

Weighted Least Squares (WLS)

INTRO

- The **Weighted Least Squares (WLS)** method is an extension of the ordinary least squares (OLS) technique. The model is formulated the same way as in OLS:

$$\hat{y}(\mathbf{x}) = \mathbf{x}^\top \hat{\beta} + \varepsilon(\mathbf{x})$$

where \mathbf{x} is the vector of features, β is the vector of parameters, $y(\mathbf{x})$ is the target variable, and $\varepsilon(\mathbf{x})$ is the error term.

- Each observation \mathbf{x} has associated weights $w(\mathbf{x})$ that reflect the importance of that particular observation.
- This method minimizes the weighted sum of squared residuals:

$$\text{RSS} = \sum_{\mathbf{x} \in X^\ell} w(\mathbf{x}) \cdot (y(\mathbf{x}) - \hat{y}(\mathbf{x}))^2 \rightarrow \min_{\beta}.$$

- The solution to this minimization problem is given by:

$$\beta^* = \underbrace{(X^\top W X)^{-1} X^\top W \mathbf{y}}_{X_W^+},$$

where W is the diagonal matrix of weights, and X_W^+ is the weighted pseudo-inverse.

FORMALISM

Weight matrix

For a weighted

$$\text{RSS} = \sum_{\mathbf{x} \in X^\ell} w(\mathbf{x}) \cdot (y(\mathbf{x}) - \hat{y}(\mathbf{x}))^2$$

let's introduce the weight matrix:

$$\begin{aligned} W &:= \text{diag}(w(\mathbf{x}_1), \dots, w(\mathbf{x}_\ell)) \\ &= \begin{pmatrix} w(\mathbf{x}_1) & & \\ & \ddots & \\ & & w(\mathbf{x}_\ell) \end{pmatrix} \end{aligned}$$

Matrix form

Thus, we can rewrite the RSS in matrix form as a quadratic form:

$$\text{RSS} = (\mathbf{y} - X\beta)^\top W(\mathbf{y} - X\beta).$$

Transformation to standard LS problem

The weighted LS problem can be easily reformulated as a standard LS problem by replacing the original variables with transformed ones:

$$\mathbf{y}' := W^{\frac{1}{2}} \mathbf{y}, \quad X' := W^{\frac{1}{2}} X, \quad \varepsilon' := W^{\frac{1}{2}} \varepsilon$$

Substituting these transformations into the original model, we get:

$$\mathbf{y}' = X' \beta + \varepsilon'$$

Heteroscedasticity can be eliminated by applying weighted LS.

For a model with non-constant variance of the error term:

$$\mathbf{y} = X\beta + \varepsilon, \quad \text{Var}[\varepsilon(\mathbf{x})] = y(\mathbf{x})^2 \cdot \sigma^2$$

To apply WLS, the weights must have a negative square unit:

$$w(\mathbf{x}) = \frac{1}{y(\mathbf{x})^2}$$

This leads to the transformations:

$$\mathbf{y}' = \frac{\mathbf{y}}{\sqrt{w}}, \quad \mathbf{x}' = \frac{\mathbf{x}}{\sqrt{w}}, \quad \varepsilon' = \frac{\varepsilon}{\sqrt{w}}$$

The weight matrix is:

$$W = \begin{pmatrix} \frac{1}{y(\mathbf{x}_1)^2} & & \\ & \ddots & \\ & & \frac{1}{y(\mathbf{x}_\ell)^2} \end{pmatrix}$$

and

$$\mathbf{y}' = \sqrt{W} \mathbf{y}, \quad \mathbf{x}' = \sqrt{W} \mathbf{x}, \quad \varepsilon' = \sqrt{W} \varepsilon$$

Now, the model can be formulated as a homoscedastic least squares problem:

$$\mathbf{y}' = X' \beta + \varepsilon', \quad \text{Var}[\varepsilon'(\mathbf{x})] = \sigma^2$$

Quadratic form is a function of the form:

$$Q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j.$$

Coefficients $a_{i,j}$ can be arranged in a symmetric matrix A , and the quadratic form can be written in matrix form as:

$$Q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}.$$

Analytical solution

Now, let's solve for β in the transformed model. Since W and $W^{\{\frac{1}{2}\}}$ are diagonal matrices, transposing them results in the same matrix:

$$\beta^* = X'^{+}y' = \left(X'^T X'\right)^{-1} X'^T y'$$

Expanding the expressions:

$$\begin{aligned} \beta^* &= \left(\left(W^{\frac{1}{2}}X\right)^T W^{\frac{1}{2}}X\right)^{-1} \left(W^{\frac{1}{2}}X\right)^T W^{\frac{1}{2}}y \\ &= \left(X^T W X\right)^{-1} X^T W y \end{aligned}$$

Therefore, the solution is:

$$\beta^* = \underbrace{\left(X^T W X\right)^{-1} X^T W y}_{X_w^{+}}$$