

Stat 511 Exam1

**February 26, 2008
Prof. Vardeman**

I have neither given nor received unauthorized assistance on this exam.

KEY

Name

Name Printed

1. Consider a problem of quadratic regression in one variable, x . In particular, suppose that $n = 5$ values of a response y are related to values $x = 0, 1, 2, 3, 4$ by a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ for

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}$$

a) A cofactor expansion of the determinant of a 3×3 matrix $\mathbf{A} = (a_{ij})$ on its first row is

$$\det(\mathbf{A}) = a_{11} \det(\mathbf{M}_{11}) - a_{12} \det(\mathbf{M}_{12}) + a_{13} \det(\mathbf{M}_{13})$$

where \mathbf{M}_{ij} is obtained from \mathbf{A} by deleting its i th row and j th column. Use this fact and argue that \mathbf{X} is of full rank.

Consider The rank of the submatrix of X consisting of its first 3 rows. This is no more than the rank of X (this submatrix has no more linearly independent rows than does X). Expanding the determinant of this submatrix on its first row, one gets $1 \cdot (4 - 2) = 2 \neq 0$. That is, this submatrix is non-singular and thus of rank 3. Hence X has rank at least 3 and is thus full rank.

b) Define

$$\mathbf{F} = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \text{ so that } \mathbf{F} \text{ is nonsingular and } \mathbf{F}^{-1} = \begin{pmatrix} 1 & 2 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

Let $\mathbf{W} = \mathbf{X}\mathbf{F}$. Argue carefully that $C(\mathbf{W}) = C(\mathbf{X})$.

$\mathbf{W} = \mathbf{X}\mathbf{F}$ implies that every column of \mathbf{W} is a linear combination of columns of \mathbf{X} , so $C(\mathbf{W}) \subset C(\mathbf{X})$. But it's also true that $\mathbf{X} = \mathbf{W}\mathbf{F}^{-1}$ so that $C(\mathbf{X}) \subset C(\mathbf{W})$. Thus $C(\mathbf{W}) = C(\mathbf{X})$.

c) Notice that $W'W$ is diagonal. Suppose that $Y' = (-2, 0, 4, 2, 2)$. Find the OLS estimate of γ in the model $Y = W\gamma + \varepsilon$ and then OLS estimate of β in the original model. (Find numerical values.)

$$W = \begin{pmatrix} 1 & -2 & 2 \\ 1 & -1 & -1 \\ 1 & 0 & -2 \\ 1 & 1 & -1 \\ 1 & 2 & 2 \end{pmatrix} \text{ so that } W'W = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 14 \end{pmatrix} \text{ and } (W'W)^{-1} = \text{diag}\left(\frac{1}{5}, \frac{1}{10}, \frac{1}{14}\right)$$

$$\text{Then } \hat{\gamma}_{OLS} = (W'W)^{-1} W'Y = \text{diag}\left(\frac{1}{5}, \frac{1}{10}, \frac{1}{14}\right) \cdot \begin{pmatrix} 6 \\ 10 \\ -10 \end{pmatrix} = \begin{pmatrix} \frac{6}{5} \\ 1 \\ -\frac{10}{14} \end{pmatrix}$$

and since $W\hat{\gamma}_{OLS} = X\hat{\beta}_{OLS}$ we must have $\hat{\beta}_{OLS} = F\hat{\gamma}_{OLS}$

$$\hat{\gamma}_{OLS} = \begin{pmatrix} 6/5 \\ 1 \\ -10/14 \end{pmatrix} \quad \hat{\beta}_{OLS} = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \hat{\gamma}_{OLS} = \begin{pmatrix} -78/35 \\ 27/7 \\ -10/14 \end{pmatrix}$$

d) As it turns out, for the set of observations in c), $SSE = 128/35$. What then is an estimated variance-covariance matrix for $\hat{\gamma}_{OLS}$? What is the corresponding estimated variance-covariance matrix for $\hat{\beta}_{OLS}$? (No need to simplify/do matrix multiplications.)

$$\text{Cov } \hat{\gamma}_{OLS} = \sigma^2 (W'W)^{-1} \text{ so } \widehat{\text{Cov}} \hat{\gamma}_{OLS} = \widehat{MSE} (W'W)^{-1}$$

$$\text{i.e. } \widehat{\text{Cov}} \hat{\gamma}_{OLS} = \frac{1}{5-3} \left(\frac{128}{35} \right) \text{diag}\left(\frac{1}{5}, \frac{1}{10}, \frac{1}{14}\right)$$

So then $\hat{\beta}_{OLS} = F\hat{\gamma}_{OLS}$ has estimated covariance matrix

$$F \widehat{\text{Cov}} \hat{\gamma}_{OLS} F' = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \left(\frac{128}{70} \right) \text{diag}\left(\frac{1}{5}, \frac{1}{10}, \frac{1}{14}\right) \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & -1 \\ 2 & -4 & 1 \end{pmatrix}$$

$\widehat{\text{Cov}} \hat{\beta}_{OLS}$

e) Based on your answer to d) give 95% confidence limits for the difference in mean responses for $x = 3$ and $x = 1$ in the normal Gauss-Markov model. (No need to simplify. If you don't have appropriate tables with you, tell me exactly what quantile(s) of what distribution you need.)

The quantity under discussion is $\underset{\sim}{c}'\underset{\sim}{\beta} = (0, 3-1, 9-1)\underset{\sim}{\beta}$.
 This has point estimate $(0, 2, 8)\hat{\beta}_{OLS}$ and confidence limits

$$(0, 2, 8)\hat{\beta}_{OLS} \pm t \sqrt{(0, 2, 8) \text{Cov} \hat{\beta}_{OLS} \begin{pmatrix} 0 \\ 2 \\ 8 \end{pmatrix}}$$

↑
↑
↑

from part c)
upper 2.5% pt of t_{5-3} dsu
from part d)

f) Consider the hypothesis

$H_0 : [x = 0 \text{ and } x = 4 \text{ have the same mean response}] \text{ and } [x = 1 \text{ and } x = 3 \text{ have the same mean response}]$

(Notice that in this model this is the hypothesis that the quadratic regression function has a critical point at $x = 2$.) Show how to test this hypothesis. (Show how to compute an appropriate test statistic and say exactly what null distribution you could use.) Be careful. Some ways of writing this hypothesis may not produce a "testable" hypothesis.

In the 2nd parameterization (W and $\underset{\sim}{\gamma}$) This is $(1 - 2 \ 2)\underset{\sim}{\gamma} = (1 \ 2 \ 2)\underset{\sim}{\gamma}$ and $(1 - 1 - 1)\underset{\sim}{\gamma} = (1 \ 1 - 1)\underset{\sim}{\gamma}$ i.e. $-2\gamma_1 = 2\gamma_1$ (i.e. $\gamma_1 = 0$) and $-1\gamma_1 = 1\gamma_1$ (i.e. $\gamma_1 = 0$).
 That is, this hypothesis is $H: \gamma_1 = 0$ in the 2nd parameterization. This can, for example, be tested

using

$$T = \frac{\hat{\gamma}_{OLS1} - 0}{\sqrt{MSE} \sqrt{\frac{1}{10}}} = \frac{1 - 0}{\sqrt{\frac{128/35}{5-3}} \sqrt{\frac{1}{10}}} = 2.34$$

and a t_{5-3} reference dsu.

2. Notice that the first characterization of the estimability of $\mathbf{c}'\boldsymbol{\beta}$ in the linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ can be rephrased as " $\mathbf{c}'\boldsymbol{\beta}$ is estimable iff there is a linear combination of the variables y_1, y_2, \dots, y_n with expected value $\mathbf{c}'\boldsymbol{\beta}$ for all $\boldsymbol{\beta}$."

a) Consider the "effects model version of the 3-group/2-observations-per-group" one-way ANOVA model used repeatedly as an example in lecture early in the term (version 2) of example b)). Using the characterization of estimability above, show that $\mu + \frac{1}{3}(\tau_1 + \tau_2 + \tau_3)$ is estimable.

This $\mathbf{c}'\boldsymbol{\beta}$ is the arithmetic average of the 3 group means (i.e. is $\frac{1}{3}((\mu + \tau_1) + (\mu + \tau_2) + (\mu + \tau_3))$ i.e. in the cell means model is $\frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$). The linear combination of group sample means

$$\frac{1}{3}(\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})$$

is clearly unbiased for this $\mathbf{c}'\boldsymbol{\beta}$ and so the parametric function is estimable.

b) In the context of part a) above, give formulas for both the BLUE of $\mu + \frac{1}{3}(\tau_1 + \tau_2 + \tau_3)$ AND an unbiased linear estimator of this quantity that is NOT the BLUE.

The Gauss-Markov thm says that $\hat{\boldsymbol{\beta}}_{OLS}$ is the BLUE here. In the cell means version of the model (that is full rank) the OLS estimator of $(\mu_1, \mu_2, \mu_3)'$ is $(\bar{y}_{1.}, \bar{y}_{2.}, \bar{y}_{3.})'$ and the BLUE of $\mathbf{c}'\boldsymbol{\mu}$ is $\mathbf{c}'(\bar{y}_{1.}, \bar{y}_{2.}, \bar{y}_{3.})'$. That is, the BLUE here is

$$\frac{1}{3}(\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.})$$

A different (not best) linear unbiased estimator of this quantity is

$$\frac{1}{3}(y_{11} + y_{21} + y_{31})$$

3. Attached to this exam is an R printout made in the analysis of data from a pilot plant chemical process. Process yield is thought to depend upon condensation temperature and amount of boron employed. A "quadratic response surface" model of the form

$$y = \beta_0 + \beta_1 t + \beta_2 B + \beta_3 t^2 + \beta_4 B^2 + \beta_5 tB + \varepsilon$$

has been fit to the yield data (and by the way, predicts a maximum yield near $t = -1$ and $B = .5$).

a) Give the value of an F statistic and degrees of freedom for testing the hypothesis that $EY \in C((1|t|B))$ in this model.

This hypothesis is the hypothesis $H_0: \beta_3 = \beta_4 = \beta_5 = 0$ and can be handled in the "Full Model / Reduced Model" paradigm.

$$F = \frac{\left(\begin{array}{l} \text{Total of "sequential sums of squares" for the} \\ \text{"last" 3 predictors} \end{array} \right) / 3}{SSE_{\text{full}} / (9 - 6)}$$

$$= \frac{(1.742 + 16.056 + 3.803) / 3}{.0940} = 76.6$$

$F = \underline{76.6}$ d.f. = 3, 3

b) Which observation (or observations) appears (appear) to be most influential in determining the character of the fitted surface? "Where" is (are) that (those) observation (observations) in (t, B) -space? (Note that there is a plot on the printout.) Provide rationale for your choice.

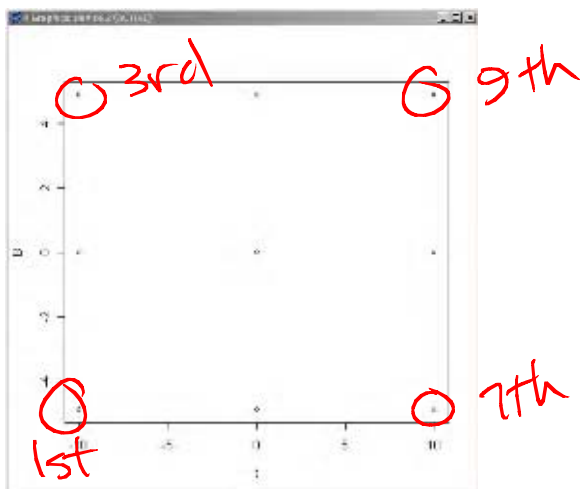
This is a question of which observations have the biggest diagonal entries in the "hat matrix" P_X . These are the 1st, 3rd, 7th and 9th observations as listed in the data set.

These are (not surprisingly) the observations "at the 4 corners" of the rectangular region in 2-space $((t, B)$ -space) where one has data.

```
> chemical
  yield time temp Bamt
1  21.1  150   90  24.4
2  23.7   10   90  29.3
3  20.7   8    90  34.2
4  21.1   35  100  24.4
5  24.1   8   100  29.3
6  22.2   7   100  34.2
7  18.4  18  110  24.4
8  23.4   8  110  29.3
9  21.9  10  110  34.2
```

```
> t<-temp-100
> B<-Bamt-29.3
```

```
> plot(t,B)
```



```
> t2<-t^2
> B2<-B^2
> tB<-t*B
```

```
> quadyield<-lm(yield~t+B+t2+B2+tB)
```

```
> summary(quadyield)
```

Call:

```
lm(formula = yield ~ t + B + t2 + B2 + tB)
```

Residuals:

```
      1      2      3      4      5      6      7      8      9
-0.06389 -0.02222  0.08611  0.27778 -0.25556 -0.02222 -0.21389  0.27778 -0.06389
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	24.355556	0.228499	106.589	1.82e-06 ***
t	-0.030000	0.012515	-2.397	0.096130 .
B	0.142857	0.025542	5.593	0.011289 *
t2	-0.009333	0.002168	-4.306	0.023061 *
B2	-0.118006	0.009028	-13.070	0.000967 ***
tB	0.019898	0.003128	6.361	0.007863 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3066 on 3 degrees of freedom
Multiple R-squared: 0.9889, Adjusted R-squared: 0.9704
F-statistic: 53.37 on 5 and 3 DF, p-value: 0.00394

```
> vcov(quadyield)
```

	(Intercept)	t	B	t2	B2	tB
(Intercept)	5.221193e-02	1.688218e-19	5.681659e-20	-3.132716e-04	-1.304755e-03	-6.622578e-20
t	1.688218e-19	1.566358e-04	5.186088e-21	-7.078389e-22	-9.409579e-21	-4.449178e-22

```

B          5.681659e-20  5.186088e-21  6.523774e-04 -6.614999e-22 -4.488606e-21  2.036626e-21
t2         -3.132716e-04 -7.078389e-22 -6.614999e-22  4.699074e-06  1.507303e-22  7.738612e-22
B2         -1.304755e-03 -9.409579e-21 -4.488606e-21  1.507303e-22  8.151321e-05  1.951499e-21
tB         -6.622578e-20 -4.449178e-22  2.036626e-21  7.738612e-22  1.951499e-21  9.785660e-06

```

```

> anova(quadyield)
Analysis of Variance Table

```

```

Response: yield
      Df Sum Sq Mean Sq F value Pr(>F)
t      1  0.5400  0.5400    5.7458 0.0961299 .
B      1  2.9400  2.9400   31.2828 0.0112892 *
t2     1  1.7422  1.7422   18.5379 0.0230607 *
B2     1 16.0556 16.0556  170.8374 0.0009672 ***
tB     1  3.8025  3.8025   40.4601 0.0078628 **
Residuals 3  0.2819  0.0940
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

> predict(quadyield, se.fit=T)
$fit
      1      2      3      4      5      6      7      8      9
21.16389 23.72222 20.61389 20.82222 24.35556 22.22222 18.61389 23.12222 21.96389

```

```

$se.fit
[1] 0.2751496 0.2284993 0.2751496 0.2284993 0.2284993 0.2284993 0.2751496 0.2284993 0.2751496

```

```

$df
[1] 3

```

```

$residual.scale
[1] 0.306564

```

```

> ones<-rep(1,9)
> X<-cbind(ones,t,B,t2,B2,tB)

```

```

> X%*%ginv(t(X)%*%X,tol=1e-15)%*%t(X)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,] 0.80555556 0.22222222 -0.02777778 0.22222222 -0.11111111 -0.11111111 -0.02777778 -0.11111111
[2,] 0.22222222 0.55555556 0.22222222 -0.11111111 0.22222222 -0.11111111 -0.11111111 0.22222222
[3,] -0.02777778 0.22222222 0.80555556 -0.11111111 -0.11111111 0.22222222 0.13888889 -0.11111111
[4,] 0.22222222 -0.11111111 -0.11111111 0.55555556 0.22222222 0.22222222 0.22222222 -0.11111111
[5,] -0.11111111 0.22222222 -0.11111111 0.22222222 0.55555556 0.22222222 -0.11111111 0.22222222
[6,] -0.11111111 -0.11111111 0.22222222 0.22222222 0.22222222 0.55555556 -0.11111111 -0.11111111
[7,] -0.02777778 -0.11111111 0.13888889 0.22222222 -0.11111111 -0.11111111 0.80555556 0.22222222
[8,] -0.11111111 0.22222222 -0.11111111 -0.11111111 0.22222222 -0.11111111 0.22222222 0.55555556
[9,] 0.13888889 -0.11111111 -0.02777778 -0.11111111 -0.11111111 0.22222222 -0.02777778 0.22222222
      [,9]
[1,] 0.13888889
[2,] -0.11111111
[3,] -0.02777778
[4,] -0.11111111
[5,] -0.11111111
[6,] 0.22222222
[7,] -0.02777778
[8,] 0.22222222
[9,] 0.80555556

```

The "hat" matrix P_X