

Robust Error

$$y_i = x_{i1}\beta_1 + \dots + x_{iK}\beta_K + \varepsilon_i = \vec{x}_i' \beta + \varepsilon_i, \quad \vec{x}_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iK} \end{bmatrix}$$

$i=1, \dots, N$ 观测值, $k=1, \dots, K$ 个自变量.

$$y = X\beta, \quad X = \begin{bmatrix} x_{11} & \dots & x_{1K} \\ \vdots & & \vdots \\ x_{N1} & \dots & x_{NK} \end{bmatrix} \xrightarrow[N]{\substack{\downarrow \\ K}} \begin{bmatrix} \vec{x}_1' \\ \vdots \\ \vec{x}_N' \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}$$

OLS

$$\min_{\hat{\beta}} (y - X\hat{\beta})'(y - X\hat{\beta}) \Rightarrow \hat{\beta} = (X'X)^{-1}X'y$$

$$X'X = [\vec{x}_1 \dots \vec{x}_N] \begin{bmatrix} \vec{x}_1' \\ \vdots \\ \vec{x}_N' \end{bmatrix} = \sum_{i=1}^N \vec{x}_i \vec{x}_i'$$

$$\begin{aligned} \hat{\beta} &= \left(\sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\sum_{i=1}^N \vec{x}_i y_i \right) = \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i y_i \right) \\ &= \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i (\vec{x}_i' \beta + \varepsilon_i) \right) \\ &= \beta + \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \varepsilon_i \right) \end{aligned}$$

$$\xrightarrow{N \rightarrow \infty} \beta + E(\vec{x}_i \vec{x}_i')^{-1} E(\vec{x}_i \varepsilon_i)$$

$$\sqrt{N} \text{ 统计量: } \sqrt{N}(\hat{\beta} - \beta) = \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\sqrt{N} \frac{1}{N} \sum_{i=1}^N \vec{x}_i \varepsilon_i \right) \equiv \sqrt{N} \frac{1}{N} \sum_{i=1}^N g_i = \sqrt{N} \bar{g}$$

$\sqrt{N} \bar{g} \xrightarrow{d} N(0, E(\varepsilon_i^2 \vec{x}_i \vec{x}_i'))$, Central Limit Theorem for Martingale Difference Sequence

$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} N(0, E(\vec{x}_i \vec{x}_i')^{-1} E(\varepsilon_i^2 \vec{x}_i \vec{x}_i') E(\vec{x}_i \vec{x}_i')^{-1})$ Robust variance.

$$\widehat{\text{Avar}}(\hat{\beta}) = \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' (y_i - \vec{x}_i' \hat{\beta})^2 \right) \left(\frac{1}{N} \sum_{i=1}^N \vec{x}_i \vec{x}_i' \right)^{-1} \quad \textcircled{1}$$

同方差:

$$E[\varepsilon_i^2 | \vec{x}_i] = \sigma^2.$$

$$E[\varepsilon_i^2 \vec{x}_i \vec{x}_i'] = E[E[\varepsilon_i^2 \vec{x}_i \vec{x}_i' | \vec{x}_i]] = E[E[\varepsilon_i^2 | \vec{x}_i] \vec{x}_i \vec{x}_i'] = \sigma^2 E[\vec{x}_i \vec{x}_i']$$

$$\hat{\sigma}^2 = \frac{1}{N-k} (y - X\hat{\beta})'(y - X\hat{\beta})$$

①式假设了无自相关, $E[\varepsilon_i^2 \vec{x}_i \vec{x}_i']$, 不同的 i, j 之间没有关系.

Newey - West Estimator

$\text{Var}(\sqrt{N} \frac{1}{N} \sum_{i=1}^N g_i)$ 若有自相关, 则 $E[g_i g_{i-j}] \neq 0$, $g_i = x_i \varepsilon_i$.

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

$$\Gamma_j = \Gamma_{-j}, \quad \text{因为 } g_i \text{ 是向量: } \text{cov}(g_i, g_{i+j}) = \text{cov} \left(\begin{bmatrix} g_i^1 \\ \vdots \\ g_i^k \end{bmatrix}, \begin{bmatrix} g_{i+j}^1 \\ \vdots \\ g_{i+j}^k \end{bmatrix} \right)$$
$$= \begin{bmatrix} \text{cov}(g_i^1, g_{i+j}^1), & \text{cov}(g_i^1, g_{i+j}^2), & \dots & \text{cov}(g_i^1, g_{i+j}^k) \\ \text{cov}(g_i^2, g_{i+j}^1), & \text{cov}(g_i^2, g_{i+j}^2), & \dots & \text{cov}(g_i^2, g_{i+j}^k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(g_i^k, g_{i+j}^1), & \text{cov}(g_i^k, g_{i+j}^2), & \dots & \text{cov}(g_i^k, g_{i+j}^k) \end{bmatrix}$$

$$\text{cov}(g_{i+j}, g_i) =$$

$$\begin{bmatrix} \text{cov}(g_{i+j}^1, g_i^1), & \text{cov}(g_{i+j}^1, g_i^2), & \dots & \text{cov}(g_{i+j}^1, g_i^k) \\ \text{cov}(g_{i+j}^2, g_i^1), & \text{cov}(g_{i+j}^2, g_i^2), & \dots & \text{cov}(g_{i+j}^2, g_i^k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(g_{i+j}^k, g_i^1), & \text{cov}(g_{i+j}^k, g_i^2), & \dots & \text{cov}(g_{i+j}^k, g_i^k) \end{bmatrix}$$

$$\begin{aligned} \text{Var}(\sqrt{N}\bar{g}) &= \text{Var}\left(\sqrt{N}\frac{1}{N}\sum_{i=1}^N g_i\right) = \frac{1}{N} \text{Var}\left(\sum_{i=1}^N g_i\right) = \frac{1}{N} \text{Var}(g_1 + \dots + g_N) \\ &= \frac{1}{N} \left(N\Gamma_0 + (N-1)\Gamma_1 + (N-2)\Gamma_2 + \dots + \Gamma_{N-1} \right. \\ &\quad \left. + (N-1)\Gamma_1 + (N-2)\Gamma_2 + \dots + \Gamma_{(N-1)} \right) \end{aligned}$$

$$\lim_{N \rightarrow \infty} \text{Var}(\sqrt{N}\bar{g}) = \sum_{j=-\infty}^{\infty} \Gamma_j = \Gamma_0 + \sum_{j=1}^{\infty} (\Gamma_j + \Gamma_j')$$

Godin's CLT 要求在足够远处自相关消失, 因此有截断位置 p ,

当 $|j| > p$ 时, $\Gamma_j = 0$.

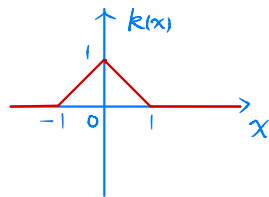
$$\text{有估计量 } \widehat{\text{Avar}}(\sqrt{N}\bar{g}) = \hat{\Gamma}_0 + \sum_{j=1}^p (\hat{\Gamma}_j + \hat{\Gamma}_j') \quad \left(\begin{array}{l} \text{Hansen \& Singleton (1982)} \\ \text{West (1986)} \end{array} \right)$$

在有限样本下不一定半正定. 调整方法: 加权重.

$$\widehat{\text{Avar}}(\sqrt{N}\bar{g}) = \sum_{j=-(p-1)}^{p-1} k\left(\frac{j}{p(N)}\right) \hat{\Gamma}_j \quad \text{若 } k(x) = \begin{cases} 1, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases} \text{ 则还原.}$$

Newey-West:

$$k(x) = \begin{cases} 1 - |x|, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}, \text{ Bartlett type kernel.}$$



例: $p(N) = 3$.

$$\begin{aligned} \widehat{\text{Avar}}(\sqrt{N}\bar{g}) &= \sum_{j=-(p-1)}^{p-1} k\left(\frac{j}{3}\right) \hat{\Gamma}_j = k\left(-\frac{2}{3}\right) \hat{\Gamma}_{-2} + k\left(-\frac{1}{3}\right) \hat{\Gamma}_{-1} + k(0) \hat{\Gamma}_0 \\ &\quad + k\left(\frac{1}{3}\right) \hat{\Gamma}_1 + k\left(\frac{2}{3}\right) \hat{\Gamma}_2 \\ &= \hat{\Gamma}_0 + \frac{2}{3} (\hat{\Gamma}_1 + \hat{\Gamma}_1') + \frac{1}{3} (\hat{\Gamma}_2 + \hat{\Gamma}_2') \end{aligned}$$

分量形式:

$$\frac{1}{N} \sum_{i=1}^N \hat{\varepsilon}_i^2 \vec{x}_i \vec{x}_i' + \frac{1}{N} \sum_{l=1}^p \sum_{i=l+1}^N k(l) \hat{\varepsilon}_i \hat{\varepsilon}_{i-l} (\vec{x}_i \vec{x}_{i-l}' + \vec{x}_{i-l} \vec{x}_i')$$