1 Matrix based approach to regression

For our small data set we have:
$y_1 = \beta_0 + \beta_1 x_{11} + \varepsilon_1$,
$y_2 = \beta_0 + \beta_1 x_{21} + \varepsilon_2$,
$y_3 = \beta_0 + \beta_1 x_{31} + \varepsilon_3$,
$y_4 = \beta_0 + \beta_1 x_{41} + \varepsilon_4$,
$y_5 = \beta_0 + \beta_1 x_{51} + \varepsilon_5$, where x_{i1}, for example, is the ith observation's value of variable x_1.

This can be rewritten as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ 1 & x_{31} \\ 1 & x_{41} \\ 1 & x_{51} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix},$$

or more briefly as:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon.$$

For the matrix-based approach to regression, the least squares estimate of β is: $\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, and the variance of the estimate is $Var(\widehat{\beta}) = \sigma_\varepsilon^2(\mathbf{X}'\mathbf{X})^{-1}$. We estimate σ_ε^2 in the simple linear regression model with $s_\varepsilon^2 = \sum_{i=1}^n (y_i - \widehat{y}_i)^2/(n-2)$, and in a mulitple regression model with k covariates with $s_\varepsilon^2 = \sum_{i=1}^n (y_i - \widehat{y}_i)^2/(n-k-1)$.

For our cereal data example we have
$110 = \beta_0 + \beta_1(0) + \varepsilon_1$,
$110 = \beta_0 + \beta_1(1) + \varepsilon_2$,
$150 = \beta_0 + \beta_1(3) + \varepsilon_3$,
$130 = \beta_0 + \beta_1(2) + \varepsilon_4$,
$120 = \beta_0 + \beta_1(1) + \varepsilon_5$, so that

$$\mathbf{Y} = \begin{bmatrix} 110 \\ 110 \\ 150 \\ 130 \\ 120 \end{bmatrix} \text{ and } \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

Then we have

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 7 & 15 \end{bmatrix},$$

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 110 \\ 110 \\ 150 \\ 130 \\ 120 \end{bmatrix} = \begin{bmatrix} 620 \\ 940 \end{bmatrix}, \text{ and}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 5 & 7 \\ 7 & 15 \end{bmatrix}^{-1} = \begin{bmatrix} 15/26 & -7/26 \\ -7/26 & 5/26 \end{bmatrix} \text{ so}$$

$$\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 15/26 & -7/26 \\ -7/26 & 5/26 \end{bmatrix} \begin{bmatrix} 620 \\ 940 \end{bmatrix} = \begin{bmatrix} 104.62 \\ 13.85 \end{bmatrix}.$$

Also $s.e.(\widehat{\beta}_1) = \sqrt{Var(\widehat{\beta}_1)} = \sqrt{\widehat{\sigma}^2_\varepsilon(5/26)} = \sqrt{(41.03)(5/26)} = 2.81.$