Weighted Least Squares

Consider the general linear model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

To obtain the least squares estimates of the unknown β parameters, recall (from Section 4.3) that we minimize the quantity

$$\text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki})]^2$$

with respect to $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$. The least squares criterion weighs each observation equally in determining the estimates of the β's. With **weighted least squares** we want to weigh some observations more heavily than others. To do this we minimize

$$\begin{aligned} \text{WSSE} &= \sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^{n} w_i [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki})]^2 \end{aligned}$$

where w_i is the weight assigned to the ith observation. The resulting parameter estimates are called **weighted least squares estimates**. [Note that the ordinary least squares procedure assigns a weight of $w_i = 1$ to each observation.]

Definition 9.1 **Weighted least squares regression** is the procedure that obtains estimates of the β's by minimizing WSSE $= \sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2$, where w_i is the weight assigned to the ith observation. The β-estimates are called **weighted least squares estimates**. minimize sum(wi(ei^2)) instead of sum(ei^2)

Definition 9.2 The **weighted least squares residuals** are obtained by computing the quantity

$$\sqrt{w_i}(y_i - \hat{y}_i)$$

for each observation, where $\hat{y}_i$ is the predicted value of y obtained using the weight w_i in a weighted least squares regression.