Math 584 2019 Exam 3 Name___________

Q9821

3.7

1) An overparameterized two way Anova model is $Y_{ijk} = \mu + \alpha_i + \beta_j + \tau_{ij} + \epsilon_{ijk}$ for $i = 1, ..., a$ and $j = 1, ..., b$ and $k = 1, ..., m$. Suppose $a = 2$, $b = 2$, and $m = 2$. Then

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \boldsymbol{X} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \tau_{11} \\ \tau_{12} \\ \tau_{21} \\ \tau_{22} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}.$$

$$X\beta = \begin{bmatrix} \mu+\alpha_1+\beta_1+\tau_{11} \\ \mu+\alpha_1+\beta_1+\tau_{11} \\ \mu+\alpha_1+\beta_2+\tau_{12} \\ \mu+\alpha_1+\beta_2+\tau_{12} \\ \mu+\alpha_2+\beta_1+\tau_{21} \\ \mu+\alpha_2+\beta_1+\tau_{21} \\ \mu+\alpha_2+\beta_2+\tau_{22} \\ \mu+\alpha_2+\beta_2+\tau_{22} \end{bmatrix}$$

a) Give the matrix $\boldsymbol{X}$.

$$\begin{array}{ccccccccc} \mu & \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \tau_{11} & \tau_{12} & \tau_{21} & \tau_{22} \end{array}$$
$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

bc can

b) Write the above model as $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$. This model is **not full rank**. What is the projection matrix $\boldsymbol{P}$ (onto the column space of $\boldsymbol{X}$)? Hint: $\boldsymbol{X}^T\boldsymbol{X}$ is singular, so use the generalized inverse.

$$\boxed{X(X'X)^{-}X'}$$

8

$\mathcal{e}$ 2) In numerical linear algebra, the least squares solution to "$\boldsymbol{Ax} = \boldsymbol{b}$" is of interest where the problem is actually the multiple linear regression model $\boldsymbol{b} = \boldsymbol{Ax} + \boldsymbol{\epsilon}$ where $\boldsymbol{A}$ has full rank p, and we will assume that $E(\boldsymbol{\epsilon}) = \boldsymbol{0}$, and $Cov(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}_n$.

a) What is the (formula for the) projection matrix $\boldsymbol{P}$ onto the column space of $\boldsymbol{A}$?

$A (A'A)^{-1} A'$

Q10)20

b) What is the OLS estimator $\hat{\boldsymbol{x}}$?

$(A'A)^{-1} A' \underset{\sim}{b}$

c) What is the vector of fitted values $\hat{\boldsymbol{b}}$?

$\underline{P}\, \underline{b}$

d) What is the residual vector $\boldsymbol{e}$?

$(\underline{I} - \underline{P})\, \underline{b}$

20

3) Consider the multivariate linear regression model $\boldsymbol{Z} = \boldsymbol{XB} + \boldsymbol{E}$.
a) What is $E(\boldsymbol{Z})$?

XB

b) What is $E(\widehat{\boldsymbol{B}})$?

B

c) What is $E(\boldsymbol{E})$?

0

d) Consider testing $H_0 : \boldsymbol{LB} = \boldsymbol{0}$. What is the limiting distribution of $(n-p)U(\boldsymbol{L})$?

$$(n-p)U(\boldsymbol{L}) \xrightarrow{D} \chi^2_{rm}$$

20

4) The data below are a sorted residuals from a lasso regression where $n = 1000$ and $p = 17$. Find shorth(997) of the residuals.

number	1	2	3	4	...	997	998	999	1000
residual	-3.28	-3.06	-3.04	-2.96	...	2.66	2.71	2.81	3.62

5.94 = 2.66 − (−3.28)

5.77 = 2.71 − (−3.06)

5.85 = 2.81 − (−3.04)

6.58 = 3.62 − (−2.96)

$\boxed{[-3.06, 2.71]}$ = shorth 997

5

$E(Y) = X\beta = \begin{pmatrix} \beta_1 - 2\beta_2 \\ \beta_1 - 2\beta_2 \end{pmatrix} \Rightarrow X = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \Rightarrow X^T = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$

$C(X^T) = \text{span}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$

5) Suppose that $\boldsymbol{Y} = (Y_1, Y_2)'$, $\text{Var}(\boldsymbol{Y}) = \sigma^2 \boldsymbol{I}$, $E(Y_1) = E(Y_2) = \beta_1 - 2\beta_2$. Show whether or not the following functions are estimable. Hint $E(\boldsymbol{Y}) = \boldsymbol{X\beta}$, so find $\boldsymbol{X}$.

a) β_1 $= (1 \;\; 0)\beta$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \notin C(X^T)$

0/2 1

no

b) β_2 $= (0 \;\; 1)\beta$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \notin C(X^T)$

no

3.8

c) $-\beta_1 + 2\beta_2$ $= (-1 \;\; 2)\beta$, $\begin{pmatrix} -1 \\ 2 \end{pmatrix} \in C(X^T)$

Yes

d) $4\beta_1 - 8\beta_2$ $= (4 \;\; -8)\beta$, $\begin{pmatrix} 4 \\ -8 \end{pmatrix} \in C(X^T)$

Yes

16

5 got $\frac{16}{16}$

6) Suppose $\boldsymbol{Y}^* = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{r}^W$ ~~where~~ where $E(\boldsymbol{r}^W) = \boldsymbol{0}$ and $Cov(\boldsymbol{r}^W) = Cov(\boldsymbol{Y}^*) = MSE\ \boldsymbol{I}_n$. Then $\hat{\boldsymbol{\beta}}^* = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}^*$. Recall that $\boldsymbol{X}$ is an $n \times p$ constant matrix. Simplify quantities when possible.

a) What is $E(\hat{\boldsymbol{\beta}}^*)$?

$$= (X^TX)^{-1}X^TEY^* = \underbrace{(X^TX)^{-1}X^TX}_{I}\underbrace{\hat{\beta}}_{PY}$$

$$= \boxed{\hat{\beta}}$$

4.9

b) What is $Cov(\hat{\boldsymbol{\beta}}^*)$? $= A\ Cov(Y^*)A^T = (X^TX)^{-1}X^T\ MSE\ I_n\ \underset{n\times p}{X}\ (X^TX)^{-1}$

$$= MSE\ (X^TX)^{-1}X^TX(X^TX)^{-1} = \boxed{MSE\ (X^TX)^{-1}}$$

c) Recall that $\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{PY}$. What is $E(\hat{\boldsymbol{\beta}}_I^*) = E[(\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\boldsymbol{Y}^*]$?

$$= (X_I^TX_I)^{-1}X_I^TX\hat{\beta} = (X_I^TX_I)^{-1}X_I^TPY =$$

$$(X_I^TX_I)^{-1}X_I^TY = \boxed{\hat{\beta}_I}$$

d) What is $Cov(\hat{\boldsymbol{\beta}}_I^*)$? $= (X_I^TX_I)^{-1}X_I^T[Cov(Y^*)]X_I(X_I^TX_I)^{-1}$

$$= (X_I^TX_I)^{-1}X_I^T\ MSE\ \underset{n\times k}{I_n}\ X_I(X_I^TX_I)^{-1}$$

$$= MSE(X_I^TX_I)^{-1}X_I^TX_I(X_I^TX_I)^{-1}$$

$$= \boxed{MSE\ (X_I^TX_I)^{-1}}$$

16

$$Cov(AY^*) = A\ Cov(Y^*)A^T$$

7) Suppose you are estimating the mean μ of losses with $T = \overline{X}$.
actual losses 14, 3, 5, 12, 20, 10, 9: $\overline{X} = 10.4286$,
a) Compute $T_1^*, ..., T_4^*$, where T_i^* is the sample mean of the i bootstrap sample.
bootstrap samples:

12, 3, 10, 14, 5, 9, 10: 63/7 = 9

10, 14, 5, 10, 10, 10, 9: 68/7 = 9.7143

20, 5, 5, 3, 5, 20, 5: 63/7 = 9

12, 20, 5, 14, 12, 14, 20: 97/9 = 13.8571

Q9D21

b) Now compute the bagging estimator which is the sample mean of the T_i^*: the bagging estimator $\overline{T}^* = \frac{1}{B}\sum_{i=1}^{B} T_i^*$ where $B = 4$ is the number of bootstrap samples.

$$\frac{9 + 9.7143 + 9 + 13.8571}{4} = \frac{41.5714}{4} = \boxed{10.3928}$$

$= \frac{291}{28} = 10.4$ c10

10

8) The output below is for forward selection and I_{min} is the minimum C_p model. Here Y = *height*, the constant $x_{i,1} \equiv 1$, $x_{i,2}$ = *height when sitting*, $x_{i,3}$ = *height when kneeling*, $x_{i,4}$ = *head length*, $x_{i,5}$ = *nasal breadth*, and $x_{i,6}$ = *span*.

robstat old Q8

Q8D21

```
          Estimate Std.Err 95% shorth CI
Intercept -42.4846 51.2863 [-192.281, 52.492]
X2               0         [   0.000,  0.268]
X3          1.1707  0.0598 [   0.992,  1.289]
X4               0         [   0.000,  0.840]
X5               0         [   0.000,  1.916]
X6          0.1467  0.0368 [  0.0747,  0.215]
  (Intercept)     a    b     c     d     e
1        TRUE FALSE TRUE FALSE FALSE FALSE
2        TRUE FALSE TRUE FALSE FALSE  TRUE
3        TRUE FALSE TRUE  TRUE FALSE  TRUE
4        TRUE FALSE TRUE  TRUE  TRUE  TRUE
5        TRUE  TRUE TRUE  TRUE  TRUE  TRUE
> tem2$cp
[1] 14.389492  0.792566  2.189839  4.024738  6.000000
```

What is the value of $C_p(I_{min})$ and what is $\hat{\beta}_{I_{min},0}$?

$C_p(I_{min}) = 0.792566$

$\hat{\beta}_{I_{min},0} = (-42.4846, 0, 1.1707, 0, 0, 0.1467)^T$

5