I. Some inference

i. OLS

We have known that the estimator for OLS can be represented in matrix form as

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

ii. Weighted Least Squares

In WLS, the loss function should be

$$J(\beta) = \ W\|Y - X\beta\|^2 = (Y - X\beta)^T W(Y - X\beta)$$

FOC: $\frac{\partial J(\beta)}{\partial \beta} = \ -2X^T WY + 2X^T WX\beta$=0.

Solution:

$$\hat{\beta} = (X^T WX)^{-1} X^T WY$$

iii. Analytical solution for restricted Least Squared.

Ruud (2000) introduced an analytical solution to least squares with linear restriction which genuinely is a linear transformation of the restriction.

With regards to the restrictions, β is expressed as

$$\beta = S\gamma + s$$

where β is a K*1 vector, S is a known K*M matrix, s is a known K*1 vector, γ is a unknown M*1 vector.

Ruud (2000) classified linear restrictions into exclusion restrictions, equality restrictions and simple linear restrictions.

To demonstrate different restrictions, suppose a regression model with 5 variables:

$$X\beta = \ \beta_1 + \ X_2\beta_2 + \ X_3\beta_3 + \ X_4\beta_4 + \ X_5\beta_5$$

For simple linear restriction, if restriction is like $\beta_3 \ + \ \beta_4 + \ \beta_5 = 1$, then

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = S\gamma + s$$

Proof:

$$y = X\beta + \epsilon$$

with simple linear restriction

$$\beta = S\gamma + s$$

We have

$$y = X\beta + \epsilon = X(S\gamma + s) + \epsilon = \ XS\gamma + Xs + \epsilon$$

$$y - Xs = XS\gamma + \epsilon$$

Assume $y_R = y - Xs$, $X_R = \ XS$, then $y_R = \ X_R\gamma + \epsilon$ become OLS.

$$\hat{\gamma} = (X_R^T X_R)^{-1} X_R^T Y_R = [(XS)^T XS]^{-1}(XS)^T(y - XS) = \ (S^T X^T XS)^{-1} S^T X^T (y - Xs)$$

$$\hat{\beta} = S\hat{\gamma} + s = S(S^T X^T XS)^{-1} S^T X^T (y - Xs) + s$$

iv. Weighted Least Squares with restriction

The loss function becomes

$$\begin{aligned} J(\beta) = &\ W\|Y_R - X_R\gamma\|^2 = (Y_R - X_R\gamma)^T W(Y_R - X_R\gamma) \\ = &\ Y_R^T WY_R - \gamma^T X_R^T WY_R - Y_R^T WX_R\gamma + \gamma^T X_R^T WX_R\gamma \end{aligned}$$

FOC: $\frac{\partial J(\beta)}{\partial \beta} = \ -2X_R^T W Y_R + 2X_R^T W X_R \beta\ = 0.$

Then

$$\hat{\gamma} = (X_R^T W X_R)^{-1} X_R^T W Y_R = (S^T X^T W X S)^{-1} S^T X^T W (y - Xs)$$

Thus,

$$\hat{\beta} = S\hat{\gamma} + s = S(S^T X^T W X S)^{-1} S^T X^T W (y - Xs) + s$$

II. Barra Model with industry neutralization restriction

We use Ricequant's factor exposures(10 risk factors, 1 comovement, and Shenwan first-level industry classification)

(i) Restriction

$$\text{s.t.} \sum_{n=1}^{N} \sum_{i} w_n X_{ni} f_i = 0$$

w_n: 第n只股票的市值权重

r_n: 第n只股票的收益率

f_c: 国家因子收益率

X_{ni}: 第n只股票在第i个行业因子上的因子暴露（0 或 1）

f_i: 第i个行业因子的收益率

X_{ns}: 第n只股票在第s个风格因子上的因子暴露

f_s: 第s个风格因子的收益率

(ii) Building restriction matrix S with ricequant factor matrix

The restriction is a simple linear restriction as is shown above. If we use the first industry as a intermediate to build the matrix S like the red box in the example, then the 12th column where the 1st industry situated would be drop finally, while to the right of which would be filled with figures of other industries.

Here, s is a 0 vector for the right side of the equation is 0.

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = S\gamma + s$$