The Vector Geometry of Linear Models

Simple Regression

In this chapter we will learn about the geometry of least-squares estimation. We start with the simple regression model $\mathbf{y} = \alpha \mathbf{1}_n + \beta \mathbf{x} + \varepsilon$ which has a fitted equation of $\mathbf{y} = A\mathbf{1}_n + B\mathbf{x} + \mathbf{e}$. Then we also have $E(\mathbf{y}) = \alpha \mathbf{1}_n + \beta \mathbf{x}$ and $\widehat{\mathbf{y}} = A\mathbf{1}_n + B\mathbf{x}$. Figures 10.1 and 10.2 in the text introduce the concept of viewing the data $\mathbf{y}$ as a vector in (at first) 3 dimensional space with the vectors $\mathbf{1}_n$ and $\mathbf{x}$ forming a subspace of this space. Recall that a subspace of this three-dimensional space that is spanned by $\mathbf{1}_n$ and $\mathbf{x}$ consists of all vectors that are linear combinations of $\mathbf{1}_n$ and $\mathbf{x}$, hence they lie in the plane illustrated in the Figures.

Mean-Deviation Form

The situation is even easier to understand by considering the simple linear regression model in mean-deviation form. We subtract the mean $\overline{Y} = A + B\overline{x}$ from the fitted model in scalar form $Y_i = A + Bx_i + E_i$ to yield $Y_i - \overline{Y} = B(x_i - \overline{x}) + E_i$. Now we write the model in matrix form using $\mathbf{y}^* = Y_i - \overline{Y}$ and $\mathbf{x}^* = x_i - \overline{x}$ to get:

$$\mathbf{y}^* = B\mathbf{x}^* + \mathbf{e}.$$

The accompanying Figure 10.3 is now two-dimensional and shows that the least-squares fit $\widehat{\mathbf{y}}^*$ is the orthogonal projection of $\mathbf{y}^*$ onto the one-dimensional subspace spanned by $\mathbf{x}^*$, (all vectors of the form $k\mathbf{x}^*$) to give $\widehat{\mathbf{y}}^* = B\mathbf{x}^*$. As shown in the text's Appendix, the orthogonal projection of $\mathbf{y}^*$ onto $\mathbf{x}^*$ is defined as:

$$B = \frac{\mathbf{x}^* \cdot \mathbf{y}^*}{\|\mathbf{x}^*\|^2} = \frac{\sum(x_i - \overline{x})(Y_i - \overline{Y})}{\sum(x_i - \overline{x})^2}$$

and leads to a right-angle relationship between $\widehat{\mathbf{y}}^*$ and $\mathbf{e}$. The expression for B is recognizable as the usual formula for the least-squares estimate of β in simple linear regression, and the right-angle relationship immediately gives us TSS = RegSS + RSS via the Pythagorean Theorem. The Figure also shows that

$$r = \sqrt{\frac{\text{RegSS}}{\text{TSS}}} = \frac{\mathbf{x}^* \cdot \mathbf{y}^*}{\|\mathbf{x}^*\| \, \|\mathbf{y}^*\|},$$

and is also recognizable as the cosine of the angle W in the Figure between $\mathbf{y}^*$ and $\widehat{\mathbf{y}}^*$. Thus as $\mathbf{y}^*$ gets closer to $\widehat{\mathbf{y}}^*$ indicating a better fit, $\cos(W)$ gets close to 1, which we know is the maximum value of the correlation r.

Degrees of Freedom

The concept of degrees of freedom is easy to understand when approached from the perspective of vector geometry, as the degrees of freedom for a sum of squares corresponds to the dimension of the subspace in which the effects vector is confined:

i) Since $\mathbf{y}$ is unconstrained, it can be anywhere in the n-dimensional observation space, therefore the uncorrected sum of squares, $\sum Y_i^2 = \|\mathbf{y}\|^2$ has n degrees of freedom.

ii) In mean-deviation form, the values $y_i^* = Y_i - \overline{Y}$ are constrained to add to zero, hence only $n-1$ of the y_i^* values are linearly independent. Thus TSS $= \|\mathbf{y}^*\|^2$ has $n-1$ degrees of freedom.

iii) For simple linear regression in mean-deviation form, $\mathbf{x}^*$ spans a one-dimensional subspace. Since $\widehat{\mathbf{y}}^*$ lies in this subspace, RegSS $= \|\widehat{\mathbf{y}}^*\|^2$ has 1 degree of freedom.

iv) For simple linear regression, the vectors $\mathbf{1}_n$ and $\mathbf{x}$ span a subspace of dimension 2. The residual vector $\mathbf{e}$ is orthogonal to the plane spanned by $\mathbf{1}_n$ and $\mathbf{x}$ (it is in the orthogonal complement of that 2 dimensional space), so it lies in a subspace of dimension $n-2$.