

Math 763 Midterm Exam, Fall Name: _____

October 25, 2011
Instructor: Dhar

Student #: _____

Must show all work for full credit!!!

I pledge that I have not violated the NJIT code of honor _____
All questions 1) through 5) are weighted equally.

- 1) Consider the linear model described by

$$\mathbf{Y} = \mathbf{X}\beta + \boldsymbol{\varepsilon},$$

where \mathbf{X} is the $n \times (p+1)$ design matrix of full rank $p+1$ and β is the $(p+1) \times 1$ vector of parameters, with $E(\boldsymbol{\varepsilon}) = \mathbf{0}$, $Cov(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$, where \mathbf{I} is the full rank n identity matrix.

- a. Derive the least squares estimator of β .
- b. Show that the least squares estimator of β is an unbiased estimator of β .
- c. Derive the covariance matrix of the least squares estimator of β .

$$(a) Q(\hat{\beta}) = (\underline{Y} - \underline{X}\hat{\beta})^T (\underline{Y} - \underline{X}\hat{\beta})$$

$$\begin{aligned} \frac{\partial Q(\hat{\beta})}{\partial \hat{\beta}} &= \frac{\partial}{\partial \hat{\beta}} [\underline{Y}^T \underline{Y} - \underline{Y}^T \underline{X} \hat{\beta} - \hat{\beta}^T \underline{X}^T \underline{Y} + \hat{\beta}^T \underline{X}^T \underline{X} \hat{\beta}] \\ &= -\underline{X}^T \underline{Y} - \frac{\partial}{\partial \hat{\beta}} (\underline{Y}^T \underline{X} \hat{\beta}) + \frac{\partial}{\partial \hat{\beta}} (\hat{\beta}^T \underline{X}^T \underline{X} \hat{\beta}), -(i) \end{aligned}$$

Note that $\frac{\partial}{\partial \underline{y}} \underline{y}^T A \underline{y} = A \underline{y} + A^T \underline{y} = 2A \underline{y}$. Asymmetric

(i) reduces to $-2\underline{X}^T \underline{Y} + 2\underline{X}^T \underline{X} \hat{\beta}$ which is set equal to zero. Hence $\underline{X}^T \underline{X} \hat{\beta} = \underline{X}^T \underline{Y}$ $\Rightarrow \hat{\beta} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}$, because \underline{X} is full rank.

$$(b) E \hat{\beta} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T E \underline{Y}. E \underline{Y} = \underline{X} \beta \text{ because } E \underline{\varepsilon} = \mathbf{0}.$$

$$E \hat{\beta} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \beta = \beta, \text{ because } \underline{X} \text{ is full rank.}$$

$$(c) Cov(\hat{\beta}) = (\underline{X}^T \underline{X})^{-1} \underline{X}^T Cov(\underline{Y}) \underline{X} (\underline{X}^T \underline{X})^{-1}. \text{ Since } Cov(\underline{\varepsilon}) = \sigma^2 \mathbf{I}, Cov(\underline{Y}) = \sigma^2 \mathbf{I}, \text{ Thus } Cov(\hat{\beta}) = \sigma^2 (\underline{X}^T \underline{X})^{-1} (\underline{X}^T \underline{X})^{-1}.$$

2) Consider the general linear model with vector of observations $\mathbf{Y}' = (Y_1, \dots, Y_n)$,

$$E(Y_i) = z_{i0}\beta_0 + z_{i1}\beta_1 + \dots + z_{ip-1}\beta_{p-1}, \quad i=1, \dots, n, \quad \Rightarrow \quad E(\mathbf{Y}) = \mathbf{Z}\beta = \mathbf{X}\beta$$

satisfying the assumptions of normality, independence, and homoscedasticity, with design matrix \mathbf{Z} of dimension $n \times p$ and the parameter vector $\beta' = (\beta_0, \dots, \beta_{p-1})$. Let β_{j_0} be one of the fixed parameter from the β' vector, $0 \leq j_0 \leq p-1$. Derive the test for the hypotheses:

$$\begin{aligned} H_0: \beta_{j_0} &= 0, \\ H_1: \beta_{j_0} &\neq 0. \end{aligned} \quad (\text{page 29, equation (2.31)})$$

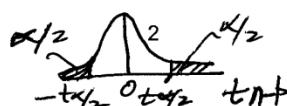
Let $\xi_{j_0} \sigma^2$ be the j_0^{th} diagonal element of $\sigma^2(\mathbf{X}^T \mathbf{X})^{-1} = \sigma^2(\mathbf{Z}^T \mathbf{Z})^{-1}$. Then,

we know $Z = \frac{\hat{\beta}_{j_0} - \beta_{j_0}}{\sigma \sqrt{\xi_{j_0} \sigma^2}}$, where Z is standard normal.

$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ is independent of $\hat{\beta}_{j_0}$ and $\frac{SSE}{\sigma^2} \sim \chi^2_{n-p}$ under the normality assumption.

Thus $T = \frac{\hat{\beta}_{j_0} - \beta_{j_0}}{\sigma \sqrt{\xi_{j_0} \sigma^2}}$ has 't' with $n-p$ degrees of freedom.

$T = \frac{\hat{\beta}_{j_0} - 0}{\sqrt{\xi_{j_0} \sigma^2} \sqrt{\frac{SSE}{n-p}}}$. Hence, Reject H_0 at level α if $|T| > t_{\alpha/2}$.



3) Consider the regression model with k regressor variables described by

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where \mathbf{X} is the $n \times p$ design matrix of full rank p and $\boldsymbol{\beta}$ is the $p \times 1$, $p = k+1$ vector of parameters, with $\boldsymbol{\epsilon}$ multivariate normal with $E(\boldsymbol{\epsilon}) = \mathbf{0}$, $\text{Cov}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}$, where \mathbf{I} is the full rank n identity matrix. Here, $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}$, $\boldsymbol{\beta}_1$ is the $r \times 1$ coefficient of r regressor

variables and $\boldsymbol{\beta}_2$ is the $(p-r) \times 1$. One would like to test $H_0: \boldsymbol{\beta}_1 = \mathbf{0}$ versus $H_1: \boldsymbol{\beta}_1 \neq \mathbf{0}$. Define $\text{SS}_R(\boldsymbol{\beta}_1 | \boldsymbol{\beta}_2)$ and MS_{res} . Use $F_0 = \{\text{SS}_R(\boldsymbol{\beta}_1 | \boldsymbol{\beta}_2)/r\} / \text{MS}_{\text{res}}$ to arrive at the

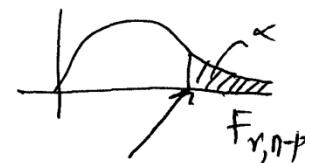
(a) test of the hypotheses. Using $S_{(i)}^2 = \frac{(n-p)\text{MS}_{\text{res}} - e_i^2 / (1-h_{ii})}{n-p-1}$ define R-student and show how one would use it to do an outlier diagnostic.

(a) $\text{SS}_R(\boldsymbol{\beta}_1 | \boldsymbol{\beta}_2)$ = sum of squares due to regression with $\boldsymbol{\beta}$ minus sum of squares due to same type of regression (under $H_0: \boldsymbol{\beta}_1 = \mathbf{0}$) with $\boldsymbol{\beta}_2$ = extra sum of squares.

$$(b) \quad \text{MS}_{\text{res}} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p},$$

$$(c) \quad H_0: \boldsymbol{\beta}_1 = \mathbf{0} \quad \text{Reject } H_0 \text{ if } F_0 > F_{r, n-p, \alpha} \text{ at } \alpha\text{-level}$$

of significance



(d) R-student is $t_i = \frac{e_i}{\sqrt{S_{(i)}^2(1-h_{ii})}}$, where e_i is the residual $y_i - \hat{y}_i$, h_{ii} is the diagonal element of $\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$. Then i^{th} data value is a potential outlier if $|t_i| > t_{\frac{\alpha}{2n}}$.

- 4) The following model has the predicted variable y as specific gravity and covariates as a result of spectrophotometer analysis as x_1 : NG (Nitroglycerine), x_2 : TA (triacetin) and x_3 : NDPA (2-nitrodiphenylamine). There is a need to estimate activity coefficients from the model

$$y = \frac{1}{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3} + \varepsilon.$$

The quantity parameters β_1, β_2 , and β_3 are ratios of activity coefficients to the individual specific gravity of the NG, TA, and 2 NDPA, respectively.

- a. Determine whether the above model is a linear model, an intrinsically linear model, or nonlinear model. Explain.
- b. Determine the starting values for the model parameters.
- c. Name the method (or its brief description) that is widely used to estimate the parameters of the above model.

(a) $EY = \text{Expected function} = \frac{1}{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3}$

$= f(\bar{x}, \beta)$ and $\frac{\partial f(\bar{x}, \beta)}{\partial \beta_i}$ depends on

β_i 's hence this model is nonlinear model.

(b) Consider the linear model $E(Y) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$.

Then use the least squares estimator

$$\hat{\beta} = (X^T X)^{-1} X^T Y^* \text{ to estimate } \beta = (\beta_1, \beta_2, \beta_3)$$

where $Y^* = \begin{pmatrix} 1/y_1 \\ \vdots \\ 1/y_n \end{pmatrix}$ and $X = \begin{pmatrix} x_{11} & x_{21} & x_{31} \\ \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & x_{3n} \end{pmatrix}$.

- (c) Since the above model is nonlinear we use the Gauss - Newton method to estimate the parameters.

- 5) Consider the non linear regression model $y_i = f(\mathbf{x}_i, \beta) + \varepsilon_i$, $i=1, 2, \dots, n$, with the errors are independent identically distributed as normal mean zero variance σ^2 . Here β is $p \times 1$ vector.
- Obtain:
- The likelihood function.
 - The score function for $\hat{\beta}$.
 - Use the appropriate form of the score function to obtain the information matrix and thus the $\text{Cov}(\hat{\beta})$, where $\hat{\beta}$ is the generalized least squares estimator of β .

$$(a) L(\beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \beta)]^2 \right\}$$

$$(b) \ln L(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \beta)]^2$$

$$\frac{\partial \ln L(\beta, \sigma^2)}{\partial \beta_j} = +\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \beta)] \frac{\partial f(\mathbf{x}_i, \beta)}{\partial \beta_j} = 0$$

$$D = \left[\frac{\partial f(\mathbf{x}_i, \beta)}{\partial \beta_j} \right]_{i=1 \dots n, j=1 \dots p}$$

$$[\underline{y} - \underline{\mu}]^T \frac{D}{\sigma^2} = 0 \quad \text{or} \quad \frac{D^T [\underline{y} - \underline{\mu}]}{\sigma^2} = 0$$

(c) Information matrix is given by

$$\frac{D^T \text{Cov}(\underline{y}) D}{\sigma^4} = \frac{D^T D}{\sigma^2} \text{ because}$$

$$\text{Cov}(\underline{y}) = \sigma^2 I \quad \text{. Thus} \quad \text{Cov}(\hat{\beta}) = \sigma^2 (D^T D)^{-1}.$$