STAT 571 — Multivariate Analysis — Fall 2014

Exam 2

November 19, 2014

Full Name: Key

- This is a 50 minute exam. There are 6 problems, worth a total of 45 points.
- You may use the textbook (Marden), any personal notes (no limit), and a standard scientific calculator. (You may *not* share these items with anyone else.) No other aids or devices are permitted!
- *Write all answers in the spaces provided.* If you require more space to write your answer, you may use the back side of the page.

Some selected formulas:

Both sides model: $\boldsymbol{Y} = \boldsymbol{x}\boldsymbol{\beta}\boldsymbol{z}' + \boldsymbol{R}$

$\boldsymbol{C_x} = (\boldsymbol{x}'\boldsymbol{x})^{-1}$

$\widehat{\boldsymbol{\Sigma}}_z = \frac{1}{n-p}\boldsymbol{Y}_z'\boldsymbol{Q_x}\boldsymbol{Y}_z$

$\boldsymbol{Y}_z = \boldsymbol{Y}\boldsymbol{z}(\boldsymbol{z}'\boldsymbol{z})^{-1}$

$\boldsymbol{Q_x} = \boldsymbol{I} - \boldsymbol{P_x}, \quad \boldsymbol{P_x} = \boldsymbol{x}(\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}'$

$\text{deviance}(M_k(\widehat{\boldsymbol{\theta}}_k); \boldsymbol{y}) = -2l_k(\widehat{\boldsymbol{\theta}}_k; \boldsymbol{y}) = -2\max_{\boldsymbol{\theta}_k \in \Theta_k} l_k(\boldsymbol{\theta}_k; \boldsymbol{y})$

$\text{AIC}(M_k; \boldsymbol{y}) = \text{deviance}(M_k(\widehat{\boldsymbol{\theta}}_k); \boldsymbol{y}) + 2d_k$

$\text{BIC}(M_k; \boldsymbol{y}) = \text{deviance}(M_k(\widehat{\boldsymbol{\theta}}_k); \boldsymbol{y}) + \log(n)\, d_k$

1. $\boldsymbol{Y}$ is a 50×2 response matrix modeled as follows:

$$\boldsymbol{Y} = \mathbf{1}_{50}\,\beta\,\mathbf{1}_2' + \boldsymbol{R}, \qquad \boldsymbol{R} \sim N(\mathbf{0}, \boldsymbol{I}_{50} \otimes \boldsymbol{\Sigma})$$

where β is an unknown scalar, and $\boldsymbol{\Sigma}$ is positive definite. Suppose

$$\mathbf{1}_{50}'\,\boldsymbol{Y}\,\mathbf{1}_2 = \sum_{ij} Y_{ij} = 230 \qquad \boldsymbol{Y}'\boldsymbol{H}_{50}\boldsymbol{Y} = \boldsymbol{Y}'(\boldsymbol{I} - \tfrac{1}{50}\mathbf{1}_{50}\mathbf{1}_{50}')\boldsymbol{Y} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

(a) Verify the conditions (on $\boldsymbol{x}$ and $\boldsymbol{z}$) for the least squares estimator of β to uniquely exist. [2 pts]

$\underline{x}'\underline{x} = \underline{1}_{50}'\underline{1}_{50} = 50$ is invertible

$\underline{z}'\underline{z} = \underline{1}_2'\underline{1}_2 = 2$ is invertible

(b) Compute the (unweighted) least squares estimate $\hat{\beta}_{LS}$ of β. [3 pts]

$$\hat{\beta}_{LS} = (\underline{x}'\underline{x})^{-1}\underline{x}'\underline{Y}\,\underline{z}(\underline{z}'\underline{z})^{-1}$$

$$= \frac{1}{50}\underline{1}_{50}'\underline{Y}\,\underline{1}_2 \cdot \frac{1}{2} = \frac{1}{100}\cdot 230 = 2.3$$

(c) Compute $\boldsymbol{C_x}$ and $\widehat{\boldsymbol{\Sigma}}_z$. [4 pts]

$$\underline{C}_x = (\underline{x}'\underline{x})^{-1} = \frac{1}{50}$$

$$\hat{\underline{\underline{\Sigma}}}_z = \frac{1}{n-p}\underline{Y}_z'\underline{Q}_x\underline{Y}_z = \frac{1}{50-1}(\underline{z}'\underline{z})^{-1}\underline{z}'\underline{Y}'(\underline{I} - \underline{P}_x)\underline{Y}\,\underline{z}(\underline{z}'\underline{z})^{-1}$$

$$= \frac{1}{49}\cdot\frac{1}{2}\cdot\underline{1}_2'\underline{Y}'(\underline{I} - \frac{1}{50}\underline{1}_{50}\underline{1}_{50}')\underline{Y}\,\underline{1}_2\cdot\frac{1}{2}$$

$$= \frac{1}{4\cdot 49}(1\ 1)\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{7}{4\cdot 49} = \frac{1}{28}$$

(d) Compute a t-statistic for testing $H_0 : \beta = 0$. Also, what are its degrees of freedom? [4 pts]

$$t = \frac{\hat{\beta}_{LS}}{\sqrt{C_{xx}\cdot\hat{\Sigma}_z}} = \frac{2.3}{\sqrt{\frac{1}{50}\cdot\frac{1}{28}}} \approx 86.06$$

$$df = 50 - 1 = 49$$

2. Clearly answer TRUE or FALSE. (Ambiguous answers receive no credit.) [1 pt each]

   (a) AIC will generally choose a model with no more parameters than the model BIC chooses (assuming $n > 7$). FALSE

   (b) A basis cannot contain the vector $\mathbf{0}$ (all zeros). TRUE

   (c) A matrix must be symmetric to have a Cholesky decomposition. TRUE

   (d) The Mallows' C_p^* statistic is always non-negative. TRUE

   (e) Wilks' Λ and Roy's maximum root are always equivalent test statistics. FALSE

   (f) Hotelling's T^2 statistic can be a special case of the Lawley-Hotelling trace statistic. TRUE

   (g) In general, the Gram-Schmidt procedure applied to vectors $\boldsymbol{d}_1$, $\boldsymbol{d}_2$ (in that order) will produce different vectors than if applied to $\boldsymbol{d}_2$, $\boldsymbol{d}_1$ (in that order). TRUE

3. Briefly answer the following:

   (a) Recall $\boldsymbol{C_x} = (\boldsymbol{x}'\boldsymbol{x})^{-1}$. Describe how to find a matrix $\boldsymbol{x}^*$ whose columns span the same linear subspace as those of $\boldsymbol{x}$, but is such that $\boldsymbol{C_{x^*}} = \boldsymbol{I}$. Explain briefly. [3 pts]

Let $\underline{x} = \underline{Q}\,\underline{R}$ be the QR decomposition of $\underline{x}$.

Then $\underline{Q}'\underline{Q} = \underline{I}$ and $\underline{Q}$ has columns that span the same lin. subspace as $\underline{x}$.

Take $\underline{x}^* = \underline{Q}$. (Then $\underline{C}_{x^*} = (\underline{I})^{-1} = \underline{I}$)

   (b) Under what condition can an orthogonal matrix be a projection matrix? Explain briefly. [3 pts]

If projection $\underline{P}$ is orthogonal, then

$\underline{I} = \underline{P}'\underline{P} = \underline{P}\,\underline{P} = \underline{P}$

so $\underline{P}$ must be an identity matrix.

4. Consider matrix

$$\begin{pmatrix} 1 & -1 & 2 \\ 1 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

(a) Of what dimension is the linear subspace spanned by its columns? 2 [1 pt]

(b) Determine an orthogonal basis for the linear subspace spanned by its columns. [3 pts]

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

(c) Would this matrix have a QR decomposition (according to the definition in the textbook)? Explain briefly. [2 pts]

No. It does not have linearly independent columns.

5. Show that the Wilks' Λ statistic $|\boldsymbol{W}|/|\boldsymbol{W}+\boldsymbol{B}|$ depends on $\boldsymbol{W}$ and $\boldsymbol{B}$ only through the matrix $\boldsymbol{W}^{-1}\boldsymbol{B}$ (assuming $\boldsymbol{W}$ is invertible). Justify each step. [4 pts]

$$\Lambda = \frac{|\underline{W}|}{|\underline{W}+\underline{B}|} = \frac{1}{|\underline{W}^{-1}|\,|\underline{W}+\underline{B}|}$$

$$= \frac{1}{|\underline{W}^{-1}(\underline{W}+\underline{B})|} = \frac{1}{|\underline{I}+\underline{W}^{-1}\underline{B}|}$$

which depends only on $\underline{W}^{-1}\underline{B}$.

6. Y is a 45×2 response matrix modeled as follows:

Model 1: $Y = \mathbf{1}_{45}\, \beta\, (1\ 1) + R$ Model 2: $Y = (\mathbf{1}_{45}\ g) \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} + R$

where g is linearly independent of $\mathbf{1}_{45}$, and $R \sim N(\mathbf{0}, I_{45} \otimes \Sigma)$.

The *observed deviance* of Model 1 is 89.18, and that of Model 2 is 81.81.

(a) What constraints on Model 2 would make it equivalent to Model 1? [2 pts]

$$\beta_{12} = \beta_{21} = \beta_{22} = 0$$

(b) Consider the null hypothesis that Model 1 is correct, versus the alternative that Model 2 is correct, but not Model 1. Compute a test statistic, based on the likelihood ratio, that has an approximate chi-square distribution under the null. Also, what would be the degrees of freedom? [3 pts]

$$2 \log(LR) = 2 \log\left(\frac{\sup_{\beta} L_2(\beta; \underline{Y})}{\sup_{\beta} L_1(\beta; \underline{Y})}\right)$$

$$= -2 \log L_1(\hat{\beta}; \underline{Y}) - \left(-2 \log L_2(\hat{\underline{\beta}}; \underline{Y})\right)$$

$$= 89.18 - 81.81 = 7.37$$

$$df = 3 \quad (= 4 - 1)$$

(c) Compute (ordinary) AIC for both models. Which would be selected? [4 pts]

Model 1:

$$AIC_1 = 89.18 + 2 \cdot (\underbrace{1}_{\beta} + \underbrace{3}_{\text{elements of } \Sigma})$$

$$= 97.18$$

Model 2:

$$AIC_2 = 81.81 + 2 \cdot (4 + 3)$$

$$= 95.81$$

Model 2 is selected (smaller AIC).