



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Daniel Borák

Kauzalita, míry kauzality

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Petr Lachout, CSc.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2023

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne

.....

Podpis autora

Na tomto místě bych rád poděkoval vedoucímu své bakalářské práce doc. RNDr. Petru Lachoutovi, CSc., za jeho trpělivost, pohotovost a celkovou vstřícnost během psaní bakalářské práce.

Název práce: Kauzalita, míry kauzality

Autor: Daniel Borák

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Petr Lachout, CSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Míry kauzality jsou užitečným nástrojem při hledání kauzality v časových řadách. Tato práce nejenom popisuje teorii za definicí měr kauzality, ale i dává vyčerpávající návod, jak míry kauzality použít k hledání možných kauzálních vztahů. Je důležité si uvědomit, že míry kauzality sice mohou poukázat na možné kauzální vztahy, nicméně je nemohou potvrdit. Kauzalita je komplexní vlastnost, kterou nelze postihnout pouze daty – je třeba také použít experimenty a fyzikální zkušenosti.

Klíčová slova: Kauzalita; Grangerova kauzalita; Korelovanost

Title: Causality, Measures of Causality

Author: Daniel Borák

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Petr Lachout, CSc., Department of Probability and Mathematical Statistics

Abstract: Causality measures are useful tools when looking for causality in time series. This thesis does not only describe the theory behind the definition of measures of causality but also gives comprehensive instructions on how to use measures of causality to search for possible causal conditions. It is important to note that although measures of causality can point to possible causal relationships, they cannot confirm them. Causality is a complex relationship that cannot be captured by data alone – experiments and physical experience must also be considered.

Keywords: Causality; Granger causality; Correlation

Obsah

Seznam často používaných zkratk	2
Úvod	3
1 Základní definice a vlastnosti	4
1.1 \mathcal{L}^2 prostor	4
1.2 Kauzalita	6
1.3 VAR(p) a VMA(q) procesy	9
2 Kauzalita VAR(p) procesu	11
3 Míry kauzality	16
4 Odhady měr kauzality	23
4.1 Nástroje	23
4.2 Bodový odhad $\Sigma[X(t+h) I_{XZ}]$	24
4.3 Bodový odhad měr kauzality	27
4.4 Intervaly spolehlivosti	31
4.5 Aplikace odhadů měr kauzality	34
Závěr	42
Zdroje	43
Přílohy	44

Seznam často používaných zkratek

E	střední hodnota
\langle , \rangle	skalární součin
s. j.	skoro jistě, až na množinu pravděpodobnostní míry 0
\mathcal{L}^2	prostor tříd ekvivalence „rovná se s. j. až na konstantu“ veličin s konečným rozptylem
var	rozptyl náhodné veličiny
Var	rozptylová matice náhodného vektoru
cov	kovariance náhodných veličin nebo kovarianční matice náhodných vektorů
1_m	jednotková matice dimenze $m \times m$
0_m	nulový vektor dimenze m
$\text{VAR}(p)$	vektorový autoregresivní proces řádu p
$\text{VMA}(q)$	vektor moving average process řádu q
π_j	j -tá autoregresivní matice (daného $\text{VAR}(p)$ procesu)
π	posloupnost autoregresivních matic (daného $\text{VAR}(p)$ procesu)
Σ_a	rozptylová matice šumu a (daného $\text{VAR}(p)$ procesu)
\mathcal{N}	množina přirozených čísel
\mathcal{Z}	množina celých čísel
\mathcal{R}	množina reálných čísel

Úvod

Mějme dvě události, které lze popsat formou diskrétní časové řady (např. kurzy měn EUR:USD a CZK:USD). Zajímá nás, jestli jedna z těchto událostí neovlivňuje tu druhou: jestli např. nárůst hodnoty eura nezvýší nárůst hodnoty české koruny.

V této práci si vysvětlíme, jak matematicky popsat kauzalitu. Toto nám přirozeně rozšíří definici i pro více událostí. Toho využijeme k definici měř kauzality; kauzalita odpovídá na otázku, jestli se události ovlivňují a jak moc. Nakonec ukážeme způsob, jak z dat vypočítat míry kauzality (bodové i intervalové odhady) a postup demonstrujeme na praktickém příkladu.

V této práci se budeme opírat mimo jiné o funkcionální analýzu, teorii pravděpodobnosti a matematické statistiky a teorii VAR procesů. Většina důkazů je vytvořena autorem této práce, ať už jsou originální, nebo přepsané a uvedené mnohem podrobněji.

1. Základní definice a vlastnosti

1.1 \mathcal{L}^2 prostor

Definice 1. *Bud' $m, n \in \mathcal{N}$. Označme 0_m nulový vektor dimenze m , 1_m jednotkovou matici dimenze $m \times m$ a $0_{m \times n}$ nulovou matici dimenze $m \times n$.*

Definice 2. *Nechť (Ω, \mathcal{F}, Q) je pravděpodobnostní prostor, kde Ω je množina všech elementárních jevů, \mathcal{F} je σ -algebra jevů z množiny Ω a Q je pravděpodobnostní míra na \mathcal{F} . Definujeme náhodnou veličinu y jako reálnou funkci na Ω splňující podmínku měřitelnosti*

$$\{\omega \in \Omega : y(\omega) \leq c\} \in \mathcal{F}$$

pro každé $c \in \mathcal{R}$. Nechť n je přirozené číslo, pak definujeme n -rozměrný náhodný vektor jako n -rozměrný vektor náhodných veličin. Budeme říkat, že náhodná veličina, resp. vektor vychází z prostoru (Ω, \mathcal{F}, Q) .

Definice 3. *Bud' n přirozené číslo. Definujeme n -rozměrný diskrétní náhodný proces jako funkci*

$$X : \mathcal{Z} \times \mathcal{R} \rightarrow \mathcal{R}^n,$$

kde $X(t, \omega)$ pro pevně zvolené $t \in \mathcal{Z}$ nabývá podoby n -rozměrného náhodného vektoru, zatímco pro dané $\omega \in \Omega$ generuje náhodný proces posloupnost, kterou nazýváme n -rozměrnou časovou řadou.

Definice 4. *V rámci této definice, relace ekvivalence $\stackrel{\text{def}}{=}$ přenáší i neexistenci, tedy je-li pravá strana nedefinovaná, je nedefinovaná i strana levá. Bud' x a y náhodné veličiny. Symbolem E budeme rozumět střední hodnotu x ,*

$$E x \stackrel{\text{def}}{=} \int_{\Omega} x(\omega) dQ.$$

Nechť x a y mají konečné střední hodnoty. Symbolem $\text{cov}(x, y) \stackrel{\text{def}}{=} E[(x - E x)(y - E y)]$ budeme rozumět kovarianci náhodných veličin x a y , symbolem $\text{var}(x) \stackrel{\text{def}}{=} \stackrel{\text{def}}{=} \text{cov}(x, x)$ rozptyl náhodné veličiny x . Všimněme si, že pro veličiny s konečným rozptylem kovariance splňuje podmínky skalárního součinu.

Bud' nyní $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ n -rozměrný náhodný vektor a $\mathbf{y} = (y_1, y_2, \dots, y_m)'$ m -rozměrný náhodný vektor. Nechť každá složka vektoru \mathbf{x} má střední hodnotu. Střední hodnotu náhodného vektoru definujeme po složkách:

$$E \mathbf{x} \stackrel{\text{def}}{=} (E x_1, E x_2, \dots, E x_n)'$$

Nechť \mathbf{x} a \mathbf{y} mají konečné střední hodnoty. Pak $\text{cov}(\mathbf{x}, \mathbf{y})$ bude značit kovarianční matici: matici $n \times m$ splňující

$$(\text{cov}(\mathbf{x}, \mathbf{y}))_{i,j} = \text{cov}(x_i, y_j), \quad \forall i \in \{1, 2, \dots, n\}, \quad \forall j \in \{1, 2, \dots, m\}.$$

Nakonec definujeme rozptylovou matici vektoru \mathbf{x} jakožto $\text{Var}(\mathbf{x}) \stackrel{\text{def}}{=} \text{cov}(\mathbf{x}, \mathbf{x})$.

Definice 5. Necht X je vektorový prostor. Necht $\|\cdot\|: X \rightarrow [0, \infty)$ je norma na X . Pak dvojici $(X, \|\cdot\|)$ nazýváme normovaným lineárním prostorem.

Necht $\langle \cdot, \cdot \rangle: X \times X \rightarrow \mathcal{R}$ je skalární součin nad X . Pak normu $\|\cdot\|$ splňující

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}, \quad \forall \mathbf{x} \in X,$$

nazýváme normou indukovanou skalárním součinem $\langle \cdot, \cdot \rangle$.

Necht $\|\cdot\|$ je norma indukovaná skalárním součinem $\langle \cdot, \cdot \rangle$ a X je úplný v metrice příslušné $(X, \|\cdot\|)$. Pak prostor $(X, \|\cdot\|)$ nazveme Hilbertův a budeme ho značit $(X, \langle \cdot, \cdot \rangle)$.

Definice 6 (faktorizace). Bud \approx relace ekvivalence na množině \mathcal{A} . Bud $a \in \mathcal{A}$. Zadefinujeme třídu ekvivalence prvku a jako $[a] \stackrel{\text{def}}{=} \{b \in \mathcal{A} : a \approx b\}$. Dovolme značit množinu $[a]$ libovolným prvkem množiny $[a]$. Označme $\mathcal{A}_{\approx} \stackrel{\text{def}}{=} \{[a] : a \in \mathcal{A}\}$.

Definice 7. Mějme dvě náhodné veličiny X a Y . Zadefinujeme relaci ekvivalence $\stackrel{\text{cov}}{=}$, kde $X \stackrel{\text{cov}}{=} Y$, pokud $X = Y + c$ Q -skoro jistě pro nějaké $c \in \mathcal{R}$.

Definice 8. Necht \mathcal{A} je prostor všech náhodných veličin s konečným rozptylem. Pak $\mathcal{L}^2 \stackrel{\text{def}}{=} (\mathcal{A}_{\stackrel{\text{cov}}{=}}, \text{cov}(\cdot, \cdot))$.

Fakt 9. \mathcal{L}^2 je dobře definovaný Hilbertův prostor.

Kovariance není schopna rozlišit veličiny, které jsou v relaci $\stackrel{\text{cov}}{=}$. To způsobí, že některé výsledky budou jednoznačné až na $\stackrel{\text{cov}}{=}$. Faktorizací tento problém odstraníme. Prostor \mathcal{L}^2 se typicky definuje se skalárním součinem $\mathbf{E}[XY]$, nicméně my potřebujeme prostor, kde je skalární součin kovariance. Norma \mathcal{L}^2 je odmocnina z rozptylu.

Definice 10. \mathcal{L}^2 veličina je veličina s konečným rozptylem. \mathcal{L}^2 vektor je vektor \mathcal{L}^2 veličin. Normou takového vektoru budeme rozumět euklidovskou normu vektoru \mathcal{L}^2 norem jeho složek (l_2 normu).

Definice 11. Řekneme, že n -rozměrný diskrétní náhodný proces X patří do \mathcal{L}^2 (je \mathcal{L}^2), pokud pro všechny přípustné hodnoty $t \in \mathcal{Z}$ je $X(t)$ \mathcal{L}^2 vektor.

Definice 12. Bud A matice.

$$\|A\|_{2,2} \stackrel{\text{def}}{=} \sqrt{\text{tr}(AA')},$$

kde tr značí součet prvků na diagonále.

1.2 Kauzalita

Domluva 13. Pokud nebude řečeno jinak, t bude značit čas, $t \in \mathcal{Z}$. Dále n, m budou přirozená čísla. X, Y a Z budou značit diskrétní náhodné \mathcal{L}^2 procesy. Vektory nebudeme vždy vyznačovat polotučně. Jednotlivé složky vektorového procesu budeme značit $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$; analogicky pro $Y(t)$ a $Z(t)$. Procesy budou mít počátek v $-\infty$.

Definice 14. Nechť \mathcal{A} je podmnožina \mathcal{L}^2 . Symbolem $\text{span}(\mathcal{A})$ budeme rozumět množinu $\{g : g = q_1 f_1 + q_2 f_2 + \dots + q_m f_m, m \in \mathcal{N}, q_i \in \mathcal{R}, f_i \in \mathcal{A}\}$. Uzávěrem množiny \mathcal{A} myslíme nejmenší uzavřenou množinu obsahující \mathcal{A} . Platí, že uzávěr podprostoru je podprostor, span podprostoru je podprostor.

Nechť proces X je dimenze n . Symbolem $X(-\infty, t]$ budeme chápat uzávěr $\text{span}(\mathcal{B})$, kde $\mathcal{B} = \{x_i(s) : i, s \in \mathcal{Z}, -\infty < s \leq t, 1 \leq i \leq n\}$.

Systém množin $I = \{I(t) : t \in \mathcal{Z}\}$ je konformní s procesem X , pokud

$$\forall t > -\infty, I(t) \subseteq X(-\infty, t].$$

Definice 15. Uvažujme systém uzavřených podprostorů $I = \{I(t) : t \in \mathcal{Z}\}$, kde jednotlivá $I(t)$ splňují následující vlastnost:

$$\forall t > -\infty, \forall s \in \mathcal{Z}, s > t : I(t) \subseteq I(s).$$

Systém I nazýváme referenční informační systém a jednotlivá $I(t)$ referenční informační množiny.

Definice 16. Budiž $f \in \mathcal{L}^2$ a A uzavřený podprostor \mathcal{L}^2 . Ortogonální projekcí f do A (značíme $P[f|A]$) chápeme takovou $g \in A$, že $\text{cov}(f - g, h) = 0, \forall h \in A$.

Budiž $\mathbf{v} \in \mathcal{L}^2$ vektor dimenze m . Ortogonální projekci vektoru \mathbf{v} do A (značíme $P[\mathbf{v}|A]$) chápeme jako vektor ortogonálních projekcí jednotlivých složek vektoru \mathbf{v} do A : $(P[v_1|A], P[v_2|A], \dots, P[v_m|A])'$.

Fakt 17 (Riesz). V Hilbertově prostoru ortogonální projekce do uzavřeného podprostoru vždy existuje a existuje právě jedna.

Tento fakt nám říká, že ortogonální projekce je dobře definovaná v \mathcal{L}^2 . Ortogonální projekce je jednoznačná až na $\overset{\text{cov}}{=}$.

Lemma 18 (linearita projekce). Budiž f, f_1 a f_2 \mathcal{L}^2 veličiny, $r \in \mathcal{R}$ a P ortogonální projekce do I — uzavřeného podprostoru \mathcal{L}^2 . Pak P je lineární operátor v \mathcal{L}^2 : nechť $f = f_1 + f_2$, pak $P(f) = P(f_1) + P(f_2)$ a $P(rf) = rP(f)$.

Důkaz. Necht h patří do I . Pak

$$\begin{aligned} \text{cov}(f - P(f_1) - P(f_2), h) &= \text{cov}(f_1 + f_2 - P(f_1) - P(f_2), h) = \\ &= \text{cov}(f_1 - P(f_1), h) + \text{cov}(f_2 - P(f_2), h) = 0 + 0 = 0. \end{aligned}$$

Tedy $P(f_1) + P(f_2)$ je ortogonální projekcí f do I .

Dále $0 = r\text{cov}(f - P(f), h) = \text{cov}(rf - rP(f), h)$. Tedy $rP(f)$ je ortogonální projekcí rf do I . □

Definice 19. Necht A, B jsou podprostory a \mathcal{A} , resp. \mathcal{B} množiny všech vektorů obsažených v A , resp. B . Pak $A+B$ značíme uzávěr $\text{span}(\mathcal{A} \cup \mathcal{B})$.

Budiž I referenční informační systém a X, Y procesy. Pak symbolem I_X rozumíme posloupnost splňující $I_X(t) = I(t) + X(-\infty, t]$. Dále umožníme zapisovat indexy indexů zprava do hlavního indexu: $I_{X_Y} = I_{XY}$.

Definice 20 ((Grangerova) nonkauzalita). Budiž $X, Y \in \mathcal{L}^2$ procesy, I referenční informační systém konformní s X a h přirozené číslo značící horizont kauzality.

1) Řekneme, že Y nezapříčiní X v horizontu h vzhledem k I , jestliže

$$\forall t > -\infty : P[X(t+h)|I(t)] = P[X(t+h)|I_Y(t)].$$

Značíme $Y \not\rightarrow_h X|I$.

2) Řekneme, že Y nezapříčiní X do horizontu h vzhledem k I , jestliže $Y \not\rightarrow_k X|I$ pro $k = 1, 2, \dots, h$. Značíme $Y \not\rightarrow_{(h)} X|I$.

3) Řekneme, že Y nezapříčiní X vzhledem k I , jestliže $Y \not\rightarrow_k X|I$ pro $k = 1, 2, 3, \dots$. Značíme $Y \not\rightarrow_\infty X|I$.

Kauzalita (příčina) je fyzikální vlastnost. Uvažujme (fiktivní) data o zásazích požárníků a vyplácení pojistného za poškození ohněm. Z dat nám vyplývá, že přítomnost požárníků je vždy následována vyplácením pojistného. Z dat se může zdát, že požárníci způsobují vyplácení pojistky. Je to až fyzikální zkušenost, která nám říká, že ve skutečnosti oba jevy mají společnou (pro nás v datech skrytou) příčinu. Data nám tedy neumožňují potvrzovat kauzalitu. Nicméně ji umožňují vyvracet — proto definujeme nonkauzalitu. Absence nonkauzality upozorňuje na možný kauzální vztah, nicméně ho negarantuje.

Lemma 21 (vnořování projekcí). Budiž I, J uzavřené podprostory \mathcal{L}^2 , I podprostor J . Necht $f \in \mathcal{L}^2$, pak $P[P[f|J]|I] = P[f|I]$.

Důkaz. $f - [P[f|J]] \perp J$, o to víc $f - [P[f|J]] \perp I$. Pak

$$\begin{aligned} P[P[f|J]|I] &= P[f - f + P[f|J]|I] = \\ &= P[f|I] - P[f - P[f|J]|I] = P[f|I] - 0 = P[f|I]. \end{aligned}$$

V druhé rovnosti využíváme linearitu projekce. □

Lemma 22 (charakteristika nonkauzality). *Bud X, Y procesy \mathcal{L}^2 , I referenční informační systém konformní s X a h přirozené číslo. Pak následující je ekvivalentní:*

- $P[X(t+h)|I(t)] = P[X(t+h)|I_Y(t)]$,
- $X(t+h) - P[X(t+h)|I(t)]$ je kolmé na $I_Y(t)$.

Důkaz.

1 \implies 2: Z předpokladu máme

$$X(t+h) - P[X(t+h)|I(t)] = X(t+h) - P[X(t+h)|I_Y(t)],$$

tedy $X(t+h) - P[X(t+h)|I(t)]$ je z definice ortogonální projekce kolmá na $I_Y(t)$.

2 \implies 1: Bud $g \in I_Y(t)$. Pak $\text{cov}(X(t+h) - P[X(t+h)|I(t)], g) = 0$ z předpokladu. Pak ale z definice ortogonální projekce je $P[X(t+h)|I_Y(t)] = P[X(t+h)|I(t)]$. □

Lemma 23. *Nechť I je referenční informační systém konformní s procesem X . X je dimenze n , Y je dimenze m . Pak následující podmínky jsou ekvivalentní:*

1. $Y \not\rightarrow_h X|I$,
2. $Y \not\rightarrow_h x_j|I, \quad \forall j \in \{1, 2, \dots, n\}$,
3. $y_i \not\rightarrow_h X|I, \quad \forall i \in \{1, 2, \dots, m\}$,
4. $y_i \not\rightarrow_h x_j|I, \quad \forall i \in \{1, 2, \dots, m\}, \forall j \in \{1, 2, \dots, n\}$.

Důkaz. Ekvivalence mezi 1, 2 a 3, 4 plyne z definice projekce vektoru jakožto projekce jednotlivých složek.

1 \implies 3: Platí $I \subset I_{y_j} \subset I_Y$. Tedy

$$\begin{aligned} P[X(t+h)|I_{y_i}(t)] &= P[P[X(t+h)|I_Y(t)]|I_{y_i}(t)] = P[P[X(t+h)|I]|I_{y_i}(t)] = \\ &= P[X(t+h)|I(t)]. \end{aligned}$$

V první rovnosti používáme vnořování projekcí a v druhé předpoklad.

3 \implies 1: Z předpokladu a charakteristiky nonkauzality je

$$X(t+h) - [X(t+h)|I(t)] \text{ kolmé na } I_{y_i}(t), \quad \forall i \in \{1, 2, \dots, m\}.$$

Je tedy také kolmé na $I_Y(t) = I + y_1(-\infty, t] + \dots + y_m(-\infty, t]$. Tudíž z charakteristiky nonkauzality je $P[X(t+h)|I_Y(t)] = P[X(t+h)|I(t)]$. □

1.3 VAR(p) a VMA(q) procesy

Definice 24 (VAR(p) proces). *Bud' $p \in \mathcal{N} \cup \infty$. (Diskrétní) \mathcal{L}^2 proces W dimenze m se nazývá VAR(p) proces, pokud splňuje následující vlastnost:*

$$W(t) = \sum_{j=1}^p \pi_j W(t-j) + a(t), \quad \forall t > -\infty, \quad (1.1)$$

kde π_j jsou matice typu $m \times m$ a $a(t)$ je m -rozměrný náhodný vektor. Budiž Σ_a rozptylová matice typu $m \times m$. Požadujeme, aby $a(t) \perp W(-\infty, t)$, $E a(t) = 0_m$, $\text{cov}(a(t), a(s)) = 0_{m \times m}$ a $\text{Var}(a(t)) = \Sigma_a$, $\forall t \in \mathcal{Z}$, $s \in \mathcal{Z}$, $t \neq s$. Navíc, je-li p nekonečno, $\sum_{j=1}^p \pi_j W(t-j)$ musí být l^2 konvergentní sumou \mathcal{L}^2 veličin. To znamená, že

$$\exists K < \infty, \forall N \in \mathcal{N} : \|\text{Var}(\sum_{j=1}^N \pi_j W(t-j))\|_{2,2} < K.$$

Nazveme matice π_j autokovariančními maticemi procesu W a matici Σ_a rozptylovou maticí šumu procesu W . Autokovarianční matice a rozptylovou maticí šumu procesu W nazýváme parametry procesu W . O (1.1) budeme také hovořit jako o VAR(p) reprezentaci procesu W .

VAR proces znamená Vektorový AutoRegresivní proces. Podíváme-li se na data z VAR(p) procesu, suma je nenáhodnou lineární kombinací p předchozích hodnot a $a(t)$ je náhodný šum. Rozptylová matice $a(t)$ nesmí záviset na t . Je-li rozptylová matice $a(t)$ regulární, $a(t)$ se nazývá bílý šum.

Definice 25 (VMA(q) proces). *Bud' $q \in \mathcal{N} \cup \infty$. (Diskrétní) \mathcal{L}^2 proces W dimenze m se nazývá VMA(q) proces, pokud splňuje následující vlastnost:*

$$W(t) = \sum_{j=0}^q \psi_j a(t-j), \quad \forall t > -\infty, \quad (1.2)$$

kde ψ_j jsou matice typu $m \times m$, ψ_0 je jednotková matice. Požadavky na $a(t)$ jsou stejné jako v předchozí definici. Je-li q nekonečno, $\sum_{j=1}^q \psi_j a(t-j)$ musí být l^2 konvergentní sumou \mathcal{L}^2 veličin. O (1.2) budeme také hovořit jaké o VMA(q) reprezentaci procesu W .

„VMA process“ v angličtině znamená „Vector Moving Average process“. Hodnota procesu v čase t je jakýmsi váženým průměrem šumu $a(t)$ v q předchozích krocích.

Definice 26. *Bud' $(b(t))_{t \in \mathcal{Z}}$ a $\pi = (\pi_j)_{j=0}^{\infty}$ maticové řady. Označme MAT prostor všech maticových řad indexovaných celým číslem. Definujeme operátor posunutí $B: \text{MAT} \rightarrow \text{MAT}$ následujícím předpisem:*

$$B^n b(t) \stackrel{\text{def}}{=} b(t-n), \quad t \in \mathcal{Z}, n \in \mathcal{N}_0.$$

Definujeme dosazení operátoru do řady: $\pi(B) \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} \pi_j B^j$.

Mějme VAR(p) proces W jako z definice 24. Zdefinujeme maticovou řadu θ následovně: $\theta_0 = 1_m$, $\theta_j = -\pi_j$, $j \in \mathcal{N}$, $1 \leq j \leq p$ a $\theta_j = 0_{m \times m}$, $j > p$. Proces $W(t)$ můžeme přepsat jako

$$\theta(B)W(t) = \sum_{j=0}^p \theta_j W(t-j) = a(t).$$

Za určitých předpokladů existuje l^2 konvergentní řada ψ taková, že $\psi(B)\theta(B) = 1_m B^0$ — identita. V takovém případě můžeme $\psi(B)$ vynásobit zleva a dostat VMA reprezentaci procesu W :

$$W(t) = \psi(B)a(t).$$

Definice 27. Buď $W(t)$ proces. Řekneme, že proces je (slabě) stacionární, pokud splňuje následující podmínky:

$$\begin{aligned} E[W(t)^2] &< \infty, & \forall t \in \mathcal{N}, \\ E[W(t)] &= E[W(s)], & \forall t, s \in \mathcal{N}, \\ \text{cov}(W(t), W(s)) &= \text{cov}(W(0), W(t-s)), & \forall t, s \in \mathcal{N}. \end{aligned}$$

Dále, buď I referenční informační systém. Necht' $\forall t, s \in \mathcal{Z}$, $t < s$: $W(\infty, t] \perp \perp I(s)$. Pak řekneme, že I je nekorelovaný s procesem W a budeme značit $I \perp W$.

Definice 28. Buď $W(t)$ VAR(p) nebo VMA(q) proces. Řekneme, že proces je regulární, pokud $\det \Sigma_a$ je nenulový. Buď nyní $W(t)$ VMA(q) proces. Řekneme, že proces $W(t)$ je invertibilní, pokud má VAR(p) reprezentaci.

2. Kauzalita VAR(p) procesu

Definice 29. (Diskrétní) \mathcal{L}^2 proces $W(t)$ dimenze m se nazývá zobecněný VAR proces, pokud splňuje následující vlastnost:

$$W(t) = \mu(t) + \sum_{j=1}^{\infty} \pi_j W(t-j) + a(t), \quad \forall t > -\infty,$$

kde π_j jsou matice $m \times m$, $a(t)$ a $\mu(t)$ m -rozměrné náhodné vektory. Označme H uzavřený podprostor \mathcal{L}^2 . Požadujeme, aby pro všechny přípustné hodnoty t bylo $\mu(t) \in H$. Dále požadujeme, aby $E a(t) = 0$, $\text{cov}(a(t), a(s)) = 0$, $a(t) \perp H + W(-\infty, t)$, $\forall t \neq s$, $t \in \mathcal{Z}$, $s \in \mathcal{Z}$. Navíc, $\sum_{j=1}^{\infty} \pi_j W(t-j)$ musí být l^2 konvergentní sumou \mathcal{L}^2 veličin.

Oproti VAR(p) procesu, v zobecněném VAR procesu povolujeme, aby $a(t)$ mělo v různém čase jinou rozptylovou matici a povolujeme existenci $\mu(t)$ — typicky konstanta. I když poznatky z této kapitoly platí i pro zobecněné VAR procesy, v této práci se omezíme na práci s VAR(p) procesy. Zdroj [2] uvádí verze vět pro zobecněné VAR procesy.

Pozorování 30. Necht f je \mathcal{L}^2 funkce a P ortogonální projekce do I , uzavřeného podprostoru \mathcal{L}^2 . Pak $\|f\| \geq \|P(f)\|$.

Důkaz. Z Pythagorovy věty $\|f\|^2 = \|P(f)\|^2 + \|f - P(f)\|^2$. Norma $f - P(f)$ je nezáporná. □

Lemma 31 (spojitost ortogonální projekce). Bud $(f_n)_{n=1}^{\infty}$ posloupnost \mathcal{L}^2 veličin s limitou f . Bud P ortogonální projekce do I , uzavřeného podprostoru \mathcal{L}^2 . Pak $\lim_{n \rightarrow \infty} P(f_n) = P(f)$.

Vzdálenost \mathcal{L}^2 funkcí měříme pomocí \mathcal{L}^2 normy.

Důkaz. Zvolme ε a k němu N takové, že $\|f - f_n\| \leq \varepsilon$ pro $n \geq N$. Pak

$$\|P(f) - P(f_n)\| = \|P(f - f_n)\| \leq \|f - f_n\| \leq \varepsilon.$$

V rovnosti používáme linearitu ortogonální projekce, v první nerovnosti pozorování 30. □

Příklad 32. Uvažujme W – m -rozměrný VAR(p) proces a referenční informační systém $I \perp W$. Bud $h \in \mathcal{N}$. Spočteme $P[W(t+h)|I_W(t)]$.

BÚNO uvažujme $p = \infty$. Jinak rozšíříme proces na VAR(∞) proces tak, že zvolíme π_j nulové matice příslušné dimenze pro $j \in \{p+1, p+2, \dots\}$.

Uvažujme $N \in \mathcal{N}$, označíme $\|\sum_{j=N}^{\infty} \pi_j W(t+h-j)\| = \varepsilon$. S rostoucím N ε konverguje k 0 (z konvergence sumy). Z pozorování 31 je

$$\|P[\sum_{j=N}^{\infty} \pi_j W(t+h-j)|I_W(t)]\| \leq \varepsilon.$$

Z nekorelovanosti $a(t+h)$ s $W(-\infty, t)$ a I dostáváme

$$P[a(t+h)|I_W(t)] = 0_m.$$

Nakonec, z konformnosti I_W vůči W je

$$P[W(t-r)|I_W(t)] = W(t-r), \quad \forall r \in \mathcal{N}_0.$$

Uvažujme $h = 1$.

$$\begin{aligned} P[W(t+1)|I_W(t)] &= P[\sum_{j=1}^{N-1} \pi_j W(t-j+1) + \sum_{j=N}^{\infty} \pi_j W(t-j+1) + \\ &+ a(t+1)|I_W(t)] = P[\sum_{j=1}^{N-1} \pi_j W(t-j+1)|I_W(t)] + P[\sum_{j=N}^{\infty} \pi_j W(t-j+1)|I_W(t)] + \\ &+ P[a(t+1)|I_W(t)] = \sum_{j=1}^{N-1} \pi_j W(t-j+1) + P[\varepsilon|I_W(t)] + 0_m. \end{aligned}$$

Pošleme-li N do nekonečna, $P[\varepsilon|I(t)]$ zkonverguje k 0_m a ze spojitosti ortogonální projekce dostáváme

$$P[W(t+1)|I_W(t)] = \sum_{j=1}^{\infty} \pi_j W(t-j+1). \quad (2.1)$$

Nyní uvažujme h libovolné. Pro $N > h$ máme

$$\begin{aligned} P[W(t+h)|I_W(t)] &= P[\sum_{j=1}^{h-1} \pi_j W(t-j+h)|I_W(t)] + P[\sum_{j=h}^{N-1} \pi_j W(t-j+h)|I_W(t)] + \\ &+ P[\sum_{j=N}^{\infty} \pi_j W(t-j+h)|I_W(t)] + P[a(t+h)|I_W(t)] = \sum_{j=1}^{h-1} \pi_j P[W(t-j+h|I_W(t)] + \\ &+ \sum_{j=h}^{N-1} \pi_j W(t-j+h) + P[\varepsilon|I_W(t)] + 0_m. \end{aligned}$$

Analogicky jako výše dostaneme

$$P[W(t+h)|I_W(t)] = \sum_{j=1}^{h-1} \pi_j P[W(t-j+h|I_W(t)] + \sum_{j=h}^{\infty} \pi_j W(t-j+h). \quad (2.2)$$

Projekce musí být (nekonečnou) lineární kombinací $W(t+1-j)$, $j = \{1, 2, \dots\}$:

$$P[W(t+h)|I_W(t)] = \sum_{j=1}^{\infty} \pi_j^{(h)} W(t+1-j). \quad (2.3)$$

Dosazením rovnice (2.3) do (2.1) a porovnáním koeficientů před $W(t+1-j)$ dostáváme $\pi_j^{(1)} = \pi_j$. Indukcí, dosazením do (2.2) a porovnáním koeficientů dostáváme $\pi_j^{(h+1)} = \pi_{j+h} + \sum_{i=1}^h \pi_{h-i+1} \pi_j^{(i)}$. Bude se hodit dodefinovat $\pi_j^{(0)} = 1_m$. Platí, že $\pi_j^{(h+1)}$ lze ekvivalentně definovat jako $\pi_j^{(h+1)} = \pi_{j+1}^{(h)} + \pi_1^{(h)} \pi_j$.

Definice 33. $W(t)$, m -rozměrný VAR(p) proces, si rozdělíme na tři podprocesy: $W(t) = (X(t)', Y(t)', Z(t)')$. Budeme zkoumat kauzalitu procesu Y na X . Proces Y může na X působit nepřímě přes Z . Označíme si dimenze procesů X, Y, Z po řadě m_1, m_2, m_3 , dále $m = m_1 + m_2 + m_3$. Označíme si bloky matice $\pi_j^{(h)}$ z příkladu 32 jako

$$\pi_j^{(h)} = \begin{pmatrix} \pi_{XXj}^{(h)} & \pi_{XYj}^{(h)} & \pi_{XZj}^{(h)} \\ \pi_{YXj}^{(h)} & \pi_{YYj}^{(h)} & \pi_{YZj}^{(h)} \\ \pi_{ZXj}^{(h)} & \pi_{ZYj}^{(h)} & \pi_{ZZj}^{(h)} \end{pmatrix}.$$

Význam bloků nám osvětlí následující věta.

Věta 34. Mějme referenční informační systém I . Necht W je VAR proces. Budiž $W \perp I$. Pak podmínka

$$\pi_{XYj}^{(h)} = 0, \quad \forall j \in \mathcal{N},$$

je postačující pro $Y \not\rightarrow_h X | I_{XZ}$.

Důkaz. Označme $J(t) = W(\infty, t]$. Aplikací postupu z příkladu 32 dostaneme

$$P[W(t+h) | J(t)] = \sum_{j=1}^{\infty} \pi_j^{(h)} W(t+1-j).$$

O to víc platí

$$P[X(t+h) | J(t)] = \left(\sum_{j=1}^{\infty} \pi_j^{(h)} W(t+1-j) \right)_X,$$

kde dolní index X znamená, že bereme pouze složky procesu X (prvních m_1 složek). Přepsáním matic π_j dle definice 33 a využitím $W(t) = (X(t)', Y(t)', Z(t)')$ získáváme

$$P[X(t+h) | J(t)] = \sum_{j=1}^{\infty} (\pi_{XXj}^{(h)} X(t+1-j) + \pi_{XYj}^{(h)} Y(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j)).$$

Z předpokladu nulovosti $\pi_{XYj}^{(h)}$ platí:

$$P[X(t+h) | J(t)] = \sum_{j=1}^{\infty} (\pi_{XXj}^{(h)} X(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j)).$$

Postupným vnořením projekcí získáme

$$\begin{aligned}
P[X(t+h)|I_{XZ}(t)] &= P[P[X(t+h)|J(t)]|I_{XZ}(t)] = \\
&= P[(\sum_{j=1}^{\infty} \pi_{XXj}^{(h)} X(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j)|I_{XZ}(t))] = \\
&= \sum_{j=1}^{\infty} \pi_{XXj}^{(h)} X(t+1-j) + \pi_{XZj}^{(h)} Z(t+1-j) = \\
&= P[X(t+h)|I_{XZY}(t)].
\end{aligned}$$

V předposlední ekvivalenci využíváme, že všechny členy sumy jsou prvky $I_{XZ}(t)$. Korektnost záměny nekonečné sumy a projekce můžeme odvodnit podobně jako v příkladu 32. □

Za předpokladu regularity procesu platí i opačná implikace.

Matice $\pi_j^{(h)}$ má (za určitých předpokladů) význam projekce hodnoty $W(t-j)$ do $W(t+h)$. Jak vidíme z předchozí věty, $\pi_{XYj}^{(h)}$ má (za určitých předpokladů) význam projekce hodnoty $Y(t-j)$ do hodnoty $X(t+h)$. Rozepíšeme-li si druhou definici $\pi_j^{(h+1)}$ pouze pro $\pi_{XYj}^{(h+1)}$ po složkách, dostaneme

$$\pi_{XYj}^{(h+1)} = \pi_{XY(j+1)}^{(h)} + \pi_{XX1}^{(h)} \pi_{XYj} + \pi_{XY1}^{(h)} \pi_{YYj} + \pi_{XZ1}^{(h)} \pi_{ZYj}. \quad (2.4)$$

Definice 35 (kauzální řetěz). *Buď $W = (X'_1, X'_2, \dots, X'_n)'$ VAR(p) proces jako z definice 24, $N \in \mathcal{N}$, $i \in \{1, 2, \dots, N\}$, $k_i \in \{1, 2, \dots, p\}$ a $j_i \in \{1, 2, \dots, n\}$. Definujeme univerzální relaci v $\{X'_1, X'_2, \dots, X'_n\}^N \times \{1, 2, \dots, p\}^{N-1}$, kterou budeme značit $X_{j_1} \xrightarrow{k_1} X_{j_2} \xrightarrow{k_2} \dots \xrightarrow{k_{N-1}} X_{j_N}$. Řekneme, že $X_{j_1} \xrightarrow{k_1} X_{j_2} \xrightarrow{k_2} \dots \xrightarrow{k_{N-1}} X_{j_N}$ je kauzální řetěz z X_{j_1} do X_{j_N} přes $X_{j_2}, \dots, X_{j_{N-1}}$ délky $\sum_{i=1}^{N-1} k_i$.*

Z definice VAR(p) procesu je vidět, že hodnota $Y(t)$ může hodnotu $X(t+h)$ ovlivňovat pouze přes autokovarianční matice. Každá „cesta“, jak může $Y(t)$ ovlivnit $X(t+h)$, odpovídá jednomu kauzálnímu řetězu z Y do X délky h . Analogicky, každý kauzální řetěz z Y do X odpovídá jedné „cestě“, jak může $Y(t)$ ovlivnit $X(t+h)$. U nonkauzality z Y do X v horizontu h bychom předpokládali, že vezme v potaz všechny kauzální řetězy z Y do X délky h . Z první definice matic $\pi_j^{(h)}$ z příkladu 32 a z definice 33 je možné nahlédnout, že Grangerova nonkauzalita se tak opravdu chová. Existuje více neekvivalentních definic nonkauzality, některé přehlížejí některé kauzální řetězce, jiné naopak požadují silnější podmínky než „neovlivňování se přes kauzální řetězy“.

Věta 36. *Nechť VAR proces $W = (X', Y', Z)'$ je regulární, buď I referenční informační systém $I \perp W$ a buď $Y \not\rightarrow_{(h)} X | I_{XZ}$. Pak $\forall j \in \mathcal{N}$ platí*

$$\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)} \pi_{ZYj}.$$

Důkaz. Zvolme $j \in \mathcal{N}$ pevně. Z regularity a nonkauzality platí, že $\pi_{XY^i}^{(k)} = 0$, $k = 1, 2, \dots, h$; $i \in \mathcal{N}$. O to víc $\pi_{XY^{(j+1)}}^{(h)} = 0$, $\pi_{XY^1}^{(h)} = 0$ a $\pi_{XY^1} = \pi_{XY^1}^{(1)} = 0$. Předposlední rovnost plyne z definice $\pi_{XY^1}^{(1)}$. Dosazením do (2.3) dostaneme dokazovanou rovnost. □

Jak vidíme, i když Y nezapříčiní X do horizontu h , nonkauzalita se může pozkatit v horizontu $h + 1$ působením přes podproces Z . Nastává otázka, jestli nonkauzalita do nějakého horizontu h už může garantovat nonkauzalitu pro všechny horizonty. Ukážeme větu, která tuto otázku zodpoví pro konečné procesy.

Věta 37. *Nechť $W = (X', Y', Z)'$ je konečný VAR(p) proces. Nechť I je referenční informační systém, $I \perp W$. Označme m_3 dimenzi podprocesu Z . Nechť dále $\pi_{XY^j}^{(h)} = 0$, $\forall j \in \mathcal{N}$, $\forall h \in \{1, 2, \dots, m_3p + 1\}$. Pak $Y \not\rightarrow_{\infty} X | I_{XZ}$.*

Důkaz je velice technický a větu nebudeme k odhadu měr kauzality potřebovat, proto důkaz vynecháme. Věta je uvedena a dokázána ve zdroji [2] pod názvem Proposition 4.5.

Věta 38. *Mějme invertibilní VAR(p) proces $W(t)$ s VMA(q) reprezentací*

$$W(t) = \sum_{j=1}^q \psi_j a(t-j).$$

Pak $\psi_j = \pi_1^{(j)}$.

Důkaz. Mějme řadu θ jako v definici 26. Ukážeme, že řada $(\pi_1^{(j)})_{j=0}^q$ je jejím inverzem. Protože inverz řady je jednoznačný, odtud plyne dokazovaná rovnost. Buď vždy $k, l \in \mathcal{N}$.

$$\begin{aligned} \sum_{j=0}^p (\theta_j B^j) \sum_{j=0}^q (\pi_1^{(j)} B^j) &= (1_m - \sum_{j=1}^p \pi_j B^j) (1_m + \sum_{j=1}^q \pi_1^{(j)} B^j) = \\ &= 1_m + \sum_{j=1}^{p+q} [\pi_1^{(j)} - \pi_j - \sum_{k+l=j} \pi_l \pi_1^{(k)}] B^j = \\ &= 1_m + \sum_{j=1}^{p+q} [\pi_1^{(j)} - \pi_j^{(1)} - \sum_{k+l=j} \pi_l^{(k+1)} - \pi_{l+1}^{(k)}] B^j = \\ &= 1_m + \sum_{j=1}^{p+q} [\pi_1^{(j)} - \pi_1^{(j)} + \pi_2^{(j-1)} - \pi_2^{(j-1)} + \dots + \pi_j^{(1)} - \pi_j^{(1)}] B^j \\ &= 1_m. \end{aligned}$$

V první rovnosti dosazujeme do sumy a využíváme fakt, že $\pi_1^{(0)} = 1_m$. V druhé rovnosti roznásobujeme sumy. Ve třetí rovnosti používáme výsledek příkladu 32, že $\pi_j = \pi_j^{(1)}$ a $\pi_1^{(h)} \pi_j = \pi_j^{(h+1)} - \pi_{j+1}^{(h)}$. □

3. Míry kauzality

Definice 39. Necht X je proces, I referenční informační množina, $t \in \mathcal{Z}$, $h \in \mathcal{N}$. Definujeme následující značení:

$$U[X(t+h)|I(t)] \stackrel{\text{def}}{=} X(t+h) - P[X(t+h)|I(t)],$$

$$\Sigma[X(t+h)|I(t)] \stackrel{\text{def}}{=} \text{Var}([U[X(t+h)|I(t)]).$$

$U[X(t+h)|I(t)]$ nazveme chybou předpovědi, $\Sigma[X(t+h)|I(t)]$ nazveme rozptylovou maticí chyby předpovědi. Dále ještě zadefinujeme

$$\sigma[X(t+h)|I(t)] \stackrel{\text{def}}{=} \text{var}([U[X(t+h)|I(t)])$$

pro jednodimenzionální X .

Příklad 40. Budiž $W = (X', Y', Z')'$ invertibilní $\text{VAR}(p)$ proces a I referenční informační systém, $W \perp I$. Chceme vyjádřit $\Sigma[X(t+h)|I_W(t)]$.

Zaobírejme se nejprve $U[W(t+h)|I_W(t)]$. Zapišeme si proces $W(t)$ v jeho MA reprezentaci:

$$W(t) = \sum_{j=0}^{\infty} \psi_j a(t-j) = \sum_{j=0}^{\infty} \pi_1^{(j)} a(t-j).$$

V druhé rovnosti používáme větu 38. Posunutím času o h kroků dozadu dostáváme

$$W(t+h) = \sum_{j=0}^{\infty} \pi_1^{(j)} a(t+h-j).$$

Necht $s \in \mathcal{Z}$, pak $a(s) \in I_W(t)$ pro $s \leq t$ a $a(s) \perp I_W(t)$ jinak. Tedy chyba předpovědi $U[W(t+h)|I_W(t)] = \sum_{j=0}^{h-1} \pi_1^{(j)} a(t+h-j)$. Nyní provedeme výpočet $\Sigma[W(t+h)|I_W(t)]$.

Z definice máme

$$\begin{aligned} \Sigma[W(t+h)|I_W(t)] &= \text{Var}(U[W(t+h)|I(t)]) = \text{Var}\left(\sum_{j=0}^{h-1} a(t+h-j)\right) = \\ &= \mathbf{E}\left(\sum_{j=0}^{h-1} \sum_{k=0}^{h-1} \pi_1^{(j)} a(t+h-j) a(t+h-k)' \pi_1^{(k)'}\right) = \\ &= \sum_{j=0}^{h-1} \sum_{k=0}^{h-1} \pi_1^{(j)} \mathbf{E}(a(t+h-j) a(t+h-k)') \pi_1^{(k)'} = \\ &= \sum_{j=0}^{h-1} \pi_1^{(j)} \Sigma_a \pi_1^{(j)'} . \end{aligned}$$

V třetí rovnosti používáme nulovost střední hodnoty $a(t)$, v poslední rovnosti používáme nekorelovanost $a(t)$ a $a(s)$ pro různá s a t .

O to víc

$$\Sigma[X(t+h)|I_W(t)] = \sum_{j=0}^{h-1} (J_X \pi_1^{(j)}) \Sigma_a (J_X \pi_1^{(j)})',$$

kde $J_X = [I_{m_1} 0_{m_2+m_3}]$ je matice, která z vektoru procesu W vrátí pouze složky podprocesu X .

Definice 41 (restringce procesu). *Bud' $N, m \in \mathcal{N}, a_1, a_2, \dots, a_m \in \{1, 2, \dots, N\}, a_1 < a_2 < \dots < a_m$. Necht' je $W = (X'_1, X'_2, X'_3, \dots, X'_N)'$ proces. Pak procesem $W|_{(X'_{a_1}, X'_{a_2}, \dots, X'_{a_m})'}$ rozumíme proces $(X'_{a_1}, X'_{a_2}, \dots, X'_{a_m})'$.*

Fakt 42. *Bud' $W(t)$ stabilní regulární invertibilní VAR proces. Pak jeho restringce je také stabilní regulární invertibilní VAR proces.*

Definice 43 (míra kauzality v horizontu h vůči I). *Necht' I je referenční informační systém, X, Y, Z procesy, $h \in \mathcal{N}$ horizont a $t \in \mathcal{Z}$. Nadefinujeme míru kauzality*

$$C_L[Y \xrightarrow{h} X|I] \stackrel{\text{def}}{=} \ln \left[\frac{\det \Sigma[X(t+h)|I_X(t)]}{\det \Sigma[X(t+h)|I_{XY}(t)]} \right],$$

kde $0/0 \stackrel{\text{def}}{=} 1$.

Uvažujeme-li X a Y jednodimenzionální podprocesy, můžeme zapsat relativní rozdíl rozptylů chyby projekce do I_{XY} naproti I_X jakožto

$$\frac{\sigma[X(t+h)|I_X(t)] - \sigma[X(t+h)|I_{XY}(t)]}{\sigma[X(t+h)|I_X(t)]} = 1 - \exp[-C_L[Y \xrightarrow{h} X|I]].$$

Míra kauzality může nabývat hodnot mezi 0 a ∞ a snaží se postihnout, jak moc nám znalost podprocesu Y pomůže s předpovědí procesu X . Ve více dimenzích determinant používáme pro měření „velikosti“ matice. Nevýhodou je, že pokud jsme schopni extrémně dobře predikovat jednu složku podprocesu X v horizontu h , determinant rozptylové matice chyby předpovědi bude velice blízko 0. Pokud nám podproces Y pomůže skoro přesně predikovat jednu složku podprocesu X v horizontu h a nepomůže s predikcí ostatních složek X , míra kauzality nám stále vyjde velice blízko ∞ . Tento problém lze ovšem odstranit počítáním měř kauzality $C_L[Y \xrightarrow{h} X_i|I_X]$ pro všechny složky X_i podprocesu X .

Lemma 44. *Bud' A symetrická pozitivně definitní matice $n \times n$ a B nenulová symetrická pozitivně semidefinitní matice $n \times n$. Pak $\det(A+B) > \det(A)$.*

Důkaz. Bud' $A = LL'$ a $B = KK'$ z Cholského rozkladu. Označme $C = L^{-1}BL'^{-1}$. Pak $(A+B) = L(1_n + C)L'$. Matice C je pozitivně semidefinitní: bud' v nenulový vektor délky n , pak

$$v'LKK'L'v = (K'L'v)'(K'L'v) \geq 0.$$

Matice C je nenulová: matice L a L' jsou regulární, tedy hodnota C je rovna hodnotě B , která je nenulová. Označme μ_1, \dots, μ_n vlastní čísla matice C . Tato čísla jsou větší nebo rovna nule. Pak vlastní čísla matice $1_n + C$ jsou větší nebo rovna jedné. Protože determinant matice se rovná součinu jejích vlastních čísel a matice C má alespoň jedno kladné vlastní číslo, $\det(1_n + C) > 1$. Tedy

$$\begin{aligned} \det(A+B) &= \det(L(1_n + C)L') = \det(L)\det(1_n + C)\det(L') = \\ &= \det(L)\det(L')\det(1_n + C) = \det(A)\det(1_n + C) > \det(A). \end{aligned}$$

□

Tvrzení 45 (charakteristika nonkauzality II). *Bud $W = (X', Y', Z)'$, $\text{VAR}(p)$ proces, I referenční informační systém, $I \perp W$ a $h \in \mathcal{N}$ horizont. Nechť platí následující podmínka*

$$\det(\Sigma[X(t+h)|I_{XY}(t)]) > 0,$$

pak

$$P[X(t+h)|I_X(t)] = P[X(t+h)|I_{XY}(t)] \iff C_L[Y \xrightarrow{h} X|I] = 0.$$

Důkaz. Začneme s implikací zleva doprava.

$$\begin{aligned} P[X(t+h)|I_X(t)] &= P[X(t+h)|I_{XY}(t)] \\ &\Downarrow \\ U[X(t+h)|I_X(t)] &= U[X(t+h)|I_{XY}(t)] \\ &\Downarrow \\ \Sigma[X(t+h)|I_X(t)] &= \Sigma[X(t+h)|I_{XY}(t)] \\ &\Downarrow \\ C_L[Y \xrightarrow{h} X|I] &= \ln(1) = 0. \end{aligned}$$

Druhou implikaci dokážeme nepřímou, nechť

$$P[X(t+h)|I_X(t)] + f(t) = P[X(t+h)|I_{XY}(t)],$$

kde $f(t) \in I_{XY}(t)$, $f(t) \neq 0$. Pak máme

$$U[X(t+h)|I_X(t)] - f(t) = U[X(t+h)|I_{XY}(t)].$$

A tedy

$$\begin{aligned} \Sigma[X(t+h)|I_{XY}(t)] &= \text{Var}(U[X(t+h)|I_X(t)] - f(t)) = \\ &= \text{Var}(U[X(t+h)|I_X(t)]) - 2\text{cov}(U[X(t+h)|I_X(t)], f(t)) + \text{Var}(f(t)) = \\ &= \Sigma[X(t+h)|I_X(t)] + \text{Var}(f(t)), \end{aligned}$$

kde

$$\text{cov}(U[X(t+h)|I_X(t)], f(t)) = 0,$$

protože $f(t) \in I_{XY}(t)$ a $U[X(t+h)|I_X(t)] \perp I_{XY}(t)$. Rozptylové matice jsou pozitivně semidefinitní a z předpoladu je $\Sigma[X(t+h)|I_X(t)]$ regulární, tedy pozitivně definitní. Dále $f(t)$ je nenulová, tedy $\text{Var}(f(t))$ je nenulová. Pak z lemmatu 44 je

$$\begin{aligned} \det(\Sigma[X(t+h)|I_{XY}(t)]) &= \det(\Sigma[X(t+h)|I_X(t)] + \text{Var}(f(t))) > \\ &> \det(\Sigma[X(t+h)|I_X(t)]) \end{aligned}$$

a tedy

$$C_L[Y \xrightarrow{h} X|I] = \ln \left[\frac{\det \Sigma[X(t+h)|I_X(t)]}{\det \Sigma[X(t+h)|I_{XY}(t)]} \right] > 0.$$

□

Příklad 46. Budiž $W = (X', Y', Z)'$ invertibilní VAR(p) proces, \tilde{Z} VAR(q) proces a I referenční informační systém, $(W, \tilde{Z})' \perp I$. Předpokládejme, že procesy W a \tilde{Z} se navzájem ovlivňují a $(W, \tilde{Z})'$ tvoří invertibilní VAR(r) proces. Dále $h \in \mathcal{N}$. Nejprve vysvětlíme, jak můžeme počítat míru kauzality v horizontu h , pak si obohatíme proces Z o proces \tilde{Z} a podíváme se, jaký to může mít efekt na hodnotu míry kauzality.

Míra kauzality v horizontu h vůči I je funkcí rozptylových matic podprocesů procesu W . Rozptylové matice můžeme získat aplikací příkladu 40 na $W \downarrow_X$ a $W \downarrow_{XY}$. Protože $W \downarrow_X = (W', \tilde{Z}')' \downarrow_X$ a na $W \downarrow_{XY} = (W', \tilde{Z}')' \downarrow_{XY}$, znalost procesu \tilde{Z} nám nezmění míru kauzality. Míra kauzality dokonce (přímo) nezávisí na procesu Z (i když změna struktury Z může změnit strukturu $W \downarrow_{XY}$ nebo $W \downarrow_X$).

Nyní uvažujme míru kauzality v horizontu h vůči I_Z .

$$C_L[Y \xrightarrow{h} X|I_Z] = \ln \left[\frac{\det \Sigma[X(t+h)|I_{XZ}(t)]}{\det \Sigma[X(t+h)|I_{XYZ}(t)]} \right].$$

Míru můžeme opět vypočítat aplikací příkladu 40, tentokrát na $W \downarrow_{XZ}$ a W . Nyní obohatíme proces Z o proces \tilde{Z} a uvažujme $C_L[Y \xrightarrow{h} X|I_{Z\tilde{Z}}]$. Ukážeme, že míra kauzality může vzrůst i klesnout.

Nejprve necht \tilde{Z} je generátor šumu procesu X . Uvažujme konkrétní procesy, dimenze X, Y a \tilde{Z} buď 1, dimenze procesu Z buď 0. Autoregresivní matice procesu $(X', Y', \tilde{Z})'$ buď

$$\begin{pmatrix} 0.5 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

a rozptylová matice šumu procesu $(X', Y', \tilde{Z})'$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Můžeme si představit, že šum procesu \tilde{Z} má rozdělení rovné normálnímu $(0, 1)$ rozdělení a hodnota šumu X má rozdělení rovné normálnímu $(0, 1)$ plus rozdělení šumu procesu \tilde{Z} v předchozím kroku.

Dopočteme, že autoregresivní matice procesu $(X', Y')'$ je

$$\begin{pmatrix} 0.5 & 1 \\ 0 & 0 \end{pmatrix},$$

procesu $(X', \tilde{Z}')'$ je

$$\begin{pmatrix} 0.5 & 1 \\ 0 & 0 \end{pmatrix}$$

a procesu X je (0.5).

Dopočteme, že rozptylová matice procesu $(X', Y')'$ je

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},$$

procesu $(X', \tilde{Z}')'$ je

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

a procesu X je (3).

Vypočteme míru kauzality v horizontu 1; využijeme, že matice $\pi_1^{(0)}$ jsou rovny jednotkové matici. Míry kauzality vyjdou

$$C_L[Y \xrightarrow{1} X | I_Z] = \ln \left[\frac{1 \cdot 3 \cdot 1}{\det \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]} \right] = \ln \left[\frac{3}{2} \right]$$

a

$$C_L[Y \xrightarrow{h} X | I_{Z\tilde{Z}}] = \ln \left[\frac{\det \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]}{\det \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]} \right] = \ln[2].$$

Nyní necht \tilde{Z} je generátor šumu procesu Y . Uvažujme konkrétní procesy, dimenze X, Y a \tilde{Z} buď 1, dimenze procesu Z buď 0.

Autoregresivní matice procesu $(X', Y', \tilde{Z}')'$ buď

$$\begin{pmatrix} 0.5 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

a rozptylová matice šumu procesu $(X', Y', \tilde{Z})'$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dopočteme, že rozptylová matice procesu $(X', Y)'$ je

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},$$

procesu $(X', \tilde{Z})'$ je

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

a procesu X je (3).

Míry vyjdou

$$C_L[Y \xrightarrow{1} X|I_Z] = \ln[3]$$

a

$$C_L[Y \xrightarrow{1} X|I_{ZZ}] = \ln[2].$$

Míra kauzality vůči I nám říká, jak moc velký vliv mají data z podprocesu Y na podproces X . Podproces Z nemá vliv na míru kauzality. V praxi nemusíme mít k dispozici parametry procesu W , jenom data z něj. V tomto případě nám data z podprocesu Z mohou zlepšit odhady parametrů procesu W . Míra kauzality vůči I_Z zanedbává kauzální řetězce ústící z podprocesu Z . Podproces Y může stále působit na podproces X přes Z , ale například působení $Z \rightarrow Y \rightarrow X$ již nepřispívá k velikosti míry kauzality.

V praxi nelze jednoznačně říci, která z těchto měř je lepší; i když míra vůči I_Z umožní odstranění řetězce $Z \rightarrow Y \rightarrow X$ (v analogii si můžeme představit Y jako přítomnost požárníků, X jako vyplácení pojistky a Z jako požár), Z může být výrazný generátor šumu X . Pak nám může vyjít míra kauzality Y na X výrazná, i když náhodnost podprocesu Z ovlivňuje hodnoty procesu X řádově více než kauzalita Y na X . Je možné obě míry kombinovat a brát míru kauzality vůči I_U , kde U je podproces Z .

Definice 47 (míra okamžité kauzality v horizontu h vůči I). *Nechť I je referenční informační systém, X, Y procesy, $h \in \mathcal{N}$ horizont a $t \in \mathcal{Z}$. Označme $R=(X', Y)'$. Nadefinujeme míru okamžité kauzality mezi X a Y jako*

$$C_L[X \xrightarrow{h} Y|I] \stackrel{\text{def}}{=} \ln \left[\frac{\det \Sigma[X(t+h)|I_{XY}(t)] \det \Sigma[Y(t+h)|I_{XY}(t)]}{\det \Sigma[R(t+h)|I_{XY}(t)]} \right].$$

Pro případ, kde X a Y jsou jednodimenzionální procesy, lze míru přepsat jako

$$C_L[X \xrightarrow{h} Y|I] = -\ln[1 - \text{cor}(U[X(t+h)|I_{XY}(t)], U[Y(t+h)|I_{XY}(t)])^2].$$

Tedy míra se snaží postihnout, jak moc jsou chyby projekce $X(t+h)$ a $Y(t+h)$ do $I_{XY}(t)$ spolu korelované.

Definice 48 (míra závislosti v horizontu h vůči I). *Nechť I je referenční informační systém, X a Y procesy, $h \in \mathcal{N}$ horizont a $t \in \mathcal{Z}$. Nadefinujeme míru závislost Y na X jako*

$$C_L^{(h)}[X, Y|I] \stackrel{\text{def}}{=} C_L[X \xrightarrow{h} Y|I] + C_L[Y \xrightarrow{h} X|I] + C_L[X \xrightarrow{h} Y|I].$$

Tato míra se snaží postihnout, jako moc jsou procesy X a Y v horizontu h propojeny. Označme $R = (X', Y)'$. Pak míru závislosti v horizontu h lze přepsat jako

$$C_L^{(h)}[X, Y|I] = \ln \left[\frac{\det \Sigma[X(t+h)|I_X(t)] \det \Sigma[Y(t+h)|I_Y(t)]}{\det \Sigma[R(t+h)|I_{XY}(t)]} \right].$$

4. Odhady měr kauzality

4.1 Nástroje

Definice 49. *Bud' A matice $m \times m$. Pak*

$$\begin{aligned} \text{vec}(A) &\stackrel{\text{def}}{=} (A_{11}, A_{21}, \dots, A_{m1}, A_{12}, A_{22}, \dots, A_{m2}, \dots, A_{mm})', \\ \text{vech}(A) &\stackrel{\text{def}}{=} (A_{11}, A_{21}, \dots, A_{m1}, A_{22}, A_{32}, \dots, A_{m2}, A_{33}, \dots, A_{mm})'. \end{aligned}$$

Operátor vec přepíše prvky matice do vektoru, operátor vech přepíše do vektoru pouze prvky na a pod diagonálou.

Definice 50. *Bud' $W(t)$ VAR(p) proces dimenze m a $j \in \mathcal{Z}$. Zadefinujeme*

$$\Gamma(j) \stackrel{\text{def}}{=} E[W(t)W(t+j)].$$

Dále Γ_k bude značit matici $km \times km$, jejíž (i, j) -tý $m \times m$ blok odpovídá $\Gamma(i-j)$.

Definice 51 (odhady nejmenších čtverců). *Bud' $W(t)$ VAR(k) proces, $T \in \mathcal{N}$ a mějme k dispozici měření $W(1), W(2), \dots, W(T)$. Zadefinujeme si následující statistiky:*

$$\begin{aligned} w(t) &\stackrel{\text{def}}{=} (W(t)', W(t-1)', \dots, W(t-k+1))', \\ \hat{\Gamma}_k &\stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=k}^T w(t)w(t)', \\ \hat{\Gamma}_{k1} &\stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=k+1}^{T-1} w(t)W(t+1)', \\ \hat{\pi}(k) &= [\hat{\pi}_{1k}, \hat{\pi}_{2k}, \dots, \hat{\pi}_{kk}] \stackrel{\text{def}}{=} \hat{\Gamma}_{k1}' \hat{\Gamma}_k^{-1}, \\ \hat{a}_k(t) &\stackrel{\text{def}}{=} W(t) - \sum_{j=1}^k \hat{\pi}_{jk} W(t-j), \\ \hat{\Sigma}_{a|k} &\stackrel{\text{def}}{=} \frac{1}{T-k} \sum_{t=k+1}^T \hat{a}_k(t) \hat{a}_k(t)'. \end{aligned}$$

Mějme VAR(∞) proces a měření z něj. Chceme odhadnout jeho parametry. Použijeme měření pro odhad parametrů VAR(k) procesu metodou nejmenších čtverců – řekneme, že proces aproximujeme procesem VAR(k). Budeme doufat, že rozptylová matice šumu a autokovarianční matice VAR(k) procesu dostatečně dobře aproximují rozptylovou matici a prvních k autokovariančních matic původního procesu. Tento nápad formalizuje následující věta.

Věta 52 ((slabá) konzistence odhadů nejmenších čtverců). *Budiž $W(t)$ invertibilní stabilní VAR(∞) proces dimenze m jako v definici 24. Budiž $T \in \mathcal{N}$ a $k \in \mathcal{N}$; T značí počet měření, k je funkcí T . Mějme k dispozici měření $W(1), \dots, W(T)$. Označme $a(t) = (a(t)_1, a(t)_2, \dots, a(t)_m)'$. Dále označme $\pi(j) = \pi_1, \pi_2, \dots, \pi_j$. Nakonec, necht jsou následující podmínky splněny:*

- $E[|a(t)_i a(t)_j a(t)_q a(t)_r|] < \infty, \quad 1 \leq i \leq j \leq q \leq r \leq m,$
- $\lim_{T \rightarrow \infty} k^3/T = 0,$
- $\lim_{T \rightarrow \infty} k^{1/2} \sum_{j=k+1}^{\infty} \|\pi_j\|_{2,2} = 0.$

Pak pro odhady nejmenších čtverců z definice 51 platí:

- $\|\hat{\pi}(k) - \pi(k)\|_{2,2} \xrightarrow{Q} 0 \quad \text{pro } T \rightarrow \infty,$
- $\|\hat{\Sigma}_{a|k} - \Sigma_a\|_{2,2} \xrightarrow{Q} 0 \quad \text{pro } T \rightarrow \infty.$

Symbol \xrightarrow{Q} značí konvergenci v pravděpodobnosti. Důkaz je uveden v Lewis, R., Reinsel, G.C., 1985. Prediction of multivariate time series by autoregressive model fitting. Journal of Multivariate Analysis 16, 393–411, Theorem 1.

Věta 53 (o spojitě transformaci). *Necht X, X_1, X_2, \dots jsou k -rozměrné náhodné vektory a funkce $g : \mathcal{R}^k \rightarrow \mathcal{R}^p$ je funkce, která je spojitá s. j. (až na množinu míry 0). Pak*

$$X_n \xrightarrow{Q} X \implies g(X_n) \xrightarrow{Q} g(X).$$

Věta 54 (Wold's Theorem). *Budiž W (slabě) stacionární proces. Pak má VMA(∞) reprezentaci.*

4.2 Bodový odhad $\Sigma[X(t+h)|I_{XZ}]$

V této podkapitole si představíme simulační metodu a ospravedlníme její použití pro výpočet parametrů podprocesu VAR procesu. Upozornujeme, že tato podkapitola je poměrně technická.

Budiž $W = (X', Y', Z)'$ stabilní invertibilní konečný VAR(p) proces dimenze m jako v definici 24 a mějme jeho hodnoty $W(1), W(2), \dots, W(n)$. Chceme odhadnout $\Sigma[X(t+h)|I_{XZ}]$. Nejprve budeme uvažovat proces $S = W \upharpoonright_{XZ}$, poté aplikujeme výsledek příkladu 40 na $\Sigma[X(t+h)|I_S] = \Sigma[X(t+h)|I_{XZ}]$. Tím získáme vyjádření $\Sigma[X(t+h)|I_{XZ}]$ pomocí parametrů podprocesu S . Označíme autokovarianční matice procesu S ρ_j a rozptylovou matici šumu Σ_ϵ . Dále označíme $\theta = ((\text{vec}(\pi_1), \text{vec}(\pi_2), \dots, \text{vec}(\pi_p), \text{vech}(\Sigma_a))'$ parametry procesu W a nakonec označíme $\psi = ((\text{vec}(\rho_1), \text{vec}(\rho_2), \dots, \text{vec}(\rho_{h-1}), \text{vech}(\Sigma_\epsilon))'$ parametry podprocesu S , které potřebujeme k výpočtu $\Sigma[X(t+h)|I_{XZ}]$ (dále relevantní parametry).

Zvolíme-li $k = p$, větu 52 lze aplikovat na konečné procesy. Parametry ψ můžeme odhadnout tak, že zahodíme Y složku našich měření a použijeme metodu nejmenších čtverců. Zahazujeme data z podprocesu Z , toto nám může zpomalit konvergenci, což může značně zhoršit odhad pro malá n . Lepší přístup by byl odhadnout parametry θ procesu W a pokusit se vypočítat parametry ψ pomocí našich odhadů parametrů θ . Označme \mathcal{C} množinu všech parametrů θ , které mohou být parametry stabilního invertibilního VAR(p) procesu, a dimenzi parametru ψ označme M .

Definujeme funkci $f(\theta) : \mathcal{C} \rightarrow \mathcal{R}^M$, která pro každé přípustné parametry θ procesu W vrátí relevantní parametry ψ podprocesu S . Funkce f je složením spojitých funkcí a vlastních autokovariančních čísel matic. Vlastní čísla závisí spojitě na prvcích matice (jsou kořenem polynomu, jehož koeficienty jsou získané násobením a sčítáním, kořeny závisí spojitě na koeficientech, viz zdroj [7] Theorem 3.2). Funkce f je tedy spojitá. Funkci f je velice těžké vyčíslit přímo. V této práci uvedeme postup navržený ve zdroji [1] – simulační metodu.

Mějme k dispozici generátor náhodných čísel (nezávislý na datech). Použijeme odhady parametrů W a generátor náhodných čísel k simulaci procesu W . Tím získáme T měření, které můžeme použít pro odhad relevantních parametrů podprocesu S metodou nejmenších čtverců.

Nejprve zformalizujeme výše zmíněný postup. Máme pravděpodobnostní prostor (Ω, \mathcal{A}, Q) , který diktuje, jaké hodnoty $W(t)$ dostaneme. Uvažujme pravděpodobnostní prostor (Ξ, \mathcal{B}, R) ; ten reprezentuje generátor náhodných čísel. Dále uvažujme zúplněný součinný pravděpodobnostní prostor $(\Omega \times \Xi, \mathcal{A} \otimes \mathcal{B}, Q \otimes R)$. Označme $\omega \in \Omega$ a $\xi \in \Xi$. Můžeme předpokládat, že proces $W(t)(\omega)$ vychází z prostoru $(\Omega \times \Xi, \mathcal{A} \otimes \mathcal{B}, Q \otimes R)$. Definujeme náhodné funkce $f_T(\theta)(\xi) : \mathcal{C} \times \Xi \rightarrow \mathcal{R}^M$ následovně: f_T nejprve vytvoří posloupnost $(w(1)(\xi), w(2)(\xi), \dots, w(T)(\xi))$, konkrétní inicializaci VAR(p) procesu příslušnému parametrům θ za předpokladu, že nastal jev ξ . Funkce poté zahodí složky podprocesu Z . Poté funkce pomocí $(w(1)(\xi), w(2)(\xi), \dots, w(T)(\xi))$ metodou nejmenších čtverců odhadne parametry ψ podprocesu S . Parametr k je zvolen tak, aby splňoval podmínky věty 52. Funkce z těchto parametrů vybere pouze relevantní parametry; pokud metoda nejmenších čtverců nevrátí některé relevantní parametry (k zvoleno moc malé), budeme předpokládat, že tyto parametry jsou rovny 0.

Označme pravou hodnotu θ jako θ_0 , analogicky pro ψ a $\Sigma[X(t+h)|I_{XZ}]$. Označme stříškou odhady získané metodou nejmenších čtverců a do závorky dolního indexu napíšeme, kolik měření bylo použito k jeho vytvoření. Tedy třeba odhad θ_0 pomocí n měření označíme $\hat{\theta}_{(n)}$ a odhad ψ_0 získaného z $\hat{\theta}_{(n)}$ pomocí T měření $\hat{\psi}_{(nT)}$. $\hat{\Sigma}[X(t+h)|I_{XZ}]_{(nT)}$ je spojitou transformací $\hat{\psi}_{(nT)}$. Je-li $\hat{\psi}_{(nT)}$ konzistentní odhad ψ_0 , z věty 53 je i $\hat{\Sigma}[X(t+h)|I_{XZ}]_{(nT)}$ konzistentní odhad $\Sigma[X(t+h)|I_{XZ}]_0$. Stačí tedy ukázat konzistenci $\hat{\psi}_{(nT)} = f_T(\hat{\theta}_{(n)}(\omega)(\xi))$.

Lemma 55. *Bud' $T \in \mathcal{N}$. Necht' $\hat{\Gamma}_{(nT)}$ je regulární R -s. j. Pak funkce f_T je spojitá $(Q \times R)$ -s. j.*

Důkaz. Stačí ukázat, že pro libovolné $\theta \in \mathcal{C}$ je f_T spojitá R -s. j. Pak je $f_T(\theta(\omega))(\xi)$ spojitá $(Q \times R)$ -s. j. (pro libovolnou hodnotu $\theta(\omega)$ je spojitá s pravděpodob-

ností 1). První část funkce f_T je spojitá, protože hodnoty stabilního invertibilního VAR procesu závisí spojitě na parametrech. Druhá část funkce f_T je složením konečně mnoha sčítání, násobení, řetězení vektorů a inverzu matice. Buď Γ čtvercová matice, pak $\Gamma^{-1} = \frac{\text{adj}(\Gamma)}{\det(\Gamma)}$. $\text{Adj}(\Gamma)$ je spojitá, $\frac{1}{\det(\Gamma)}$ spojitá kromě nuly. Ale z předpokladu je naše matice regulární, tedy determinant je nenulový R -s.j. \square

Předpoklad regulárnosti $\hat{\Gamma}_{(nT)}$ R -s.j. je rozumný:

- Pokud matice vyjde singulární, dozvíme se to, protože výpočet selže.
- Pravděpodobnost, že vyjde determinant nula, musí konvergovat k nule, jinak by věta 52 nefungovala na proces $W \downarrow_{XZ}$, kde W má parametry θ (metoda nejmenších čtverců by v tomto případě vždy selhala s kladnou pravděpodobností, protože by nebyla schopna spočítat inverz singulární matice).
- Kauzalita procesu nezávisí na rozdělení šumu, jen na jeho rozptylové matici. Můžeme předpokládat, že šum pochází z normálního rozdělení. Determinant $\hat{\Gamma}_{(nT)}$ je poměrně hezkou funkcí šumu příslušnému ξ . Pokud cestou k $\hat{\Gamma}_{(nT)}$ nedojde k degeneraci, tedy determinant $\hat{\Gamma}_{(nT)}$ vyjde nulový jen pro množinu ξ (Lebesgueovy) míry 0, pravděpodobnost, že determinant vyjde 0, musí být nulová ze spojitosti normálního rozdělení.

Důsledek. Necht $\theta_{(n)}(\omega) \xrightarrow{Q} \theta_0$. Pak $f_T(\theta_{(n)}(\omega))(\xi) \xrightarrow{Q \times R} f_T(\theta_0)(\xi)$.

Důkaz. Z $\theta_{(n)}(\omega) \xrightarrow{Q} \theta_0$ máme $(\theta_{(n)}(\omega), \xi) \xrightarrow{Q \times R} (\theta_0, \xi)$. Z lemmatu 55 f_T je spojitá $(Q \times R)$ -s. j., tedy z věty 53 $f_T(\theta_{(n)}(\omega))(\xi) \xrightarrow{Q \times R} f_T(\theta_0)(\xi)$. \square

Lemma 56. *Pro pevné $\theta \in \mathcal{C}$ funkce f_T bodově konvergují k funkci f v pravděpodobnosti R .*

Důkaz. Zvolme $T \in \mathcal{N}$ a $\theta \in \mathcal{C}$ pevně. Funkce $f_T(\theta)(\xi)$ je odhad podmnožiny parametrů procesu $W \downarrow_{XZ}$ metodou nejmenších čtverců za předpokladu, že W má parametry θ . Odhady všech parametrů konvergují z věty 52 k jejím skutečným hodnotám v pravděpodobnosti, o to víc konverguje jejich podmnožina. Tedy $f_T(\theta)$ konverguje ke $f(\theta)$ v pravděpodobnosti R . \square

Věta 57 (částečná konzistence $\hat{\psi}_{(nT)}$). *Buď $\hat{\theta}_n$ konzistentní odhad θ_0 . Pak existuje funkce r na přirozených číslech taková, že je-li $T > r(n)$, pak $f_T(\hat{\theta}_n(\omega))(\xi)$ je konzistentní odhad $f(\theta_0)$.*

Důkaz. Mějme ε kladné a menší než 1. Zvolíme N takové, že pro $n > N$ je

$$Q\{\omega : \|f(\hat{\theta}_n) - f(\theta_0)\| > \varepsilon\} < \varepsilon.$$

To můžeme díky spojitosti f a konvergenci $\hat{\theta}_n$ k θ_0 v pravděpodobnosti Q . Pro každé $n \in \mathcal{N}$ definujeme hodnotu $r(n)$ tak, že pro každé $T \geq r(n)$ je

$$R\{\xi : \|f_T(\hat{\theta}_n)(\xi) - f(\hat{\theta}_n)\| > \varepsilon\} < \varepsilon.$$

To můžeme díky bodové konvergenci f_T k f v pravděpodobnosti R . Pak pro $n > N$ a $T \geq r(n)$ máme

$$\begin{aligned} (Q \times R)\{(\omega, \xi) : \|f_T(\hat{\theta}_n(\omega))(\xi) - f(\theta_0)\| > 2\varepsilon\} &\leq \\ &\leq (Q \times R)\{(\omega, \xi) : \|f_T(\hat{\theta}_n(\omega))(\xi) - f(\hat{\theta}_n(\omega))\| > \varepsilon\} + \\ &\quad + (Q \times R)\{(\omega, \xi) : \|f(\hat{\theta}_n(\omega)) - f(\theta_0)\| > \varepsilon\} < 2\varepsilon. \end{aligned}$$

Tedy $f_T(\hat{\theta}_n)$ je konzistentní odhad relevantních parametrů podprocesu S . Dokonce platí, že s rostoucím T při pevném n se asymptoticky zmenšuje chyba odhadu. □

Pro odhady ostatních rozptylových matic předpovědi potřebných pro míry kauzality lze použít analogický postup. Míry kauzality jsou spojité funkce rozptylových matic předpovědí. Jsou-li tedy odhady rozptylových matic konzistentní, jsou konzistentní i odhady měř.

Nyní buď W VAR(∞) proces. Budeme předpokládat, že se vlastně jedná o VAR(p) proces \tilde{W} , kde p roste s počtem měření do ∞ a splňuje podmínky z věty 52. Na proces \tilde{W} poté aplikujeme postup popsany výše. Tento předpoklad je rozumný: lze stále odhadnout parametry ψ z původních dat. Použijeme-li k „výpočtu“ ψ simulační metodu, kde T je řádově větší než n , odhad ψ by měl být přesnější, protože k jeho odhadu používáme i data z podprocesu Y .

4.3 Bodový odhad měř kauzality

Mějme proces dimenze m $W(t)^*$ a n měření z něj: $W(1)^*, W(2)^*, \dots, W(n)^*$. Buď I referenční informační systém, $I \perp W$. Chceme spočítat míru kauzality v horizontu h . Postupujeme takto:

0) Kdykoliv budeme používat generátor náhodných čísel pro generování měření z VAR procesu, budeme požadovat nezávislost generátoru šumu na měřeních z $W(t)^*$. Definujeme pomocnou funkci $g(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1, m_2, m_3, p)$ takto:

$$\begin{aligned} m &= m_1 + m_2 + m_3, \\ \pi_1^{(0)} &= I_m, \quad \pi_j^{(1)} = \hat{\pi}_j, & j = 1, 2, \dots, p, \\ \pi_j^{(1)} &= 0_{m \times m}, & j > p, \\ \pi_j^{(i+1)} &= \pi_{j+1}^{(i)} + \pi_1^{(i)} \pi_j^{(1)}, & i = 1, 2, \dots, h-2; j = 1, 2, \dots, p, \\ J &= [I_{m_1}, 0_{m_1 \times m_2}, 0_{m_1 \times m_3}], \\ r &= \det\left[\sum_{k=0}^{h-1} (J\pi_1^{(k)} \hat{\Sigma}_{a|p} (J\pi_1^{(k)})')\right], \\ &\text{return } r. \end{aligned}$$

a pomocnou funkci $h(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1, m_2, m_3, p)$ takto:

$$\begin{aligned}
m &= m_1 + m_2 + m_3, \\
\pi_1^{(0)} &= I_m, \quad \pi_j^{(1)} = \hat{\pi}_j, & j = 1, 2, \dots, p, \\
\pi_j^{(1)} &= 0_{m \times m}, & j > p, \\
\pi_j^{(i+1)} &= \pi_{j+1}^{(i)} + \pi_1^{(i)} \pi_j^{(1)}, & i = 1, 2, \dots, h-2; j = 1, 2, \dots, p, \\
J &= [0_{m_2 \times m_1}, I_{m_2}, 0_{m_2 \times m_3}], \\
r &= \det \left[\sum_{k=0}^{h-1} (J \pi_1^{(k)} \hat{\Sigma}_{a|p} (J \pi_1^{(k)})') \right], \\
&\text{return } r.
\end{aligned}$$

1) Nejprve otestujeme stabilitu každé složky procesu pomocí modifikovaného Dickeyova-Fullerova testu. Pokud i -tá složka test nesplní, vezmeme difference jejích dat ($W(t)_i^{*'} = W_i(t)^* - W_i(t-1)^*$, zahodíme $W(1)^{*'}$) a test pustíme znovu. Opakujeme tak dlouho, než všechny složky test splní. Posuneme čas o $d \in \mathcal{N}_0$, abychom v čase 1 měli měření všech složek. Získáme nový proces $W(t)$ a nový vektor dat. Dále budeme pracovat s procesem $W(t)$ a novými daty $W(1), W(2), \dots, W(n-d)$.

2) Předpokládáme, že $W(t)$ je stabilní, protože splňuje modifikovaný Dickeyův-Fullerův test. Dle Woldova tvrzení má $W(t)$ VMA reprezentaci. Budeme předpokládat, že $W(t)$ splňuje podmínky pro invertibilitu a tedy má VAR reprezentaci.

3) Použijeme Akaikeho informační kritérium a zvolíme p , které ho minimalizuje. Budeme předpokládat, že $W(t)$ je VAR(p) proces.

4) Pomocí metody nejmenších čtverců odhadneme $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$.

Nyní se postup liší podle toho, kterou míru kauzality odhadujeme.

$$\boxed{C_L[Y \xrightarrow{h} X | I_Z]}$$

5) Vygenerujeme T měření VAR(p) procesu odpovídajícím parametrům $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$.

6) V měřeních z předchozího kroku zahodíme složky podprocesu Y a použijeme je pro odhad parametrů podprocesu $S = (X', Z)'$ metodou nejmenších čtverců pro $k = p$. Odhady označíme $\tilde{\pi}(p)$ a $\tilde{\Sigma}_{u|p}$.

7) Označme náš odhad míry kauzality $\hat{C}_L[Y \xrightarrow{h} X|I_Z]$. Pak

$$\hat{C}_L[Y \xrightarrow{h} X|I_Z] = \ln \left[\frac{g(\tilde{\pi}(p), \tilde{\Sigma}_{u|p}, h, m_1, m_3, 0, p)}{g(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1, m_2, m_3, p)} \right].$$

$$\boxed{C_L[X \xrightarrow{h} Y|I_Z]}$$

5) Označme náš odhad míry kauzality $\hat{C}_L[X \xrightarrow{h} Y|I_Z]$. Pak

$$\hat{C}_L[X \xrightarrow{h} Y|I_Z] = \ln \left[\frac{g(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1, m_2, m_3, p)h(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1, m_2, m_3, p)}{g(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1 + m_2, 0, m_3, p)} \right].$$

$$\boxed{C_L^{(h)}[X, Y|I_Z]}$$

5) Vygenerujeme T měření VAR(p) procesu odpovídajícímu parametrům $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$.

6) Uděláme kopii měření z kroku 5) a v ní zahodíme složky podprocesu Y . Tuto kopii použijeme pro odhad parametrů podprocesu $S_1 = (X', Z)'$ metodou nejmenších čtverců pro $k = p$. Odhady označíme $\tilde{\pi}(p)$ a $\tilde{\Sigma}_{u|p}$.

7) V měřeních z kroku 5) zahodíme složky podprocesu X . Použijeme je pak pro odhad parametrů podprocesu $S_2 = (Y', Z)'$ metodou nejmenších čtverců pro $k = p$. Odhady označíme $\bar{\pi}(p)$ a $\bar{\Sigma}_{u|p}$.

8) Označme náš odhad míry kauzality $\hat{C}_L^{(h)}[X, Y|I_Z]$. Pak

$$\hat{C}_L^{(h)}[X, Y|I_Z] = \ln \left[\frac{g(\tilde{\pi}(p), \tilde{\Sigma}_{u_1|p}, h, m_1, 0, m_3, p)g(\bar{\pi}(p), \bar{\Sigma}_{u_2|p}, h, m_2, 0, m_3, p)}{g(\hat{\pi}(p), \hat{\Sigma}_{a|p}, h, m_1 + m_2, 0, m_3, p)} \right].$$

$$\boxed{C_L[Y \xrightarrow{h} X|I]}$$

5) Vygenerujeme T měření VAR(p) procesu odpovídajícímu parametrům $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$. Zahodíme složky podprocesu Z .

6) Uděláme kopii měření z kroku 5) a v ní zahodíme složky podprocesů Y a Z .

7) Použijeme Akaikeho informační kriterium pro měření z 5) a 6). Označme hodnotu kriteria pro měření z 5) pro k stupňů volnosti A_{k1} a hodnotu kriteria pro měření z 6) pro k stupňů volnosti A_{k2} . Zvolíme k , které minimalizuje $\max \{A_{k1}, A_{k2}\}$.

8) Měření z kroku 5) použijeme pro odhad parametrů podprocesu $S_1 = (X', Y)'$ metodou nejmenších čtverců pro k . Odhady označíme $\tilde{\pi}(k)$ a $\tilde{\Sigma}_{u_1|k}$.

9) Měření z kroku 6) použijeme pro odhad parametrů podprocesu $S_2 = X$ metodou nejmenších čtverců pro k . Odhady označíme $\bar{\pi}(k)$ a $\bar{\Sigma}_{u_2|k}$.

10) Označme náš odhad míry kauzality $\hat{C}_L[Y \xrightarrow{h} X|I]$. Pak

$$\hat{C}_L[Y \xrightarrow{h} X|I] = \ln \left[\frac{g(\bar{\pi}(k), \bar{\Sigma}_{u_2|k}, h, m_1, 0, 0, k)}{g(\tilde{\pi}(k), \tilde{\Sigma}_{u_1|k}, h, m_1, m_2, 0, k)} \right].$$

$$\boxed{C_L[X \xrightarrow{h} Y|I]}$$

5) Vygenerujeme T měření $\text{VAR}(p)$ procesu odpovídajícímu parametrům $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$. Zahodíme složky podprocesu Z .

6) Použijeme Akaikeho informační kriterium a zvolíme k , které ho minimalizuje.

7) Měření z kroku 5) použijeme na odhad parametrů podprocesu $S = (X', Y)'$ metodou nejmenších čtverců pro k . Odhady označíme $\tilde{\pi}(k)$ a $\tilde{\Sigma}_{u|k}$.

8) Označme náš odhad míry kauzality $\hat{C}_L[X \xrightarrow{h} Y|I]$. Pak

$$\hat{C}_L[X \xrightarrow{h} Y|I] = \ln \left[\frac{g(\tilde{\pi}(k), \tilde{\Sigma}_{u|k}, h, m_1, m_2, 0, k)h(\tilde{\pi}(k), \tilde{\Sigma}_{u|k}, h, m_1, m_2, 0, k)}{g(\tilde{\pi}(k), \tilde{\Sigma}_{u|k}, h, m_1 + m_2, 0, 0, k)} \right].$$

$$\boxed{C_L^{(h)}[X, Y|I]}$$

5) Vygenerujeme T měření $\text{VAR}(p)$ procesu odpovídajícímu parametrům $\hat{\pi}(p)$ a $\hat{\Sigma}_{a|p}$. Zahodíme složky podprocesu Z .

6) Uděláme kopii měření z kroku 5) a v ní zahodíme složky podprocesů X a Z .

7) Uděláme kopii měření z kroku 5) a v ní zahodíme složky podprocesů Y a Z .

8) Použijeme Akaikeho informační kriterium pro měření z 5), 6) a 7). Označme hodnotu kriteria pro měření z 5) pro k stupňů volnosti A_{k1} , hodnotu kriteria pro měření z 6) pro k stupňů volnosti A_{k2} a hodnotu kriteria pro měření z 7) pro k stupňů volnosti A_{k3} . Zvolíme k , které minimalizuje $\max \{A_{k1}, A_{k2}, A_{k3}\}$.

9) Měření z kroku 5) použijeme pro odhad parametrů podprocesu $S_1 = (X', Y)'$ metodou nejmenších čtverců pro k . Odhady označíme $\hat{\pi}(k)$ a $\hat{\Sigma}_{u_1|k}$.

10) Měření z kroku 6) použijeme pro odhad parametrů podprocesu $S_2 = Y$ metodou nejmenších čtverců pro k . Odhady označíme $\bar{\pi}(k)$ a $\bar{\Sigma}_{u_2|k}$.

11) Měření z kroku 7) použijeme pro odhad parametrů podprocesu $S_3 = X$ metodou nejmenších čtverců pro k . Odhady označíme $\dot{\pi}(k)$ a $\dot{\Sigma}_{u_3|k}$.

12) Označme náš odhad míry kauzality $\hat{C}_L^{(h)}[X, Y|I]$. Pak

$$\hat{C}_L^{(h)}[X, Y|I] = \ln \left[\frac{g(\dot{\pi}(k), \dot{\Sigma}_{u_3|k}, h, m_1, 0, 0, k)g(\bar{\pi}(k), \bar{\Sigma}_{u_2|k}, h, m_2, 0, 0)}{g(\hat{\pi}(k), \hat{\Sigma}_{u_1|k}, h, m_1 + m_2, 0, 0, k)} \right].$$

Poznámky: Aby výpočet byl dostatečně přesný, T by mělo být dostatečně velké, např. $T = 1\,000\,000$. $\text{VAR}(p)$ proces má stupeň p . Důvod, proč chceme, aby v každém případě měly všechny procesy stejný stupeň, je ten, že v míře kauzality dělíme determinanty matice. Je lepší porovnávat determinanty matic stejného stupně.

Míry vůči I lze počítat tak, že zahodíme hodnoty podprocesu Z a dále počítáme míry vůči I_Z , kde Z má dimenzi 0. Pokud Z skoro neovlivňuje X a Y , vyhneme se tím jednomu odhadu parametrů podprocesu simulační metodou (nepřesnost) a Akaikeho informační kriterium nebude ovlivňováno hodnotami procesu Z . Nicméně, pokud Z ovlivňuje X nebo Y a my máme k dispozici malé množství dat, zahození dat z podprocesu Z může značně znekválit naše odhady.

O tom, jestli se rozhodneme zahodit hodnoty podprocesu Z , se můžeme rozhodnout z fyzikálních zkušeností. Pokud nevíme, jestli Z ovlivňuje X nebo Y , můžeme porovnat prvky matic $\hat{\pi}_j$, které transformují hodnoty podprocesu Z do X a Y s ostatními prvky těchto matic – pokud vyjdou výrazně menší, podproces Z můžeme vyškrtnout. Toto lze udělat i pro jednotlivé složky podprocesu Z . Pokud použijeme stejná data pro rozhodnutí, jestli ponecháme hodnoty Z , a pro výpočet míry kauzality, může být narušena konvergence; především chceme-li data použít pro intervalové odhady.

4.4 Intervaly spolehlivosti

V této kapitole popíšeme, jak získat interval spolehlivosti pro odhad měr kauzality. Označíme míru kauzality, která nás zajímá, C . Mějme $\text{VAR}(p)$ proces $W(t)$ (jako v bodě 4 v předchozí sekci), odhady jeho autoregresivních matic

(označíme $\hat{\pi}$), spolehlivost $\alpha \in (0,1)$, bodový odhad míry kauzality (označme \hat{C} , musí být nezáporný, jinak volíme $\hat{C} = 0$) a metodu G, která nám spočítá bodový odhad \hat{C} (např. metoda z předchozí podkapitoly).

1) Zvolme $B \in \mathcal{N}$, aby $\frac{\alpha}{2}B$ bylo přirozené číslo (nebo alespoň při „malém“ zvětšení α byl výraz přirozené číslo).

2) Zvolme $K \in \mathcal{N}$ dostatečně velké.

3) Spočteme rezidua $u(t)$ procesu $W(t)$:

$$\begin{aligned} u(t) &= 0, & t \in \{1, 2, \dots, p\}, \\ u(t) &= W(t) - \sum_{i=1}^p \hat{\pi}_{t-i} W(t-i), & t \in \{p+1, p+2, \dots, T\}. \end{aligned}$$

4) Opakujeme kroky 5) až 7) K -krát. Opakování indexujeme indexem k .

5) Vygenerujeme bootstrapová rezidua $u^*(t)$, $t \in \{p+1, p+2, \dots, T\}$, nezávislým výběrem s vrácením z $u(s)$, $s \in \{p+1, p+2, \dots, T\}$, tedy

$$P[\hat{u}^*(t) = u(s)] = \frac{1}{T-p}, \quad s \in \{p+1, p+2, \dots, T\}.$$

6) Vygenerujeme bootstrapový proces $W^*(t)$, tak, že

$$\begin{aligned} W^*(t) &= W(t), & t \in \{1, 2, \dots, p\}, \\ W^*(t) &= \sum_{i=1}^p \pi_{t-i} W^*(t-i) + u^*(t), & t \in \{p+1, p+2, \dots, T\}. \end{aligned}$$

7) Budeme předpokládat, že $W^*(t)$ je VAR(p) proces. Odhadneme autoregresivní matice π_k^* tohoto procesu metodou nejmenších čtverců.

8) Označme

$$\widehat{Bias} = \frac{1}{K} \left(\sum_{k=1}^K \pi_k^* \right) - \hat{\pi}.$$

9) Spočteme upravená rezidua $a(t)$ procesu $W(t)$:

$$\begin{aligned} a(t) &= 0, & t \in \{1, 2, \dots, p\}, \\ a(t) &= W(t) - \sum_{i=1}^p [\hat{\pi}_{t-i} - \widehat{Bias}] W(t-i), & t \in \{p+1, p+2, \dots, T\}. \end{aligned}$$

10) Opakujeme kroky 11) až 15) B -krát. Opakování indexujeme indexem l .

11) Vygenerujeme bootstrapová residua $a^*(t), t \in \{p+1, p+2, \dots, T\}$ nezávislým výběrem s vrácením z $a(s), s \in \{p+1, p+2, \dots, T\}$.

12) Vygenerujeme bootstrapový proces $W^*(t)$, tak, že

$$W^*(t) = W(t), \quad t \in \{1, 2, \dots, p\},$$

$$W^*(t) = \sum_{i=1}^p (\hat{\pi}_{t-i} - \widehat{Bias}) W^*(t-i) + a^*(t), \quad t \in \{p+1, p+2, \dots, T\}.$$

13) Budeme předpokládat, že $W^*(t)$ je VAR(p) proces. Odhadneme autoregresivní matice $\hat{\pi}^*$ tohoto procesu metodou nejmenších čtverců. Od těchto odhadů odečteme \widehat{Bias} :

$$\hat{\pi}_j^* = \hat{\pi}_j^* - \widehat{Bias}.$$

Tím získáme upravený odhad $\hat{\pi}^*$.

14) Použijeme upravený odhad nejmenších čtverců pro odhad autokovarianční matice procesu $W^*(t)$:

$$\hat{a}^*(t) \stackrel{\text{def}}{=} W^*(t) - \sum_{j=1}^p \hat{\pi}_{p-j}^* W^*(t-j),$$

$$\nu \stackrel{\text{def}}{=} \frac{1}{T-p} \sum_{t=p+1}^T \hat{a}^*(t),$$

$$\tilde{a}^*(t) \stackrel{\text{def}}{=} \hat{a}^*(t) - \nu,$$

$$\hat{\Sigma}_a \stackrel{\text{def}}{=} \frac{1}{T-p} \sum_{t=p+1}^T \tilde{a}^*(t) \tilde{a}^*(t)'$$

15) Použijeme odhady $\hat{\pi}^*$ a $\hat{\Sigma}_a$ a znalost p k odhadu míry kauzality metodou G. Odhad označíme C_l . Víme, že míra musí být nezáporná; pokud by její odhad vycházel menší než 0, budeme klást $C_l = 0$.

16) Označme

$$\widehat{bias} = \frac{1}{B} \sum_{l=1}^B C_l - \hat{C}.$$

17) Uvažujme následující odhad měr kauzality:

$$C_i^* = \max\{0, C_i - \widehat{bias}\}.$$

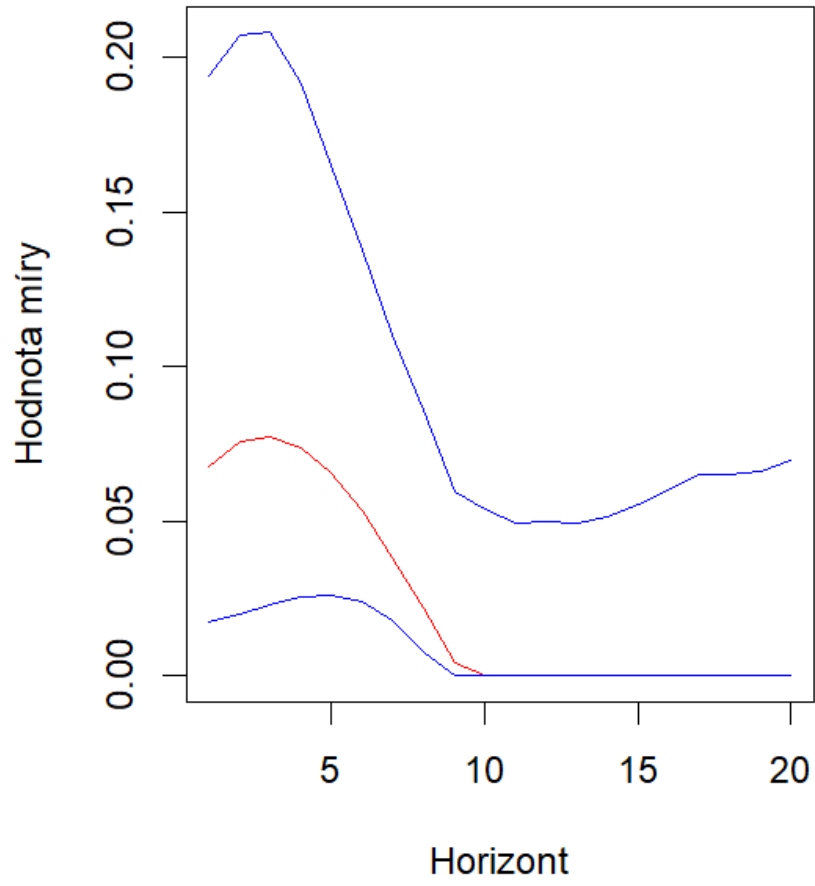
Zahodíme $\frac{\alpha}{2}B$ nejmenších a $\frac{\alpha}{2}B$ největších odhadů C_i^* . Ve zbytku najdeme maximum C_{max} a minimum C_{min} . Požadovaný interval spolehlivosti je (C_{min}, C_{max}) .

Poznámka: K a B mohou být různá. Růst B zvyšuje výpočetní náročnost mnohem více než růst K , proto doporučujeme volit K větší než B .

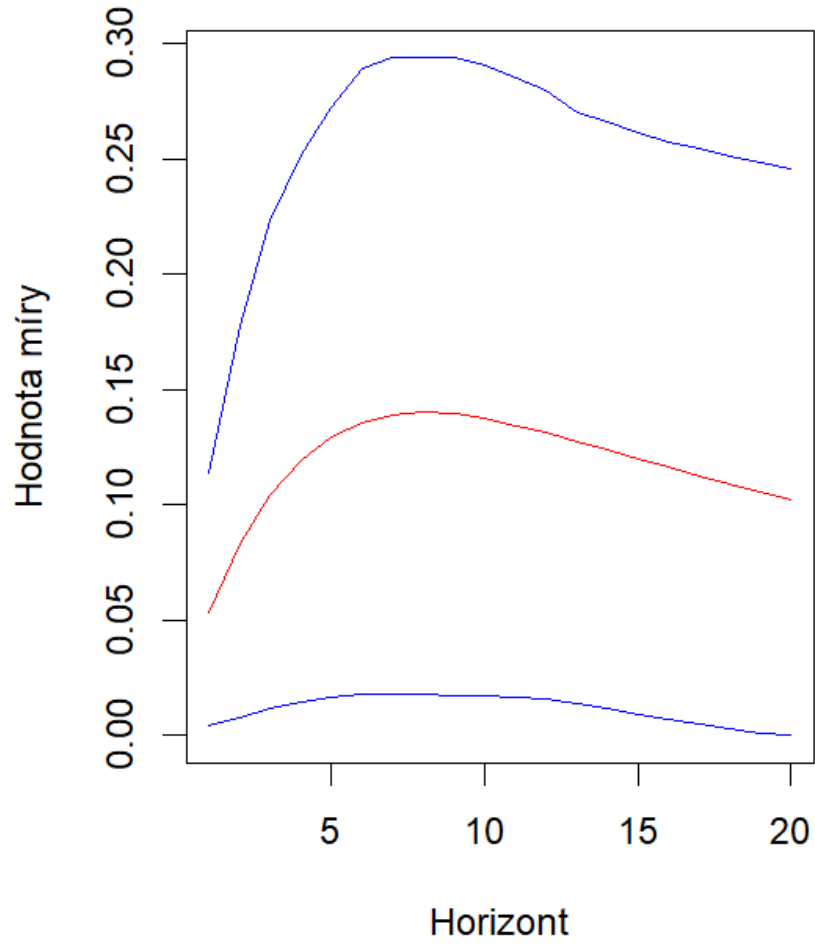
4.5 Aplikace odhadů měr kauzality

Z období 1. 1. 2010 až 1. 1. 2011 vezmeme data o hodnotách české koruny a eura v USD. Budeme zjišťovat míry kauzality hodnoty eura na hodnotu české koruny. Data získáme z webu <https://www.ofx.com/en-au/forex-news/historical-exchange-rates/>. Dále uvažujeme vektor Z hodnot měn zemí, které významně obchodují s Českou republikou nebo Evropskou Unií (Švýcarsko, Polsko, Velká Británie, Čína), opět v USD. Puštěním modifikovaného Dickeyova-Fullerova testu na hodnoty každé měny zjistíme, že je třeba brát veškerá data s první diferencí. Akaikeho informační kritérium říká, že je vhodné brát model VAR(1). Počítali jsme hodnoty měr kauzality v horizontech 1 až 20. Viz následující grafy.

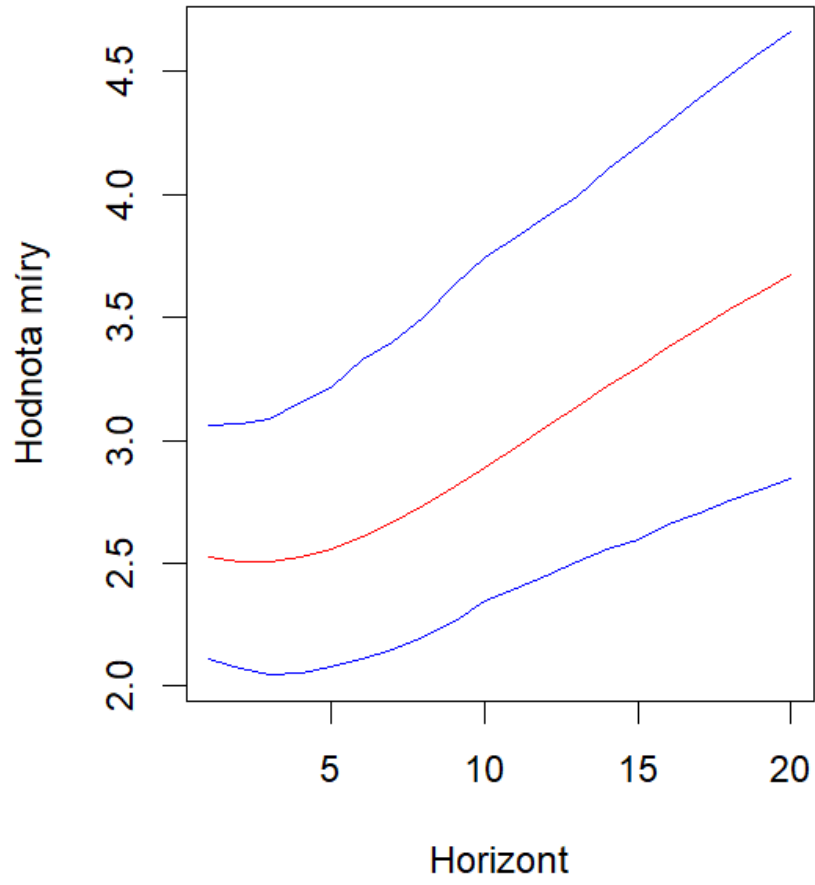
$$C_L[Y \xrightarrow{h} X | I_Z]$$



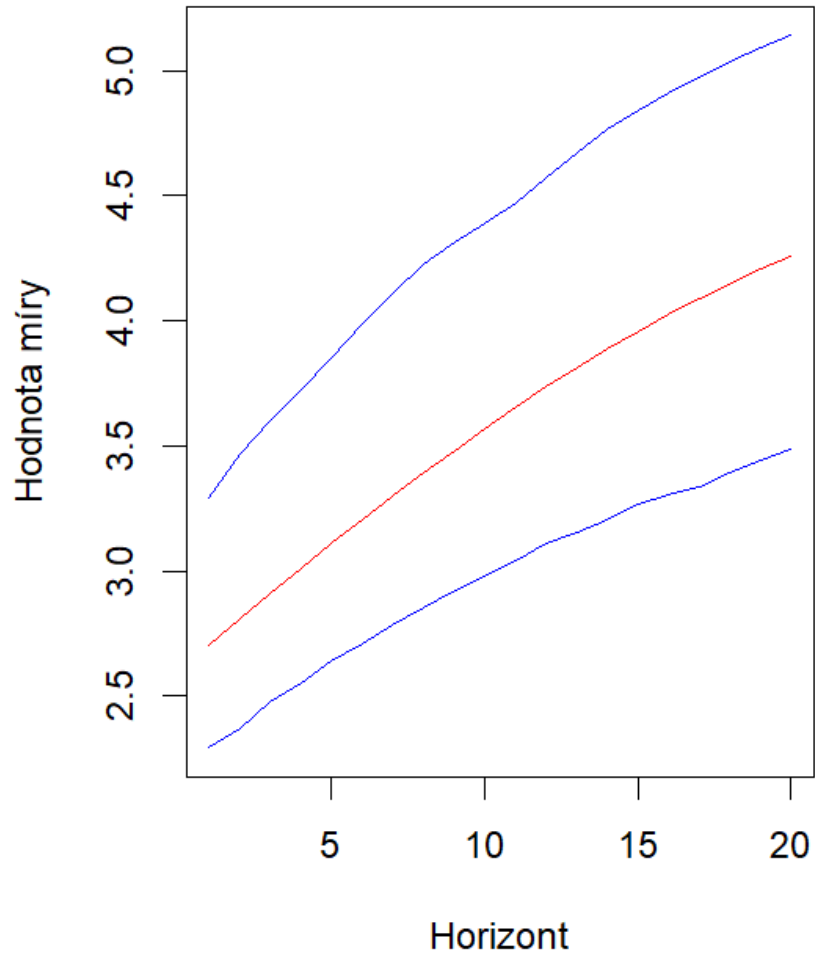
$$C_L[Y \xrightarrow{h} X|I]$$



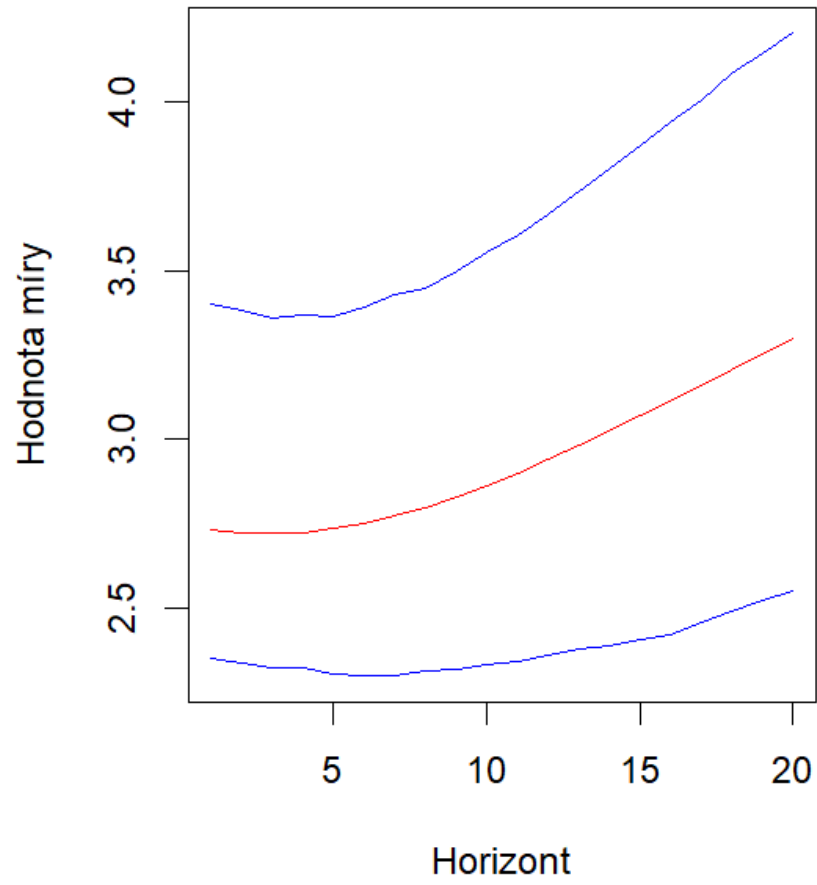
$$C_L[X - Y|I_Z]$$



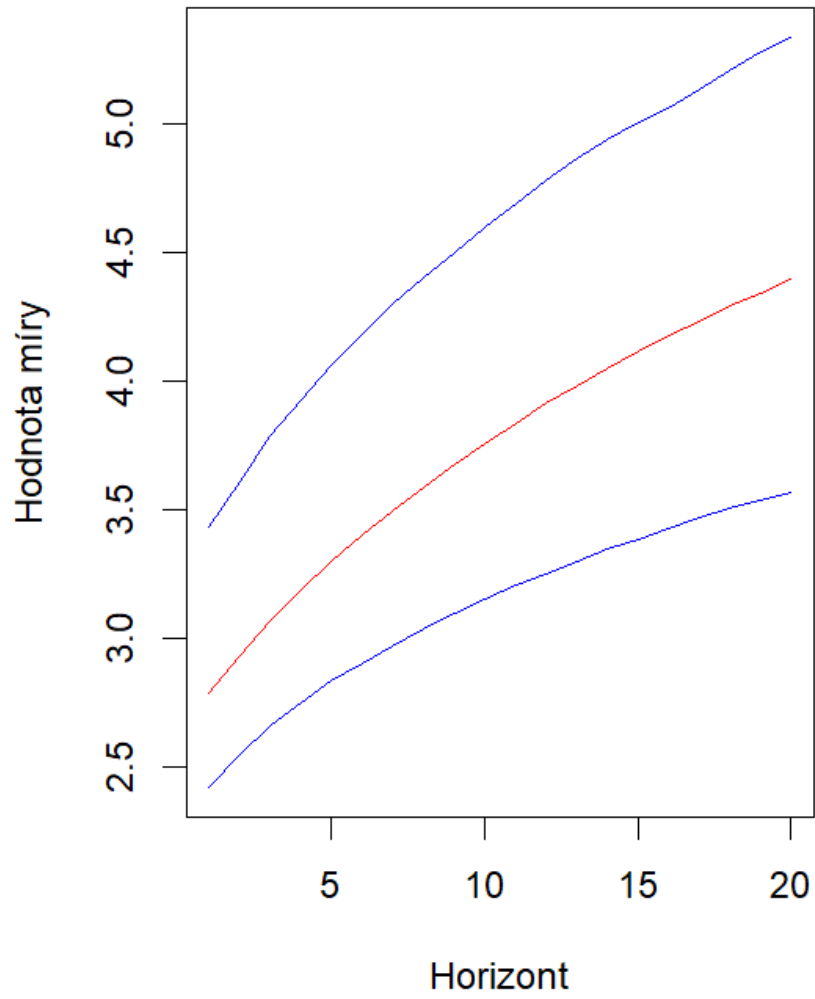
$$C_L[X - Y|I]$$



$$C_L^{(h)}[X, Y|I_Z]$$



$$C_L^{(h)}[X, Y|I]$$



Modré čáry označují interval o spolehlivosti 95 %, červená čára uprostřed je bodový odhad. Bylo zvoleno $B = 600$, $K = 100\,000$, $T = 100\,000$ a I prázdná množina. Odhad míry kauzality je víceméně nulový ve všech horizontech. Z duality intervalového odhadu a testování hypotéz (a z malého vychýlení našich intervalových odhadů) ovšem plyne, že (s chybou 2,5 %) kauzalita v malém horizontu existuje. Odhad míry okamžité kauzality roste a je velký; pro srovnání hodnota míry 2 odpovídá absolutní hodnotě korelace mezi $P[X|I(t+h)]$ a $P[Y|I(t+h)]$ zhruba 0,93. Hodnota 4 odpovídá už absolutní hodnotě korelace 0,99. Toto znamená, že odhad rychlosti růstu eura je (po přeškálování ne nutně nezápornou konstantou) také dobrým odhadem rychlosti růstu české koruny. Odhad míry závislosti je skoro roven odhadu míry okamžité závislosti, protože míra kauzality rychlosti růstu EUR na rychlost růstu CZK a míra kauzality rychlosti růstu CZK na rychlost růstu EUR je skoro nulová.

Vysoká hodnota okamžité míry kauzality a existence míry kauzality v horizontu 1 napovídá, že (významná) kauzalita existovat může, nicméně pouze v horizontu menším než jeden den. Je rozdíl mezi mírami vůči I a I_Z , hodnoty jsou větší pro míry vůči I , především pro větší horizonty. Toto může znamenat, že část informace o rychlosti růstu CZK dává jak rychlost růstu EUR, tak i rychlost růstu jiných měn.

Závěr

Uvažovali jsme stabilní, invertibilní VAR procesy. Nejprve jsme definovali lineární projekci těchto procesů, poté jsme definovali Grangerovu kauzalitu jakožto zlepšení této projekce. Ukázali jsme souvislost mezi lineární projekcí a VMA reprezentací VAR procesu. Poté jsme ukázali, že za drobných předpokladů lze zlepšení lineární projekce chápat jako snížení rozptylu chyby této projekce. Toto nám umožnilo vyčíslit míry kauzality.

Ukázali jsme způsob, jak vypočítat míry kauzality jen ze znalosti parametrů VAR procesu. Platí, že většina procesů, se kterými se v praxi setkáváme, lze (po dostatečném množství diferencování) dobře aproximovat $\text{VAR}(p)$ modelem, $p \in \mathcal{N}$. Kauzalita takového procesu také dobře aproximuje kauzalitu předchozího procesu. Nutné množství diferencí se dá zjistit modifikovaným Dickeyovým-Fullerovým testem a p lze zjistit z Akaikeho informačního kritéria. Nyní můžeme odhadnout parametry tohoto $\text{VAR}(p)$ procesu metodou nejmenších čtverců, odtud už lze vypočítat míry kauzality. Někdy potřebujeme vypočítat parametry podprocesu $\text{VAR}(p)$, uvedli jsme způsob výpočtu simulační metodou. Interval spolehlivosti jsme získali bootstrapovou metodou. Bootstrapová metoda vrací vychýlené odhady, námi uvedená metoda obsahuje korekce na nestrannost, které značně snižují vychýlení.

Míry kauzality mají široká uplatnění: nejenom nám umožní lépe zkoumat a vyčíslit příčinné vztahy, ale dokáží detekovat prekurzorní vztahy (= detekována kauzalita, i když příčinný vztah nemusí existovat; náhlá změna prekurzoru může varovat před náhlou změnou jiného procesu). Míra závislosti může ukázat na možné propojení procesů.

Zdroje

- [1] Dufour, J.-M., Pelletier, D., Renault, É.: *Short Run and Long Run Causality in Time Series: Inference*. Journal of Econometrics 132,2 (2009).
- [2] Dufour, J.-M., Renault, É.: *Short Run and Long Run Causality in Time Series: Theory*. Econometrica 66,5 (1998).
- [3] Dufour, J.-M., Taamouti, A.: *Short and Long Run Causality Measures: Theory and Inference*. Journal of Econometrics 154,1 (2010).
- [4] Granger, C.W.J.: *Investigating Causal Relations by Econometric Models and Cross-Spectral Methods*. Econometrica 37,3 (1969), 424-438.
- [5] Korda, J.: *Kauzalita jako metodologický problém ekonomie*. Electronic Journal For Philosophy, (2007), 1-11.
- [6] Marčiny, J.: *Grangerova kauzalita ve finančních časových řadách*. Bachelor Thesis of KPMS MFF UK, Praha, iv+32, (2011).
- [7] Alexandrian, A.: *On continuous dependence of roots of polynomials on coefficients*. The North Carolina State University, (2022).
- [8] Lewis, R., Reinsel, G.-C.: *Prediction of multivariate time series by autoregressive model fitting*. Journal of Multivariate Analysis 16, (1985), 393–411.

Přílohy

Program na výpočet měr kauzality v R.