Chapter 11: Building the Regression Model III: Remedial Measures

STAT 3240

Michael McIsaac

UPEI

11.1: Remedial Measures for Unequal Error Variances: Weighted Least Squares

Suppose $Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_{p-1} X_{i,p-1} + \varepsilon_i;$　　　　　$\varepsilon_i \sim N(0, \sigma_i^2)$
so

$$\sigma^2\{\varepsilon\} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix}$$

Then instead of finding parameter estimates that minimize the least squares criterion $\sum_{i=1}^{n}(Y_i - (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_{p-1} X_{i,p-1}))^2$
we would be better off minimizing the *weighted least squares criterion*:

$$\sum_{i=1}^{n} w_i(Y_i - (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_{p-1} X_{i,p-1}))^2, \qquad \text{where } w_i = \frac{1}{\sigma_i^2}.$$

Consider

$$\mathbb{W} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & w_n \end{bmatrix}$$

where, again, $w_i = \frac{1}{\sigma_i^2}$ would account for the unequal variances.

The weighted least squares estimators are

$$\mathbf{b_w} = (\mathbb{X}'\mathbb{W}\mathbb{X})^{-1}\mathbb{X}'\mathbb{W}\mathbb{Y}$$

with variance-covariance matrix

$$\sigma^2\{\mathbf{b_w}\} = (\mathbb{X}'\mathbb{W}\mathbb{X})^{-1}$$

We can also think about this as the OLS estimator in the scaled model

$$(\mathbb{W}^{1/2}\mathbb{Y}) = (\mathbb{W}^{1/2}\mathbb{X})\beta + \mathbb{W}^{1/2}\varepsilon$$

3 / 8

Of course, we won't know σ_i^2 $(i = 1, \ldots, n)$ in practice, so we will need to estimate these in order to estimate the $w_i = 1/\sigma_i^2$.

In some experimental settings, we may have sufficient replicates to estimate all of the σ_i directly.

More generally, we try to get a sense of how σ^2 changes as X_j or $\hat{Y}$ changes (just like we do when we attempt to identify non-constant variance).

4 / 8

Weighted least squares estimation process:

1. Fit the regression model by unweighted least squares and analyze the residuals.

2. Estimate the variance function or the standard deviation function by regressing either the squared residuals or the absolute residuals on the appropriate predictor(s).

3. Use the fitted values from the estimated variance or standard deviation function to obtain the weights w_i.

4. Estimate the regression coefficients using these weights.

For *iteratively reweighted least squares*, we would iterate this process until convergence.

We need to do a little more work to correctly estimate standard errors of the parameter estimates. Bootstrapping can again help with this.

11.3: Remedial Measures for Influential Cases: Robust Regression

What to do when an influential case is found:

- examine whether an outlying case is the result of a recording error.

    - if erroneous data can be corrected, this should be done. Otherwise, truly erroneous data should be discarded.

- Many times, it is not possible to tell for certain whether the observations for an outlying case are erroneous.

    - Such cases should usually not be discarded.

- If an outlying influential case is not clearly erroneous, the next step should be to examine the adequacy of the model.

    - Examination of these outlying cases may provide important clues as to how the model needs to be modified.

6 / 8

- Discarding of outlying influential cases that are not clearly erroneous and that cannot be accounted for by model improvement should be done only rarely, such as when the model is not intended to cover the special circumstances related to the outlying case.

- An alternative to discarding outlying cases that is less severe is to dampen the influence of these cases. That is the purpose of robust regression.

One potential method of *robust regression* is to, again, employ *iteratively reweighted least squares* to minimize

$$\sum_{i=1}^{n} w_i (Y_i - (\beta_0 + \beta_1 X_{i1} + \cdots + \beta_{p-1} X_{i,p-1}))^2.$$

Here, instead of using weights that measure our uncertainty around a given point $(w_i = 1/\sigma_i^2)$, we use a weight that reduces the influence of outlying cases (so w_i decreases as the size of the residual e_i increases).

This can be accomplished by the following steps:

1. Choose a weight function for weighting the cases.

2. Obtain starting weights for all cases.

3. Use the starting weights in weighted least squares and obtain the residuals from the fitted regression function.

4. Use the residuals in step 3 to obtain revised weights.

5. Continue the iterations until convergence is obtained.

Robust regression, and weighted least squares, are often useful for confirming reasonableness of ordinary least squares results.