1. Consider the linear regression with data

\[ y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \cdots, n, \]

with the assumptions that \( E(\epsilon) = 0 \) and \( \text{var}(\epsilon) = \sigma^2 I \), where \( \epsilon = (\epsilon_1, \cdots, \epsilon_n)^T \).

(a) Write down the model in matrix form, defining the design matrix \( X \), the response vector \( y \), and the parameter \( \beta \).

(b) Calculate \( X^T X \), and \( (X^T X)^{-1} \). Hence using \( \hat{\beta} = (X^T X)^{-1} X^T y \), find \( \hat{\beta}_1 \).

(c) Now consider the previous model, but with the covariate \( x \) centered:

\[ y_i = \beta_0 + \beta_1 (x_i - \bar{x}) + \epsilon_i, \quad i = 1, \cdots, n. \]

Redo (a) and (b). Find the variance of \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) in this model, and the covariance between them.

2. Consider the general multiple linear regression model \( y = X\beta + \epsilon \), with normality assumption on \( \epsilon \), i.e., we assume \( \epsilon \sim N(0, \sigma^2 I) \).

(a) Write down the likelihood for the data. Hence show that the log-likelihood \( \ell(\beta, \sigma^2) \) is as given in lecture 1.

(b) Show that

\[ \frac{\partial \ell(\beta, \sigma^2)}{\partial \beta} = \frac{X^T y}{\sigma^2} - \frac{X^T X \beta}{\sigma^2}, \]

by first expanding \( \|y - X\beta\|^2 \) in linear and quadratic terms in \( \beta \).

(c) With \( \sigma^2 \) assumed known, the derivative in (b) is called the score vector for the parameter \( \beta \), denoted by \( U(\beta) \).

Write \( U(\beta) = (U_1(\beta), \cdots, U_p(\beta))^T \), the information matrix for \( \beta \) is defined as

\[ I(\beta) = E\left( -\frac{\partial U(\beta)}{\partial \beta^T} \right), \]

with

\[ \frac{\partial U(\beta)}{\partial \beta^T} = \begin{pmatrix} \left( \frac{\partial U_1(\beta)}{\partial \beta} \right)^T \\ \vdots \\ \left( \frac{\partial U_p(\beta)}{\partial \beta} \right)^T \end{pmatrix}. \]

Find \( I(\beta) \) for our multiple regression model given \( \sigma^2 \).

3. Consider the multiple linear regression model in question 2.

(a) Show that

\[ \text{Total corrected SS} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - n \bar{y}^2. \]
Hence, by writing $\bar{y} = n^{-1}1_n^T y$ where $1_n$ is a vector of ones of length $n$, show that

$$\text{Total corrected SS} = y^T y - n(n^{-1}1_n^T y)^2 = y^T (I - n^{-1}1_n1_n^T) y.$$  

(b) Show that $HX = X$, where $H$ is the hat matrix. Assuming our model has a constant term, so that $X$ has a column of ones (i.e., $1_n$ is a column in $X$), show in particular that $H1_n = 1_n$.

(c) Using part (b), show that

$$\sum \hat{y}_i = n\bar{y}.$$  

(Hint: Write $\sum \hat{y}_i = 1_n^T \hat{y}$.)

(d) Hence, show that

$$\text{SS(reg)} = \sum (\hat{y}_i - \bar{y})^2 = \sum \hat{y}_i^2 - n\bar{y}^2.$$  

(e) By writing $\sum \hat{y}_i^2 = ||\hat{y}||^2$, using (d), show that

$$\text{SS(reg)} = y^T (H - n^{-1}1_n1_n^T)y.$$  

(f) Show that $\text{RSS} = ||y - \hat{y}||^2 = y^T (I - H)y$. Hence show that

$$\text{Total corrected SS} = \text{SS(reg)} + \text{RSS}.$$  

4. You have shown in 3(f) that

$$\text{RSS} = y^T (I - H)y.$$  

(a) Show that $(I - H)X = 0$.

(b) By writing $y = X\beta + \epsilon$, show that

$$\text{RSS} = \epsilon^T (I - H)\epsilon.$$  

(c) Hence show that

$$E(\text{RSS}) = \sigma^2 (n - p),$$  

so that $S^2 = \text{RSS}/(n-p)$ is unbiased for $\sigma^2$. (Hint: Observe that $\epsilon^T (I - H)\epsilon$ is a scalar, so that $\epsilon^T (I - H)\epsilon = \text{tr}(\epsilon^T (I - H)\epsilon) = \text{tr}(\epsilon\epsilon^T (I - H))$. Now take expectation on both sides and swap the order of trace and expectation on the right hand side. What is $E(\epsilon\epsilon^T)$?)

5. Consider the linear regression model

$$y = X\beta + \epsilon,$$

where $\epsilon \sim N(0, \sigma^2 I)$, $y$ is the response vector of length $n$, $\beta$ is the vector of parameters of length $p$.

(a) Show that $X(\hat{\beta} - \beta) = H\epsilon$, and hence

$$||X(\hat{\beta} - \beta)||^2 = \epsilon^T H\epsilon.$$
(b) Show that an idempotent matrix \( W \) has only 0 or 1 as its eigenvalues. (Hint: If \( \lambda \) is an eigenvalue for \( W \), then \( \lambda^2 \) is an eigenvalue for \( W^2 \).)

(c) For a real symmetric matrix \( W \), we can always decompose \( W \) as
\[
W = QDQ^T,
\]
where \( Q \) is an orthogonal matrix (i.e., \( Q^TQ = QQ^T = I \)), and \( D \) is a diagonal matrix containing all the eigenvalues of \( W \).

Apply this to \( H \), show that
\[
\epsilon^T He = z^T Dz,
\]
where \( D \) is a diagonal matrix containing all the eigenvalues of \( H \), and \( z = Q\epsilon \) with \( Q \) an orthogonal matrix.

What is the distribution of \( z \)?

(d) Find \( \text{tr}(H) \), the trace of \( H \) (Hint: Write \( H \) in terms of \( X \) first, then use \( \text{tr}(AB) = \text{tr}(BA) \)). Hence using (b), find the exact number of eigenvalues for \( H \) which are equal to 1.

(e) Using (c) and (d), show that
\[
\|X(\hat{\beta} - \beta)\|^2 = \frac{\epsilon^T H \epsilon}{\sigma^2} \sim \chi^2_p.
\]

(f) From question 3(e), we have the sum of squares due to regression
\[
SS(\text{reg}) = y^T (H - n^{-1}1_n1_n^T)y.
\]

Under the null hypothesis \( H_0 : \beta_1 = \cdots = \beta_k = 0 \), then \( y = \beta_0 1_n + \epsilon \).
Show that then
\[
SS(\text{reg}) = \epsilon^T (H - n^{-1}1_n1_n^T) \epsilon.
\]
(hint: \( H1_n = 1_n \) if \( 1_n \) is a column in \( X \))

Show that the matrix \( H - n^{-1}1_n1_n^T \) is also symmetric idempotent, and hence using (c), (d) and (e), replacing \( H \) by \( H - n^{-1}1_n1_n^T \), show that under the null hypothesis \( H_0 : \beta_1 = \cdots = \beta_k = 0 \),
\[
\frac{SS(\text{reg})}{\sigma^2} \sim \chi^2_{p-1}.
\]

6. Using similar technique as in question 5, show that
\[
\frac{RSS}{\sigma^2} = \frac{\epsilon^T (I-H) \epsilon}{\sigma^2} \sim \chi^2_{n-p}.
\]

7. With same model as in question 1, Consider the two random vectors \( w_1 = He \) and \( w_2 = (I-H)e \).

(a) Find the mean and covariance matrix for each of them.
(b) Consider the joint random vector
\[
w = \left( \begin{array}{c} w_1 \\ w_2 \end{array} \right).
\]
Show that \( \mathbf{w} \) has a normal distribution, and find its mean and covariance matrix. (Hint: Write \( \mathbf{w} \) as \( \mathbf{A}\mathbf{\epsilon} \), and find \( \mathbf{A} \))

(c) Hence, using part (b), show that \( \mathbf{w}_1 \) is independent of \( \mathbf{w}_2 \), i.e., any elements in \( \mathbf{w}_1 \) is independent of all the elements in \( \mathbf{w}_2 \) and vice versa.

(d) Using (c), show that \( \|\mathbf{X}(\hat{\mathbf{\beta}} - \mathbf{\beta})\|^2 \) and RSS from questions 5 and 6 are independent. (You can use the same technique to prove that \( \hat{\mathbf{\beta}} \) and \( S^2 \) are independent, as well as the independence of SS(reg) and RSS under the null hypothesis \( H_0 : \mathbf{\beta} = 0 \).)