Robust Error

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ik}\beta_k + \varepsilon_i = x_i'\beta + \varepsilon_i$$

OLS 估计量

$$\hat{\beta} = (X'X)^{-1}X'y, \quad X = \begin{array}{cc} \multicolumn{1}{c}{\text{k个变量}\rightarrow} & \\ \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ & \ddots & \\ x_{n1} & \cdots & x_{nk} \end{bmatrix} & \downarrow \text{n个观测值} \end{array} \qquad X = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

$$X'X = [x_1 \cdots x_n]\begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = \sum_{i=1}^n x_i x_i'$$

$$\hat{\beta} = \left(\sum_{i=1}^n \vec{x}_i \vec{x}_i'\right)^{-1}\left(\sum_{i=1}^n \vec{x}_i y_i\right) = \left(\frac{1}{n}\sum_{i=1}^n \vec{x}_i \vec{x}_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \vec{x}_i y_i\right)$$

$$= \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i(x_i'\beta + \varepsilon_i)\right)$$

$$= \beta + \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i \varepsilon_i\right)$$

小样本时

$$\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$$

$$E(\hat{\beta} - \beta) = E((X'X)^{-1}X'\varepsilon)$$

此时要有无偏性则需 $E(\varepsilon|X) = 0$

大样本理论 可以拆开是由于Slusky's Theorem

$$\xrightarrow{p} \beta + E(x_i x_i')^{-1}E(x_i\varepsilon_i)$$

大样本理论假设 $E(x_i\varepsilon_i) = 0$，x_i和ε_i同期

也即 contemporaneous exogeneity

$\hat{\beta}$的方差显然，也依概率收敛于0.

所以讨论的是$\sqrt{n}$统计量

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\underbrace{\left(\sqrt{n}\,\frac{1}{n}\sum_{i=1}^n x_i\varepsilon_i\right)}_{\equiv \sqrt{n}\bar{g}}$$

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, Avar(\hat{\beta}))$$

这里假设了中间这个S存在，同时 $\sqrt{n}\bar{g} \sim N(0, S)$

$$Avar(\hat{\beta}) = E(x_i x_i')^{-1}\underbrace{E(\varepsilon_i^2 x_i x_i')}_{\equiv S}E(x_i x_i')^{-1}$$

此即为"夹心"估计量，robust variance.

如果是同方差也即 $E(\varepsilon_i^2|x_i) = \sigma^2$ ←没有下标，则

$$E(\varepsilon_i^2 x_i x_i') = E\left[E[\varepsilon_i^2 x_i x_i' | x_i]\right] = E\left[E[\varepsilon_i^2|x_i]\,x_i x_i'\right] = \sigma^2 E[x_i x_i']$$

此时 $Avar(\hat{\beta}) = \sigma^2 E(x_i x_i')^{-1}$，也即一般OLS的方差.

Robust 的意思：没有假设特别的形式，都能适用.

注意：上面的 Robust 方差仍然假设了无自相关！

因为 $\sqrt{n}\bar{g}$ 的方差被写成了 $S \equiv E(\varepsilon_i^2 x_i x_i')$，只有同期的 i 出现

Newey-West Adjustment for Heteroscedasticity & Autocorrelation (HAC)

问题的关键是 $Var(\sqrt{n}\bar{g})$. 若有序列相关，则 $E[g_i g_{i-j}] \neq 0$, $g_i = x_i \varepsilon_i$

$$cov(g_i, g_{i+j}) \equiv \Gamma_j, \quad cov(g_{i+j}, g_i) = \Gamma_{-j}, \quad Var(g_i) = \Gamma_0$$

注意有：$\Gamma_j = \Gamma_{-j}'$. 因为 $cov(g_i, g_{i+j}) = cov\left(\begin{bmatrix} g_{i1} \\ \vdots \\ g_{iK} \end{bmatrix}, \begin{bmatrix} g_{i+j,1} \\ \vdots \\ g_{i+j,K} \end{bmatrix}\right)$

$cov(g_{i+j1}, g_{i1})$ $\quad cov(g_{i+j1}, g_{i2})$

$$= \begin{bmatrix} cov(g_{i1}, g_{i+j1}), & cov(g_{i1}, g_{i+j2}) & \cdots & cov(g_{i1}, g_{i+jK}) \\ cov(g_{i2}, g_{i+j1}), & cov(g_{i2}, g_{i+j2}) & \cdots & cov(g_{i2}, g_{i+jK}) \\ & & \ddots & \\ cov(g_{iK}, g_{i+j1}) & & \cdots & cov(g_{iK}, g_{i+jK}) \end{bmatrix}$$

$$Var(\sqrt{n}\bar{g}) = Var\left(\sqrt{n}\frac{1}{n}\sum_{i=1}^{n} g_i\right) = \frac{1}{n} Var\left(\sum_{i=1}^{n} g_i\right) = \frac{1}{n} Var(g_1 + \cdots + g_n)$$

$$= \frac{1}{n}\Big(n\Gamma_0 + (n-1)\Gamma_1 + (n-2)\Gamma_2 + \cdots + \Gamma_{n-1}$$
$$+ (n-1)\Gamma_{-1} + (n-2)\Gamma_{-2} + \cdots + \Gamma_{-(n-1)}\Big)$$

当 $n \to \infty$ 时，$\lim\limits_{n\to\infty} Var(\sqrt{n}\bar{g}) = \sum_{j=-\infty}^{\infty} \Gamma_j$

$$= \Gamma_0 + \sum_{j=1}^{\infty} \Gamma_j + \sum_{j=1}^{\infty} \Gamma_j'$$

理想的情况：$n \to \infty$，有无穷多个协方差矩阵，但显然无实际可能.
事实上，在有 Autocorrelation 时的 Central Limit Theorem (Gordin's CLT) 必须假设自相关在足够远处消失，也即 $\Gamma_j \to 0$ when $j \to \infty$.

因此 $\Gamma_j=0$ for some $|j|>p$，p已知，即为滞后阶数，用

$$\hat{S}=\sum_{j=-p}^{p}\hat{\Gamma}_j=\hat{\Gamma}_0+\sum_{j=1}^{p}\hat{\Gamma}_j+\hat{\Gamma}_j'$$

(Hansen-White (1982)) 但这个估计量可能不半正定.

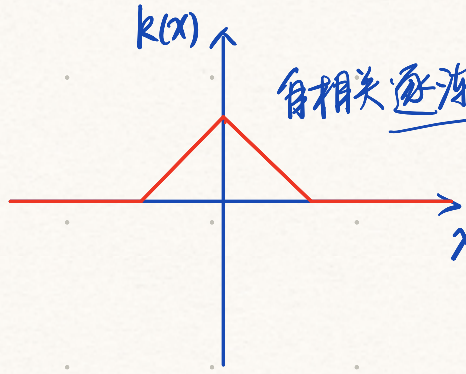
这时候用 kernel-based estimator (加权重)

$$\hat{S}=\sum_{j=-(n-1)}^{n-1}k\left(\frac{j}{p(n)}\right)\hat{\Gamma}_j$$

上面的统计量可看成是 $p(n)=p$，$k(x)=\begin{cases}1, & \text{if } |x|\le 1\\ 0, & \text{if } |x|>1\end{cases}$

Newey-West:

$$k(x)=\begin{cases}1-|x|, & |x|\le 1\\ 0, & |x|>1\end{cases}$$



自相关逐渐消失.

例: $p(n)=3$

$$\hat{S}=\sum_{j=-(n-1)}^{n-1}k\left(\frac{j}{3}\right)\hat{\Gamma}_j=k\left(-\frac{2}{3}\right)\hat{\Gamma}_{-2}+k\left(-\frac{1}{3}\right)\hat{\Gamma}_{-1}+k(0)\hat{\Gamma}_0+k\left(\frac{1}{3}\right)\hat{\Gamma}_1+k\left(\frac{2}{3}\right)\hat{\Gamma}_2$$

$$=\hat{\Gamma}_0+\frac{2}{3}\left(\hat{\Gamma}_1+\hat{\Gamma}_1'\right)+\frac{1}{3}\left(\hat{\Gamma}_2+\hat{\Gamma}_2'\right)$$

写成分量形式:

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2x_ix_i'+\frac{1}{n}\sum_{l=1}^{L}\sum_{i=l+1}^{n}w_t\hat{\varepsilon}_i\hat{\varepsilon}_{i-l}\left(x_ix_{i-l}'+x_{i-l}x_i'\right)$$