1. The figure below shows a plot of the quantity of beer (in millions of barrels) produced in the U.S. each quarter (i.e., every three months), from the first quarter of 1975 to the fourth quarter of 1982.

Given the periodic behavior apparent in the data, a so-called trigonometric regression model of the form

\[ y_t = \beta_0 + \beta_1 \cos(2\pi t/4) + \beta_2 \sin(2\pi t/4) + \beta_3 \cos(\pi t) + e_t \]  

may be appropriate, where \( t = 1, \ldots, 32 \), the vector \((\beta_0, \beta_1, \beta_2, \beta_3)\) contains the unknown coefficients, ‘cos’ and ‘sin’ refer to the cosine and sine functions, respectively, and the \( e_t \) are taken to be i.i.d. errors with mean zero and constant variance \( \sigma^2 > 0 \).

a. Since the only unknowns in our model are the coefficients in \( \beta = (\beta_0, \beta_1, \beta_2, \beta_3) \), and the model is linear in these coefficients, we can rewrite the model as a linear regression model i.e., in the form \( Y = X\beta + \epsilon \). Show that the model
matrix $\mathbf{X}$ is given by

$$
\mathbf{X} = \begin{bmatrix}
1 & 0 & 1 & -1 \\
1 & -1 & 0 & 1 \\
1 & 0 & -1 & -1 \\
1 & 1 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 0 & 1 & -1 \\
1 & -1 & 0 & 1 \\
1 & 0 & -1 & -1 \\
1 & 1 & 0 & 1 \\
\end{bmatrix},
$$

(2)

where the pattern in the upper $4 \times 4$ block is replicated 8 times.

b. Argue that an ordinary least squares regression for $\beta$ under this model is in fact an orthogonal regression. That is, argue that the columns of $\mathbf{X}$ are mutually orthogonal. What if we had measurements only for, say, the second two quarters of 1975 through the first two quarters of 1982. Would we still have an orthogonal regression? Explain.

c. The coefficients $(\beta_1, \beta_2)$ correspond to the strength of the periodic quarterly component, and the coefficient $\beta_3$, to the periodic semi-annual component. Give a formal statement of the null hypothesis that there is no quarterly component in the data. Show that an appropriate $F$ statistic to test this hypothesis takes the form

$$
F = \frac{16 \left( \hat{\beta}_1^2 + \hat{\beta}_2^2 \right) / 2}{\hat{\sigma}^2},
$$

(3)

where $\hat{\sigma}^2$ is the usual estimate of $\sigma^2$. For the data above, what are the corresponding degrees of freedom of this $F$ statistic?

d. Note that the data, as shown in the plot on the previous page, have an increasing trend to them, as a function of time. Do you expect the model in (1) to be appropriate for this data? Explain why or why not. If not, propose and justify an alternative model that you suspect would be better. Does your model still represent an orthogonal regression? If not, what are the implications on the fitting of the coefficients? On the interpretation of the coefficients?
2. Consider the usual multiple linear regression model i.e.,
\[ y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + e_i, \]
for \( i = 1, \ldots, n \), where the \( e_i \) are assumed i.i.d., with \( E[e_i|X] = 0 \) and \( \text{Var}(e_i|X) = \sigma^2 \). An algorithm for obtaining the ordinary least squares estimate \( \hat{\beta}_p \) of \( \beta_p \) is as follows.

Step 1. Initialize \( z_0 = x_0 = 1 \).
Step 2. For \( j = 1, 2, \ldots, p \),
Regress \( x_j \) on \( z_0, \ldots, z_{j-1} \), and call the residual vector \( z_j \).
Step 3. Regress \( Y \) on \( z_p \), and set \( \hat{\beta}_p \) equal to the resulting slope estimate.

This procedure is known as ‘regression by successive orthogonalization’, and is reminiscent of the Gram-Schmidt procedure in linear algebra.

a. For each iteration \( j \) in Step 2, argue that the vectors \( z_0, \ldots, z_{j-1} \) are orthogonal.

b. Show that the result of the algorithm is the estimate
\[ \hat{\beta}_p = \frac{z_p^T Y}{z_p^T z_p}, \] 
where \( Y = (y_1, \ldots, y_n)^T \). Furthermore, show that this estimate is indeed equal to the OLS estimate of \( \beta_p \).

(Hint: One approach to arguing the latter half of this question is to (i) argue that \( \{z_0, \ldots, z_p\} \) form a basis for the column space of \( X \), (ii) note that by definition \( z_j = x_j - \sum_{k=0}^{j-1} \hat{\gamma}_{kj} z_k \), for regression coefficients \( \hat{\gamma}_{kj} \), and (iii) exploit the fact that among the vectors \( x_j \), only \( x_p \) involves \( z_p \).)

c. Show that
\[ \text{Var}(\hat{\beta}_p) = \frac{\sigma^2}{z_p^T z_p}. \]
d. Alternatively, it is known that the variance of the OLS estimate of $\beta_p$ can be written in the form
\[
\frac{\sigma^2}{1 - c} \cdot \frac{1}{SX_pX_p},
\]
for some $c$, where $SX_pX_p = \sum_{i=1}^{n} (x_{pi} - \bar{x}_p)^2$.

Equating (5) and (6), solve for $c$ and, importantly, interpret your answer.