

Třístupňová metoda nejmenších čtverců
3SLS (3-stage least squares)
v soustavě simultánních regresních rovnic
autoři metody: Henri Theil , Arnold Zellner [1962]

Třístupňová metoda nejmenších čtverců je jednodušší ze dvou odhadových metod poskytujících konzistentní a asymptoticky vydatné odhady strukturních parametrů v interdependentních soustavách simultánních regresních rovnic.

Název metody však nevyjadřuje přesně výpočetní "algoritmus" metody. Metoda sestává sice ze tří kroků, avšak nejde o trojnásobné nasazení metody LS-nejmenších čtverců. Vztah 3SLS k 2SLS je analogický vztahu GLS k OLS - přesnější název by tedy byl **Aitkenovo zobecnění 2SLS**. Algoritmus metody lze formálně zapsat jedním explicitním vzorcem, takže k výpočtu parametrů není třeba použít numerické iterační metody.

Rozdíl oproti 2SLS i jiným metodám s omezenou informací je mj. to, že výpočet se uskutečňuje paralelně pro všechny rovnice soustavy. Nejdříve se odhadnou pomocí 2SLS strukturní parametry všech modelových rovnic. Poté se z 2SLS-reziduí odhadne souhrnná kovarianční matice 3SLS odhadové funkce a způsobem podobným GLS se spočtou nejprve 3SLS-odhady parametrů a poté 3SLS-asymptotická kovarianční matice.

Formální popis metody (uvažujeme libovolnou i-tou regresní rovnici):

Při zápisu libovolné i-té rovnice ve tvaru

$$1) \quad y_i = Y_i \cdot \beta_i + X_i \cdot \gamma_i + \varepsilon_i$$

můžeme psát celou soustavu m rovnic v souhrnném tvaru:

$$1a) \quad w = Q \cdot \delta + r$$

kde $w = (w'_1, w'_2, \dots, w'_m)'$ je $q \cdot m$ -rozměrný vektor tvořený $q \cdot m$ -rozměrnými vektory $w_i = R^{-1} X'_i y_i$ transformovaných pozorování vysvětlovaných proměnných i – té rovnice

$\delta = (\delta_1, \delta_2, \dots, \delta_m)'$ je $\sum(m_i + q_i)$ - rozměrný vektor parametrů všech m regresních rovnic, $\delta_i = (\beta'_i, \gamma'_i)'$.

$r = (r'_1, r'_2, \dots, r'_m)'$ je $q \cdot m$ -rozměrný vektor tvořený $q \cdot m$ -rozměrnými vektory transformovaných náhodných složek $r_i = R^{-1} X'_i \varepsilon_i$ i – té rovnice.

$Q = \text{diag}(Q_1, Q_2, \dots, Q_m)$ je blokově diagonální matice složená z m matic rozměrů $[q, m_i + q_i]$ transformovaných pozorování vysvětlujících (běžných endogenních i predeterminovaných) proměnných i – té rovnice.

$$2) \quad Q_i = (R^{-1} X'_i Y_i, R^{-1} X'_i X_i)$$

Vztah mezi vektory původních náhodných složek ε_i a transformovaných náhodných složek r_i lze vyjádřit souhrnně jako

$$3) \quad r = (I_m \otimes R^{-1} X') \varepsilon$$

kde I_m je jednotková matice řádu m a symbol \otimes označuje **Kroneckerův součin matic**.

Souhrnný výraz pro **3SLS-odhadovou funkci** (daný jako Aitkenův estimátor v soustavě rovnic) je při zápisu modelu ve tvaru o m stochastických rovnicích ve tvaru

$$\mathbf{w} = \mathbf{Q}\delta + \mathbf{r}$$

dán výrazem

$$4) \quad {}_{3\text{SLS}}\hat{\delta} = (\mathbf{Q}'\Phi^{-1}\mathbf{Q})^{-1}\mathbf{Q}'\Phi^{-1}\mathbf{w}$$

v němž $\Phi = \Sigma_m \otimes \mathbf{I}_q$ je kovarianční matice transformovaných náhodných složek \mathbf{r} .

Pokud by matice Φ (nebo, což je totéž, matice Σ) byla známá, mohli bychom výraz 4) již přímo převzít jako výraz pro $\hat{\delta}$ a mít tak 3SLS-odhadovou funkci. Situace, kdy známe Σ , jsou ovšem výjimečné a tak je zpravidla nutné při praktickém uplatnění 3SLS-estimátoru nahradit Σ nějakým jejím konzistentním odhadem. Místo vzorce 3) se proto zpravidla uplatňuje výraz

$$4a) \quad {}_{3\text{SLS}}\hat{\delta} = (\mathbf{Q}'\hat{\Phi}^{-1}\mathbf{Q})^{-1}\mathbf{Q}'\hat{\Phi}^{-1}\mathbf{w}$$

v němž $\hat{\Phi}$ je příslušný konzistentní odhad matice Φ . Za odhad Σ se zpravidla vezme estimátor ${}_{2\text{SLS}}\hat{\Sigma}_m$ určený **dvoustupňovou metodou nejmenších čtverců**, tj. estimátor daný výběrovou kovarianční maticí reziduí z 2SLS, tedy

$$\hat{\Sigma}_m = \{\hat{\sigma}_{ij}\} = \frac{\mathbf{e}'_i \mathbf{e}_j}{\mathbf{T}}$$

kde $\mathbf{e}'_i = \mathbf{y}_i - \mathbf{W}_{i2\text{SLS}}\hat{\delta}_i$ pro $i, j = 1, 2, \dots, m$

Shrnutí První stupeň 3SLS odhadové metody tedy představuje (stejně jako u 2SLS) očištění (všech) vysvětlujících běžných endogenních proměnných od jejich stochastických komponent (jejich náhradou maticí $\mathbf{X}\hat{\Pi}$). Druhý krok představuje získání konzistentního odhadu kovarianční matice Σ pomocí nahrazení maticí ${}_{2\text{SLS}}\hat{\Sigma}$ získanou návazně na první krok. Hledané 3SLS odhady parametrů jsou získány ve třetím kroku dosazením $\hat{\Sigma}$ resp. $\hat{\Phi}$ do výrazu 4a).

Poznámka Bez ohledu na název lze 3SLS-odhady parametrů všech m rovnic pořídit pomocí jediné operace (znamená to ovšem nahradit prvky v Σ veličinami vyjádřenými v transformovaných pozorovaných proměnných). Každé σ_{ij} lze totiž zapsat jako

$$\hat{\sigma}_{ij} = \frac{\mathbf{y}'_i \mathbf{A}_{ij} \mathbf{y}_j}{\mathbf{T}}$$

kde \mathbf{A}_{ij} je matice

$$\mathbf{A}_{ij} = [\mathbf{I}_T - \mathbf{W}_i(\mathbf{Q}'_i \mathbf{Q}_i)^{-1} \mathbf{Q}'_i \mathbf{R}^{-1} \mathbf{X}'_i] [\mathbf{I}_T - \mathbf{W}_j(\mathbf{Q}'_j \mathbf{Q}_j)^{-1} \mathbf{Q}'_j \mathbf{R}^{-1} \mathbf{X}'_j]$$

přičemž $\mathbf{W}_i = (\mathbf{Y}_i, \mathbf{X}_i)$ $\mathbf{Q}_i = (\mathbf{R}^{-1} \mathbf{X}'_i \mathbf{Y}_i; \mathbf{R}^{-1} \mathbf{X}'_i \mathbf{X}_i)$, kde $\mathbf{X}'\mathbf{X} = \mathbf{R}\mathbf{R}'$.

Připomeňme, že matice \mathbf{R} rozměrů $[\mathbf{q}, \mathbf{q}]$ je regulární a platí $\mathbf{X}'\mathbf{X} = \mathbf{R}\mathbf{R}'$ (\mathbf{X} má rozměry $[\mathbf{T}, \mathbf{q}]$ a nemůže být tedy – až na ojedinělý případ, kdy $\mathbf{T} = \mathbf{q}$ – regulární).

Kovarianční matice transformovaných náhodných složek má tvar

$$\text{Cov}(\mathbf{r}) = (\mathbf{I}_m \otimes \mathbf{R}^{-1} \mathbf{X}') (\Sigma \otimes \mathbf{I}_T) (\mathbf{I}_m \otimes \mathbf{R}^{-1} \mathbf{X}')' = \Sigma \otimes \mathbf{I}_q$$

Tato matice (ani $\mathbf{E}(\mathbf{r})$) nezávisí na matici predeterminovaných proměnných \mathbf{X} , transformace snižuje dimenzi kovarianční matice celé soustavy z $m\mathbf{T}$ na $m\mathbf{q}$. Lze se tedy důvodně domnívat, že aplikace Aitkenova přístupu poskytne ve srovnání s 2SLS vydatnější odhad.

Vlastnosti 3SLS -odhadové funkce

Lze ukázat, že 3SLS-odhadová funkce má tyto vlastnosti:

1) **Odhady parametrů** ${}_{3SLS}\hat{\beta}_i, {}_{3SLS}\hat{\gamma}_i$ $i = 1, 2, \dots, m$ jsou konzistentní, tj. platí

$${}_{\mathbf{p}} \lim_{T \rightarrow \infty} \begin{pmatrix} {}_{3SLS}\hat{\beta}_i \\ {}_{3SLS}\hat{\gamma}_i \end{pmatrix} = \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}$$

2) **Odhady parametrů** ${}_{3SLS}\hat{\beta}_i, {}_{3SLS}\hat{\gamma}_i$ nejsou neustranné (podobně jako u 2SLS, ILS, LIML, IV)

3) **Odhady parametrů** ${}_{3SLS}\hat{\beta}_i, {}_{3SLS}\hat{\gamma}_i$ jsou vydatné i v rámci metod s úplnou informací (asymptotická rozdělení 3SLS a FIML jsou totožná)

Rozdělení 3SLS-odhadové funkce pro konečný rozsah výběru je sice (opět díky P.C.B. Phillipsovi) známo (cca od roku 1984), ale je příliš komplikované, než aby ho bylo možno prakticky použít pro testovací účely. V úlohách, které jsou spojeny s konstrukcí intervalů spolehlivosti a postupy testování hypotéz se proto omezujeme na rozdělení asymptotické (tj. rozdělení, které má 3SLS-odhadová funkce při neomezeně rostoucím rozsahu výběru $T \rightarrow \infty$).

4) Dále lze ukázat, že za poměrně předpokladů totožných jako byly zavedeny u 2SLS-odhadové funkce je **rozdělení 3SLS-odhadové funkce normální**. Platí

$$\text{Cov}({}_{3SLS}\hat{\delta}) = T^{-1} \cdot {}_{\mathbf{p}} \lim_{T \rightarrow \infty} \{ \mathbf{W}' [\Sigma^{-1} \otimes \mathbf{X}(\mathbf{X}\mathbf{X})^{-1} \mathbf{X}'] \mathbf{W} \}^{-1}$$

Za podmínek 1) - 4) platí, že asymptotické rozdělení 3SLS-estimátoru je normální se střední hodnotou δ a kovarianční maticí ${}_{\mathbf{p}} \lim_{T \rightarrow \infty} (\mathbf{Q}' \Phi^{-1} \mathbf{Q})^{-1}$.

Obvyklý zápis tohoto rozdělení je

$$\sqrt{T}({}_{3SLS}\hat{\delta} - \delta) \approx \mathbf{N} \left[0, {}_{\mathbf{p}} \lim_{T \rightarrow \infty} \left(\frac{\mathbf{Q}' \Phi^{-1} \mathbf{Q}}{T} \right)^{-1} \right]$$

Statistické testy budou tedy založeny na normálním $\mathbf{N}(0,1)$ – rozdělení.

Poznámka Ve dvou situacích dochází ke shodě 3SLS a 2SLS odhadových funkcí. Těmito situacemi jsou :

a) Příklad, kdy jsou **náhodné složky různých rovnic vzájemně nekorelovány**, tzn., kdy platí : $\sigma_{ij} = 0$ pro všechna $i \neq j$ a Σ je tedy diagonální matice.¹ Pak

$${}_{3SLS}\hat{\delta} = (\mathbf{Q}' \hat{\Phi}^{-1} \mathbf{Q})^{-1} \mathbf{Q}' \hat{\Phi}^{-1} \mathbf{w} = \text{diag} \left[\hat{\sigma}_{11} \left(\mathbf{Q}'_1 \mathbf{Q}_1 \right)^{-1}, \dots, \hat{\sigma}_{mm} \left(\mathbf{Q}'_m \mathbf{Q}_m \right)^{-1} \right] \cdot \text{diag} \left[\mathbf{Q}'_1 \hat{\sigma}_{11}^{-1}, \dots, \mathbf{Q}'_m \hat{\sigma}_{mm}^{-1} \right] \mathbf{w} = (\mathbf{Q}' \mathbf{Q})^{-1} \mathbf{Q}' \mathbf{w}$$

¹ Všimněme si rozdílu oproti jednorovnicovému modelu: Zobecněná metoda nejmenších čtverců GLS ztrácí vůči OLS přednost (ve vydatnosti) pouze tehdy, jsou-li náhodné složky homoskedastické a neautokorelované. Přednost 3SLS oproti 2SLS se ztrácí již tehdy, jsou-li náhodné složky „neautokorelované“. Zde však nemá smysl uvažovat o „homoskedasticitě“, protože není žádný důvod předpokládat, že by náhodné složky **různých rovnic** měly mít stejný rozptyl (závisle proměnné rovnic nemají k sobě velikostí svých hodnot vůbec žádný vztah).

Tuto okolnost lze za určitých okolností využít takto: Pokud je možné analyzovaný ekonometrický model rozložit do několika submodelů takových, že náhodné složky jednoho submodelu jsou nekorelované s náhodnými složkami ostatních submodelů, stačí se nasadit 3SLS proceduru vždy jen na individuální submodel (rozšíření na celý model již neznamena žádný přínos ve vydatnosti).

b) Příklad, kdy jsou **všechny rovnice modelu přesně identifikované**, tzn. platí $m_i + q_i = q$ pro všechna $i = 1, 2, \dots, m$. Pak jsou totiž všechny matice Q_i čtvercové matice řádu (i hodnot) q a lze tedy provést invertování všech matic ve vzorci pro 3SLS estimátor následovně:

$${}_{3SLS} \hat{\delta} = (Q' \hat{\Phi}^{-1} Q)^{-1} Q' \hat{\Phi}^{-1} w = Q^{-1} \hat{\Phi} Q'^{-1} \cdot Q' \hat{\Phi}^{-1} w = Q^{-1} w$$

rovněž tak pro 2SLS estimátor

$${}_{2SLS} \hat{\delta} = (Q' Q)^{-1} Q' w = Q^{-1} Q'^{-1} \cdot Q' w = Q^{-1} w$$

V tomto případě mají obě odhadové funkce zjednodušený tvar $Q^{-1} w$. Tuto okolnost lze také využít pro případ, kdy by model obsahoval více **přesně identifikovaných rovnic**. V takovém případě stačí použít 3SLS-estimátor pouze na **předidentifikované rovnice**, zatímco (stejně kvalitní) odhad přesně identifikovaných rovnic zajistíme nasazením jednodušší 2SLS - metody. Obvykle ovšem v běžných modelech převažují předidentifikované rovnice.² Že by model sestával jen z přesně identifikovaných rovnic

V jiných případech bude rozdíl asymptotických kovariančních matic 2SLS- a 3SLS-odhadových funkcí pozitivně definitní matice. Platí vztah

$$\text{plim}_{T \rightarrow \infty} \left[\left(\frac{Q' Q}{T} \right)^{-1} \left(\frac{Q' Q}{T} \right) \left(\frac{Q' Q}{T} \right)^{-1} \right] = \text{plim}_{T \rightarrow \infty} \left[\left(\frac{Q' \Phi^{-1} Q}{T} \right)^{-1} \right] + D \Phi D'$$

Všimněme si, že odhad parametrů **metodou 3SLS** představuje vždy **simultánní výpočet všech** $\sum m_i + q_i$ **parametrů současně**, zatímco **metodou 2SLS** odhadujeme zpravidla **parametry jednotlivých rovnic odděleně**³.

² Že by model obsahoval jen přesně identifikované rovnice, lze snad v realitě připustit jen u velmi malých modelů (cca do počtu 3-4 rovnic).

³ Ovšem také metodou 2SLS můžeme – pokud je model zapsán v kompaktním tvaru - spočítat odhady všech parametrů všech rovnic najednou.

Připojíme ještě souhrnné vyjádření 3SLS - odhadové funkce :

$$\begin{pmatrix} \hat{\sigma}^{11}Q_1'Q_1 & \hat{\sigma}^{12}Q_1'Q_2 & \hat{\sigma}^{13}Q_1'Q_3 & \dots & \hat{\sigma}^{1m}Q_1'Q_m \\ \hat{\sigma}^{21}Q_2'Q_1 & \hat{\sigma}^{22}Q_2'Q_2 & \hat{\sigma}^{23}Q_2'Q_3 & \dots & \hat{\sigma}^{2m}Q_2'Q_m \\ \hat{\sigma}^{31}Q_3'Q_1 & \hat{\sigma}^{32}Q_3'Q_2 & \hat{\sigma}^{33}Q_3'Q_3 & \dots & \hat{\sigma}^{3m}Q_3'Q_m \\ \dots & \dots & \dots & \dots & \dots \\ \hat{\sigma}^{m1}Q_m'Q_1 & \hat{\sigma}^{m2}Q_m'Q_2 & \hat{\sigma}^{m3}Q_m'Q_3 & \dots & \hat{\sigma}^{mm}Q_m'Q_m \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_{j=1}^m \hat{\sigma}^{1j}W_1'X(X'X)^{-1}X'w_{.1} \\ \sum_{j=1}^m \hat{\sigma}^{2j}W_2'X(X'X)^{-1}X'w_{.2} \\ \sum_{j=1}^m \hat{\sigma}^{3j}W_3'X(X'X)^{-1}X'w_{.3} \\ \dots \\ \sum_{j=1}^m \hat{\sigma}^{mj}W_m'X(X'X)^{-1}X'w_{.m} \end{pmatrix}$$

kde jak víme $w_i = R^{-1}X'y_i$ $Q_i = (R^{-1}X'Y_i; R^{-1}X'X_i)$ $W_i = (Y_i, X_i)$

σ^{ij} je prvek ležící na v i-tém řádku a j-tém sloupci matice Σ^{-1}
 X je matice všech predeterminovaných proměnných soustavy .