University of California, Los Angeles
Department of Statistics

Statistics 100C **Instructor: Nicolas Christou**

Homework 4

Exercise 1
Consider the following simple regression model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, for which $E(\epsilon_i) = 0$, $E(\epsilon_i \epsilon_j) = 0$ for $i \neq j$, and $\text{var}(\epsilon_i) = \sigma^2$. The normal equations discussed earlier in class are:

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i \\ \hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i y_i \end{aligned}$$

In matrix form this system of two equations with two unknowns can be expressed as follows:

$$\begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}$$

a. Use matrix algebra to find the solution for the vector $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)'$.

b. Use matrix algebra to find the variance covariance matrix of the vector $\hat{\boldsymbol{\beta}}$, i.e.

$$\begin{pmatrix} var(\hat{\beta}_0) & cov(\hat{\beta}_0, \beta_1) \\ cov(\hat{\beta}_1, \beta_1) & var(\hat{\beta}_1) \end{pmatrix}.$$

Exercise 2
Consider the following simple regression model for which $\epsilon_i \sim N(0, \sigma)$.

$$\begin{aligned} y_1 &= \beta_0 + 0.5\beta_1 + \epsilon_1 \\ y_2 &= \beta_0 - \beta_1 + \epsilon_2 \\ y_3 &= \beta_0 + 0.5\beta_1 + \epsilon_3 \end{aligned}$$

a. Write the above model in matrix form.

b. Find the least squares estimates using vectors and matrices.

c. Find the variance-covariance matrix of $\hat{\beta}$.

d. Find the hat matrix. Verify that the sum of the diagonal elements of the hat matrix is equal to 2 ($\sum_{i=1}^n h_{ii} = k + 1$).

e. Generate your own data with $n = 3$ based on this model and verify that the estimates of β_0 and β_1 are those given by part (b).

Exercise 3
Suppose that you need to fit the multiple regression model $y_i = \beta_0 + \beta_1 x1_i + \beta_2 x2_i + \epsilon_i$, where $E(\epsilon_i) = 0$, $E(\epsilon_i \epsilon_j) = 0$ for $i \neq j$, and $\text{var}(\epsilon_i) = \sigma^2$, to the following data:

y	x_1	x_2
-43.6	27	34
3.3	33	30
-12.4	27	33
7.6	24	11
11.4	31	16
5.9	40	30
-4.5	15	17
22.7	26	12
-14.4	22	21
-28.3	23	27

It turns out that

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{pmatrix} 1.97015 & -0.05623 & -0.01572 \\ -0.05623 & 0.00289 & -0.00091 \\ -0.01572 & -0.00091 & 0.00174 \end{pmatrix} \text{ and } \boldsymbol{X}'\boldsymbol{Y} = \begin{pmatrix} -52.3 \\ -1076.3 \\ -2220.2 \end{pmatrix}$$

a. Find the least squares estimator of $\beta = (\beta_0, \beta_1, \beta_2)'$.

b. Find the variance-covariance matrix of the previous estimator.

c. Compute the estimate s_e^2 of σ^2.

d. Using your answers to parts (b) and (c) find the variances of $\hat{\beta}_0, \hat{\beta}_1$, and $\hat{\beta}_2$.

e. Find the fitted value $\hat{y}_1$ a nd its variance.

f. What is the variance of the first residual ($\text{var}(e_i)$)?

Exercise 4
Show that the residuals are orthogonal to the matrix $\mathbf{X}$ as well as to the fitted values $\hat{\mathbf{Y}}$. This is true for simple or multiple regression models.

a. $\mathbf{e}'\mathbf{X} = \mathbf{0}$.

b. $\mathbf{e}'\hat{\mathbf{Y}} = \mathbf{0}$.

c. Use part (a) to show the already known result that $\sum_{i=1}^{n} e_i = 0$.