Chapter 11 Remedial Measures

When an appropriate regression relationship has been found but errors are non-constant we can use weighted least squares

ie: We have model: $Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_{p-1} X_{i,p-1} + \epsilon_i \quad \ni \epsilon_i \sim N(0, \sigma_i^2)$

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

Error Variance Known

Let $W_{n\times n} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}$ $\qquad w_i = \frac{1}{\sigma_i^2}$

Then the normal equation: $(\underline{X}'\underline{W}\underline{X})\underline{b}_w = \underline{X}'\underline{W}\underline{Y}$

Thus

$$\underset{p\times 1}{\underline{b}_w} = (\underline{X}\underline{W}\underline{X})^{-1}\underline{X}'\underline{W}\underline{Y}$$

$$\sigma^2\{\underline{b}_w\} = (\underline{X}'\underline{W}\underline{X})^{-1}$$

Error Variance Known up to a Proportionality constant

- Relative magnitudes of the variances are known

$$w_i = k\left(\frac{1}{\sigma^2}\right)$$

$$\sigma^2\{b_{w_{p\times p}}\} = k(\underline{X}'\underline{W}\underline{X})^{-1}$$

$$s^2\{b_{w_{p\times p}}\} = MSE_w(\underline{X}'\underline{W}\underline{X})^{-1}$$

$$MSE_w = \frac{\Sigma w_i (Y_i - \hat{Y}_i)^2}{n-p} = \frac{\Sigma w_i e_i^2}{n-p}$$

Unknown Variance $\quad \sigma_i^2 = E(\epsilon_i^2) - (E(\epsilon_i))^2 = (E(\epsilon_i))^2$

$w_i = \frac{1}{v_i} = 1/(\hat{s}_i)^2$ where $\hat{s}_i$ is fitted value from st. dev function

$$b_w = (\underline{X}'\underline{W}\underline{X})^{-1}X'WY$$

1) Fit the regression model w/o least squares
2) Estimate variance or st deviation by regressing either the squared or the absolute residuals on the appropriate predictors
   - A residual plot against X_1 exhibits a megaphone shape. Regress the absolute residuals against X_1
   - A residual plot against $\hat{Y}$ exhibits a megaphone shape. Regress the absolute residuals against $\hat{Y}$
   - A plot of the squared residuals against X_3 exhibits an upward tendency. Regress the squared residuals against X_3
   - A plot of the residuals against X_2 suggests that the variance increases rapidly w/ increases in X_2 up to a point an then more slowly. Regress absolute residuals against X_2 and X_2^2
3) Use the fitted values from the estimated variance or st. deviation function to obtain weights w_i
4) Estimate the regression coefficient using these weights

$\boxed{SAS}$

①
```
proc reg;
  model y= x1;
  output out=temp r=residual;
  plot y*x r.*age;
run;
```

②
```
data temp;
  set temp;
  absr = abs(residual);
run;
proc sgscatter;
  plot absr*age;
run;
proc reg;
model absr= age;
output out=templ p=s_hat;
run;
```

③
```
data templ;
  set templ;
  w = 1/(s_hat**2); run;
```

④
```
proc reg data=templ;
  weight w;
  model y=x/clb;
  output out= temp r=residual;
  plot y*x r.*age;
run;
```

Multicollinearity Remedial Measures

1) Restrict the use of data to predictor variables that follow the same multicollinearity
2) We can center data $x_i = X_i - \bar{X}$
3) One or several predictor variables may be dropped from the model
4) Add some cases that break the pattern of multicollinearity

Ridge Regression Modify OLS to allow biased estimators in cases where a little bias lets us be much more precise

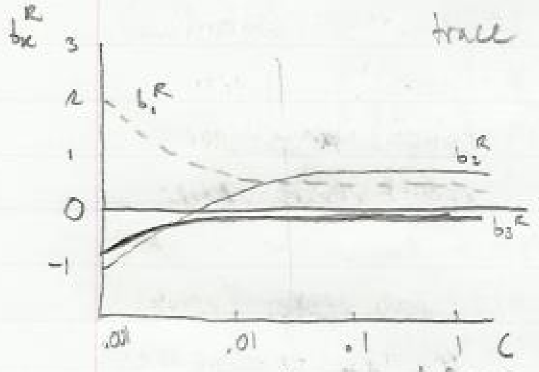
r_{XX} = correlation matrix of X variables

r_{YX} = vector of coefficients of correlation between Y and each X

$$(\underline{r}_{XX} + c\,\underline{I})\underline{b}^R = \underline{r}_{YX}$$

$$\underset{(p-1)\times 1}{\underline{b}^R} = \begin{bmatrix} b_1^R \\ b_2^R \\ \vdots \\ b_{p-1}^R \end{bmatrix} = (r_{XX} + cI)^{-1} r_{YX}$$

Choice of c: · optimal value varies across applications
· commonly used method uses VIF and ridge regression trace



.02 is used for c as the graphs level out i.e. changes less

SAS
```
proc reg outest=ridge outvif ridge=.01 to .5 by .01;
    model y = x1 x2 x3;
    plot / ridgeplot;
run;
* choose c
proc reg outest=ridge ridge=.2
    model y=x1 x2 x3;
run;
proc print data=ridge; run;
```

Remedial Measures for Influential Cases - Robust Regression

- Robust Regression dampens the influence of outlying cases compared to OLS
- For use when there is no time for rigerous identification of outlying cases

L_1 Regression → Median, LAR or LAD Regression

- Insensitive to both outlying data and inadequacies of the model employed
- minimizes the sum of absolute deviations

SAS
```
proc quantreg;
    model y=x / quantile=.25;
    output out=o1 p=p1; run;
proc quantreg;
    model y=x / quantile=.5;
    output out=o2 p=p2; run;
proc quant reg;
    model y=x / quantile=.75;
    output out=o3 p=p3; run;
data o4; set o1 o2 o3; proc sort data=o4;
by quantile eject;
```

```
goptions reset=all;
symbol1 color=black value=dot interpol=join;
symbol2 color=black value=none l=3 interpol=join;
symbol3 color=black value=none l=1 interpol=join;
symbol4 color=black value=none l=3 interpol=join;
legend1 value=("data" "25%" "50%" "75%");
proc gplot data=o4;
    plot y*x p1*x p2*x p3*x / overlay
        legend=legend1;
```

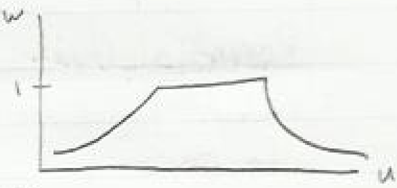
IRLS Robust Regression

- Uses weighted least squares to dampen the influence of outlying observations.

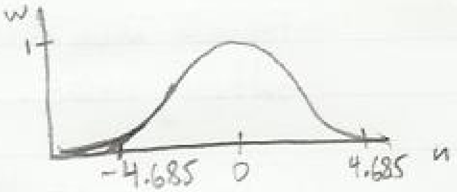
1) Choose a weight function for weighting the cases
2) Obtain starting weights for all cases
3) Use the starting weights in the weighted least squares and obtain the residuals from the fitted regression function
4) Use the residuals in step 3 to obtain revised weights
5) Continue the iterations until convergence is obtained

Weight functions

$$\ast \text{ Huber: } w = \begin{cases} 1 & |u| \le 1.345 \\ \dfrac{1.345}{|u|} & |u| > 1.345 \end{cases}$$



$$\text{Bisquare: } w = \begin{cases} \left[1-\left(\dfrac{u}{4.685}\right)^2\right]^2 & |u| \le 4.685 \\ 0 & |u| > 4.685 \end{cases}$$



SAS

```
proc robustreg method=m (wf=huber);
    model  y= x1  x2  x3;
    test   x2  x3 ;
    id     x4;
    output out=out p=p  sr=sr;
run;
goptions reset=all;
symbol1 v=dot pointlabel=("# x4")
proc gplot;
plot sr*(p x1 x2 x3);
run;
```