WEIGHTED LEAST SQUARES

Parastoo Bybordi

Regression2

Spring 2021

INTRODUCTION

- Ordinary least squares: To estimate the regression coefficients, the sum of the second power of the fitting errors is zero. Normal equations are obtained and coefficients are estimated.
- When can we not use OLS?

One of the common assumptions in the modeling process(Least Square), both linear and nonlinear, is that each data point provides equally accurate information about the determinant of the whole process.. In other words, it is assumed that the standard deviation of the error term is constant on all values of the predictive or explanatory variables. But this assumption does not apply to some models.

In the typical regression model, $Y = \beta 0 + \beta 1Xi + \varepsilon i$, random errors can be evenly distributed and independent with constant variance. When this assumption is violated, then OLS estimators of regression coefficients lose the property of least variance among all unbiased linear estimators (according to Gauss-Markov theorem, they are not BLUE).

INTRODUCTION

Violation of such an assumption can occur in any of the following situations:

1. The variance of the random error components is not constant.
2. Random error components are not independent.
3. Random error components have no fixed variance and are not independent.

In such cases, the covariance matrix of the random error components does not remain in the form of the identity matrix but can be considered as any positive-definite matrix. Under such an assumption, the OLS estimation does not remain as efficient as the identity covariance matrix. In such cases, the GLS or WLS method is used to estimate the model parameters to estimate the model parameters.

WEIGHTED LEAST SQUARES

Weighted Least Squares is a method that deals with observations which have non-constant variance. Suppose $Y = a + bx_i + \varepsilon_i$ and the variances are not constant, and the observed variance of i is equal to σ_i^2, then:

$$w_i = \frac{1}{\sigma_i^2}$$

$$\sum w_i e_i^2 = \sum \frac{1}{\sigma_i^2} e_i^2 = \sum \left(\frac{e_i}{\sigma_i}\right)^2$$

$$L = \sum w_i e_i^2$$

$$\Longrightarrow \quad L = \sum w_i (y_i - \alpha - \beta x_i)^2$$

WEIGHTED LEAST SQUARES

$$\frac{dL}{d\alpha} = -2\sum w_i(y_i - a - bx_i) = 0$$

$$-2\sum w_i(y_i - a - bx_i) = 0$$

$$\sum w_i(y_i - a - bx_i) = 0$$

$$\sum w_i y_i - \sum a w_i - \sum b w_i x_i = 0$$

$$\sum w_i y_i - a\sum w_i - b\sum w_i x_i = 0$$

$$a\sum w_i = \sum w_i y_i - b\sum w_i x_i$$

$$\boldsymbol{a} = \frac{\sum \boldsymbol{w_i y_i} - \boldsymbol{b}\sum \boldsymbol{w_i x_i}}{\sum \boldsymbol{w_i}}$$

$$\frac{dL}{d\beta} = -2\sum w_i(y_i - a - bx_i)x_i = 0$$

$$-2\sum w_i(x_i y_i - a x_i - b x_i^2) = 0$$

$$\sum w_i(x_i y_i - a x_i - b x_i^2) = 0$$

$$\sum w_i x_i y_i - \sum a w_i x_i - \sum b w_i x_i^2 = 0$$

$$\sum w_i x_i y_i - a\sum w_i x_i - b\sum w_i x_i^2 = 0$$

$$b\sum w_i x_i^2 = \sum w_i x_i y_i - a\sum w_i x_i$$

$$b = \frac{\sum w_i x_i y_i - a\sum w_i x_i}{\sum w_i x_i^2}$$

WEIGHTED LEAST SQUARES

$$b = \frac{\sum w_i x_i y_i - \left[\frac{\sum w_i y_i - b\sum w_i x_i}{\sum w_i}\right]\sum w_i x_i}{\sum w_i x_i^2}$$

$$= \frac{\sum w_i x_i y_i - \left[\frac{\sum w_i x_i \sum w_i y_i - b\left(\sum w_i x_i\right)^2}{\sum w_i}\right]}{\sum w_i x_i^2}$$

$$= \frac{\frac{\sum w_i \sum w_i x_i y_i - \left[\sum w_i x_i \sum w_i y_i - b\left(\sum w_i x_i\right)^2\right]}{\sum w_i}}{\sum w_i x_i^2}$$

$$= \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i + b\left(\sum w_i x_i\right)^2}{\sum w_i \sum w_i x_i^2}$$

WEIGHTED LEAST SQUARES

$$b\sum w_i \sum w_i x_i^2 = \sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i + b\left(\sum w_i x_i\right)^2$$

$$b\sum w_i \sum w_i x_i^2 - b\left(\sum w_i x_i\right)^2 = \sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i$$

$$b\left[\sum w_i \sum w_i x_i^2 - \left(\sum w_i x_i\right)^2\right] = \sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i$$

$$\boldsymbol{b} = \frac{\sum \boldsymbol{w_i} \sum \boldsymbol{w_i x_i y_i} - \sum \boldsymbol{w_i x_i} \sum \boldsymbol{w_i y_i}}{\sum \boldsymbol{w_i} \sum \boldsymbol{w_i x_i^2} - \left(\sum \boldsymbol{w_i x_i}\right)^2}$$

DIFFERENCES BETWEEN OLS AND WLS

$$\boldsymbol{y_i = a + bx_i}$$

Ordinary Least Square (OLS)	*Weighted Least Square*
$b = \dfrac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$	$b = \dfrac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i}{\sum w_i \sum w_i x_i^2 - (\sum w_i x_i)^2}$
$a = \bar{y} - b\bar{x}$	$a = \dfrac{\sum w_i y_i - b\sum w_i x_i}{\sum w_i}$

If the weights, w_i, are all the same constant, then we have ordinary least squares (OLS) regression.

MULTIVARIATE COEFFICIENT ESTIMATORS

Assume that the covariance matrix of variance is :

$$\begin{pmatrix} \sigma_{۱}^{۲} & ۰ & \cdots & ۰ \\ ۰ & \sigma_{۲}^{۲} & \cdots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \cdots & \sigma_{n}^{۲} \end{pmatrix}$$

$$،W_i = \frac{۱}{\sigma_i^{۲}} \;, i = ۱, ۲ \ldots, n$$

In this case, suppose:

$$\mathbf{W} = \mathbf{W}^{1/2}\mathbf{W}^{1/2}, \qquad \text{and} \qquad \mathbf{W}^{1/2}\mathbf{W}^{-1/2} = \mathbf{W}^{-1/2}\mathbf{W}^{1/2} = \mathbf{I};$$

$$(\mathbf{W}^{1/2})_{ii} := \sqrt{w_i}, \qquad \text{and} \qquad (\mathbf{W}^{-1/2})_{ii} := \frac{1}{\sqrt{w_i}}.$$

MULTIVARIATE COEFFICIENT ESTIMATORS

If we multiply the matrix $W^{1/2}$ on the equation from the left, we have:

$$\mathbf{W}^{1/2}\mathbf{y} = \mathbf{W}^{1/2}\mathbf{X}\beta + \mathbf{W}^{1/2}\mathbf{e},$$

We have:

$$\mathbf{z} := \mathbf{W}^{1/2}\mathbf{y}, \qquad \mathbf{M} := \mathbf{W}^{1/2}\mathbf{X}, \qquad \text{and} \qquad \mathbf{d} := \mathbf{W}^{1/2}\mathbf{e}.$$

So:

$$\mathbf{z} = \mathbf{M}\beta + \mathbf{d}.$$

$$\begin{aligned}
\mathrm{RSS}(\beta; \mathbf{W}) &= (\mathbf{z} - \mathbf{M}\beta)^T(\mathbf{z} - \mathbf{M}\beta) \\
&= [\mathbf{W}^{1/2}(\mathbf{y} - \mathbf{X}\beta)]^T[\mathbf{W}^{1/2}(\mathbf{y} - \mathbf{X}\beta)] \\
&= (\mathbf{y} - \mathbf{X}\beta)^T\mathbf{W}(\mathbf{y} - \mathbf{X}\beta).
\end{aligned}$$

MULTIVARIATE COEFFICIENT ESTIMATORS

$$\begin{aligned}\mathbb{V}\mathrm{ar}[\mathbf{d}|\mathbf{X}] &= \mathbb{V}\mathrm{ar}[\mathbf{W}^{1/2}\mathbf{e}|\mathbf{X}] \\ &= \mathbf{W}^{1/2}\,\mathbb{V}\mathrm{ar}[\mathbf{e}|\mathbf{X}](\mathbf{W}^{1/2})^T \\ &= \mathbf{W}^{1/2}\sigma^2\mathbf{W}^{-1}(\mathbf{W}^{1/2})^T \\ &= \sigma^2\mathbf{W}^{1/2}\mathbf{W}^{-1/2}\mathbf{W}^{-1/2}(\mathbf{W}^{1/2}) \\ &= \sigma^2\mathbf{I}.\end{aligned}$$

$$\mathbf{z} = \mathbf{M}\beta + \mathbf{d}, \qquad \text{and} \qquad \mathbb{V}\mathrm{ar}[\mathbf{d}|\mathbf{X}] = \sigma^2\mathbf{I}.$$

WLS estimator is equal to:

$$\begin{aligned}\widehat{\beta}_{\mathrm{WLS}} &= (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{z} \\ &= ((\mathbf{W}^{1/2}\mathbf{X})^T\mathbf{W}^{1/2}\mathbf{X})^{-1}(\mathbf{W}^{1/2}\mathbf{X})^T\mathbf{z} \\ &= (\mathbf{X}^T\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{y} \\ &= (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{y}.\end{aligned}$$

MULTIVARIATE COEFFICIENT ESTIMATORS

And the vector of observations and variables is as follows:

$$\mathbf{M} = \begin{vmatrix} \sqrt{w_1} & \sqrt{w_1}x_{11} & \dots & \sqrt{w_1}x_{1p} \\ \sqrt{w_2} & \sqrt{w_2}x_{21} & \dots & \sqrt{w_2}x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{w_n} & \sqrt{w_n}x_{n1} & \dots & \sqrt{w_n}x_{np} \end{vmatrix}, \qquad \text{and} \qquad \mathbf{z} = \begin{vmatrix} \sqrt{w_1}y_1 \\ \sqrt{w_2}y_2 \\ \vdots \\ \sqrt{w_n}y_n \end{vmatrix}$$

The regression problem simply involves finding the WLS fitted values $\widehat{\mathbf{z}} := \mathbf{M}\widehat{\beta}$.

Note: The first assumption of regression ($\mathbb{E}\left[\widehat{\beta}_{WLS}|\mathbf{x}\right] = \beta$)was established and with these transformations the problem of variance changes and the second assumption is established. Under these two assumptions, according to Gauss-Markov theorem, this estimator is also BLUE.

FINDING WEIGHTS

- If the i-th response is an average of n_i equal to the observed variables.,then:

$$\text{Var}(y_i) = \sigma^2/n_i, \text{ and } w_i = n_i.$$

- If y_i is the sum of n_i observations, then:

$$\text{Var}(y_i) = n_i\sigma^2, \text{ and } w_i = 1/n_i \ .$$

- If the variance is proportional to some of the predictor variables xi, then:

$$\text{Var}(y_i) = x_i\sigma^2, \text{ then } w_i = 1/x_i.$$

REFERENCES

- https://towardsdatascience.com/when-and-how-to-use-weighted-least-squares-wls-models-a68808b1a89d
- http://www.statpower.net/Content/313/Lecture%20Notes/WLS.pdf
- http://www.philender.com/courses/linearmodels/notes3/wls.html
- linear regression analysis theory and computing(Xin Yan, Xiaogang Su)
- https://www.stat.cmu.edu/~cshalizi/mreg/15/lectures/24/lecture-24--25.pdf
- https://nptel.ac.in/content/storage2/courses/111104074/Module7/Lecture25.pdf
- https://www.stat.cmu.edu/~cshalizi/350/lectures/18/lecture-18.pdf
- http://www.eviews.com/help/helpintro.html#page/content/Regress2-Weighted_Least_Squares.html
- Application of Weighted Least Squares Regression in Forecasting.(Sulaimon Mutiu O.

Department of Statistics & Mathematics Moshood Abiola Polytechnic, Abeokuta, Ogun State, Nigeria)