STAT 3008 Applied Regression Analysis
Tutorial 7.

XU Yongze & DONG Fangyuan

Nov 4 & 5 & 7, 2013

Contents

1 Weighted Least Square

1.1 Model introduction

In chapter 5, the assumption that errors have equal variance is relaxed.

$$\text{Var}(y_i|\mathbf{x_i}) = \text{Var}(e_i) = \frac{\sigma^2}{w_i} \tag{1.1}$$

Where i = 1, $\cdots$, n

The model is:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\beta + \mathbf{e} \\ E(\mathbf{e}) &= \mathbf{0} \\ \text{Var}(\mathbf{e}) &= \sigma^2\mathbf{W}^{-1} \end{aligned} \tag{1.2}$$

$$\text{Where} \quad \mathbf{W} = \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{pmatrix} \quad \text{and} \quad \mathbf{W}^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{w_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{w_n} \end{pmatrix}$$

1.2 Least Square Estimation

- Standardized by the variance of each observation

$$\begin{aligned} \text{RSS}(\beta) &= \sum \frac{(y_i\text{-}\hat{y}_i)^2}{Var(y_i)} = \frac{1}{\sigma^2}\sum w_i(y_i\text{-}\hat{y}_i)^2 \\ &\propto \sum w_i(y_i\text{-}\hat{y}_i)^2 \\ &= \begin{pmatrix} y_1\text{-}\hat{y}_1 & \cdots & y_n\text{-}\hat{y}_n \end{pmatrix} \begin{pmatrix} w_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_n \end{pmatrix} \begin{pmatrix} y_1\text{-}\hat{y}_1 \\ \vdots \\ y_n\text{-}\hat{y}_n \end{pmatrix} \\ &= (Y - X\beta)'W(Y - X\beta) \end{aligned}$$

How is RSS defined? RSS in this model is changed.

$$\begin{aligned} RSS(\beta) &= (\mathbf{Y} - \mathbf{X}\beta)^t\mathbf{W}(\mathbf{Y} - \mathbf{X}\beta) \\ &= \sum w_i(y_i - \mathbf{x}_i^t\beta) \\ &= \mathbf{Y}^t\mathbf{W}\mathbf{Y} - \beta^t\mathbf{X}^t\mathbf{W}\mathbf{Y} - \mathbf{Y}^t\mathbf{W}\mathbf{X}\beta + \beta^t\mathbf{X}^t\mathbf{W}\mathbf{X}\beta \end{aligned}$$

Solve $\frac{\partial RSS}{\partial \beta} = \mathbf{0}$
we have $\hat{\beta} = (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{WY}$.

♣ $E(\hat{\beta}) = (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{WX}\beta = \beta$.

♣ $Var(\hat{\beta}) = (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{W}(\sigma^2\mathbf{W}^{-1})\mathbf{WX}(\mathbf{X^t W X})^{-1} = \sigma^2(\mathbf{X^t W X})^{-1}$.

1.3 Model transformation

Suppose $\mathbf{W} = \mathbf{W^{\frac{1}{2}}W^{\frac{1}{2}}}$

$$\mathbf{W^{\frac{1}{2}}} = \begin{pmatrix} w_1^{\frac{1}{2}} & 0 & \cdots & 0 \\ 0 & w_2^{\frac{1}{2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n^{\frac{1}{2}} \end{pmatrix}$$

Clearly, $\mathbf{W^{-1}} = \mathbf{W^{-\frac{1}{2}}W^{-\frac{1}{2}}}$

Then, we can derive
$\mathrm{Var}(\mathbf{W}^{\frac{1}{2}}\mathbf{e}) = \sigma^2\mathbf{I}$

Multiply $\mathbf{W}^{\frac{1}{2}}$ to the regression model (1.2):

$$\mathbf{W}^{\frac{1}{2}}\mathbf{Y} = \mathbf{W}^{\frac{1}{2}}\mathbf{X}\beta + \mathbf{W}^{\frac{1}{2}}\mathbf{e}$$

Let $\mathbf{Z} = \mathbf{W}^{\frac{1}{2}}\mathbf{Y}, \quad \mathbf{M} = \mathbf{W}^{\frac{1}{2}}\mathbf{X}, \quad \mathbf{d} = \mathbf{W}^{\frac{1}{2}}\mathbf{e}$

Now we have the regular multiple linear regression model:

$$\mathbf{Z} = \mathbf{M}\beta + \mathbf{d}$$

Using Ordinary Least Square,

$$\begin{aligned} RSS(\beta) &= (Z - M\beta)'(Z - M\beta) \\ &= (W^{\frac{1}{2}}Y - W^{\frac{1}{2}}X\beta)'(W^{\frac{1}{2}}Y - W^{\frac{1}{2}}X\beta) \\ &= (\mathbf{Y} - \mathbf{X}\beta)^t\mathbf{W}(\mathbf{Y} - \mathbf{X}\beta) \end{aligned}$$

We have $\hat{\beta} = (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{WY}$

♣ $E(\hat{\beta}) = (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{WX}\beta = \beta$.

♣ $$\begin{aligned} Var(\hat{\beta}) &= (\mathbf{X^t W X})^{-1}\mathbf{X}^t\mathbf{W}(\sigma^2\mathbf{W}^{-1})\mathbf{WX}(\mathbf{X^t W X})^{-1} \\ &= \sigma^2(\mathbf{X^t W X})^{-1} \\ &= \sigma^2(\mathbf{X^t W^{\frac{1}{2}} W^{\frac{1}{2}} X})^{-1} \\ &= \sigma^2(\mathbf{M'M})^{-1} \end{aligned}$$

1.4 WLS summary

All you need to know

- Estimators
  $$\hat{\beta}=(M'M)^{-1}M'Z=(X'WX)^{-1}X'WY$$
  $$\sigma^2=RSS/(n-p-1)=(Z-M\hat{\beta})'(Z-M\hat{\beta})/(n-p-1)$$
- Distribution of estimators
  $$\hat{\beta}\sim N(\beta,\sigma^2(M'M)^{-1}) \qquad \frac{(n-p-1)\hat{\sigma}^2}{\sigma^2}=\chi^2_{n-p-1}$$
- F-test
- T-test
  $$T=\frac{\hat{\beta}_k}{sd(\hat{\beta}_k)}\sim t(n-p-1) \qquad F=\frac{(RSS_{NH}-RSS_{AH})/(df_{NH}-df_{AH})}{RSS_{AH}/df_{AH}}=\sim F(df_{NH}-df_{AH},df_{AH})$$
- Prediction Interval.
  $$\hat{y}_{\cdot}\pm t\left(\frac{\alpha}{2},n-p-1\right)\hat{\sigma}\sqrt{1+\mathbf{x}_{\cdot}'(M'M)^{-1}\mathbf{x}_{\cdot}}, \quad \hat{y}_{\cdot}=\mathbf{x}_{\cdot}\hat{\beta}$$
- C.I. for fitted value
  $$\hat{y}\pm t\left(\frac{\alpha}{2},n-p-1\right)\hat{\sigma}\sqrt{\mathbf{x}'(M'M)^{-1}\mathbf{x}}, \quad \hat{y}=\mathbf{x}\hat{\beta}$$
- C.B. for fitted value
  $$\hat{y}\pm\sqrt{(p+1)F(\alpha,p+1,n-p-1)}\hat{\sigma}\sqrt{\mathbf{x}'(M'M)^{-1}\mathbf{x}}$$
- Confidence Ellipse
  $$\frac{(\hat{\beta}-\beta)'(M'M)(\hat{\beta}-\beta)}{(p+1)\hat{\sigma}^2}\le F(\alpha,p+1,n-p-1)$$

What is equivalent? Transforming WLS to OLS v.s. WLS

Estimators;

Distribution of estimators;

F-test;

(remark: $H_0: Z=\beta_0\widetilde{int}+e \leftrightarrow H_A: Z=M\beta+e$ v.s. $H_0: y=\beta_0+e \leftrightarrow H_A: y=\beta_0+\beta_1 x$ under weight)

T-test;

Confidence Ellipse for coefficents.

1.5 WLS v.s OLS

OLS: minimizing $RSS(\beta)=(Y-X\beta)'(Y-X\beta)$

WLS: minimizing $RSS(\beta)=(Y-X\beta)'W(Y-X\beta)$, where $Var(\mathbf{e})=\sigma^2W^{-1}$.

1.6 Determine w_i

- The i-th Observations is an average of n_i variables
  $$Y_i=\frac{z_{i1}+z_{i2}+\dots+z_{in_i}}{n_i}, \quad Var(Y_i\mid X=x_i)=\frac{Var(z_{i1}\mid X=x_i)}{n_i}=\frac{\sigma^2}{n_i}\Rightarrow w_i=n_i$$
- The i-th Observations is a sum of n_i variables
  $$Y_i=z_{i1}+z_{i2}+\dots+z_{in_i}, \quad Var(Y_i\mid X=x_i)=n_iVar(z_{i1}\mid X=x_i)=n_i\sigma^2\Rightarrow w_i=1/n_i$$

2 Exercise

Consider:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad \text{where } \mathbf{e} \sim N(\mathbf{0}, \sigma^{\mathbf{2}}\mathbf{I})$$

Let $\mathbf{X}$ be n $\times$ p, and let y_i denote the elements of y. Also assume $\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}}$

(a) Calculate $\sum_{i=1}^{n} \hat{y}_i(y_i - \hat{y}_i)$

(b) Calculate $\sum_{i=1}^{n} \mathrm{Var}(\hat{y}_i)$

(a)

$$
\begin{aligned}
&\sum_{i=1}^{n} \hat{y}_i(y_i - \hat{y}_i) \\
&= \hat{Y}'(Y - \hat{Y}) \\
&= \hat{Y}'Y - \hat{Y}'\hat{Y} \\
&= Y'X(X'X)^{-1}X'Y - Y'X(X'X)^{-1}X'X(X'X)^{-1}X'Y \\
&= Y'X(X'X)^{-1}X'Y - Y'X(X'X)^{-1}X'Y \\
&= \mathbf{0}.
\end{aligned}
$$

(b)

$$
\begin{aligned}
&\sum_{i=1}^{n} Var(\hat{y}_i) \\
&= tr[Var(\hat{Y})] \\
&= tr[Var(X(X'X)^{-1}X'Y)] \\
&= tr[X(X'X)^{-1}X'\sigma^2 IX(X'X)^{-1}X'] \\
&= tr[\sigma^2 X(X'X)^{-1}X'] \\
&= \sigma^2 tr[(X'X)^{-1}X'X] \\
&= \sigma^2 tr[I_{p\times p}] \\
&= \sigma^2 p.
\end{aligned}
$$