Midterm I—Statistics 232A
29 October 2013

NAME: SOLUTION SKETCHES **SID:** ____________

Instructions: Solve the problems in the blank space below each. **Justify your answers** by brief argument. **Closed Book**. You may consult two, double-sided, letter-size sheets of personal notes.

1. Let A be a nonzero $m \times n$ matrix of rank $r \geq 1$. Let $P = AA^+$.

   a) (4 points) Express the singular value decomposition of P in terms of the singular value decomposition of A. Find numerically the singular values of P.

A has svd $\underset{m\times r}{U}\,\underset{r\times r}{L}\,\underset{r\times n}{V'}$ $\quad U'U = V'V = I_r$, $L = \operatorname{diag}\{l_i \quad 1 \le i \le r\}$

$l_1 \geq l_2 \geq \dots \geq l_r > 0$

$P = AA^+ = ULV' \cdot VL^{-1}U' = UU' = \underset{m\times r}{U}\,I_r\,\underset{r\times m}{U'}$ = svd of $\underset{m\times m}{P}$

The r singular values of P each equal 1

   b) (2 points) Find numerically the eigenvalues of P.

Let $\underset{m\times m}{O} = (U \; \bar{U})$ where $\underset{m\times(m-r)}{\bar{U}}$ is chosen so that O is an orthogonal matrix: $O'O = OO' = I_m$

Then $P = (U \; \bar{U})\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} U' \\ \bar{U}' \end{pmatrix} = O\Lambda O'$ where $\Lambda = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

This is a spectral representation of P

The eigenvalues of P are 1 (r times) and 0 ($(m-r)$ times)

   c) (2 points). Find the trace of P and the rank of P.

From the svd of P in part a:

- $\operatorname{tr}(P) = \operatorname{tr}(UU') = \operatorname{tr}(U'U) = \operatorname{tr}(I_r) = r$
- $\operatorname{rank}(P) = \operatorname{rank}(I_r) = r$

d) (4 points) Find P^+ as a function of A.

From the svd of P,

$P^+ = U I_r^{-1} U' = UU' = P = AA^+$

e) (4 points) Suppose that $A = a$, a nonzero column vector with m components. Find a^+ as a function of a.

Since $a \neq 0$, $\text{rank}(a) = 1$

By Lab #1, $a^+ = (a'a)^{-1} a' = \dfrac{a'}{|a|^2}$

f) (4 points) The **null space** $\mathcal{N}(A) = \{x \in R^n \colon Ax = 0\}$. Show that $\mathcal{N}(A) = \{(A^+A - I_n)c \colon c \in R^n\}$.

The equation $Ax = 0$ is consistent because $x = 0$ is a solution

Hence $\mathcal{N}(A)$ = solution set of $Ax = 0$

$= \{A^+ 0 + (A^+A - I_n)c : c \in R^n\}$

$= \{(A^+A - I_n)c : c \in R^n\}$

using class results on linear equations

SID: ____________

2. Consider the Gaussian linear model in which $y = (y_1, y_2, y_3, y_4)'$ has a $N(\eta, \sigma^2 I_4)$ distribution and $\eta = (\eta_1, \eta_2, \eta_3, \eta_4)'$ has the structure

$$\eta_1 = m + a_1, \qquad \eta_2 = m + a_1$$
$$\eta_3 = m + a_2, \qquad \eta_4 = m + a_2.$$

In this model, m, a_1, a_2 and $\sigma^2 > 0$ are unknown scalar parameters.

a) (4 points) Let $\beta = (m, a_1, a_2)'$. Write out the matrix X such that $\eta = X\beta$. Find and state the rank of X.

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Col 2 $\perp$ Col 3 $\rightarrow$ rank$(X) \geq 2$

Col 1 = Col 2 + Col 3 $\rightarrow$ rank$(X) \leq 2$

Hence rank$(X) = 2$

b) (4 points) Show that m is **not** linearly estimable.

$m = \lambda'\beta$ with $\lambda' = (1,0,0)$

m is linearly estimable iff $\exists \underset{4\times1}{a}$ such that $a'X = \lambda'$

i.e. iff
$$\left.\begin{aligned} a_1 + a_2 + a_3 + a_4 &= 1 \\ a_1 + a_2 &= 0 \\ a_3 + a_4 &= 0 \end{aligned}\right\}$$
No such a exists because equations 2 and 3 contradict equation 1

Hence m is <u>not</u> linearly estimable

c) (4 points) Show that $\psi = a_1 - a_2$ is linearly estimable.

$E(y_1 - y_3) = (m + a_1) - (m + a_2) = a_1 - a_2$

Hence ψ has a linear unbiased estimator

d) (4 points) Show that $\hat{\beta} = (1/2)(0, y_1 + y_2, y_3 + y_4)'$ is a least squares estimator of β. Are there other least squares estimators of β in this model? [Hint: Consider the normal equation].

Normal equation $X'X\beta = X'y$ is here

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 + y_3 + y_4 \\ y_1 + y_2 \\ y_3 + y_4 \end{pmatrix}$$

$\hat{\beta} = \frac{1}{2}(0, y_1 + y_2, y_3 + y_4)'$ is one solution,

<u>not unique</u> because $\text{rank}(X) = 2 < 3$

e) (4 points) Show that $\hat{\eta} = (1/2)(y_1 + y_2, y_1 + y_2, y_3 + y_4, y_3 + y_4)'$ is the **unique** least squares estimator of η.

$\hat{\eta} = X\hat{\beta} = \frac{1}{2}(y_1 + y_2, y_1 + y_2, y_3 + y_4, y_3 + y_4)'$

is the <u>unique</u> LSE of η, using class theory