

ALTERNATIVNÍ POSTUPY K OLS / MNČ

Kromě odhadové metody OLS lze uplatnit v podstatě tyto alternativní postupy

- metodu maximální věrohodnosti (pro aplikaci nutno specifikovat hustotu rozdělení náhodných složek)
- metodu momentů (využívá se znalosti empirických momentů modelových veličin a porovnání s vypočtenými teoretickými hodnotami)
- metodu nejmenších absolutních odchylek (LAD, LAR, LAE, MAD) - uplatňuje se kritérium minimalizace součtu absolutních hodnot odchylek pozorovaných od vyrovnaných hodnot.

metoda LAD v jednorovnicovém modelu

kritérium LAD je v jistém směru přirozenější než OLS : původní formulace snad již Galton a Edgeworth (1888), podstatně nižší frekvence jejího uplatnění vyplývá z toho, že odhad parametrů nelze zde vyjádřit v explicitním tvaru, takže je nutno uplatnit iterační postupy.

minimalizační kritérium:

$$(1) \quad \text{Min} \left[\sum_{t=1}^T |y_t - \hat{y}_t| \right] = \text{Min} \left[\sum_{t=1}^T \left| y_t - \sum_{j=1}^k X_{tj} \hat{\beta}_j \right| \right]$$

Úlohou je nalézt hodnoty strukturních parametrů, při kterých je výraz (1) minimalizován.

Součet absolutních hodnot reziduí lze psát jako výraz/kritérium odpovídající vážené metodě nejmenších čtverců WLS, kde za váhy bereme převrácené hodnoty absolutních odchylek, tedy

$$(2) \quad \sum_{t=1}^T |e_t| = \sum_{t=1}^T \frac{e_t^2}{|e_t|} = \sum_{t=1}^T w_t z_t, \quad \text{kde } w_t = \frac{1}{|e_t|}, \quad z_t = e_t^2.$$

Minimalizace $\sum_{t=1}^T |e_t|$ takto vede k nasazení vážené metody nejmenších čtverců

WLS, přičemž váhy jsou dány převrácenými hodnotami reziduálních hodnot. Výraz pro odhadnuté parametry je dán jako

$$(3) \quad \text{WLS } \hat{\beta} = (\mathbf{X}' \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}' \Sigma^{-1} \mathbf{y} \quad \text{s maticí } \Sigma \text{ definovanou jako}$$

$$\Sigma = \begin{pmatrix} |e_1|^{-1} & 0 & \dots & 0 \\ 0 & |e_2|^{-1} & \dots & 0 \\ \dots & \dots & |e_i|^{-1} & \dots \\ 0 & 0 & \dots & |e_T|^{-1} \end{pmatrix}$$

Protože na počátku neznáme potřebné váhy (jmenovitě v nich obsažené reziduální hodnoty), musíme si je nějakým vhodným způsobem opatřit. Nabízí se k tomu metoda nejmenších čtverců, kterou na počátku uplatníme k získání počátečních reziduálních hodnot.

Postup vedoucí k získání LAD-odhadů parametrů je tedy aplikací vážené metody nejmenších čtverců WLS, při které postupně měníme (zpřesňujeme váhy) \mathbf{W}_t , přičemž jejich počáteční hodnoty jsou $\mathbf{W}_t^{(0)} =_{\text{OLS}} \mathbf{W}_t$. Dále postupujeme iteračně:

krok 0

Spočteme vektor parametrů $_{\text{OLS}} \hat{\beta}$ obyčejnou metodou nejmenších čtverců. Určíme vektor vyrovnaných hodnot $_{\text{OLS}} \hat{y}$ a odvodíme rezidua $_{\text{OLS}} \mathbf{e} = \mathbf{y} -_{\text{OLS}} \hat{y}$. Převrácené hodnoty jeho složek v následujícím užijeme k definici vah $_{\text{LAD}} \mathbf{w}_t^{(0)} = \frac{1}{|_{\text{OLS}} \mathbf{e}_t|}$.

kroky $r = 1, 2, \dots$

Aplikací vážené metody nejmenších čtverců s užitím vzorců

$$_{\text{LAD}_1} \hat{\beta} = (\mathbf{X}' \Sigma_1^{-1} \mathbf{X})^{-1} \mathbf{X}' \Sigma_1^{-1} \mathbf{y}, \quad \Sigma_1 = \text{diag}({}_{\text{LAD}} \mathbf{w}_1^{(0)}, {}_{\text{LAD}} \mathbf{w}_2^{(0)}, \dots, {}_{\text{LAD}} \mathbf{w}_t^{(0)})$$

získáme (první odhad) vektoru parametrů $_{\text{LAD}} \hat{\beta}^{(1)}$ a následně vektor vyrovnaných hodnot $_{\text{LAD}} \hat{y} = \mathbf{X} \cdot_{\text{LAD}} \hat{\beta}^{(1)}$. Opět odvodíme vektor reziduí $_{\text{LAD}} \mathbf{e}^{(1)} = \mathbf{y} -_{\text{LAD}} \hat{y}^{(1)}$, který dosadíme do (2) a pokračujeme váženou metodou nejmenších čtverců dle (3). Máme

$$_{\text{LAD}_2} \hat{\beta} = (\mathbf{X}' \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}' \Sigma^{-1} \mathbf{y} \quad \Sigma_2 = \text{diag}({}_{\text{LAD}} \mathbf{w}_1^{(1)}, {}_{\text{LAD}} \mathbf{w}_2^{(1)}, \dots, {}_{\text{LAD}} \mathbf{w}_t^{(1)})$$

Postup opakujeme do té doby, než se hodnoty parametrů přiblíží na dostatečně těsnou vzdálenost, např. na 10^{-5} nebo 10^{-6} . Zpravidla bereme za míru odchýlení

$$d = \max_i \left| {}_{\text{LAD}} \hat{\beta}_i^{(r+1)} - {}_{\text{LAD}} \hat{\beta}_i^{(r)} \right| \quad \text{nebo} \quad d = \max_i \frac{\left| {}_{\text{LAD}} \hat{\beta}_i^{(r+1)} - {}_{\text{LAD}} \hat{\beta}_i^{(r)} \right|}{\left| {}_{\text{LAD}} \hat{\beta}_i^{(r)} \right|}$$

Ray C. Fair (a nejen on) zjistil, že při tomto postupu stačí nanejvýš 4 iterace, aby se zpřesňované odhady již téměř nelišily. Obvykle dokonce již rozdíly mezi odhady získanými v 2. a 3. iteraci jsou velmi malé¹.

Problém: Dost často se stává, že se velmi brzy některá (j-tá) vyrovnaná hodnota těsně přiblíží pozorované, takže příslušné j-té reziduum $_{\text{LAD}} \mathbf{e}_j^{(s)} = y_j -_{\text{LAD}} \hat{y}_j^{(s)}$ v s-tém kroku bude malinké a váha jemu odpovídající veličině bude naopak obrovská (to je nežádoucí, protože se setře vliv jiných reziduálních hodnot.)

Řešení: Pokud se stane, že v některé iteraci bude hodnota některého rezidua nulová, pak se takové malinké reziduum nahradí nějakou malou hodnotou, např. 10^{-5} .

Uvedený postup nazval R.C.Fair metodou *WLS-I*.

¹ Výsledky lze nalézt v : Taylor, L.D.: Estimation by Minimizing the Sum of Absolute Errors. In Zarembka: Frontiers in Econometrics.

2) Variantní postup může spočívat v tom, že se aplikuje kritérium *OLS pro malá rezidua*, zatímco *LAD pro velká rezidua*. Tzn., že váhy jsou vzaty jako

$$w_t = \frac{1}{|e_t|} \text{ pro } |e_t| \geq h \quad \text{zatímco } w_t = \frac{1}{h} \text{ pro } |e_t| < h$$

Doporučení pro volbu konstanty k je $\frac{m}{0,6745}$, kde m je medián v (seřazených)

absolutních hodnotách reziduí. Jako počáteční hodnoty se doporučuje volit získané hodnoty z procedury *WLS-I*. Konstanta h se přepočítává po každé iteraci znova.² Postup se nazývá metoda *WLS-II*.

3) Ještě další postup představuje užití vah ve tvaru

$$w_i = \begin{cases} 1 - \left(\frac{z_i}{h}\right)^2 & , \text{ jestliže } |z_i| < h \\ w_i = 0 & \text{v opačném případě} \end{cases}$$

Hodnota z_i se volí jako $z_i = \frac{e_i}{p}$, přičemž p bylo voleno jako 6 a h opět jako

$\frac{m}{0,6745}$. Hodnota h byla proměnlivá od iterace k iteraci. Postup byl nazván

metoda *WLS-III*. Pokud byly posuzovány predikce, pak tato metoda dávala z uvedené trojice nejlepší výsledky.

Přestože výpočet pomocí metody *LAD* nepředstavuje žádný větší problém, rozdělení těchto odhadových funkcí nejsou dostatečně známa. Jsou známy podmínky, za kterých jsou konzistentní a nestranné, ale nejsou známy ani přesné tvary výrazů pro směrodatné odchylky parametrů (takže nelze přesně konstruovat intervaly spolehlivosti*)

R.Blattberg a T.Sargent [1971] ukázali, že jestliže se náhodné složky ε_t řídí druhým *Laplaceovým zákonem* (dvoustranné exponenciální rozdělení) s hustotou

$$f(\varepsilon_t) = \frac{1}{2\lambda} \cdot \exp\left(-\frac{|\varepsilon_t|}{\lambda}\right), \text{ pro kterou } \text{var}(\varepsilon) = 2\lambda^2,$$

Pak je maximalizace věrohodnostní funkce totožná s minimalizací součtu absolutních odchylek $\sum |\varepsilon_t|$ a nasazená **LAD- odhadová funkce je ML- odhadová funkce (pro toto rozdělení)**

Ověření:


V případě nezávislých *dvojexponenciálně rozdělených* náhodných složek lze jejich sdruženou hustotu $f(\varepsilon, \lambda)$, resp. věrohodnostní funkci $L(\lambda, \varepsilon)$ zapsat jako

$$f(\varepsilon, \lambda) = \prod_{t=1}^T f(\varepsilon_t) = \prod_{t=1}^T \frac{1}{2\lambda} \cdot \exp\left(-\frac{|\varepsilon_t|}{\lambda}\right) = (2\lambda)^{-T} \cdot \prod_{t=1}^T \exp\left(-\frac{|\varepsilon_t|}{\lambda}\right) = L(\lambda, \varepsilon), \text{ přičemž}$$

² Není důvod očekávat, že se seřazení reziduálních hodnot udrží od iterace k iteraci stejně.

po logaritmování dostáváme

$$\log L(\lambda, \varepsilon) = -T \cdot \log(2\lambda) - \frac{1}{\lambda} \cdot \sum_{t=1}^T |\varepsilon_t| .$$

Vzhledem k tomu, že první člen tohoto (záporného) součtu nezávisí na ε_t , je maximalizace věrohodnostní funkce rovnocenná s minimalizací $\sum |\varepsilon_t|$. 

Dvoustranné exponenciální rozdělení je ve srovnání s normálním u vrcholu daleko více špičaté a má užší konce (ve větší vzdálenostech od střední hodnoty)