Obecná metoda instrumentálních proměnných
(G)IV (General Instrumental Variables method)
v soustavě simultánních regresních rovnic

Metoda instrumentálních proměnných je jistým zobecněním dvoustupňové metody nejmenších čtverců 2SLS. Poskytuje, stejně jako 2SLS, vždy (přinejmenším) **konzistentní odhady strukturních parametrů regresních rovnic v interdependentních ekonometrických modelech.**

Základní motivací metody je nalézt určité pomocné proměnné - tzv. instrumentální proměnné - které sehrají stejnou úlohu , jako má transformace $R^{-1}X'$ při odvození odhadové funkce 2SLS (viz druhý postup odvození 2SLS)

Hledají se tedy takové proměnné - jejich matici ve vztahu k i-té rovnici označme jako P_i - které budou vyhovovat vztahu

$$P_i' y_i = P_i' W_i \delta_{.i} + P_i' \varepsilon_i$$

kde $$W_i = (Y_i, X_i); \quad \delta_{.i} = (\beta_{.i}', \gamma_{.i}')'$$

a přitom takové, že
a) budou nekorelované s náhodnými složkami i-té strukturní rovnice
b) budou co nejvíce korelované s vysvětlujícími proměnnými i-té rovnice

Podmínka a) **je nutná k tomu, aby byl odhad takto pořízený konzistentní.**
Podmínka b) **je potřebná k tomu, aby proměnné-instrumenty zastupující vysvětlující veličiny v rovnici je nahrazovaly co nejvýstižněji**

Z podmínek je zřejmé, že instrumentální proměnné lze vybírat (pouze) z predeterminovaných proměnných modelu (běžné endogenní jsou korelované s náhodnými složkami). **Problém nespočívá v tom, čím nahradit v i-té rovnici přítomné predeterminované proměnné, ale čím nahradit přítomné běžné endogenní veličiny.**

Zbývá tedy provést co nejvhodnější výběr z predeterminovaných proměnných modelu. Je tedy zřejmé, že instrumentální proměnné budou definované pomocí maticového vztahu

$$P_i = X.A_i$$

kde

- A_i **je určující matice definující instrumentální proměnné** (matice tzv.instrumentů)
- P_i **je matice instrumentálních proměnných pro i-tou rovnici**
- (X **je matice všech predeterminovaných proměnných modelu) .**

Volba instrumentálních proměnných (matice P_i) **je tedy rovnocenná určení matice instrumentů A_i. Index příslušnosti k rovnici lze vynechat, pokud pro odhad každé rovnice modelu použijeme tutéž skupinu instrumentálních proměnných** (je to obvyklé, nikoliv nutností).

V tomto případě bychom psali

$$P = X.A$$

kde

- A **je matice instrumentů definujících instrumentální proměnné pro odhad parametrů všech rovnic.**
- Z **požadavků, které byly na instrumenty položeny, plyne, že IV-odhadová funkce strukturních parametrů modelu má tvar**

$$_{IV}\hat{\delta}_{.i} = (P_i' W_i)^{-1}.P_i' y_{.i} = (P_i' W_i)^{-1}.P_i' (W_i.\delta_{.i} + \varepsilon_i) = \delta_{.i} + (P_i' W_i)^{-1}.P_i' \varepsilon_i$$

Poznámka **podmínkou existence IV-estimátoru je, aby byly existovala inverzní matice k matici** $P_{i\,[q,T]}'.W_{i\,[T;m_i+q_i]}$ **K tomu je opět přinejmenším nutné, aby byla splněna podmínka** $m_i + q_i = q$ **: jinak by matice** $P_{i\,[q,T]}'.W_{i\,[T;m_i+q_i]}$ **nemohla být ani čtvercová (tím méně ne regulární).**[1] **(obvykle předpokládáme** $\mathbf{q} \le \mathbf{T}$ **)**

Vlastnosti IV-odhadové funkce:

Lze ukázat, že IV-estimátor strukturních parametrů modelu má tyto vlastnosti:

1) Odhady parametrů $\delta_{.i}$ **(tj.** $\beta_{.i}, \gamma_{.i}$ **) jsou konzistentní, neboť platí**

$$p\lim_{T\to\infty} \hat{\delta}_i - \delta_i = p\lim_{T\to\infty}(P_i' W_i / T)^{-1}.p\lim_{T\to\infty}(P_i' \varepsilon_i / T) = 0$$

v důsledku (asymptotické) nekorelovanosti proměnných P_i **a náhodných složek** ε_i

2) Odhady parametrů δ_i **i (neboli** β_i, γ_i **) nejsou nestranné, protože**

$$E\hat{\delta}_i = E\left[\delta_i + (P_i' W_i)^{-1}.P_i' \varepsilon_i\right] = \delta_i + E(P_i' W_i)^{-1}.P_i' \varepsilon_i$$

ale výraz $E(P_i' W_i)^{-1} P_i' \varepsilon_j \ne E(P_i' W_i)^{-1} E(P_i' \varepsilon_j)$ **vzhledem k možné závislosti běžných endogenních proměnných přítomných ve** W_i **a náhodných složek** ε_i **.**

3) Odhady parametrů δ_i (tj. β_i, γ_i) **nejsou, až na výjimku, kdy metoda IV přechází v 2SLS, obecně vydatné** (ani v rámci metod s omezenou informací).

4) **Odhady parametrů** δ_i (tj. β_i, y_i i) **jsou** (za stejných předpokladů (e), (f), (g), (h) jako u 2SLS) **vždy asymptoticky normální , tedy platí**

$$\sqrt{T}.(\hat{\delta}_{.i} - \delta_{.i}) \approx N\left(0, \sigma_{ii}.p\lim_{T\to\infty}\left[\left(\frac{P_i' W_i}{T}\right)^{-1}\left(\frac{P_i' P_i}{T}\right)\left(\frac{W_i' P_i}{T}\right)^{-1}\right]\right)$$

[1] *Počet instrumentů potřebných k odhadu i-té rovnice musí být tedy roven počtu vysvětlujících proměnných této rovnice. Podrobněji v části pojednávající o identifikačním problému.*

Konzistentní odhad prvků $_{IV}\sigma_{ij}$ pro jednotlivé rovnice získáme obvyklým způsobem:

$$_{IV}\hat{\sigma}_{ij} = \frac{_{iv}\boldsymbol{e}_{.i}'\ _{iv}\boldsymbol{e}_{.j}}{\boldsymbol{T}}$$

kde za rezidua $\boldsymbol{e}_{.i},\boldsymbol{e}_{.j}$ vezmeme odhady náhodných složek $\varepsilon_{.i},\varepsilon_{.j}$ získané metodou IV.

Je tedy zřejmé, že otázka nejlepšího výběru (poskytujícího nejvydatnější IV-odhad) **mezi různými IV-estimátory spočívá v optimální definici matice $\boldsymbol{A}$. Jinými slovy, vyšetřujeme, pro jakou volbu matice $\boldsymbol{A}$ nastává maximální možná korelace mezi instrumenty v $\boldsymbol{A}$** (resp. mezi instrumentálními proměnnými v $\boldsymbol{P}$) **a vysvětlujícími proměnnými i-té rovnice $\boldsymbol{W}_i$?**

Pro měření korelace mezi dvěma skupinami náhodných veličin (majících stejný počet pozorování) **se užívá vektorový korelační koeficient definovaný jako**:

$$r_C(\boldsymbol{W}_i;\boldsymbol{P}_i) = (-1)^{m_i+q_i}.\frac{\begin{vmatrix} 0 & \boldsymbol{W}_i'\boldsymbol{P}_i \\ \boldsymbol{P}_i'\boldsymbol{W}_i & \boldsymbol{P}_i'\boldsymbol{P}_i \end{vmatrix}}{\left|\boldsymbol{W}_i'\boldsymbol{W}_i\right|.\left|\boldsymbol{P}_i'\boldsymbol{P}_i\right|} = \frac{\left|(\boldsymbol{W}_i'\boldsymbol{P}_i)(\boldsymbol{P}_i'\boldsymbol{P}_i)^{-1}(\boldsymbol{P}_i'\boldsymbol{W}_i)\right|}{\left|\boldsymbol{W}_i\boldsymbol{W}_i\right|}$$

Hodnota tohoto koeficientu se pohybuje mezi 0 (nezávislost) **a 1** (přesná závislost) **.**
Výraz, který v kovarianční maticí IV-estimátoru $(\boldsymbol{P}_i'\boldsymbol{W}_i/\boldsymbol{T})^{-1}(\boldsymbol{P}_i'\boldsymbol{P}_i/\boldsymbol{T})(\boldsymbol{W}_i'\boldsymbol{P}_i/\boldsymbol{T})^{-1}$ v sobě obsahuje fragment výrazu pro tzv. zobecněný rozptyl. Ten je definován jako

$$\left|\boldsymbol{GVar}\hat{\delta}_{.i}\right| = \sigma_{ii}^{\ m_i+q_i}.\left|(\boldsymbol{P}_i'\boldsymbol{W}_i)^{-1}.(\boldsymbol{P}_i'\boldsymbol{P}_i).(\boldsymbol{W}_i'\boldsymbol{P}_i)^{-1}\right|$$

Mezi vektorovým korelačním koeficientem a zobecněným rozptylem platí tedy vztah

$$\left|\boldsymbol{GVar}\hat{\delta}_{.i}\right| = \frac{\sigma_{ii}^{\ m_i+q_i}}{\left|\boldsymbol{W}_i'\boldsymbol{W}_i\right|}.(r_C(\boldsymbol{W}_i;\boldsymbol{P}_i))^{-1}$$

z čehož je patrné, že pro taková $\boldsymbol{P}_i$, pro která je minimalizována hodnota $\left|\boldsymbol{GVar}\hat{\delta}_{.i}\right|$ je právě maximalizována korelace mezi $\boldsymbol{W}_i$ a $\boldsymbol{P}_i$.

Vyšetříme, kdy taková korelace nabude maximální možné hodnoty; v tomto případě poskytne IV-odhadová funkce δ_i nejvydatnější odhad. Lze přitom ukázat, že platí:

$$r_C(\boldsymbol{W}_i,\boldsymbol{P}) = r_C(\boldsymbol{W}_i,\boldsymbol{X})$$

Znamená to tedy, že nemůže být překročena horní hranice daná (vektorovou) **korelací mezi množinou instrumentálních proměnných a množinou všech predeterminovaných proměnných.**
Této maximální korelovanosti je dosaženo pro volbu

$$\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}_i$$

Při této volbě matice $\boldsymbol{A}$ dostaneme :

$$\boldsymbol{P} = \boldsymbol{X}.\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}_i = \left[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_i;\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}_i\right] = \left[\boldsymbol{X}\hat{\Pi}_i;\boldsymbol{X}_i\right]$$

Pak je IV- odhadová funkce rovna

$$_{IV}\hat{\delta}_i = (\boldsymbol{P}_i'\boldsymbol{W}_i)^{-1}\boldsymbol{P}'\boldsymbol{y}_{.i} =$$

$$\begin{pmatrix} \boldsymbol{Y}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_i & \boldsymbol{Y}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_i \\ \boldsymbol{X}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_i & \boldsymbol{X}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}_i \end{pmatrix}^{-1} . \begin{pmatrix} \boldsymbol{Y}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}_i \\ \boldsymbol{X}_i'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}_i \end{pmatrix} = {}_{2SLS}\hat{\delta}_{.i}$$

Znamená to tedy, že :

1) 2SLS-odhadová funkce je speciálním případem IV-odhadové funkce při volbě matice instrumentů jako $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}_i$
2) 2SLS-odhadová funkce poskytuje ve srovnání s jakoukoliv jinou volbou matice $\boldsymbol{A}$ nejvydatnější odhad. tj. ve smyslu asymptotické vydatnosti je 2SLS-odhadová funkce dominantní vůči všem ostatním IV-estimátorům.

Skutečnost, že aplikací techniky IV nelze překonat metodu 2SLS může být jistým zklamáním. V nelineárních modelech tomu tak není, zde můžeme za instrumenty vzít též nelineární kombinace z predeterminovaných proměnných. Ani *NL2S* estimátor (nelineární dvoustupňová metoda nejmenších čtverců) není zde definován jednoznačně : existují např. *BNL2S* (best) a *MNLS* (minimal) estimátor .

Počet instrumentálních proměnných n musí být v rozmezí mezi $m_i + q_i$ a q, tedy

$$m_i + q_i \le n \le q$$

Pokud uplatníme *instrumentální proměnné v maximálním možném počtu* q tj. jako všechny predeterminované proměnné, pak

- využijeme maximum informace obsažené v modelových proměnných, což povede k vydatnému odhadu , ale
- budeme pracovat s obsažnějšími maticemi a případně nižší spolehlivostí výsledku

Pokud uplatníme *instrumentální proměnné v minimálním přípustném počtu* $m_i + q_i$ tj. jako výběr $m_i + q_i$ predeterminovaných proměnných, pak

- nevyužijeme všechnu potřebnou informaci obsaženou v modelových proměnných, což bude mít za následek méně kvalitní (byť konzistentní) odhadu , ale
- výpočet bude úspornější a počet stupňů volnosti modelu vyšší.

Kompromisem může být vzetí instrumentálních proměnných v podobě lineární kombinace sestávající z prvních $m_i + q_i$ hlavních komponent momentové matice $\boldsymbol{X}'\boldsymbol{X}$.

Při řešení konkrétních úloh se uplatňují tyto přístupy k volbě instrumentálních proměnných (*definujících matici A*):

a) prostý výběr počtu $m_i + q_i$ z celkem q predeterminovaných proměnných . Matice instrumentů bude zde mít tvar $A[q, m_i + q_i]$, přičemž v této obdélníkové matici budou jedničkové prvky pouze v hlavní "pseudodiagonále" $a_{ii}, i = 1,2,\ldots, m_i + q_i$.

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

U predeterminovaných proměnných, které jsou vzaty jako instrumentální, je v příslušném *sloupci* A₁ jednička – vynechávaným odpovídají nulové sloupce.

**b) $m_i + q_i$ – členná lineární kombinace složená z predeterminovaných proměnných
V tomto případě má příslušná matice tvar**

$$A_2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1,m_i+q_i} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2,m_i+q_i} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3,m_i+q_i} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{q,1} & a_{q,2} & a_{q,3} & \ldots & a_{q,m_i+q_i} \end{pmatrix}$$

Koeficienty lineární kombinace jsou obsaženy ve *sloupcích* této matice.

c) prvních $m_i + q_i$ hlavních komponent sestrojených z matice predeterminovaných proměnných

$$A_3 = \begin{pmatrix} m_{11} & m_{12} & m_{13} & \ldots & m_{1,m_i+q_i} \\ m_{21} & m_{22} & m_{23} & \ldots & m_{2,m_i+q_i} \\ m_{31} & m_{32} & m_{33} & \ldots & m_{3,m_i+q_i} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ m_{q,1} & m_{q,2} & m_{q,3} & \ldots & m_{q,m_i+q_i} \end{pmatrix}$$

Koeficienty této lineární kombinace (opět obsažené ve sloupcích matice A_3) **představují prvky vlastních vektorů příslušných momentové matici $X'X$. Z celkem q hlavních komponent se omezujeme na „největších" $m_i + q_i$ z nich.**