta} &= \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{Y}] = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{d}'\boldsymbol{\beta} \\
\mathbf{c}'\hat{\boldsymbol{\theta}} &= \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\theta}} = \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\
&= \mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{d}'\hat{\boldsymbol{\beta}}
\end{aligned}
$$

So by the theorem, $\mathbf{d}'\hat{\boldsymbol{\beta}}$ (which equals $\mathbf{c}'\hat{\boldsymbol{\theta}}$) is BLUE for $\mathbf{d}'\boldsymbol{\beta}$ (which equals $\mathbf{c}'\boldsymbol{\theta}$).

5. (10 points) Since $\text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$, one might wonder whether the fitted residuals can also be uncorrelated with the same variance. That is, one might wonder whether $\text{cov}(\hat{\boldsymbol{\varepsilon}})$ can have the form $\tau^2\mathbf{I}$. Prove that $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \tau^2\mathbf{I}$ for some $\tau^2 \geq 0$ if and only if $\hat{\mathbf{Y}} = \mathbf{Y}$.

$\Leftarrow$: If $\hat{\mathbf{Y}} = \mathbf{Y}$ then $\hat{\boldsymbol{\varepsilon}} = \mathbf{0}$ so $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \mathbf{0} = \tau^2\mathbf{I}$ for τ=0.

$\Rightarrow$ Solution 1: Suppose $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \tau^2\mathbf{I}$. In general $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \sigma^2(\mathbf{I} - \mathbf{P})$. If $\sigma^2(\mathbf{I} - \mathbf{P}) = \tau^2\mathbf{I}$, then $\mathbf{P} = \frac{\sigma^2 - \tau^2}{\sigma^2}\mathbf{I}$. Then $\mathbf{P}\mathbf{Y} = \hat{\mathbf{Y}} = \frac{\sigma^2-\tau^2}{\sigma^2}\mathbf{Y} \Rightarrow \frac{\sigma^2-\tau^2}{\sigma^2}\mathbf{Y} \in \mathcal{R}(\mathbf{X}) \Rightarrow \mathbf{Y} \in \mathcal{R}(\mathbf{X}) \Rightarrow \hat{\mathbf{Y}} = \mathbf{Y}$.

$\Rightarrow$ Solution 2: Suppose $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \tau^2\mathbf{I}$. In general $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \sigma^2(\mathbf{I} - \mathbf{P})$. If $\sigma^2(\mathbf{I} - \mathbf{P}) = \tau^2\mathbf{I}$, then $\mathbf{P} = \frac{\sigma^2-\tau^2}{\sigma^2}\mathbf{I}$. Since $\mathbf{P} = \mathbf{P}^2$, either $\frac{\sigma^2-\tau^2}{\sigma^2} = 0$ or $\frac{\sigma^2-\tau^2}{\sigma^2} = 1$. We can rule out $\mathbf{P} = 0\mathbf{I}$ since then $\mathbf{X}$ has rank 0. ($\mathbf{X}$ would be a matrix of 0's and you would not actually have a model for your data.) Therefore, $\mathbf{P} = \mathbf{I}$ and $\mathbf{P}\mathbf{Y} = \hat{\mathbf{Y}} = \mathbf{Y}$.

6. (15 points) Circle true or false after each statement

$\hat{\mathbf{Y}}$, the least-squares estimate, is always unique

TRUE

$\hat{\boldsymbol{\beta}}$, the least-squares estimate, is always unique

FALSE

The result that $\hat{\mathbf{Y}}$ is the BLUE of $E[\mathbf{Y}]$ requires the assumption that $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$

FALSE