1. First-order model with two predictor variables (and no interaction)

• model \( Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \quad i = 1, \ldots, n \)

• regression function: \( E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \)
  here the regression surface (or response surface) is a plane

• \( \beta_1 \) is the change in mean response per unit increase in \( X_1 \) when \( X_2 \) is held constant
• \( \beta_2 \) is the change in mean response per unit increase in \( X_2 \) when \( X_1 \) is held constant

\( \beta_1 \) and \( \beta_2 \) are sometimes called partial regression coefficients

• note that \( \frac{\partial E\{Y\}}{\partial X_1} = \beta_1 \) and \( \frac{\partial E\{Y\}}{\partial X_2} = \beta_2 \)

• When the effect of \( X_1 \) on the mean response does not depend on the level of \( X_2 \) and
correspondingly the effect of \( X_2 \) does not depend on the level of \( X_1 \), the two
independent variables are said to have additive effects or not to interact

2. General linear model in matrix terms

\[
Y = X\beta + \epsilon \quad \text{where} \quad Y_{n \times 1} \quad \text{is a vector of responses}
\]

\[
\beta_{p \times 1} \quad \text{is a vector of parameters}
\]

\[
X_{n \times p} \quad \text{is a matrix of constants}
\]

\[
\epsilon_{n \times 1} \quad \text{is vector of iid normal rv's}
\]

with \( E\{\epsilon\} = 0 \) and \( \sigma^2\{\epsilon\} = \sigma^2 I \)

Example.

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
Y_3 \\
Y_4 \\
Y_5 \\
Y_6
\end{bmatrix} = 
\begin{bmatrix}
1 & X_{11} & X_{12} & X_{13} \\
1 & X_{21} & X_{22} & X_{23} \\
1 & X_{31} & X_{32} & X_{33} \\
1 & X_{41} & X_{42} & X_{43} \\
1 & X_{51} & X_{52} & X_{53} \\
1 & X_{61} & X_{62} & X_{63}
\end{bmatrix} 
\begin{bmatrix}
\beta_0 \\
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix} + 
\begin{bmatrix}
\epsilon_1 \\
\epsilon_2 \\
\epsilon_3 \\
\epsilon_4 \\
\epsilon_5 \\
\epsilon_6
\end{bmatrix}
\]
• regression function: \( \mathbb{E}\{Y\} = X \beta \); also \( \sigma^2 \{Y\} = \sigma^2 I \)

• normal equations: \( X'Xb = X'Y \) whose solution gives least squares estimators

\[
b = (X'X)^{-1} X'Y = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{bmatrix}
\]

• fitted values \( \hat{Y} = Xb \) or \( \hat{Y} = HY \) where hat matrix \( H = X(X'X)^{-1} X' \)

• residuals \( e = Y - \hat{Y} = Y - Xb = Y - HY = (I - H)Y \)

• \( \sigma^2 \{e\} = \sigma^2(I - H) \) is estimated by \( s^2 \{e\} = \text{MSE}(I - H) \)

• inferences about regression parameters

\( b \) is an unbiased estimator of \( \beta \)

variance-covariance matrix of \( b \): \( \sigma^2 \{b\} = \sigma^2 (X'X)^{-1} \) estimated by \( s^2 \{b\} = \text{MSE} (X'X)^{-1} \)

\( 1 - \alpha \) confidence limits for \( \beta_k \): \( \beta_k \pm t(1 - \alpha/2; \ n - p) \cdot s\{b_k\} \) since

\[
\frac{b_k - \beta_k}{s\{b_k\}} \sim t(n - p)
\]

Bonferroni joint confidence intervals can also be obtained;

Hypothesis test for \( H_0: \beta_k = 0 \) vs \( H_a: \beta_k \neq 0 \) uses \( t^* = \frac{b_k}{s\{b_k\}} \)

reject \( H_o \) if \( |t^*| > t(1 - \alpha/2; \ n - p) \)