MATH 204 - EXERCISES 3

These exercises are not for assessment

Note: Also try exercises from Chapter 11, especially 11.96 - 11.119, from McClave & Sincich.

1. The form of the least-squares estimates of β_0 and β_1 for the regression model are obtained by minimizing $SSE(\beta_0, \beta_1)$ for data $\{(x_i, y_i), i = 1, \ldots, n\}$, where

$$SSE(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

To do this we take (partial) derivatives with respect to β_0 and β_1, and solve the equations that result when the derivatives are set equal to zero. That is, the two equations are

$$\begin{aligned} \text{For } \beta_0 &: \sum_{i=1}^{n} -(y_i - \beta_0 - \beta_1 x_i) = 0 \\ \text{For } \beta_1 &: \sum_{i=1}^{n} -(y_i - \beta_0 - \beta_1 x_i) x_i = 0 \end{aligned}$$

(a) Manipulate these equations to get two equations in the parameters that can be solved simultaneously, and hence derive the forms for estimates $(\widehat{\beta}_0, \widehat{\beta}_1)$ given in lectures.

(b) If the residuals $\{\hat{e}_i, i = 1, \ldots, n\}$ derived from a fit of the model with parameter estimates $(\widehat{\beta}_0, \widehat{\beta}_1)$ are defined by $\hat{e}_i = y_i - \hat{y}_i = y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i$, show that

$$\sum_{i=1}^{n} \hat{e}_i = 0.$$

2. The following data correspond to the relationship between Population and Gross National Product (GNP) in the US between 1947 and 1962

Year	Population (mil) x	GNP (bil \$) y
1947	107.608	234.289
1948	108.632	259.426
1949	109.773	258.054
1950	110.929	284.599
1951	112.075	328.975
1952	113.270	346.999
1953	115.094	365.385
1954	116.219	363.112
1955	117.388	397.469
1956	118.734	419.180
1957	120.445	442.769
1958	121.950	444.546
1959	123.366	482.704
1960	125.368	502.601
1961	127.852	518.173
1962	130.081	554.894

Fit a simple linear regression using least-squares to these data (available as LONGLEY.SAV from the course website).

Reference: J. W. Longley (1967) An appraisal of least-squares programs from the point of view of the user. *Journal of the American Statistical Association*, **62**, 819-841.