

ZOBECNĚNÝ LINEÁRNÍ REGRESNÍ MODEL. METODA ZOBECNĚNÝCH NEJMENŠÍCH ČTVERCŮ

V následujícím textu se podíváme na to, co dělat, když jsou porušeny některé GM předpoklady. Nejprve si připomeňme, o jaké předpoklady jde.¹

- 1) $E(\mathbf{u}) = 0$, tedy střední hodnota náhodné složky ve všech výběrech bude nulová.
- 2) $E(\mathbf{u}\mathbf{u}') = \sigma^2 \mathbf{I}_n$, tedy náhodné složky jsou sériově nezávislé a mají konstantní rozptyl.
- 3) \mathbf{X} je nestochastická matice.
- 4) \mathbf{X} má plnou hodnost.

PROČ VADÍ, KDYŽ JSOU TYTO PŘEDPOKLADY PORUŠENY?

Pokud je porušen první GM předpoklad, promítne se nenulová střední hodnota náhodné složky do odhadu úroňové konstanty. To znamená, že odhad úroňové konstanty bude **vychýlený**.

Druhý GM předpoklad v sobě obsahuje hned dva požadavky: náhodné složky musí mít jednak stejný rozptyl, jednak na sobě náhodné složky nesmí být závislé. Při porušení prvního požadavku mluvíme o **heteroskedasticitě**, při porušení druhého požadavku pak o **autokorelaci**. Odhady parametrů sice zůstanou nestranné a konzistentní, ale nebudou vydatné ani asymptoticky vydatné (tedy nebudou mít nejmenší rozptyl). Navíc budou odhady směrodatných chyb parametrů a odhad rozptylu náhodné složky σ^2 vychýlené, takže se na ně nemůžeme spolehnout při standardních statistických testech a při konstrukci intervalů spolehlivosti.

Třetí GM předpoklad bývá porušen zejména v případě simultánních rovnic, takže je podrobněji rozabráán v příslušných otázkách.

Čtvrtý GM předpoklad říká, že v matici vysvětlujících proměnných nesmí být lineárně závislé sloupce. Pokud tam jsou, mluvíme o perfektní kolinearitě, pokud jsou „jen“ silně závislé, jde o (multi)kolinearitu.

JAK FUNGUJE METODA ZOBECNĚNÝCH NEJMENŠÍCH ČTVERCŮ?

Podstatou MZNC (nazývána také Aitkenův postup) je upravit původní model tak, aby byly splněny GM požadavky a mohla se použít MNČ. Při porušení druhého GM předpokladu neplatí vztah $E(\mathbf{u}\mathbf{u}') = \sigma^2 \mathbf{I}_n$, nýbrž platí, že $E(\mathbf{u}\mathbf{u}') = \sigma^2 \mathbf{V}$, kde σ^2 je neznámý skalár a \mathbf{V} je známá symetrická pozitivně definitní matice řádu n . Matici \mathbf{V} lze vyjádřit také jako součin dvou regulárních vzájemně transponovaných matic: $\mathbf{V}^{-1} = \mathbf{T}\mathbf{T}'$. Pokud model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ pronásobíme maticí \mathbf{T} , dostaneme upravený model $\mathbf{T}\mathbf{y} = \mathbf{T}\mathbf{X}\boldsymbol{\beta} + \mathbf{T}\mathbf{u}$, čili $\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta} + \mathbf{u}^*$, který již splňuje druhý GM předpoklad, protože $E(\mathbf{u}^*\mathbf{u}^{*\prime}) = E(\mathbf{T}\mathbf{u}\mathbf{u}'\mathbf{T}') = \sigma^2\mathbf{T}\mathbf{V}\mathbf{T}' = \sigma^2\mathbf{I}_n$. Na takto upravený model už můžeme použít klasickou MNČ. Pro odhadovou funkci \mathbf{b}^* platí $\mathbf{b}^* = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ a jeho kovarianční matice je $\sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$. Nestranný odhad rozptylu σ^2 získáme jako $s^{2*} = \frac{\mathbf{e}^{*\prime}\mathbf{e}^*}{n-k}$, kde \mathbf{e}^* je vektor reziduí z MZNC modelu. Otázkou je, jak má matice \mathbf{T} , kterou model chceme pronásobit, vypadat. Tato matice se liší pro heteroskedasticitu a pro autokorelaci.

¹ Pozor: Zouhar pracuje ve své knize s $k+1$ odhadovanými parametry, Hušek s k , tedy včetně úroňové konstanty

HETEROSKEDASTICITA

Jak bylo řečeno výše, heteroskedasticita znamená, že pro různé hodnoty vysvětlujících proměnných má náhodná složka, a tedy i vysvětlovaná proměnná, jiný rozptyl. Setkáme se s ní většinou:

- v průřezových datech, zejm. mikroekonomických analýzách (příklad: výdaje domácností budou mít větší rozptyl u bohatších domácností než u chudších; objem produkce firmy bude mít větší rozptyl v případě větších než menších firem)
- při chybné specifikaci modelu (jde o tzv. „kvaziheteroskedasticitu“), například pokud vynecháme významnou proměnnou;
- kvůli chybám měření (s rostoucí hodnotou vysvětlované proměnné se kumulují chyby a zvětšuje se rozptyl reziduí);
- při použití skupinových průměrů či jiných agregovaných údajů.

Proč nám vadí heteroskedasticita?

- odhad pomocí MNČ bude sice nestranný a konzistentní, ale nebude vydatný ani asymptoticky vydatný;
- odhad směrodatných chyb parametrů a odhad rozptylu modelu bude vychýlený, takže se nelze spolehnout na zkonstruované intervaly spolehlivosti ani závěry testů hypotéz.

Jak zjistit, jestli je v modelu heteroskedasticita?

- I. **Spermanův test** korelace pořadí zjišťuje, zda se v modelu vyskytuje **lineární závislost** směrodatné odchylky náhodné složky na vysvětlující proměnné. Jak na to? Jak absolutní hodnoty reziduí, tak hodnoty vysvětlujících proměnných uspořádáme vzestupně a spočítáme párový koeficient korelace pořadí ze vzorce: $r_{ex} = 1 - \frac{6 \sum (\text{pořadí } |e_i| - \text{pořadí } |X_i|)^2}{n(n^2-1)}$.
 Statistickou významnost testujeme pomocí $t = r_{ex} \frac{\sqrt{n-k}}{\sqrt{1-r_{ex}^2}}$. Pokud je $|t| > t_{1-\alpha/2}(n-k-1)$, pak na α procentní hladině významnosti zamítáme nulovou hypotézu o homoskedasticitě. Zjišťujeme tedy vlastně, jestli je mezi pořadím hodnoty reziduí a pořadím hodnoty vysvětlující proměnné statisticky významná závislost, čili jestli pro větší hodnoty vysvětlující proměnné jsou i rezidua, a tudíž rozptyl náhodné složky, větší.
- II. **Golfeldův-Quandtův test** je použitelný tehdy, když je směrodatná odchylka monotónní funkcí některé vysvětlující proměnné. Jak na to? Uspořádáme všech n pozorování vzestupně podle velikosti, pak g pozorování uprostřed vynecháme a zbytek rozdělíme na poloviny. Pro větší výběry nad 30 by g mělo být mezi čtvrtinou až třemi osminami, pro menší výběry alespoň jedna čtvrtina. Chceme zjistit, jestli je rozptyl shodný v obou těchto polovinách. Odhadneme regresi pro jednotlivé poloviny a zjistíme součet čtverců reziduí s_1, s_2 (kde $s_2 > s_1$). Spočteme testovou statistiku: $F = s_2/s_1$. Jestliže je testová statistika větší než kritická hodnota Fisherova rozdělení pro $(\frac{n-g}{2} - k - 1, \frac{n-g}{2} - k - 1)$ stupňů volnosti, zamítáme nulovou hypotézu o homoskedasticitě.

- III. **Glejserův test** postupuje tak, že odhaduje závislost absolutní hodnoty reziduí na vysvětlující proměnné. Nejčastěji mají tyto **pomocné regrese** tvar:
- $$|e_i| = \beta_0 + \beta_1 X_i + \eta_i$$
- $$|e_i| = \beta_0 + \beta_1 \sqrt{X_i} + \eta_i$$
- $$|e_i| = \beta_0 + \beta_1 (1/X_i) + \eta_i$$
- kde η_i je náhodná složka. Pak testujeme významnost parametru β_1 klasickým t-testem. Je-li parametr statisticky významný, vyskytuje se v modelu heteroskedasticita. Výhodou testu je, že nám pomůže určit i konkrétní formu heteroskedasticity, a tedy i poradí, jak ji odstranit.
- IV. **Parkův test** je podobný jako Glejserův test: také sestavuje **pomocnou regresi**, a to ve tvaru $\ln e_i^2 = \beta_0 + \beta_1 \ln X_i + \eta_i$. Nulovou hypotézu o homoskedasticitě zamítneme tehdy, pokud je parametr β_1 v modelu dle t-testu statisticky významný.
- V. **Whiteův test** je obecnější, protože není potřeba vycházet z žádného předpokladu o tvaru heteroskedasticity ani z předpokladu závislosti rozptylu reziduí jen na jedné proměnné. Sestaví se pomocná regrese čtverců reziduí na exogenních proměnných, jejich druhých mocninách a jejich násobcích, tedy např. pro dvě vysvětlující proměnné bude mít tvar:
- $$e_i^2 = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{1i}^2 + \beta_4 X_{2i}^2 + \beta_5 X_{1i} X_{2i} + \eta_i$$
- Zjistí se koeficient determinace z této pomocné regrese. Testová statistika má pak tvar $LM = nR_e^2$. Tuto statistiku porovnáme s kritickou hodnotou rozdělení chí-kvadrát pro q stupňů volnosti, kde q značí počet vysvětlujících proměnných v pomocné regresi zmenšený o jednotku. Je-li $LM >$ kritická hodnota, znamená to, že alespoň jeden parametr v pomocné regresi je statisticky významně odlišný od nuly, a zamítáme tak nulovou hypotézu o homoskedasticitě.
- Test vyžaduje, aby byl původní model správně specifikován, jinak se může projevit heteroskedasticita jen zdánlivě, právě kvůli špatné specifikaci. Potřebuje dostatek pozorování, alespoň 30. Při menším počtu pozorování lze vynechat součiny vysvětlujících proměnných z pomocné regrese.
- VI. **Breusch-Paganův test** sestavuje pomocnou regresi čtverců reziduí původního modelu na vysvětlujících proměnných (některých či všech) či jejich funkcích (třeba druhých mocninách). Tento funkční vztah si musíme sami určit, podle toho, které proměnné a v jaké formě mají nejspíš vliv na rozptyl náhodné složky. Pak testujeme statistikou významnost koeficientu determinace z této pomocné regrese. Spočítáme testovou statistiku: nR^2 , která má chí-kvadrát rozdělení s počtem volnosti $p - 1$, kde p značí počet parametrů v této pomocné regresi. Je-li testová statistika větší než kritická hodnota chí-kvadrát rozdělení, pak zamítáme nulovou hypotézu o homoskedasticitě.

Jak z modelu odstranit heteroskedasticitu?

Je potřeba najít transformační matici **T**.

- pokud **známe rozptyly reziduí**, můžeme jimi všechny proměnné v modelu vydělit a odhadnout upravený model. Rozptyl takto odhadnutého modelu bude již konstantní. Tato metoda se nazývá metoda vážených nejmenších čtverců.

$$\frac{Y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{X_{i2}}{\sigma_i} + \dots + \beta_k \frac{X_{ik}}{\sigma_i} + \frac{u_i}{\sigma_i}$$

$$Y_i^* = \beta_1 X_{i1}^* + \beta_2 X_{i2}^* + \dots + \beta_k X_{ik}^* + u_i^*$$

Rozptyl takto upraveného modelu je už konstantní, protože platí:

$$\text{var}(u_i^*) = \text{var}(u_i/\sigma_i) = \text{var}(u_i)/\sigma_i^2 = \sigma_i^2/\sigma_i^2 = 1$$

- pokud jsou **rozptyly reziduí neznámé**, jako je tomu ve většině případů, můžeme vyjít jen z předpokladu o určitém tvaru heteroskedasticity. Tedy určíme s pomocí některého z výše uvedených testů funkční formu závislosti rozptylu na vysvětlujících proměnných.
- v případě lineární závislosti, tzn. $E(u_i^2) = \sigma^2 X_i$, můžeme původní model vydělit odmocninou z X_i . Transformační matice má pak tvar:

$$\mathbf{T} = \begin{pmatrix} 1/\sqrt{X_1} & 0 & 0 & 0 \\ 0 & 1/\sqrt{X_2} & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1/\sqrt{X_n} \end{pmatrix}$$

- v případě kvadratické závislosti, tzn. $E(u_i^2) = \sigma^2 X_i^2$, můžeme původní model vydělit X_i . Transformační matice má pak tvar:

$$\mathbf{T} = \begin{pmatrix} 1/X_1 & 0 & 0 & 0 \\ 0 & 1/X_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1/X_n \end{pmatrix}$$

Náhodná složka v tomto modelu bude již homoskedastická. Po odhadu parametrů přejdeme zpětnou transformací k původnímu modelu. Mezi transformovanými proměnnými se totiž může objevit zdánlivá korelace a je lepší pracovat s původními proměnnými.

- V určitých případech může pomoci **logaritmická transformace**. Ta snižuje heteroskedasticitu modelu, protože snižuje stupnici, v níž jsou proměnné měřeny.
- Můžeme odhadnout tzv. **robustní směrodatné chyby** odhadovaných parametrů, které zůstávají u velkých výběrů konzistentní i v případě heteroskedasticity (a lze je tak použít k vytváření intervalů spolehlivosti a testům hypotéz).

AUTOKORELACE

Autokorelace znamená závislost mezi různými hodnotami jedné proměnné, tzn. $cov(u_t, u_s) = E(u_t u_s) \neq 0$. Prvky mimo diagonálu kovarianční matice náhodných složek v tom případě nejsou nulové. Autokorelace se nejčastěji vyskytuje:

- v časových řadách: hodnoty různých ekonomických veličin apod. vykazují obvykle pozitivní autokorelaci kvůli setrvačnosti, takže nezahrnutí zpožděných proměnných do modelu se projeví autokorelací náhodné složky (kvaziautokorelace);
- chybná specifikace (např. lineární funkce místo kvadratické) se též může projevit autokorelací;
- chyby měření;
- zprůměrovaná, extrapolovaná či jinak upravená data;
- zahrnutí zpožděných endogenních či exogenních proměnných do modelu (náhodné složky pak mohou být sériově závislé).

Autokorelace způsobuje obdobné problémy jako heteroskedasticita: odhady parametrů nejsou vydatné a standardní chyby a odhad rozptylu modelu jsou vychýlené. Při pozitivní autokorelaci jsou obvykle odhady rozptylu podhodnocené.

Při **autokorelaci prvního řádu** je náhodná složka generována vztahem $u_t = \rho u_{t-1} + \varepsilon_t$ (AR(1) proces). Pro kovariance (mimodiagonální prvky) platí: $E(u_t, u_{t-1}) = \rho \sigma^2$. Kovariance nezávisí na období (například mezi roky 2013 a 2012 musí být stejná jako mezi roky 2012 a 2011 apod., protože jde vždy o dvě těsně po sobě jdoucí pozorování). Dále musí platit, že $|\rho| < 1$, protože pro $|\rho| > 1$ by AR proces nebyl stacionární. Při kladném ρ jde o pozitivní autokorelaci, při záporném ρ jde o negativní autokorelaci.

Při **autokorelaci druhého řádu** je náhodná složka generována vztahem $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$. V případě **čtvrtletních dat** může mít AR proces tvar např. $u_t = \rho u_{t-4} + \varepsilon_t$. Nemusí jít nutně o AR proces, ale i o MA proces.

Jak zjistit, jestli je v modelu autokorelace?

- v případě autokorelace prvního řádu se můžeme podívat na odhad koeficientu autokorelace ρ (měl by se blížit nule a bývá automaticky součástí výstupu softwaru)
- V případě autokorelace prvního řádu je častou metodou **Durbin Watsonova statistika**, která má tvar:

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$$

V čitateli je součet čtverců rozdílů sousedních reziduí a ve jmenovateli je reziduální součet čtverců. Spočtenou statistiku d srovnáme s kritickými hodnotami z DW tabulky pro daný počet pozorování a vysvětlujících proměnných: najdeme dolní mez d_L a horní mez d_U . DW statistika má symetrické rozdělení od 0 do 4 se střední hodnotou rovnou 2. Je-li spočtená statistika nižší než dolní mez nebo vyšší než horní mez, zamítáme nulovou hypotézu o neexistenci autokorelace. Ideálně by se hodnota d měla blížit 2. Při hodnotách blízkých nule bude v modelu nejspíš pozitivní autokorelace (protože rozdíl dvou po sobě jdoucích reziduí se bude blížit nule, takže čítec zlomku bude také blízký nule). Při hodnotách blízko 4 tam bude negativní autokorelace. Aby bylo možné DW statistiku použít, nesmí být autokorelace nelineární, v modelu nesmí být zpožděná endogenní proměnná a musí jít o autokorelaci prvního řádu.

pozitivní autokorelace	pásma neprůkaznosti	nulová autokorelace	pásma neprůkaznosti	negativní autokorelace
0	d_L	d_U	2	4
		$4 - d_U$		$4 - d_L$
				4

- pokud je v modelu zpožděná hodnota vysvětlující proměnné, můžeme použít **Durbinovo h**. Testová statistika má tvar: $h = (1 - 0,5d) \sqrt{\frac{T}{T - s_b^2}}$, kde d je DW statistika, s_b^2 je odhad rozptylu bodového odhadu u zpožděné endogenní proměnné, T je počet pozorování. Tuto hodnotu porovnáme s tabulkami normálního rozdělení.
- Breusch-Godfrey test lze použít i pro autokorelaci vyšších řádů. Nulovou hypotézou je, že jsou všechny autokorelační koeficienty $\rho_1 \dots \rho_q$ rovny 0. Odhadneme původní model a uložíme si rezidua. Pak odhadneme pomocnou regresi: vysvětlovanou proměnnou bude e_t , vysvětlujícími pak všechny proměnné z původního modelu a zpožděná rezidua $e_{t-1} \dots e_{t-q}$. Zjistíme koeficient determinace R^2 z této pomocné regrese. Testová statistika má tvar $(T - q)R^2$ a má chí-kvadrát rozdělení s q stupni volnosti, kde q je počet řádů autokorelace. Překročí-li testová statistika kritickou hodnotu, zamítneme nulovou hypotézu o neexistenci autokorelace.
- Pro autokorelaci vyšších řádů (jen pro AR procesy) lze také použít Boxův-Piercův test, který také vychází z odhadu reziduí pro $t = 1, 2 \dots T$. Nejprve se odhadne koeficient autokorelace j -tého řádu jako r_j pro $j = 1, 2 \dots p$ a z se pak spočítá testová statistika jako $Q = T \sum_{j=1}^p r_j^2$ a ta se srovnává s kritickými hodnotami chí-kvadrát rozdělení s p stupni volnosti. Modifikací pro menší vývěry je Ljung-Box test, který je součástí řady programů.

Jak odstranit z modelu autokorelaci?

Nejprve musíme zkontrolovat, že jsme model správně specifikovali (že by nebyla třeba vhodnější kvadratická funkce či logaritmický model) a že jsme nezapomněli zahrnout nějakou vysvětlující proměnnou (třeba některou zpožděnou hodnotu). Pokud je specifikace modelu správná, nezbude nic jiného než model opět vhodně transformovat pomocí nějaké transformační matice.

- Pro autokorelaci prvního řádu můžeme použít **Prais-Winstonovu transformaci**.

- Mějme model $Y_t = \beta_0 + \beta_1 X_t + u_t$, kde $u_t = \rho u_{t-1} + \varepsilon_t$
- Vyjádříme si model v čase $t - 1$ a vynásobíme ho koeficientem autokorelace ρ , takže dostaneme model: $\rho Y_{t-1} = \rho \beta_0 + \rho \beta_1 X_{t-1} + \rho u_{t-1}$
- Tento model odečteme od původního modelu a dostaneme model:

$$Y_t - \rho Y_{t-1} = \beta_0(1 - \rho) + \beta_1(X_t - \rho X_{t-1}) + u_t - \rho u_{t-1}$$

$$Y_t - \rho Y_{t-1} = \beta_0(1 - \rho) + \beta_1(X_t - \rho X_{t-1}) + \varepsilon_t$$

$$Y_t^* = \beta_0^* + \beta_1^* X_t^* + \varepsilon_t$$

Tento model lze odhadnout MNČ, protože náhodná složka modelu ε_t již splňuje GM předpoklady. Problém je s prvním pozorováním: abychom ho neztratili, transformuje se vztahem $Y_1^* = Y_1 \sqrt{1 - \rho^2}$ a $X_1^* = X_1 \sqrt{1 - \rho^2}$. V tom případě bude mít transformační matice tvar:

$$\mathbf{T} = \frac{1}{\sqrt{1-\rho^2}} \begin{bmatrix} v(1-\rho^2) & 0 & 0 & 0 & 0 \\ -\rho & 1 & 0 & 0 & 0 \\ 0 & -\rho & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & -\rho & 1 \end{bmatrix}$$

Výraz před maticí je konstanta a lze ji vynechat. Prvky upraveného modelu kromě prvního sloupce mají podobu tzv. **zobecněných diferencí**. Koeficient autokorelace ρ většinou neznáme, můžeme jej ale nahradit odhadem koeficientu autokorelace r , pro který platí, že je roven přibližně $1 - 0,5d$ (d značí DW statistiku).

- Metoda **Cochrane-Orcutt** je iterativní metoda použitelná i pro autokorelaci vyšších řádů. Nejprve odhadneme původní model $Y_t = \beta_0 + \beta_1 X_t + u_t$, uložíme si rezidua a odhadneme koeficient autokorelace. V případě autokorelace prvního řádu jsou náhodné složky (a tedy i rezidua) generována vztahem $e_t = \rho e_{t-1} + \eta_t$. Konzistentní odhadová funkce pro koeficient autokorelace prvního řádu je $r = \frac{\sum_{t=2}^T e_t e_{t-1}}{\sum_{t=2}^T e_{t-1}^2}$. Model se stejně jako výše převede na tvar zobecněných diferencí s pomocí takto odhadnutého koeficientu r : $Y_t - rY_{t-1} = \beta_0(1-r) + \beta_1(X_t - rX_{t-1}) + u_t - ru_{t-1} \rightarrow Y_t^* = \beta_0^* + \beta_1 X_t^* + u^*$. Protože ale výše odhadnutý koeficient r nemusí být moc přesný, dosadíme nové odhady parametrů získané z této regrese do původního modelu a spočítáme nová rezidua. Uložíme si nově získaná rezidua a celý postup opakujeme. Pokračujeme tak dlouho, až se odhadnuté parametry ve dvou po sobě jdoucích krocích moc neliší.

Viz též http://www.nyu.edu/classes/nagler/quant2/notes/cochran_orcutt.pdf

- **Durbinova dvoustupňová metoda** je založena na podobném principu. Těmto dvoustupňovým procedurám se říká **metoda odhadnutých zobecněných nejmenších čtverců**.
- Lze použít i jiné metody vycházející z metody nelineárních nejmenších čtverců či MMV.

MULTIKOLINEARITA

Perfektní kolinearita znamená, že sloupce v matici vysvětlujících proměnných jsou perfektně lineárně závislé. Pak bude determinant této matice nulový a nebude ji možné invertovat, a tedy ani použít MNČ. Pokud jsou některé sloupce „jen silně lineárně závislé,“ jde o multikolinearitu. Determinant matice vysvětlujících proměnných tak sice bude také blízky nule, ale inverzní matici spočítat dokážeme.² To může být ale také problém. Odhady sice budou nestranné i vydatné, ale budou velmi nestabilní (tzn. budou citlivé i na malé změny v matici vysvětlujících proměnných). Tedy při jiném výběru z téže populace se odhady mohou velmi lišit. Druhým problémem jsou velké směrodatné chyby odhadnutých parametrů, takže některé parametry se mohou jevit jako statisticky nevýznamné, a pak nevíme, jestli některé proměnné vyřadit a případně které. Příčinou může být to, že kvůli silné vzájemné závislosti proměnných často nedokážeme určit vliv jedné konkrétní proměnné na vysvětlovanou proměnnou, proto se může jevit jako nevýznamná. Dokážeme určit jen společný vliv všech proměnných.

Příčinou multikolinearity je:

- podobný vývoj makroekonomických veličin (při analýze časových řad);
- souvislost veličin (v průřezové analýze), např. kdybychom vysvětlovali výdaje domácnosti výší příjmu a výší bohatství (tyto dvě proměnné souvisí), nebo objem výstupu počtem pracovníků a hodnotou fixního kapitálu;
- zahrnutí zpožděných hodnot, ty bývají totiž zkorelovány;
- zahrnutí nesprávného počtu umělých proměnných (pak se vyskytne perfektní kolinearita).

Problém multikolinearity se týká jednoho konkrétního výběru, takže mluvíme o jejím zjišťování v daném výběru, nikoli testování.

Jak zjišťujeme, zda je ve výběru multikolinearita?

- „podezřelá“ je situace, když je R^2 vysoký, ale některé proměnné se v modelu jeví jako nevýznamné. Nemusí to být však způsobeno pouze multikolinearitou.
- Pokud má model jen dvě vysvětlující proměnné, tak spočítáme jejich párový korelační koeficient. Když překročí hodnotu 0,8, popř. 0,9, pak multikolinearita již není v modelu únosná.
- Pokud má model více vysvětlujících proměnných, pak tento postup nestačí (závislost nemusí být mezi žádnou dvojicí proměnných, ale mezi všemi jako celkem). V tomto případě spočítáme pomocné regrese každé z vysvětlujících proměnných na zbylých vysvětlujících proměnných modelu a podíváme se na R^2 z těchto dílčích regresí. Pokud je mezi proměnnými více než jedna lineární závislost, tak nám to ale stejně moc nepomůže.
 - empirickým pravidlem je, že by žádný z koeficientů determinace z těchto regresí neměl překročit koeficient determinace z původní regrese
 - významnost R^2 z dílčích regresí lze ověřit i F -testem
- Nebo lze matici vysvětlujících proměnných normovat, tzn. spočítat matici $X^* = (X_{ji} - \bar{X}_j)(s_{bj}/\sqrt{n})$, a pak měřit odchylku této matice od ortogonality. Jsou-li všechny sloupce matice X^* nezávislé, pak by determinant matice $X^{*'}X^*$ měl být roven 1, při perfektní lineární závislosti naopak nule. Otázka je, jaká je kritická hodnota.

² Opakem k perfektní kolinearitě je případ ortogonálních vysvětlujících proměnných, kdy jsou sloupce matice zcela nezávislé, tzn. jejich kovariance je nulová.

Co dělat, když je v modelu multikolinearita?

- **zvětšit rozsah výběru**, pokud to jde, například zahrnout i extrémní pozorování (nové hodnoty se musí lišit od těch v původním výběru);
- přidat **dodatečnou apriorní informaci** (apriorní omezení), díky kterému pak snížíme směrodatné chyby odhadových funkcí (např. omezení parametrů produkčních či poptávkových funkcí atd.);
- použít tzv. „**smíšený odhad modelu**“ – zkombinovat průřezová data a údaje z časových řad;
- **změnit specifikaci**, tedy vypustit některé proměnné, které se jeví jako nevýznamné, ale pozor na to, že tím může vzniknout specifikační chyba;
- **transformovat původní proměnné** (například pomocí prvních diferencí nebo podílů, ale často tak do modelu zase dostaneme autokorelaci nebo heteroskedasticitu);
- použít **metody vícerozměrné statistiky** (metodu hlavních komponent, faktorovou analýzu, hřebenovou regresi).

První tři řešení bychom měli preferovat, protože měnit specifikaci modelu může být riskantní a nesprávná specifikace modelu má horší dopady než multikolinearita. Ta v podstatě ani moc nevadí, pokud chceme model použít k prognózování, v případě, že vztahy mezi proměnnými se nebudou nějak dramaticky měnit a proměnné budou i nadále lineárně závislé.

ZDROJE

Hušek, R: Ekonometrická analýza. Nakladatelství Oeconomica, Praha 2007.

Krkošková, Š., Ráčková, A., Zouhar, J.: Základy ekonometrie v příkladech. Nakladatelství Oeconomica, Praha 2010.