5. Consider the fitted values that result from performing linear regression without an intercept. In this setting, the ith fitted value takes the form

$$\hat{y}_i = x_i\hat{\beta},$$

where

$$\hat{\beta} = \left(\sum_{i=1}^{n} x_i y_i\right) \Big/ \left(\sum_{i'=1}^{n} x_{i'}^2\right). \tag{3.38}$$

Show that we can write

$$\hat{y}_i = \sum_{i'=1}^{n} a_{i'} y_{i'}.$$

What is $a_{i'}$?

By the above equation, $\hat{y}_i = x_i\hat{\beta}$, where $\hat{\beta} = \frac{\Sigma x_i y_i}{\Sigma x_i^2}$.

Then, $\hat{y}_i = x_i \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2} = \frac{\sum_{j=1}^{n} x_i x_j y_j}{\sum_{j=1}^{n} x_j^2}$. Since x_j are fixed for $\forall j \in \{1, \cdots, n\}$, $\sum_{j=1}^{n} x_j$ is a constant, say $\frac{1}{S} := \sum_{j=1}^{n} x_j$.

Thus, $\hat{y}_i = \frac{\sum_{j=1}^{n} x_i x_j y_j}{\sum_{j=1}^{n} x_j^2} = S \sum_{j=1}^{n} x_i x_j y_j = \sum_{j=1}^{n} (S x_i x_j) y_j$. $\therefore$ $\underline{a_j = \frac{x_i x_j}{\sum_{k=1}^{n} x_k^2}}$

Also, we can write $\underline{\hat{y}_i} = \sum_{i'=1}^{n} a_{i'} y_{i'} = \underline{\sum_{i'=1}^{n} \left\{ \left(\frac{x_i}{\sum_{j=1}^{n} x_j^2} \right) x_{i'} \right\} y_{i'}}$

Ex. 3.4 Show how the vector of least squares coefficients can be obtained from a single pass of the Gram–Schmidt procedure (Algorithm 3.1). Represent your solution in terms of the QR decomposition of $\mathbf{X}$.

Suppose $X = QR$, where $Q^TQ = QQ^T = I$, and R is the upper triangular matrix.

Define a linear regression model $y = f(x) + \varepsilon = X\beta + \varepsilon = (QR)\beta + \varepsilon$, where $\varepsilon \sim N(0, \sigma^2 I_n)$.

Then. $\varepsilon^T\varepsilon = \sum_{i=1}^{n} \varepsilon_i^2 = (y-(QR)\beta)^T(y-(QR)\beta) = (y^T - \beta^T R^T Q^T)(y - QR\beta) = y^Ty - y^TQR\beta - \beta^T R^T Q^T y + \beta^T R^T Q^T Q R \beta$

Since $(y^TQR\beta)^T = y^TQR\beta$, and $Q^TQ = I$, $\varepsilon^T\varepsilon = y^Ty - 2\beta^T R^T Q^T y + \beta^T R^T R \beta$

$\therefore$ By the normal equation, $-2R^TQ^Ty + 2R^TR\hat{\beta} = 0$. $\therefore R^TR\hat{\beta} = R^TQ^Ty$

$\therefore \hat{\beta} = (R^TR)^{-1}R^TQ^Ty$

By our assumption, X has full rank. $\therefore$ R is invertible since there is no zero element on the diagonal.

$\therefore (R^TR)^{-1} = R^{-1}(R^T)^{-1}$ $\therefore \hat{\beta} = R^{-1}(R^T)^{-1}R^TQ^Ty = R^{-1}Q^Ty$