STAT 571 — Multivariate Analysis — Fall 2015

Exam 2

November 18, 2015

Full Name: ____

- This is a 50 minute exam. There are 5 problems, worth a total of 35 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: $\boldsymbol{Y} = \boldsymbol{x}\boldsymbol{\beta}\boldsymbol{z}' + \boldsymbol{R}$

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

$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} = \dfrac{1}{n-p}\boldsymbol{Y}_{\boldsymbol{z}}'\boldsymbol{Q_x}\boldsymbol{Y_z}$

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

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

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

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

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

1. (a) Use Gram-Schmidt on the columns of the matrix below:

$$\begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & -2 \end{pmatrix}$$

(3pts)

Ans: First dot $\mathbf{d}_1 = (1,1,1)'$ out of $\mathbf{d}_2 = (3,0,0)'$:

$$\mathbf{d}_{2.1} = \mathbf{d}_2 - \frac{\mathbf{d}_1'\mathbf{d}_2}{\mathbf{d}_1'\mathbf{d}_1}\mathbf{d}_1 = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} - \frac{(1,1,1)(3,0,0)'}{(1,1,1)(1,1,1)'}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} - \frac{3}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$$

You can stop now, since $(1,1,1)'$, $(2,-1,-1)'$ and $(0,2,-2)'$ are mutually orthogonal. Or you can proceed, but dotting $(1,1,1)'$ out of $(0,2,-2)'$ gives you back $(0,2,-2)'$, then dotting $(2,-1,-1)'$ out of $(0,2,-2)'$ gives you back $(0,2,-2)'$.

(b) Find an orthogonal basis for the span of the columns of the following matrix:

$$\begin{pmatrix} 1 & -3 & 5 \\ 1 & 0 & 0 \\ 1 & 3 & -5 \end{pmatrix}$$

(2pts)

Answer: You can use Gram-Schmidt again. Then $\mathbf{1}$ vector is orthogonal to the other two, so dotting it out does not change anything. Then note that the second the third vectors are proportional, so if you dot the second out of the third, you end up with $\mathbf{0}$. Throw that out, and your orthogonal basis is just the first two vectors.

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

(a) AIC$_c$ will generally choose a model with no less parameters than the model AIC chooses (assuming $n > p + q + 1$). (F)

(b) Let D be a $n \times p$ matrix and A be a $p \times p$ matrix. If the columns of D are linearly independent, then the columns of DA is also linear independent provided that A is invertible. (T)

(c) $\{\mathbf{x} = (x_1, x_2) \in R^2 | x_1 > 0\}$ is a linear subspace. (F)

(d) Any symmetric idempotent matrix is a projection matrix for some subspace. (T)

(e) In the both-sides model $\mathbf{Y} = \mathbf{x}\beta\mathbf{z}' + \mathbf{R}$, where $\mathbf{R} \sim N(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{\Sigma}_R)$, and $\mathbf{x}'\mathbf{x}$, $\mathbf{z}'\mathbf{z}$ and $\mathbf{\Sigma}_R$ are assumed to be invertible. Then the weighted least square estimate of β and the least square estimator are equivalent when $\mathbf{z}$ is a square and invertible matrix. (T)

(f) For any $q \times q$ symmetric positive definite matrix $\mathbf{S}$, there exists an invertible matrix $\mathbf{R}$, such that $\mathbf{S} = \mathbf{R}'\mathbf{R}$, and furthermore, $\mathbf{R}$ is unique. (F)

(g) Suppose $(\mathbf{D}_1\ \mathbf{D}_2)$ is $N \times K$, where $\mathbf{D}_1$ is $N \times K_1$ and $\mathbf{D}_2$ is $N \times K_2$ and $\mathcal{W} = span\{columnsof\ (\mathbf{D}_1\ \mathbf{D}_2)\}$. Suppose $\mathbf{D}_1'\mathbf{D}_1$ is invertible, and let $\mathbf{D}_{2\cdot 1} = \mathbf{Q}_{\mathbf{D}_1}\mathbf{D}_2$. Then the columns of $\mathbf{D}_1$ and $\mathbf{D}_{2\cdot 1}$ are orthogonal and $\mathcal{W} = span\{columnsof(\mathbf{D}_1\ \mathbf{D}_{2\cdot 1})\}$. (T)

(h) Suppose $(\mathbf{D}_1\ \mathbf{D}_2)$ is $N \times K$, where $\mathbf{D}_1$ is $N \times K_1$ and $\mathbf{D}_2$ is $N \times K_2$, and $\mathbf{D}_1'\mathbf{D}_2 = \mathbf{0}$. Then $\mathbf{P}_{\mathbf{D}} = \mathbf{P}_{\mathbf{D}_1} + \mathbf{P}_{\mathbf{D}_2}$. (T)

3. Consider the multivariate regression model

$$\mathbf{Y} = \mathbf{x}\beta + \mathbf{R},\ \mathbf{R} \sim N(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{\Sigma}_R),$$

where $\mathbf{Y}$ is $n \times q$, and $\mathbf{x}$ is $n \times n$ and invertible.

(a) Show that the least squares estimate of β is

$$\hat{\beta} = \mathbf{x}^{-1}\mathbf{Y}.$$

You can start with the usual formula for the estimator. (2pts)

Ans: There is no $\mathbf{z}$, so

$$\hat{\beta} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{Y}.$$

Since $\mathbf{x}$ itself is invertible, $(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \mathbf{x}^{-1}(\mathbf{x}')^{-1}\mathbf{x}' = \mathbf{x}^{-1}$, the last two matrices canceling.

(b) Find $cov[\hat{\beta}]$. (2pts)

Ans: From the usual formula, without $\mathbf{z}$ (or with $\mathbf{z} = \mathbf{I}_q$),

$$cov(\hat{\beta}) = \mathbf{C}_x \otimes \mathbf{\Sigma}_z = (\mathbf{x}'\mathbf{x})^{-1} \otimes \mathbf{\Sigma}_R.$$

(c) Find $\mathbf{P}_x$ and $\mathbf{Q}_x$ explicitly. (All the entries in these matrices are integers. What are these integers?) What are $\hat{\mathbf{Y}}$ and $\hat{\mathbf{R}}$? (4pts)

Ans: Using the formula, then writing out the inverses:

$$\mathbf{P}_x = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \mathbf{x}\mathbf{x}^{-1}(\mathbf{x}')^{-1}\mathbf{x}' = \mathbf{I}_n.$$

Then

$$\mathbf{Q}_x = \mathbf{I}_n - \mathbf{P}_x = \mathbf{I}_n - \mathbf{I}_n = \mathbf{0}.$$

So $\hat{\mathbf{Y}} = \mathbf{P}_x\mathbf{Y} = \mathbf{Y}$ and $\hat{\mathbf{R}} = \mathbf{0}$. Or, use $\hat{\mathbf{Y}} = \mathbf{x}\hat{\beta} = \mathbf{x}\mathbf{x}^{-1}\mathbf{Y} = \mathbf{Y}$.

4. In the prostaglandin data set, measurements were taken at six four-hour intervals over the course of a day for 10 individuals. The measurements are prostaglandin contents in their urine. The data matrix $\mathbf{Y}$ is then 10×6. We assume the same sine/cosine curve is fit to the measurements for each man, so that $\mathbf{x} = \mathbf{1}_{10}$, and

$$\mathbf{z} = \begin{pmatrix} 1 & \cos(2\pi(1/6)) & \sin(2\pi(1/6)) \\ 1 & \cos(2\pi(2/6)) & \sin(2\pi(2/6)) \\ 1 & \cos(2\pi(3/6)) & \sin(2\pi(3/6)) \\ 1 & \cos(2\pi(4/6)) & \sin(2\pi(4/6)) \\ 1 & \cos(2\pi(5/6)) & \sin(2\pi(5/6)) \\ 1 & \cos(2\pi(6/6)) & \sin(2\pi(6/6)) \end{pmatrix}.$$

The model is the both-sides model,

$$\mathbf{Y} = \mathbf{x}\beta\mathbf{z}' + \mathbf{R},\ \mathbf{R} \sim N(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{\Sigma}_R).$$

The least squares estimate of β is

$$\hat{\beta} = (\hat{\beta}_1\ \hat{\beta}_2\ \hat{\beta}_3) = (188.5\ \ -53.18\ 4.186),$$

and the estimate of $\mathbf{\Sigma}_z$ is

$$\hat{\Sigma}_z = \begin{pmatrix} 2076 & -454 & 273 \\ -454 & 1283 & 235 \\ 273 & 235 & 281 \end{pmatrix}.$$

(a) Show that $\mathbf{C}_x = \frac{1}{10}$. (1pts)

Ans: $\mathbf{C}_x = (\mathbf{x}'\mathbf{x})^{-1}$, and $\mathbf{x} = \mathbf{1}_{10}$, so

$$\mathbf{C}_x = (\mathbf{1}_{10}'\mathbf{1}_{10})^{-1} = \frac{1}{10}.$$

(b) Find the estimated standard error of $\hat{\beta}_2$. Is the $\hat{\beta}_2$ statistically significant? (3pts)

Ans: Here,

$$cov(\hat{\beta}) = (\mathbf{x}'\mathbf{x})^{-1} \otimes \mathbf{\Sigma}_z = \frac{1}{10}\mathbf{\Sigma}_z,$$

so that variances of $\hat{\beta}_j$ is the jth diagonal of that, $\sigma_{zjj}/10$. The estimates of the variances of the three $\hat{\beta}_j$'s are, respectively, 207.6, 128.3 and 28.1. Square-root those to get the se's:

$$se(\hat{\beta}_2) = \sqrt{128.3} = 11.33, \ se(\hat{\beta}_3) = \sqrt{28.1} = 5.30.$$

The t's are

$$\hat{\beta}_2/se(\hat{\beta}_2) = -53.18/11.33 = -4.69 \text{ and } \hat{\beta}_3/se(\hat{\beta}_3) = 4.186/28.1 = 0.79.$$

So $\hat{\beta}_2$ is statistically significant, and $\hat{\beta}_3$ is not. (You had to do only one.)

(c) Consider testing the null hypothesis that $\beta_2 = \beta_3 = 0$. The $T^2 = 30.44$. Find the F version of the statistic. What are its degrees of freedom? Do you reject the null hypothesis? (You can use 4 as the cutoff point.) (4pts)

Answer: You need the various dimensions. So $\mathbf{x}$ is $n \times p = 10 \times 1$, $\mathbf{Y}$ is $n \times q = 10 \times 6$, and $\beta^* = (\beta_2 \ \beta_3)$ is $p^* \times l^* = 1 \times 2$. This means that $v = n - p = 9$, so that

$$F = \frac{v - l^* + 1}{vp^*l^*}T^2 = \frac{9 - 2 + 1}{9 \cdot 1 \cdot 2}T^2 = \frac{4}{9}(30.44) = 13.53.$$

Under the null, the F is $F_{p^*l^*, v-l^*+1} = F_{2,8}$, so the $df = (2, 8)$. You can reject the null.

5. Consider the mouth-size data, where the $\mathbf{x} = (\mathbf{1}_n, \mathbf{g})$, where g is the vector with 1's for the girls and 0's for the boys, and the $\mathbf{z}$ has the constant, linear, and quadratic orthogonal polynomial vectors for the growth curves. The $\mathbf{Y}$ is 27×4, so

$$\mathbf{Y} = \mathbf{x}\beta\mathbf{z}' + \mathbf{R} = \begin{pmatrix} \mathbf{1}_{11} & \mathbf{1}_{11} \\ \mathbf{1}_{16} & \mathbf{0}_{16} \end{pmatrix} \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 \\ \delta_1 & \delta_2 & \delta_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \end{pmatrix} + \mathbf{R}.$$

(a) Three submodels are considered in the table below. The table has the patterns, deviances and BIC-based estimates of the probabilities of the models. You are to fill in the dimensions, and the BIC for the first model. (You don't have to fill in the other two BIC's). (2pts)

Ans: The dimensions are the number of parameters in the $\mathbf{\Sigma}$, which is $4 \times 5/2 = 10$, plus the non-zero β_{ij}'s, which is $3, 4$ and 6 for the three models. Since the number of $\mathbf{\Sigma}$ parameters is the same for the three models, you could also just use the $3, 4, 6$. Then the BIC=deviance+log(n)(dimension), where $n = 27$. The complete table is

(b) From the table in part (a), find the estimated probability that in the true model, the boys' and girls' curves are not parallel. (4pts)

Ans: The second row in β gives the differences between the boys and girls for the intercept, slopes, and quadratic terms, respectively. Thus the curves are not parallel if δ_2 or δ_3 are not 0. So add the probabilities of the second and third models, to get $75.59 + 12.55 = 88.14\%$.

Find the estimated Probability that the curves for both the boys and girls are actually straight lines.

Ans: They are both straight lines as long as there are no quadratic terms, so you leave out the third model: $100 - 12.55 = 87.45\%$.

Pattern	$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$
deviance	228	221	218
Dimension			
BIC		***	***
$\widehat{Prob}$	11.86	75.59	12.55

Pattern	$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$
deviance	228	221	218
Dimension	13	14	16
BIC	270.85	267.14	270.73
$\widehat{Prob}$	11.86	75.59	12.55