Problem 1
Suppose Z is an n-dimensional random vector with expectation $E(Z)$ and variance-covariance matrix:

$$Var(Z) = Cov(Z, Z) = \Sigma$$

Let A be an $s \times n$ nonrandom matrix and B a $t \times n$ nonrandom matrix. Show the following:

(a) $E(AZ) = AE(Z)$

(b) $Cov(AZ, BZ) = A\Sigma B^T$. In particular, $Var(AZ) = A\Sigma A^T$.

Solution:

(a) Observe, for $E[AZ] = \{c_{i,j}\}_{(i,j)\in s\times n}$, then we see that

$$c_{i,j} = E\left[\sum_{k=1}^{n} a_{i,k} z_{k,j}\right] = \sum_{k=1}^{n} E\left[a_{i,k} z_{k,j}\right] = \sum_{k=1}^{n} a_{i,k} E\left[z_{k,j}\right]$$

but this is equivalent to saying

$$\{c_{i,j}\}_{(i,j)\in s\times n} = AE[Z]$$

Therefore, $E[AZ] = AE[Z]$.

(b) Recall,

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])^T]$$

Observe,

$$\begin{aligned}
Cov(AZ, BZ) &= E[(AZ - E[AZ])(BZ - E[BZ])^T] \\
&= E[(AZ - AE[Z])(BZ - BE[Z])^T] && \text{by Part (a)} \\
&= E[A(Z - E[Z])(Z - E[Z])^T B^T] \\
&= AE[(Z - E[Z])(Z - E[Z])^T B^T] \\
&= AE[(Z - E[Z])(Z - E[Z])^T]B^T \\
&= ACov(Z, Z)B^T \\
&= A\Sigma B^T
\end{aligned}$$

In particular,

$$Var(AZ) = cov(AZ, AZ) = ACov(Z, Z)A^T = A\Sigma A^T$$

Problem 2
Derive the following:

(a) $\sum_{i=1}^{n}(X_i - \overline{X}) = 0, \sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n}(X_i - \overline{X})X_i = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$

(b) $\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^{n}(X_i - \overline{X})Y_i = \sum_{i=1}^{n} X_iY_i - n\overline{X}\ \overline{Y}$

Solution:

(a) Since $\overline{X} = \frac{1}{n}\left(\sum_{i=1}^n X_i\right)$

$$\sum_{i=1}^n (X_i - \overline{X}) = \left(\sum_{i=1}^n X_i\right) - n\overline{X} = \left(\sum_{i=1}^n X_i\right) - \left(\sum_{i=1}^n X_i\right) = 0$$

and

$$\begin{aligned}
\sum_{i=1}^n (X_i - \overline{X})^2 &= \sum_{i=1}^n X_i^2 - 2X_i\overline{X} + \overline{X}^2 \\
&= \sum_{i=1}^n X_i(X_i - \overline{X}) - \overline{X}(X_i - \overline{X}) \\
&= \sum_{i=1}^n X_i(X_i - \overline{X}) - \overline{X}\underbrace{\sum_{i=1}^n (X_i - \overline{X})}_{=0} \\
&= \sum_{i=1}^n X_i(X_i - \overline{X}) \\
&= \sum_{i=1}^n X_i^2 - \sum_{i=1}^n X_i\overline{X} \\
&= \sum_{i=1}^n X_i^2 - \overline{X}\sum_{i=1}^n X_i \\
&= \sum_{i=1}^n X_i^2 - n\overline{X}^2
\end{aligned}$$

(b)

$$\begin{aligned}
\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) &= \sum_{i=1}^n (X_i - \overline{X})Y_i - \overline{Y}(X_i - \overline{X}) \\
&= \sum_{i=1}^n (X_i - \overline{X})Y_i - \overline{Y}\underbrace{\sum_{i=1}^n (X_i - \overline{X})}_{=0} \\
&= \sum_{i=1}^n (X_i - \overline{X})Y_i \\
&= \sum_{i=1}^n X_iY_i - \overline{X}\sum_{i=1}^n Y_i \\
&= \sum_{i=1}^n X_iY_i - n\overline{X}\,\overline{Y}
\end{aligned}$$

Problem 3
Least-squares principle.

(a) State the least-squares principle.

(b) Derive the LS estimators for simple linear regression model.

(c) Assume the observations follow:

$$Y_i = exp(a + bX_i) + \epsilon_i, \ i = 1, \dots, n$$

where $a, b \in \mathbb{R}$ are unknown parameters and ϵ_i's are uncorrelated random variables with $E(\epsilon_i) = 0, Var(\epsilon_i) = \sigma^2$. Describe how to estimate the regression function (equivalently, a, b) by the least-squares principle. *Notes:* You only need to provide a description. This is an example of a nonlinear regression model.

Solution:

(a) The least squares principle is the best fit to a set of observed data $\{(X_i, Y_i)\}_{i=1}^n$ by a line is equivalent to solving the minimization problem

$$\min_{b_0, b_1} \sum_{i=1}^n (Y_i - (b_0 + b_1 X_i))^2$$

(b) Differentiating with respect to b_0, b_1, we get

$$\frac{\partial}{\partial b_0} \sum_{i=1}^n (Y_i - (b_0 + b_1 X_i))^2 = \frac{\partial}{\partial b_0} \left(b_0^2 + 2b_0 b_1 X_i - 2b_0 Y_i + b_1^2 X_i^2 - 2b_1 X_i Y_i + Y_i^2 \right) = \sum_{i=1}^n (2b_0 + 2b_1 X_i - 2Y_i)$$

$$\frac{\partial}{\partial b_1} \sum_{i=1}^n (Y_i - (b_0 + b_1 X_i))^2 = \frac{\partial}{\partial b_1} \left(b_0^2 + 2b_0 b_1 X_i - 2b_0 Y_i + b_1^2 X_i^2 - 2b_1 X_i Y_i + Y_i^2 \right) = \sum_{i=1}^n (2b_0 X_i + 2b_1 X_i^2 - 2X_i Y_i)$$

So identifying the estimators is equivalent to solving the system of equations:

$$\begin{cases} \sum_{i=1}^n (b_0 + b_1 X_i - Y_i) \\ \sum_{i=1}^n (b_0 X_i + b_1 X_i^2 - X_i Y_i) = 0 \end{cases}$$

Solving the first equation for β_0, we get:

$$0 = \sum_{i=1}^n (b_0 + b_1 X_i - Y_i) = n b_0 + b_1 \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i$$

$$\implies n b_0 = \sum_{i=1}^n Y_i - b_1 \sum_{i=1}^n X_i$$

$$\implies b_0 = \overline{Y} - b_1 \overline{X}$$

Solving the second equation for b_0, we get:

$$0 = \sum_{i=1}^n (b_0 X_i + b_1 X_i^2 - X_i Y_i) = b_0 \sum_{i=1}^n X_i + b_1 \sum_{i=1}^n X_i^2 - \sum_{i=1}^n X_i Y_i$$

$$\implies b_0 \sum_{i=1}^n X_i = \sum_{i=1}^n X_i Y_i - b_1 \sum_{i=1}^n X_i^2$$

$$\implies b_0 = \frac{\sum_{i=1}^n X_i Y_i - b_1 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i}$$

Setting the two equations as equal:

$$\overline{Y} - b_1 \overline{X} = \frac{\sum_{i=1}^n X_i Y_i - b_1 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i}$$

Solving for b_1, we get:

$$\begin{aligned}
\overline{Y} - b_1\overline{X} &= \frac{\sum_{i=1}^n X_iY_i}{n\overline{X}} - b_1\frac{\sum_{i=1}^n X_i^2}{n\overline{X}} \\
\implies b_1\overline{X} - b_1\frac{\sum_{i=1}^n X_i^2}{n\overline{X}} &= \overline{Y} - \frac{\sum_{i=1}^n X_iY_i}{n\overline{X}} \\
\implies b_1\overline{X}^2 - b_1\frac{1}{n}\sum_{i=1}^n X_i^2 &= \overline{X}\overline{Y} - \frac{1}{n}\sum_{i=1}^n X_iY_i \\
\implies b_1 &= \frac{\overline{X}\overline{Y} - \frac{1}{n}\sum_{i=1}^n X_iY_i}{\overline{X}^2 - \frac{1}{n}\sum_{i=1}^n X_i^2}
\end{aligned}$$

Now, simplifying, we see that

$$\begin{aligned}
b_1 &= \frac{\overline{X}\overline{Y} - \frac{1}{n}\sum_{i=1}^n X_iY_i}{\overline{X}^2 - \frac{1}{n}\sum_{i=1}^n X_i^2} \\
&= \frac{n\overline{X}\overline{Y} - \sum_{i=1}^n X_iY_i}{n\overline{X}^2 - \sum_{i=1}^n X_i^2} \\
&= \frac{\sum_{i=1}^n X_iY_i - n\overline{X}\overline{Y}}{\sum_{i=1}^n X_i^2 - n\overline{X}^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2} && \text{by Problem 2}
\end{aligned}$$

(c) The least squares principle is the best fit to a set of observed data $\{(X_i, Y_i)\}_{i=1}^n$ by a line is equivalent to solving the minimization problem

$$\min_{a,b\in\mathbb{R}} \sum_{i=1}^n (Y_i - exp(a + bX_i))^2$$

Problem 4

Tell true or false (with a brief explanation) of the following statements with regard to simple linear regression.

(a) The least squares line always passes the center of the data $(\overline{X}, \overline{Y})$.

(b) If $\overline{X} = 0, \overline{Y} = 0$, then $\hat{\beta}_0 = 0$ no matter what is $\hat{\beta}_1$.

(c) Given the sample size, the larger the sample variance of X_i's, the smaller the standard errors of $\hat{\beta}_0, \hat{\beta}_1$ tend to be.

Solution:

(a) True. This is forced by the constraint that $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$.

(b) True. $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X} = 0 - \hat{\beta}_1 \cdot 0 = 0$.

(c) True. Observe,

$$\sigma^2(\hat{\beta}_0) = \sigma^2\left[\frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right] \to \frac{\sigma^2}{n} \text{ as } \sum_{i=1}^n (X_i - \bar{X})^2 \to \infty$$

$$\sigma^2(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \to 0 \text{ as } \sum_{i=1}^n (X_i - \bar{X})^2 \to \infty$$

Since the variance in their distributions tend to zero, we conclude their errors must also tend to zero.

Problem 5
Under the simple linear regression model, recall that the residuals

$$\begin{aligned}\epsilon_i &= Y_i - \hat{Y}_i \\ &= Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i) \\ &= (Y_i - \overline{Y}) - \hat{\beta}_1 (X_i - \overline{X}).\end{aligned}$$

for $i = 1, l \ldots, n$. Show that:

(a) $\sum_{i=1}^n e_i = 0$

(b) $\sum_{i=1}^n X_i e_i = 0$

(c) $\sum_{i=1}^n \hat{Y}_i e_i = 0$

Solution:

(a)

$$\begin{aligned}\sum_{i=1}^n e_i &= \sum_{i=1}^n [(Y_i - \overline{Y}) - \hat{\beta}_1 (X_i - \overline{X})] \\ &= \sum_{i=1}^n (Y_i - \overline{Y}) - \hat{\beta}_1 \sum_{i=1}^n (X_i - \overline{X}) \\ &= 0 + \hat{\beta}_1 0 && \text{Problem 1(a)} \\ &= 0\end{aligned}$$

(b)

$$\begin{aligned}\sum_{i=1}^n X_i e_i &= \sum_{i=1}^n (X_i Y_i - X_i \overline{Y}) - \hat{\beta}_1 X_i (X_i - \overline{X}) \\ &= \sum_{i=1}^n X_i (Y_i - \overline{Y}) - \hat{\beta}_1 X_i (X_i - \overline{X}) \\ &= \sum_{i=1}^n X_i (Y_i - \overline{Y}) - \hat{\beta}_1 \sum_{i=1}^n X_i (X_i - \overline{X}) \\ &= \sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) - \hat{\beta}_1 \sum_{i=1}^n (X_i - \overline{X})^2 && \text{Problem 2} \\ &= \sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) - \left[\frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2}\right] \sum_{i=1}^n (X_i - \overline{X})^2 && \text{Problem 3(b)} \\ &= 0\end{aligned}$$

(c) Observe

$$\begin{aligned}
\sum_{i=1}^{n} \hat{Y}_i e_i &= \sum_{i=1}^{n} (\beta_0 + \beta_1 X_i) e_i \\
&= \sum_{i=1}^{n} (\beta_0 e_i + \beta_1 X_i e_i) \\
&= \beta_0 \underbrace{\sum_{i=1}^{n} e_i}_{(a) \implies 0} + \beta_1 \underbrace{\sum_{i=1}^{n} X_i e_i}_{(b) \implies 0} \\
&= 0 + 0 = 0
\end{aligned}$$

Problem 6
Under the simple linear regression model, show that the LS estimator $\hat{\beta}_0$ is an unbiased estimator of β_0.

Solution: We calculate

$$\begin{aligned}
E[\hat{\beta}_1 | X_1, \ldots, X_n] &= E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} | X_1, \ldots, X_n\right] \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2} E\left[\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) | X_1, \ldots, X_n\right] \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2} E\left[\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) | X_1, \ldots, X_n\right] \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2} E\left[\sum_{i=1}^{n}(X_i - \overline{X}) Y_i | X_1, \ldots, X_n\right] && \text{Problem 2(b)} \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \sum_{i=1}^{n}(X_i - \overline{X}) E\left[Y_i | X_1, \ldots, X_n\right] \\
&= \frac{1}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \sum_{i=1}^{n}(X_i - \overline{X})(\beta_0 + \beta_1 X_i) \\
&= \frac{\beta_0 \sum_{i=1}^{n}(X_i - \overline{X}) + \beta_1 \sum_{i=1}^{n}(X_i - \overline{X}) X_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \\
&= \frac{\beta_1 \sum_{i=1}^{n}(X_i - \overline{X}) X_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2} && \text{Problem 2(a)} \\
&= \frac{\beta_1 \sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} && \text{Problem 2(a)} \\
&= \beta_1
\end{aligned}$$

$$
\begin{aligned}
E[\hat{\beta}_0|X_1, \dots, X_n] &= E[\overline{Y} - \beta_1 \overline{X}|X_1, \dots, X_n] \\
&= E[\frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{\beta_1}{n}\sum_{i=1}^{n} X_i|X_1, \dots, X_n] \\
&= E[\frac{1}{n}\sum_{i=1}^{n}(Y_i - \beta_1 X_i)|X_1, \dots, X_n] \\
&= \frac{1}{n}\sum_{i=1}^{n} E[(Y_i - \beta_1 X_i)|X_1, \dots, X_n] \\
&= \frac{1}{n}\sum_{i=1}^{n}(E[Y_i|X_1, \dots, X_n] - \beta_1 X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 X_i - \beta_1 X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \beta_0 \\
&= \beta_0
\end{aligned}
$$

Problem 7
Submitted as a Markdown file.