

# O modelowaniu graficznym w kontekście analizy szeregów czasowych\*

Piotr Zwiernik

październik 2006 r.

## Streszczenie

W pracy na początku opisujemy podstawy modeli graficznych, które stanowią dużą i popularną rodzinę modeli statystycznych. Ich zaletą jest możliwość efektywnego modelowania nawet bardzo skomplikowanych systemów. Uogólniają one również bardzo wiele znanych modeli statystycznych, jak np. sieci bayesowskie, filtry Kalmana, ekonometryczne modele strukturalne. Następnie prezentujemy, jak rozszerzyć modele graficzne na przypadek szeregów czasowych. Jest to kierunek badań, który objawił się kilka lat temu i rozwija się bardzo dynamicznie. Korzystając z prezentowanych metod przedstawimy również model dla gospodarki amerykańskiej, który pozwala między innymi ocenić zdolność wskaźników gospodarczych do opisywania obecnej i przyszłej sytuacji ekonomicznej.

*Słowa kluczowe: modelowanie graficzne, koherencja cząstkowa, niezależność szeregów czasowych.*

---

\*Praca została sfinansowana w ramach badań własnych Instytutu Ekonometrii Szkoły Głównej Handlowej w Warszawie, nr badania 03/E/0009/06.

# Spis treści

<b>1</b>	<b>Wprowadzenie</b>	<b>2</b>
<b>2</b>	<b>Modele graficzne</b>	<b>4</b>
2.1	Grafy i relacja separowalności . . . . .	4
2.2	Niezależność warunkowa . . . . .	6
2.3	Modele grafów nieskierowanych . . . . .	8
2.4	Modele grafów skierowanych . . . . .	10
2.5	Wybór najlepszego modelu . . . . .	11
2.6	Przypadek gaussowskich modeli graficznych . . . . .	12
<b>3</b>	<b>Modele dla szeregów czasowych</b>	<b>13</b>
3.1	Modele interakcji . . . . .	15
3.2	Własności Markowa w grafach interakcji . . . . .	18
3.3	Estymacja modeli interakcji . . . . .	19
<b>4</b>	<b>Model dla gospodarki amerykańskiej</b>	<b>20</b>
<b>5</b>	<b>Zakończenie</b>	<b>23</b>

## 1 Wprowadzenie

Założmy, że  $n$ -elementowy zbiór wierzchołków  $V$  pewnego grafu  $\mathcal{G} = (V, E)$  odpowiada zmiennym losowym w analizowanym systemie  $X_V = (X_1, \dots, X_n)$  o wartościach w  $\mathcal{X} = \times_{i=1}^n \mathcal{X}_i$ . Niech zbiór krawędzi  $E$  grafu  $\mathcal{G}$  odpowiada pewnym zależnościom i niezależnościom warunkowym pomiędzy zmiennymi lub ich grupami. Wówczas - na razie mało formalnie - modelem graficznym będziemy nazywać trójkę  $\{\mathcal{X}, \mathcal{G}, \mathcal{P}_{\mathcal{G}}\}$ , gdzie  $\mathcal{G}$  jest grafem o zbiorze wierzchołków  $V$ , a  $\mathcal{P}_{\mathcal{G}}$  jest rodziną rozkładów prawdopodobieństwa dla  $X_V$ , które spełniają formuły zależności i niezależności warunkowych, które reprezentowane są przez zbiór krawędzi  $E$ .

Statystyczne modele graficzne stały się popularną metodą analizowania skomplikowanych systemów stochastycznych. Wyrazem tego było powstanie trzech ważnych monografii: Joe Whittakera [Whi90], Steffena Lauritzena [Lau96] oraz Davida Edwardsa [Edw00], a także setek różnych artykułów. Istnieje kilka powodów rosnącej popularności tej metody. Przede wszystkim, jak zostało wykazane, aksjomatyka niezależności warunkowych zmiennych losowych odpowiada w sposób jednoznaczny aksjomatyce relacji separowalności w grafach nieskierowanych. Oznacza to, że grafy

stanowią w pewnym sensie zupełną formę zapisu relacji niezależności zmiennych losowych. Istotne, że jest to również forma niezwykle zwięzła, gdyż liczba wszystkich możliwych relacji niezależności opisywanych przez graf jest niezwykle duża nawet dla małej liczby wierzchołków grafu. Po drugie, rozważanie grafów skierowanych pozwala na interpretowanie strzałki pomiędzy dwoma wierzchołkami, jako kierunek wpływu stochastycznego jednej zmiennej na drugą. Prowadzi to do pojęcia przyczynowości (patrz monografia Judea Pearla [Pea00]), które stanowi silną motywację dla stosowania modeli graficznych na przykład w naukach społecznych. Poznanie kierunków przyczynowych w analizowanym systemie jest nie tylko sposobem na pogłębienie wiedzy o nim, ale wspomaga także procesy decyzyjne. Po trzecie, co stanie się jasne w następnych sekcjach, modele graficzne pozwalają na podzielenie wielkiego systemu opisywanego przez model graficzny na jego prostsze podsystemy. W ten sposób uzyskujemy możliwość analizy skomplikowanego problemu poprzez badanie prostszych podproblemów. Po czwarte, jak się okazuje, modele graficzne stanowią uogólnienie ogromnej liczby modeli statystycznych istniejących w praktyce. W tym kontekście warto wymienić sieci bayesowskie, łańcuchy Markowa, ekonometryczne modele strukturalne, modele hierarchiczne, itd. To ogromna zachęta do poznania tej szerokiej rodziny modeli – jeżeli zrozumiemy coś w przypadku modeli graficznych, to najprawdopodobniej zrozumiemy to również dla szczególnych przypadków. Ostatecznie, modele graficzne zaopatrują nas w wygodny sposób wizualizacji wyników w postaci grafu. Pozostanie to tajemnicą ludzkiego umysłu dlaczego grafy tak intuicyjnie potrafią przedstawić naszą wiedzę o zależnościach pojawiających się w badanym układzie.

Modelowanie graficzne jest również wykorzystywane w analizie ekonometrycznej. Wiele modeli szeregów czasowych takich, jak na przykład przełącznikowe modele Markowa lub dynamiczne układy liniowe (zwane czasem filtrami Kalmana) mają swoją reprezentację za pomocą grafu. Ponad to, jak stwierdził w swoim artykule z 1996 roku guru analizy szeregów czasowych David Brillinger [Bri96] „można się spodziewać, że modele tego typu okażą się użyteczne w analizie najważniejszych problemów, które pojawiają się w strukturalnym modelowaniu ekonomicznym oraz w analizie związków przyczynowych zachodzących w systemach ekonomicznych”<sup>1</sup>. Jednak właśnie w kontekście przyczynowości napotkano na istotną trudność. Polegała ona na tym, że oryginalnie teoria modeli graficznych nie obejmowała przypadku, gdy wierzchołkami grafu były całe procesy stochastyczne (nieskończenie wymiarowe wektory losowe). Taka sytuacja jest wymagana, gdy chcemy budować hipotezy postaci: „proces stochastyczny  $\{X_t\}_{t=-\infty}^{\infty}$  wpływa na proces stochastyczny  $\{Y_t\}_{t=-\infty}^{\infty}$ ”. Niezależnie kilku badaczy zaczęło jednak pracę nad wkomponowaniem modeli graficznych w analizę szeregów czasowych. Ogólne podejście zaproponował David Brillinger [Bri96]. W tym duchu Rainer Dahlhaus [Dah00] przedstawił graficzne modele interakcji, które są analogonem

---

<sup>1</sup>Originalne zdanie brzmi: „It can be anticipated that such models will prove useful in addressing basic questions arising in economics concerning structural and causal modelling.”

klasycznych modeli grafów nieskierowanych. Były doktorant Dahlhausa – Michael Eichler – w serii swoich prac (patrz np. [Eic99], [Eic03]) zaprezentował próbę wykorzystania mieszanych modeli graficznych w analizie szeregów czasowych. Obecnie nasza wiedza o modelach graficznych dla szeregów czasowych jest nieporównywalnie większa niż jeszcze 5 lat temu, ale to wciąż wierzchołek góry lodowej.

Głównym celem tej pracy jest przedstawienie podstawowych wyników dotyczących przeniesienia metod modelowania graficznego na grunt analizy szeregów czasowych. Obejmuje to głównie wykorzystanie analizy spektralnej i spektralnych koherencji cząstkowych. Drugim celem jest przedstawienie głównych kierunków badawczych, ważnych w celu lepszego poznania opisywanych modeli, efektywniejszego sposobu ich estymacji i pełniejszego procesu wnioskowania na ich podstawie. W tym kontekście szczególnie obiecujące wydaje się wykorzystanie metodologii bayesowskiej, która stała się niezwykle popularna w przypadku klasycznych modeli graficznych.

Na początku w sekcji 2 przedstawimy podstawowe pojęcia związane z modelami graficznymi. W podsekcji 2.1 zaprezentujemy wykorzystywane dalej elementy teorii grafów. Następnie w podsekcji 2.2 w miarę szczegółowo zaprezentujemy pojęcie warunkowej niezależności w grupie zmiennych losowych, które jest podstawowe w dyskutowanej teorii. W podsekcjach 2.3 oraz 2.4 w możliwie ogólny sposób opiszemy idee modeli graficznych dla grafów nieskierowanych i skierowanych.

W sekcji 3 przedstawimy podstawowe rozszerzenia modeli graficznych, które umożliwiają wkomponowanie w tę metodologię również szeregów czasowych. Skupimy się na modelach interakcji opisywanych przez Dahlhausa. Następnie w sekcji 3.3 zaprezentujemy, jak w praktyce wygląda estymacja takich modeli. Rozważania obu powyższych sekcji pozwolą oszacować i zaprezentować model dla gospodarki amerykańskiej, w którym między innymi uzyskujemy ciekawe wnioski dotyczące użyteczności wskaźników ekonomicznych do opisu bieżącej i przyszłej sytuacji gospodarczej.

## 2 Modele graficzne

### 2.1 Grafy i relacja separowalności

W tej sekcji wprowadzimy podstawowe definicje związane z teorią grafów, które będą wykorzystywane w dalszej części tekstu. Polska nomenklatura zaprezentowana poniżej jest zgodna z zaprezentowaną w polskim tłumaczeniu książki Robina Wilsona [Wil04].

*Grafem* będziemy nazywać parę  $\mathcal{G} = (V, E)$ , gdzie  $V$  jest skończonym zbiorem wierzchołków, a  $E$  stanowi zbiór krawędzi i jest podzbiorem zbioru  $V \times V$ . Ograniczamy się dodatkowo do grafów bez pętli (krawędzi o początku i końcu w tym samym wierzchołku), które nie posiadają wielokrotnych krawędzi między tą samą parą wierzchołków. Krawędź  $(\alpha, \beta)$  będziemy nazywać

*nieskierowaną* jeżeli zarówno  $(\alpha, \beta)$ , jak i  $(\beta, \alpha)$  należą do  $E$ . Krawędź nieskierowaną będziemy na grafie zaznaczać odcinkiem. Gdy  $\alpha$  i  $\beta$  połączone są krawędzią nieskierowaną, będziemy pisać  $\alpha \sim \beta$  i mówić, że wierzchołki sąsiadują ze sobą. Zbiór wierzchołków sąsiadujących z wierzchołkiem  $\alpha$  będziemy oznaczać  $ne(\alpha)$ .

*Graf nieskierowany* to graf, którego wszystkie krawędzie są nieskierowane. Krawędź  $(\alpha, \beta)$  nazywamy *skierowaną* jeśli  $(\alpha, \beta) \in E$  oraz  $(\beta, \alpha) \notin E$ . Krawędź skierowaną będziemy zaznaczać na grafie strzałką. *Graf skierowany* to graf, którego wszystkie krawędzie są skierowane. Jeżeli w grafie istnieje strzałka od  $\alpha$  do  $\beta$ , to będziemy pisać  $\alpha \rightarrow \beta$  i mówić, że  $\alpha$  jest rodzicem  $\beta$ , a  $\beta$  jest dzieckiem  $\alpha$ . Zbiór rodziców wierzchołka  $\alpha$  będziemy oznaczać  $pa(\alpha)$ , a zbiór dzieci  $ch(\alpha)$ .

Graf, w którym każda para wierzchołków jest połączona albo linią albo strzałką nazywamy *pełnym*. Podzbiór  $A \subseteq V$  nazywamy pełnym, jeśli indukuje pełny podgraf grafu  $\mathcal{G}$ . Maksymalny (w sensie inkluzji) zbiór wierzchołków, takich, że indukowany podgraf jest pełny nazywamy *kliką*. Obrzeżem  $bd(A)$  podzbioru  $A$  zbioru wierzchołków nazywamy zbiór wierzchołków w  $V \setminus A$ , które są rodzicami lub sąsiadują z wierzchołkiem  $\alpha$ , tzn.  $bd(A) = pa(A) \cup ne(A)$ . Domknięcie zbioru  $A$  to  $cl(A) = A \cup bd(A)$ .

Ścieżką długości  $n$  z  $\alpha$  do  $\beta$  nazywamy sekwencję różnych wierzchołków  $\alpha = \alpha_0, \alpha_1, \dots, \alpha_n = \beta$  taka, że dla każdego  $i = 1, \dots, n$  zachodzi  $(\alpha_{i-1}, \alpha_i) \in E$ . Mówimy wtedy, że ścieżka wiedzie z  $\alpha$  do  $\beta$  i piszemy  $\alpha \mapsto \beta$ . Wierzchołki  $\alpha$  takie, że  $\alpha \mapsto \beta$  oraz  $\beta \not\mapsto \alpha$  nazywamy przodkami  $\beta$  oznaczając ich zbiór przez  $an(\beta)$ . Natomiast wierzchołki  $\alpha$  takie, że  $\beta \mapsto \alpha$  oraz  $\alpha \not\mapsto \beta$  nazywamy potomkami wierzchołka  $\beta$  i oznaczamy  $de(\beta)$ . Jeśli  $bd(\alpha) \subseteq A$  dla wszystkich  $\alpha \in A$ , mówimy, że  $A$  jest zbiorem rodowym<sup>2</sup> (ang. ancestral set). W grafach skierowanych zbiór  $A$  jest zbiorem rodowym wtedy i tylko wtedy, gdy  $an(\alpha) \subseteq A$  dla wszystkich  $\alpha \in A$ . W grafach nieskierowanych, zbiory rodowe są sumami spójnych składowych grafu. Przecięcie zbiorów rodowych jest zbiorem rodowym, a zatem dla każdego zbioru wierzchołków istnieje najmniejszy zbiór rodowy zawierający  $A$ , który będziemy oznaczać  $An(A)$ .

Interesującą klasę grafów, z punktu widzenia zastosowań w modelowaniu statystycznym, stanowią grafy łańcuchowe (ang. chain graphs). Są to grafy, dla których zbiór wierzchołków  $V$  można podzielić na ponumerowane podzbiory  $\{V(i)\}_{i=1}^T$  formujących tzw. łańcuch zależności  $V = V(1) \cup \dots \cup V(T)$  taki, że wszystkie krawędzie pomiędzy wierzchołkami w każdym z podzbiorów są nieskierowane oraz wszystkie krawędzie pomiędzy podzbiorami są skierowane i idą z podzbioru o niższym indeksie do podzbioru o wyższym indeksie. Można pokazać, że grafy takie odpowiadają

---

<sup>2</sup>Wedle naszej najlepszej wiedzy, w polskiej literaturze nie pojawiła się dotąd żadna publikacja dotycząca modeli graficznych. W związku z tym czasami stoimy wobec konieczności zaproponowania polskiego odpowiednika dla jakiegoś pojęcia w języku angielskim. W takich sytuacjach w nawiasach będziemy podawali brzmienie oryginału. Z nadzieją oczekujemy na wszelkie komentarze.

grafom bez cykli skierowanych (tzn. takich, że nie da się idąc wzdłuż strzałek wrócić do punktu wyjścia).

Dla grafu mieszanego  $\mathcal{G}$  definiujemy graf obyczajny (ang. moral graph)  $\mathcal{G}^m$  jako graf nieskierowany o takim samym zbiorze wierzchołków oraz zbiorze krawędzi takim, że  $\alpha$  i  $\beta$  sąsiadują w  $\mathcal{G}^m$  wtedy i tylko wtedy, gdy albo  $\alpha \rightarrow \beta$  albo  $\beta \rightarrow \alpha$  albo jeśli są  $\gamma_1, \gamma_2$  w tym samym podzbiorze  $V(i)$  takie, że  $\alpha \rightarrow \gamma_1$  oraz  $\beta \rightarrow \gamma_2$ . W przypadku skierowanych grafów acyklicznych (sieci Bayesowskich) oznacza to, że strzałki grafu zamieniamy na linie oraz dla każdego wierzchołka  $\alpha$  łączymy liniami wszystkie wierzchołki ze zbioru  $\text{pa}(\alpha)$ <sup>3</sup>.

Kluczowym pojęciem teorii grafów, z którego będziemy korzystać w kontekście rozważań probabilistycznych dla modeli graficznych, jest pojęcie separowalności. Wiąże się ono w sposób bezpośredni z pojęciem niezależności warunkowej, której napiszemy w podsekcji 2.2. Niech  $A$ ,  $B$ , i  $C$  będą trzema rozłącznymi podzbiórmi zbioru wierzchołków grafu  $\mathcal{G}$ . Wówczas będziemy mówić, że zachodzi warunek separowalności  $J(A, B|C)$  dla grafu  $\mathcal{G}$ , jeśli każda ścieżka pomiędzy wierzchołkiem w zbiorze  $A$  a wierzchołkiem w  $B$  zawiera wierzchołek z  $C$ . Równoważnie mówimy, że  $\mathcal{G}$  spełnia  $J(A, B|C)$  jeśli zbiory wierzchołków  $A$  i  $B$  są separowane przez zbiór wierzchołków  $C$ .

## 2.2 Niezależność warunkowa

Pojęcie niezależności warunkowej jest podstawowe w przedstawianej teorii. Niech  $X$ ,  $Y$  i  $Z$  będą trzema zmiennymi losowymi o łącznym rozkładzie  $P$ . Mówimy, że  $X$  jest warunkowo niezależny od  $Y$  pod warunkiem  $Z$  względem  $P$  i piszemy  $X \perp\!\!\!\perp Y|Z[P]$ , jeśli dla każdego mierzalnego zbioru  $A$  w przestrzeni wartości zmiennej  $X$  istnieje wersja prawdopodobieństwa warunkowego  $P(A|Y, Z)$ , która jest funkcją samego  $Z$ . Zwykle  $P$  będzie ustalony i pomijany w notacji.

Kiedy  $X$ ,  $Y$  i  $Z$  są dyskretnymi zmiennymi losowymi, warunek na  $X \perp\!\!\!\perp Y|Z$  upraszcza się do:

$$P(X = x, Y = y|Z = z) = P(X = x|Z = z) \cdot P(Y = y|Z = z),$$

gdzie równanie zachodzi dla wszystkich  $z$  spełniających  $P(Z = z) > 0$ . Jeżeli  $X$ ,  $Y$  i  $Z$  mają łączną gęstość względem miary produktowej  $\mu$ , mamy

$$X \perp\!\!\!\perp Y|Z \iff f_{XY|Z}(x, y|z) = f_{X|Z}(x|z) \cdot f_{Y|Z}(y|z), \quad (1)$$

gdzie równość zachodzi prawie na pewno względem  $P$ . Jeśli wszystkie gęstości są funkcjami ciągłymi, równanie 1 musi zachodzić dla wszystkich  $z$  spełniających  $f_Z(z) > 0$ . Warunek 1 można przepisać jako

$$X \perp\!\!\!\perp Y|Z \iff f_{XYZ}(x, y, z) \cdot f_Z(z) = f_{XZ}(x, z) \cdot f_{YZ}(y, z), \quad (2)$$

---

<sup>3</sup>Czyli żenimy rodziców, skąd tłumaczenie „graf obyczajny”.

gdzie równość musi zachodzić dla wszystkich  $z$ .

Istotną zaletą modeli graficznych jest, że istnieje bezpośredni związek pomiędzy relacją separowalności w grafie, a relacją warunkowej niezależności w sensie przedstawionym powyżej. Geiger i Pearl [GP93] w sposób systematyczny badali warunkowe niezależności w zbiorze zmiennych losowych. By być zgodnymi z ich notacją będziemy pisać  $I(X, Y|Z)$  zamiast  $X \perp\!\!\!\perp Y|Z$ . W cytowanej pracy Geiger i Pearl zdefiniowali aksjomat warunkowej niezależności  $I(\cdot, \cdot|\cdot)$ , jako pewien warunek dotyczący zestawu niezależności warunkowych

$$I(X_1, Y_1|Z_1), I(X_2, Y_2|Z_2), \dots, I(X_k, Y_k|Z_k) \implies I(X_{k+1}, Y_{k+1}|Z_{k+1}),$$

który musi być spełniony dla wszystkich modeli probabilistycznych zdefiniowanych na zbiorze wszystkich zmiennych. Oczywiście aksjomaty te wiążą się bezpośrednio z definicją warunkowej niezależności. Przykładem aksjomatów warunkowej niezależności są:

- Trywialna niezależność:  $I(X, \emptyset|Y)$ .
- Symetria:  $I(X, Y|Z) \implies I(Y, X|Z)$ .
- Dekompozycja:  $I(X, Y \cup W|Z) \implies I(X, Y|Z)$ .
- $I(X \cup W, Y|Z), I(X, W|Z \cup Y) \implies I(X, Y \cup W|Z)$ .
- $I(X, Y \cup W|Z) \implies I(X, Y|Z \cup W)$ .

Kluczowym dla nas wynikiem pracy Geigera i Pearla jest stwierdzenie, którego nie będziemy dowodzić, że każdy aksjomat warunkowej niezależności jest aksjomatem separowalności w grafie nieskierowanym. Jeżeli  $A, B, C$  są rozłącznymi zbiorami zbioru wierzchołków  $V$  oraz  $X_A = (X_\alpha)_{\alpha \in A}$ , to powyższy wynik oznacza, że

$$I(X_A, X_B|X_C) \iff J(A, B|C).$$

Zauważmy, że dzięki temu każdy graf możemy traktować jako pełną specyfikację systemu zależności i niezależności warunkowych wśród zmiennych reprezentowanych przez wierzchołki tego grafu. Jeśli zachodzi odpowiedni warunek separowalności w grafie, zapewniony jest również odpowiedni warunek warunkowej niezależności. Wynik ten stał się główną motywacją dla stosowania modeli graficznych. W dalszej perspektywie uzasadnił również potrzebę badania związków relacji separowalności z warunkowymi niezależnościami w przypadku grafów skierowanych i mieszanych. Tej części teorii nie będziemy opisywać w niniejszej pracy, jednak warto podkreślić, że badania te były niezwykle ważne jeśli chodzi o wykorzystanie modeli graficznych w analizie przyczynowości.

## 2.3 Modele grafów nieskierowanych

Niech  $\mathcal{G} = (V, E)$  będzie grafem nieskierowanym. Z grafem tym zwiążemy zbiór zmiennych losowych  $X_V = (X_1, \dots, X_n)$  o wartościach w przestrzeniach  $\mathcal{X} = \times_{i=1}^n \mathcal{X}_i$ , które będą albo skończeniowymi wymiarowymi przestrzeniami wektorowymi lub skończonymi zbiorami. Dla podzbioru  $A$  zbioru  $V$  niech  $\mathcal{X}_A = \times_{i \in A} \mathcal{X}_i$ . Elementy przestrzeni  $\mathcal{X}_A$  będziemy zwykle oznaczać przez  $x_A$ . Zamiast  $X_A \perp\!\!\!\perp X_B | X_C$  będziemy skrótowo pisać  $A \perp\!\!\!\perp B | C$ .

Kluczowym pojęciem wiążącym teoriografowe pojęcie separowalności oraz pojęcie warunkowej niezależności zmiennych losowych jest własność Markowa grafu. Mówimy, że miara probabilistyczna  $P$  na  $\mathcal{X}$  spełnia:

**(P)** wzajemną własność Markowa (ang. pairwise Markov property) względem  $\mathcal{G}$ , jeśli dla każdej pary  $(\alpha, \beta)$  nie sąsiadujących ze sobą wierzchołków

$$\alpha \perp\!\!\!\perp \beta | V \setminus \{\alpha, \beta\}.$$

**(L)** lokalną własność Markowa (ang. local Markov property) względem  $\mathcal{G}$ , jeśli dla każdego wierzchołka  $\alpha \in V$

$$\alpha \perp\!\!\!\perp \{V \setminus \text{cl}(\alpha)\} | \text{bd}(\alpha).$$

**(G)** globalną własność Markowa (ang. global Markov property) względem  $\mathcal{G}$ , jeśli dla każdej trójki  $(A, B, S)$  rozłącznych podzbiorów zbioru  $V$  takich, że  $S$  separuje  $A$  od  $B$  w  $\mathcal{G}$

$$A \perp\!\!\!\perp B | S.$$

Oczywiście rozważania poprzedniego paragrafu pokazują, że w przypadku nieskierowanych modeli graficznych mamy spełnioną globalną własność Markowa. Ale stosunkowo łatwo zobaczyć, że zachodzi ciąg implikacji:  $(G) \implies (L) \implies (P)$ . Pearl i Paz [PP87] pokazali, że dla grafów nieskierowanych wszystkie trzy własności Markowa są równoważne, jeśli dla wszystkich rozłącznych czwórek podzbiorów  $A, B, C$  i  $D$  zachodzi:

$$\left( A \perp\!\!\!\perp B | (C \cup D) \text{ oraz } A \perp\!\!\!\perp C | (B \cup D) \right) \implies A \perp\!\!\!\perp (B \cup C) | D.$$

Warunek ten jest spełniony na przykład, gdy  $P$  posiada dodatnią i ciągłą gęstość względem miary produktowej  $\mu$ .

Jak to już zostało przedstawione wcześniej, warunkowa niezależność w przypadku zmiennych ciągłych łączy się w sposób bezpośredni z faktoryzacją gęstości łącznej. W związku z tym należy również oczekiwać takiego związku między faktoryzacją i własnościami Markowa. Mówimy, że



miara probabilistyczna  $P$  na  $\mathcal{X}$  faktoryzuje się względem  $\mathcal{G}$ , jeśli dla wszystkich podzbiorów  $A \subseteq V$ , które indukują pełne podgrafy grafu  $\mathcal{G}$ , istnieją nieujemne funkcje  $\psi_A$ , które zależą od  $x \in \mathcal{X}$  tylko poprzez  $x_A$  oraz istnieje miara produktowa  $\mu = \otimes_{i=1}^n \mu_i$  na  $\mathcal{X}$  taka, że  $P$  ma gęstość  $f$  względem  $\mu$ , gdzie  $f$  jest postaci

$$f(x) = \prod_{A \text{ pełny}} \psi_A(x).$$

Analogiczną definicję przeprowadzamy dla rozkładów dyskretnych.

Funkcje  $\psi_A$  nie są zadane jednoznacznie. Często w praktyce zakłada się, że w produkcie występują jedynie kliki grafu  $\mathcal{G}$ , to znaczy łączny rozkład prawdopodobieństwa dla  $X_V$  ma postać

$$f(x) = \prod_{c \in \mathcal{C}} \psi_c(x), \quad (3)$$

gdzie produkt przebiega po wszystkich klikach grafu  $\mathcal{G}$ . Na przykład w przypadku, gdy  $f$  jest gęstością wielowymiarowego rozkładu normalnego, sposób w jaki faktoryzuje się  $f$  zadaje zerowe warunki na pewne parametry tego rozkładu (patrz paragraf 2.6).

Jeżeli  $P$  się faktoryzuje, za Lauritzenem [Lau96] będziemy mówić, że  $P$  spełnia własność (F). Rodzinę takich miar probabilistycznych będziemy oznaczać  $M_F(\mathcal{G})$ .

**Twierdzenie 2.1** *Dla każdego nieskierowanego grafu  $\mathcal{G}$  i każdego rozkładu prawdopodobieństwa na  $\mathcal{X}$  zachodzi*

$$(F) \implies (G) \implies (L) \implies (P)$$

*Dowód:* patrz [Lau96], twierdzenie 3.8.□

Okazuje się, że jak wcześniej przy spełnieniu pewnych warunków można uzyskać równoważność wszystkich czterech własności. Mówi o tym poniższe, bardzo ważne twierdzenie.

**Twierdzenie 2.2 (Hammersley-Clifford)** *Rozkład prawdopodobieństwa  $P$  o dodatniej i ciągłej gęstości  $f$  względem miary produktowej  $\mu$  spełnia wzajemną własność Markowa względem nieskierowanego grafu  $\mathcal{G}$  wtedy i tylko wtedy, gdy faktoryzuje się względem  $\mathcal{G}$ .*

*Dowód:* patrz [Lau96], twierdzenie 3.9.□

Powyższy wynik, rozważania tej sekcji oraz wynik Geigera i Pearla zaprezentowany w poprzedniej sekcji, pokazują jak głęboka jest zależność między grafami nieskierowanymi oraz modelami statystycznymi, które spełniają system warunkowych niezależności zakodowanych w tym grafie. Okazuje się jednak, że tych nietrywialnych zależności jest znacznie więcej. Pomijamy ich opis odsyłając do monografii Lauritzena.

## 2.4 Modele grafów skierowanych

W tym opracowaniu, tak jak to się z reguły dzieje w praktyce, w przypadku grafów skierowanych będziemy się ograniczać głównie do grafów acyklicznych, tzn. takich, w których nie występują cykle skierowane. Jest to bardzo ważny przypadek z punktu widzenia wnioskowania statystycznego, gdyż wówczas każda strzałka w grafie może być interpretowana jako istnienie oraz kierunek związku stochastycznego pomiędzy dwiema zmiennymi. Bardziej praktycznym wymiarem ograniczania się do acyklicznych grafów skierowanych jest możliwość ponumerowania wierzchołków tak, że strzałki w grafie będą zawsze od wierzchołka o niższym indeksie do wierzchołka o wyższym indeksie. Podstawowe definicje związane z teorią grafów skierowanych zostały przedstawione w paragrafie 2.1.

Rozważając modele grafów nieskierowanych ograniczaliśmy się do rozkładów prawdopodobieństwa, które podlegają odpowiedniej faktoryzacji wyznaczanej przez graf. Podobnie będzie w przypadku grafów skierowanych. Mówimy, że rozkład prawdopodobieństwa  $P$  podlega rekursywnej faktoryzacji (ang. recursive factorization) względem  $\mathcal{G}$ , jeśli istnieją nieujemne funkcje  $k^i(\cdot, \cdot)$ ,  $i \in V$  zadane na  $\mathcal{X}_i \times \mathcal{X}_{\text{pa}(i)}$  takie, że

$$\int k^i(y_i, x_{\text{pa}(i)}) \mu_a(dy_i) = 1,$$

oraz  $P$  ma gęstość  $f$  względem  $\mu$ , gdzie

$$f(x) = \prod_{i \in V} k^i(x_i, x_{\text{pa}(i)}).$$

Warunki te mówią, że w istocie funkcje  $k^i(\cdot, x_{\text{pa}(i)})$  są gęstościami rozkładu warunkowego zmiennej  $X_i$  pod warunkiem  $X_{\text{pa}(i)} = x_{\text{pa}(i)}$ . Analogiczną definicję możemy sformułować dla zmiennych dyskretnych.

**Lemat 2.3** *Jeśli  $P$  podlega rekursywnej faktoryzacji względem skierowanego acyklicznego grafu  $\mathcal{G}$ , wówczas faktoryzuje się względem grafu obyczajnego  $\mathcal{G}^m$  a zatem spełnia globalną własność Markowa względem  $\mathcal{G}^m$ .*

*Dowód:* Z samej definicji grafu obyczajnego (patrz paragraf 2.1) oraz z faktu, że  $\mathcal{G}$  jest skierowany wynika, że zbiory  $\{i\} \cup \text{pa}(i)$  indukują pełne podgrafy w grafie  $\mathcal{G}^m$ . Niech zatem  $\psi_{\{i\} \cup \text{pa}(i)} = k^i$ . Z twierdzenia 2.1 dostajemy, że (F) implikuje (G), a zatem tezę.  $\square$

Dowód powyższego lematu pozwala sformułować następujące twierdzenie.

**Twierdzenie 2.4** *Jeśli  $P$  podlega rekursywnej faktoryzacji względem skierowanego acyklicznego grafu  $\mathcal{G}$  oraz  $A$  jest zbiorem rodowym, wówczas rozkład brzegowy  $P_A$  podlega rekursywnej faktoryzacji względem indukowanego podgrafu  $\mathcal{G}_A$ .*

A z tego natychmiast wynika.

**Wniosek 2.5** *Niech  $P$  podlega rekursywnej faktoryzacji względem  $\mathcal{G}$ . Wówczas*

$$A \perp\!\!\!\perp B \mid S$$

*zawsze, gdy  $S$  separuje  $A$  od  $B$  w grafie  $(\mathcal{G}_{\text{An}(A \cup B \cup S)})^m$*

Własność z powyższego wniosku będziemy nazywać skierowaną globalną własnością Markowa (DG). Ma ona podobną rolę jak globalna własność Markowa w przypadku grafów nieskierowanych. Jak widać, w przypadku skierowanym, inaczej niż to było w poprzedniej podsekcji, globalnej własności Markowa nie daje się bezpośrednio odczytać z grafu  $\mathcal{G}$ . W praktyce jednak nie jest to żadnym ograniczeniem, gdyż w sposób natychmiastowy potrafimy uzyskać żądany graf obyczajny.

W literaturze często formułuje się skierowaną globalną własność Markowa w alternatywny sposób używając konceptu  $d$ -separacji (patrz Pearl [Pea00], s. 16) i nie definiując wówczas grafów obyczajnych. W tym podejściu aby sprawdzić, czy zbiory zmiennych  $X_A$  i  $X_B$  są warunkowo niezależne pod warunkiem zmiennych  $X_C$  musimy sprawdzić, czy w grafie  $\mathcal{G}$  (skierowanym, acyklicznym) wszystkie ścieżki od wierzchołków w  $A$  do wierzchołków w  $B$  (lub na odwrót) są blokowane przez wierzchołki w  $C$ . Mówimy, że ścieżka  $p$  jest  $d$ -separowana (blokowana) przez zbiór wierzchołków  $C$  wtedy i tylko wtedy, gdy:

- (i)  $p$  zawiera łańcuch  $\alpha \rightarrow m \rightarrow \beta$  lub rozwidlenie (ang. fork)  $\alpha \leftarrow m \rightarrow \beta$  takie, że środkowy wierzchołek  $m$  należy do zbioru  $C$  lub
- (ii)  $p$  zawiera zwidlenie (ang. collider)  $\alpha \rightarrow m \leftarrow \beta$  takie, że  $m$  nie należy do zbioru  $C$  i dodatkowo żaden z potomków wierzchołka  $m$  nie należy do  $C$ .

Można pokazać, że obie charakteryzacje są równoważne. Warto je obie znać, gdyż w różnych zastosowaniach różne się przydają.

Wyprowadzenie własności Markowa dla grafów skierowanych pokazuje, że jednak jest zmiana jakościowa w porównaniu z prostotą analizy w przypadku grafów nieskierowanych. Pokazuje to również, że badanie własności Markowa jest głównym problemem w rozszerzaniu teorii na grafy dowolnego rodzaju. Piszemy o tym fakcie, gdyż jest on również istotnym ograniczeniem w przypadku modeli graficznych dla szeregów czasowych.

## 2.5 Wybór najlepszego modelu

W przypadku wnioskowania statystycznego niezwykle istotne jest by dysponować wiarygodną metodą wyboru właściwego modelu graficznego opisującego badany układ. Metoda wyboru najlep-

szego modelu obejmuje wybór dobrego kryterium oceny modelu, oraz efektywnej procedury przeszukiwania zadanej podprzestrzeni przestrzeni modeli.

Jeżeli chodzi o kryteria, to można w tym momencie przywołać elementy bardzo ogólnej dyskusji na temat optymalnej oceny modelu, która aplikuje się nie tylko do modeli graficznych. Zwyczajowo stosuje się jakieś kryterium informacyjne, którego ideaą jest nakładanie na wartość funkcji wiarygodności modelu dodatkowej kary związanej z jego wielkością. Dwa najczęściej stosowane kryteria to kryterium AIC (Akaike Information Criterion) oraz kryterium BIC (Bayesian Information Criterion).

Istnieją różne procedury przeszukiwania dostępnej przestrzeni modeli graficznych. W przypadku modeli nieskierowanych najpopularniejsze w praktyce są metody sekwencyjne, które polegają na wyrzucaniu (lub dodawaniu) kolejnych krawędzi z jakiegoś wyjściowego grafu – np. grafu pełnego. Decyzja o odrzuceniu (lub dodaniu) podejmowana jest na podstawie wartości statystyki ilorazu wiarygodności dla dwóch modeli reprezentowanych przez graf z krawędzią oraz bez niej. Statystyka ta przy założeniu prawdziwości modelu o większej liczbie krawędzi ma asymptotyczny rozkład  $\chi^2$ . Szczegóły opisane są w szóstym rozdziale książki Edwardsa [Edw00]. Tam również opisane są inne stosowane w praktyce metody wyboru możliwie najlepszego modelu graficznego dla zaobserwowanych danych.

Zastanówmy się teraz nad trochę innym problemem. Załóżmy, że postać modelu graficznego jest dana. Możemy wówczas zastanawiać się nad siłą powiązania związaną z konkretną krawędzią. Brillinger [Bri96] zaproponował stosowanie statystyki wzajemnej informacji (ang. mutual information). Rozważmy dwie ciągłe zmienne losowe  $X$  i  $Y$  z gęstością łączną  $p(x, y)$  oraz gęstościami brzegowymi  $p_X(x)$  i  $p_Y(y)$ . Statystyka wzajemnej informacji pomiędzy zmiennymi  $X$  i  $Y$  wyrażona jest wzorem (patrz np. Cover, Thomas [CT91], s. 6):

$$I_{XY} = \iint p(x, y) \log \frac{p(x, y)}{p_X(x)p_Y(y)} dx dy. \quad (4)$$

Statystyka ta przyjmuje wartość zero, jeśli  $X$  i  $Y$  są niezależne i jest dodatnia w pozostałych przypadkach. Brillinger zaproponował również drugą użyteczną miarę, która równa jest zmianie w wartości funkcji wiarygodności spowodowanej wykasowaniem badanej krawędzi. W tym wypadku miara ta jest równa podwojonej statystyce wzajemnej informacji, a zatem istotnie stanowi ona intuicyjnie dobrą miarę siły zależności związanej z daną krawędzią.

## 2.6 Przypadek gaussowskich modeli graficznych

W tym paragrafie opiszemy prosty przypadek modelu graficznego. Ograniczymy się do grafów nieskierowanych i zmiennych o rozkładach gaussowskich. Jest to ważny przykład, gdyż w dalszej

części tej pracy pomoże nam również zarysować podobieństwa między modelami graficznymi a modelami graficznymi dla szeregów czasowych.

Niech  $X = [X_1, \dots, X_n]$  będzie  $n$ -wymiarową zmienną losową o wielowymiarowym rozkładzie normalnym ze średnią

$$\mu = [\mu_1, \dots, \mu_n]'$$

oraz macierzą kowariancji

$$\Sigma = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_{nn} \end{bmatrix}.$$

Niech dodatkowo  $\Omega = \Sigma^{-1} = [\omega_{ij}]$  będzie odwrotnością macierzy kowariancji zwaną macierzą precyzji.

W paragrafie 2.3 pokazaliśmy, że w przypadku nieskierowanym brak krawędzi w grafie pomiędzy wierzchołkami  $i$  i  $j$  odpowiada warunkowej niezależności zmiennych  $X_i$  i  $X_j$  pod warunkiem wszystkich innych zmiennych w układzie. Bez straty ogólności rozważań możemy założyć, że  $i = 1$  i  $j = 2$ . Warunkowy rozkład  $(X_1, X_2)|(X_3, \dots, X_n)$  jest dwuwymiarowym rozkładem normalnym o macierzy kowariancji

$$\begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}^{-1} = \frac{1}{\omega_{11}\omega_{22} - (\omega_{12})^2} \cdot \begin{bmatrix} \omega_{22} & -\omega_{21} \\ -\omega_{12} & \omega_{11} \end{bmatrix}$$

oraz współczynnika korelacji

$$\rho_{12|3\dots n} = \frac{-\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}},$$

który będziemy nazywać współczynnikiem korelacji cząstkowej (patrz Whittaker [Whi90]). W przypadku zmiennych gaussowskich zachodzi

$$\rho_{12|3\dots n} = 0 \quad \Leftrightarrow \quad \omega_{12} = 0 \quad \Leftrightarrow \quad Y_1 \perp\!\!\!\perp Y_2 | (Y_3, \dots, Y_n),$$

a zatem wszystkie informacje o zbiorze wierzchołków grafu zapisane są w macierzy precyzji lub w macierzy korelacji cząstkowych  $R = [\rho_{ij|reszta}]$ . Jasne jest, że w praktyce sprawdzamy, które elementy estymowanej macierzy  $\hat{R}$  nie są istotnie różne od zera. Alternatywnie można badać zmianę w wartości funkcji wiarygodności pomiędzy modelami, w których  $\rho_{ij|reszta}$  jest zerowe lub dowolne.

### 3 Modele dla szeregów czasowych

Najczęściej modele graficzne szeregów czasowych zdefiniowane są w domenie czasowej, tzn. definiują one prawdopodobieństwo przejścia zdefiniowane na przestrzeni zmiennych stanu warunkowe względem ich wartości w przeszłych okresach. Modele te w literaturze często nazywa się dynamicznymi

sieciami Bayesowskimi (patrz Murphy [Mur02]). W modelach tych wierzchołki grafu reprezentowane są przez zmienne losowe reprezentujące wartość procesu w danym okresie  $t$ . Najpopularniejsze przykłady to łańcuchy Markowa, przełącznikowe modele Markowa, czy filtry Kalmana.

W tej sekcji będziemy się zajmować takimi modelami graficznymi dla szeregów czasowych, w których grafy są nieskierowane a wierzchołki reprezentują całe procesy. Jest to najprostszy przypadek. Z punktu widzenia analizy ekonometrycznej jest to również przypadek najmniej ciekawy, gdyż rezygnując z krawędzi skierowanych, ignorujemy związki przyczynowe w badanym systemie. Jednak jak się później okaże, niżej opisywana metoda ciągle pozwala wyciągnąć z danych wiele ciekawych informacji o badanym układzie. Poza tym, jest to historycznie pierwsze podejście do traktowania wierzchołków w grafach modeli graficznych jako całych procesów. Zatem podejście to ma również zaletę taką, że buduje intuicję i mówi, z jakimi nowymi problemami się spotkamy rozszerzając analizę na szeregi czasowe.

W przypadku grafów nieskierowanych analiza znacznie się upraszcza jeśli będziemy operować w dziedzinie częstotliwości. Niech  $X(t) = [X_1(t), \dots, X_k(t)]$  będzie  $k$ -wymiarowym szeregiem czasowym. Siłę związku pomiędzy  $i$ -tym a  $j$ -tym komponentem mierzy funkcja korelacji krzyżowej (patrz Brillinger [Bri96], s. 10)

$$\rho_{ij}(s) = \text{corr}(X_i(t+s), X_j(t))$$

lub równoważnie funkcja koherencji

$$R_{ij}(\omega) = \text{corr}\left(\sum_{t=-\infty}^{\infty} e^{-i\omega t} X_i(t), \sum_{t=0}^{n-1} e^{-i\omega t} X_j(t)\right),$$

jako funkcja częstotliwości  $\omega$ . Obie te funkcje są zerowe jeśli szeregi  $X_i(t)$  oraz  $X_j(t)$  są niezależne. Macierz koherencji jest blisko związana z macierzą korelacyjną, składa się ona z korelacji komponentów związanych z częstotliwością  $\omega$  dla obu szeregów  $X_i(t)$  i  $X_j(t)$ .

W praktyce jednak, badając złożony system i rozważając parę szeregów z tego systemu, jesteśmy zainteresowani w pewnym sensie „czystym” związkiem pomiędzy tymi szeregami. Często bowiem zdarza się w rzeczywistości, że dwa szeregi związane są ze sobą tylko dlatego, że istnieje jakiś trzeci proces, który na obie wpływa. Podobnie jak w podsekcji 2.6 badanie korelacji cząstkowych pozwalało taki „czysty” efekt uzyskać, tak w przypadku szeregów cząstkowych będziemy mówić o koherencjach cząstkowych, które mówią o związku pomiędzy dwoma szeregami po eliminacji liniowych efektów wszystkich innych szeregów. Rozważanie tych wielkości prowadzi do modeli interakcji dla szeregów czasowych.

### 3.1 Modele interakcji

Pokażemy, jak w najprostszy sposób przenieść metody modelowania graficznego na przypadek szeregów czasowych. Będziemy w tym miejscu wzorować się na podejściu zaprezentowanym w pierw przez Brillingera [Bri96] a sformalizowanym później przez Dahlhausa [Dah00].

Mówimy, że szeregi czasowe  $X_i$  i  $X_j$  są niezależne wtedy i tylko wtedy, gdy nieskończone wektory losowe  $\{X_i(t) : t \in \mathbb{Z}\}$  oraz  $\{X_j(t) : t \in \mathbb{Z}\}$  są niezależne. Podobnie szeregi czasowe  $X_i$  i  $X_j$  są niezależne pod warunkiem szeregu czasowego  $X_k$  wtedy i tylko wtedy, gdy nieskończone wektory losowe  $\{X_i(t) : t \in \mathbb{Z}\}$  oraz  $\{X_j(t) : t \in \mathbb{Z}\}$  są niezależne pod warunkiem  $\{X_k(t) : t \in \mathbb{Z}\}$  (porównaj paragraf 2.2).

Dla klasycznych modeli graficznych, w których dodatkowo zakładamy, że zmienne reprezentowane przez wierzchołki mają rozkład gaussowski, niezależności brzegowe i warunkowe mogą być odczytane z macierzy precyzji (patrz paragraf 2.6). Jak się okaże, daje się to rozszerzyć w sposób naturalny na stacjonarne procesy gaussowskie, jeśli zastąpimy macierz kowariancji na macierz gęstości spektralnej. Mówiąc niezbyt formalnie: stacjonarny proces gaussowski  $X(t) = (X_1(t), \dots, X_k(t))$  posiada graf warunkowej niezależności  $\mathcal{G} = (V, E)$ , gdzie  $V = \{1, \dots, k\}$ , jeśli brak krawędzi  $(i, j)$  odpowiada temu, że składniki  $X_i$  oraz  $X_j$  są nieskorelowane pod warunkiem informacji o wszystkich innych składnikach procesu  $X$ . Formalniej<sup>4</sup>: Niech  $Y_{ij}(t) = (X_k(t) : k \neq i, j)$ . Będziemy eliminować efekt liniowy procesu  $Y_{ij}$  z  $X_i(t)$  poprzez zadania optymalnego  $\mu_i$  oraz optymalnego filtru  $1 \times (k - 2)$  postaci  $\{d_i(u)\}$  takiego, że

$$\mathbb{E}\left(X_i(t) - \mu_i - \sum_{u=-\infty}^{\infty} d_i(t-u)Y_{ij}(u)\right)^2$$

jest minimalne. Reszty będziemy oznaczać przez  $\epsilon_i(t)$ , a dokładniej

$$\epsilon_i(t) := \epsilon_{i|\{i,j\}^c}(t) := X_i(t) - \mu_i^{\text{opt}} - \sum_{u=-\infty}^{\infty} d_i^{\text{opt}}(t-u)Y_{ij}(u), \quad (5)$$

gdzie  $\{i, j\}^c = V \setminus \{i, j\}$ . W ten sam sposób definiujemy

$$\epsilon_j(t) := \epsilon_{j|\{i,j\}^c}(t) := X_j(t) - \mu_j^{\text{opt}} - \sum_{u=-\infty}^{\infty} d_j^{\text{opt}}(t-u)Y_{ij}(u). \quad (6)$$

Niech  $X_i = (X_i(t) : t \in \mathbb{Z})$  oraz  $Y_{ij} = (Y_{ij}(t) : t \in \mathbb{Z})$ . Zdefiniujmy relację

$$X_i \perp\!\!\!\perp X_j | Y_{ij} \iff \text{cov}\left(\epsilon_{i|\{i,j\}^c}(t), \epsilon_{j|\{i,j\}^c}(t+u)\right) = 0 \text{ dla wszystkich } u \in \mathbb{Z},$$

która pozwoli nam zdefiniować graf korelacji cząstkowych.

---

<sup>4</sup>Korzystamy częściowo z notacji zaproponowanej przez Dahlhausa [Dah00].

**Definicja 3.1** Niech  $X(t) = (X_1(t), \dots, X_k(t))$  będzie wieloczynnikowym szeregiem czasowym oraz niech  $V = \{1, \dots, k\}$  będzie odpowiadającym zbiorem wierzchołków. Niech  $(i, j) \notin E$  wtedy i tylko wtedy, gdy  $X_i \perp\!\!\!\perp X_j | Y_{ij}$ . Wówczas  $\mathcal{G} = (V, E)$  nazywamy grafem korelacji cząstkowych dla szeregów czasowych.

**Uwaga 3.2** Dla zmiennych gaussowskich  $X_i$  oraz  $X_j$  oczywiście  $\epsilon_i$  oraz  $\epsilon_j$  są również gaussowskie. Wówczas najlepszy liniowy predyktor  $X_i(t)$  pod warunkiem  $Y_{ij}$  jest również najlepszym predyktorem. Dodatkowo z definicji wynika, że  $X_i \perp\!\!\!\perp X_j | Y_{ij}$  wtedy i tylko wtedy, gdy  $\epsilon_i$  i  $\epsilon_j$  są niezależne, a zatem w tym wypadku graf z definicji 3.1 jest grafem warunkowych niezależności dla szeregów czasowych czynników  $X$ .

Bardzo użyteczny warunek dla krawędzi grafu możemy uzyskać stosując narzędzia analizy spektralnej. Kluczowe, że estymator koherencji cząstkowej może posłużyć do identyfikacji grafu na podstawie zaobserwowanych danych. Niech

$$\Gamma_{ij}(u) = \text{cov}(X_i(t+u), X_j(t))$$

będzie funkcją autokowariancji procesu. Jeśli  $\sum_{u=-\infty}^{\infty} |\Gamma_{ij}(u)| < \infty$  to kospektrum pomiędzy  $X_i(t)$  a  $X_j(t)$  zdefiniowane jest jako

$$f_{ij}(\omega) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} \Gamma_{ij}(u) e^{-i\omega u}.$$

Niech  $f(\omega) = (f_{ij}(\omega))_{i,j=1,\dots,k}$  oraz  $\Gamma(u) = (\Gamma_{ij}(u))_{i,j=1,\dots,k}$ . Znany jest wynik, że odwrotna transformacja Fouriera pozwala z macierzy gęstości spektralnej odzyskać funkcję autokowariancji

$$\Gamma(u) = \int_{-\pi}^{\pi} f(\omega) e^{i\omega u} d\omega.$$

Zauważmy w szczególności, że czynniki  $X_i(t+u)$  oraz  $X_j(t)$  są nieskorelowane dla wszystkich  $u \in \mathbb{Z}$  wtedy i tylko wtedy, gdy  $f_{ij} = 0$ .

Jak już anonsowaliśmy wcześniej, interesujący jest „czysty” związek pomiędzy dwoma czynnikami. Miarą zależności pomiędzy  $X_i$  a  $X_j$  pod warunkiem  $Y_{ij}$  jest kospektrum cząstkowe (patrz Dahlhaus [Dah00], s. 160) postaci

$$f_{ij|\{i,j\}^c}(\omega) := f_{\epsilon_i \epsilon_j}(\omega).$$

Przeskalowując otrzymujemy koherencję cząstkową

$$R_{ij|\{i,j\}^c}(\omega) = \frac{f_{ij|\{i,j\}^c}(\omega)}{\sqrt{f_{ii|\{i,j\}^c}(\omega) f_{jj|\{i,j\}^c}(\omega)}}.$$



Okazuje się, że możemy uniknąć wyliczania współczynników w równaniach 5 oraz 6. Brillinger w swojej książce<sup>5</sup> [Bri81] wykazał, że zachodzi następująca równość

$$f_{X_a X_b | Y}(\omega) = f_{X_a X_b}(\omega) - f_{X_a Y}(\omega) f_{Y Y}(\omega)^{-1} f_{Y X_b}(\omega). \quad (7)$$

Skoro  $f_{\epsilon_i \epsilon_j}(\cdot) = 0$  wtedy i tylko wtedy, gdy  $\text{cov}(\epsilon_i(t), \epsilon_j(t+u)) = 0$  dla wszystkich  $u \in \mathbb{Z}$ , dostajemy następujący rezultat.

**Twierdzenie 3.3** *Niech  $\mathcal{G} = (V, E)$  będzie grafem korelacji cząstkowych dla wielowymiarowego szeregu czasowego. Wówczas*

$$(i, j) \notin E \text{ wtedy i tylko wtedy, gdy } R_{ij|\{i,j\}^c}(\omega) \equiv 0.$$

Relację niezależności dla szeregów czasowych zdefiniowaliśmy dla niezależności dwóch jedno-czynnikowych szeregów pod warunkiem wszystkich innych. W istocie możemy uogólnić tę definicję na dowolne trójki podzbiorów zbioru czynników. Niech  $X_A = (X_i(t) | i \in A, t \in \mathbb{Z})$ . Wówczas:

$$\begin{aligned} X_A \perp\!\!\!\perp X_B | X_C &\iff \text{cov}(\epsilon_{A|C}(t), \epsilon_{B|C}(t+u)) = 0 \text{ dla wszystkich } u \in \mathbb{Z} \\ &\iff f_{X_A X_B | X_C}(\cdot) \equiv 0 \\ &\iff R_{X_A X_B | X_C}(\cdot) \equiv 0 \end{aligned} \quad (8)$$

dla rozłącznych zbiorów  $A, B, C \subset V$ , gdzie  $f_{X_A X_B | X_C}$  oraz  $R_{X_A X_B | X_C}$  definiujemy analogicznie jak w przypadku, gdy  $A$  i  $B$  są singletonami.

To co istotne jest w praktyce, to fakt, że cząstkowe koherencje spektralne można uzyskać jako ujemne wartości przeskalowanej odwrotności macierzy spektralnej. Niech

$$g(\omega) := f(\omega)^{-1}$$

oraz

$$d(\omega) := \begin{bmatrix} \frac{1}{\sqrt{g_{11}(\omega)}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{g_{kk}(\omega)}} \end{bmatrix} g(\omega) \begin{bmatrix} \frac{1}{\sqrt{g_{11}(\omega)}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{g_{kk}(\omega)}} \end{bmatrix}.$$

**Twierdzenie 3.4** *Niech  $X(t) = (X_i(t) | i \in V)$  będzie wielowymiarowym szeregiem czasowym z macierzą spektralną  $f(\omega)$ . Jeśli  $f(\omega)$  jest nieosobliwa, wówczas zachodzi*

$$d_{ij}(\omega) = -R_{ij|\{i,j\}^c}(\omega) \quad \text{dla } i \neq j.$$

---

<sup>5</sup>Cytat za Dahlhausem [Dah00].

Oraz

$$g_{ii}(\omega) = \frac{1}{f_{ij|\{i\}^c}(\omega)}.$$

Jeśli  $X_A(t)$  oraz  $X_B(t)$  są procesami wektorowymi, to  $g_{AB} = 0$  (równanie macierzowe) wtedy i tylko wtedy, gdy  $f_{X_A X_B | Y_{AB}}(\omega) = 0$  dla wszystkich  $\omega \in \mathbb{R}$ .

*Dowód:* Dahlhaus [Dah00], s. 171.  $\square$

Waga tego twierdzenia ma przede wszystkim wymiar praktyczny. Po pierwsze na mocy twierdzenia 3.3 dostajemy, że brak krawędzi w grafie korelacji cząstkowej może jednoznacznie być zidentyfikowany przez zera w przeskalowanej odwrotności macierzy spektralnej. Na podstawie tego warunku możemy na przykład w prosty sposób uzyskać ograniczenia na parametry w parametrycznych modelach szeregów czasowych (popularne jest podejście wykorzystujące strukturalne modele VAR – patrz na przykład [Mon03]).

Po drugie twierdzenie 3.4 jest podstawą do nieparametrycznej identyfikacji grafu dla szeregu czasowego na podstawie oszacowanej macierzy spektralnej. To podejście jest znacznie bardziej efektywne obliczeniowo niż estymacja wszystkich składników losowych  $\epsilon_{i|\{i,j\}^c}(t)$  użytych w definicji grafu korelacji cząstkowej, co wymagałoby szacowania  $2 \cdot \binom{k}{2}$  filtrów liniowych (patrz Dahlhaus [Dah00]).

## 3.2 Własności Markowa w grafach interakcji

Tak jak w przypadku grafów dla zmiennych losowych, tak również w przypadku szeregów czasowych będziemy chcieli sformułować analogon globalnej własności Markowa.

**Lemat 3.5** *Niech  $X(t) = (X_i(t)|i \in V)$  będzie wielowymiarowym szeregiem czasowym o nieosobliwej macierzy spektralnej. Niech  $X_A, X_B, X_C$  i  $X_D$  będą wielowymiarowymi szeregami czasowymi dla  $A, B, C, D$  rozłącznych podzbiorów zbioru  $V$ . Wówczas*

(i)  $X_A \perp\!\!\!\perp (X_B, X_C) | X_D \implies X_A \perp\!\!\!\perp X_B | X_D,$

(ii)  $X_A \perp\!\!\!\perp (X_B, X_C) | X_D \iff X_A \perp\!\!\!\perp X_B | (X_C, X_D)$  oraz  $X_A \perp\!\!\!\perp X_C | (X_B, X_D).$

*Dowód:* Dahlhaus [Dah00].  $\square$

Tak jak w sekcji 2 dla rozłącznych podzbiorów  $A, B, S \subset V$  mówimy, że  $S$  rozdziela  $A$  i  $B$ , jeśli każda ścieżka od wierzchołka w  $A$  do wierzchołka  $B$  musi przechodzić przez wierzchołek w  $S$ .

**Twierdzenie 3.6** *Niech  $X(t)$  będzie wielowymiarowym szeregiem czasowym o nieosobliwej macierzy spektralnej z odpowiadającym jej grafem korelacji cząstkowych  $\mathcal{G} = (V, E)$ . Niech  $A, B, S \subset V$ , gdzie  $S$  rozdziela  $A$  i  $B$ . Zachodzi wówczas  $X_A \perp\!\!\!\perp X_B | X_S$ .*

*Dowód:* Dahlhaus [Dah00].□

A zatem dla modeli interakcji spełniona jest globalna własność Markowa, a zatem również wszystkie inne – dowód analogiczny, jak w przypadku klasycznym opisanym w paragrafie 2.3. Warto podkreślić, że główna trudność w zdefiniowaniu modeli graficznych dla szeregów czasowych leży właśnie w udowodnieniu, że zachodzą odpowiednie własności Markowa. Możliwe rozwiązanie tego problemu w przypadku grafów mieszanych zaproponował Eichler (patrz np. [Eic99]).

### 3.3 Estymacja modeli interakcji

Założmy, że zaobserwowaliśmy  $k$ -wymiarowy szereg czasowy  $(x_t)_{t=1}^T$  długości  $T$  i chcemy zidentyfikować graf dla modelu graficznego opisującego ten system. Z twierdzenia 3.3 istnienie danej krawędzi odpowiada nie znikaniu odpowiedniej cząstkowej koherencji spektralnej. W procesie estymacji będziemy również korzystać z odpowiedniości z twierdzenia 3.4. Procedura podobna do opisanej poniżej (właściwą procedurę przedstawił Dahlhaus [Dah00]) jest zaimplementowana przez Michaela Eichlera w programie Spectrum<sup>6</sup>.

Funkcja autokowariancji z próby zadana jest wzorem:

$$\hat{\Gamma}(h) = \frac{1}{T} \sum_{t=0}^{T-h-1} (X(t+h) - \bar{X})(X(t) - \bar{X})'$$

dla  $h = 0, \dots, T-1$ , co z symetrii możemy rozszerzyć również na ujemne wartości  $h$  oraz jest równe zeru dla  $|h| \geq T$  (wektor  $\bar{X} = \frac{1}{T} \sum_{t=0}^{T-1} X(t)$  jest średnią dla danych). Funkcja autokowariancji z próby jest estymatorem zgodnym i asymptotycznie normalnym przy spełnieniu stosunkowo słabych warunków (patrz [BD88]).

Periodogram jest okresową funkcją macierzową o okresie  $2\pi$  określoną dla częstotliwości  $\omega_k = \frac{2\pi k}{T}$  dla  $\omega_k \in [0, 2\pi]$  i zdefiniowaną jako dyskretna transformata Fouriera funkcji autokowariancji z próby. Niech  $d(0), \dots, d(T-1)$  będzie dyskretną transformatą Fouriera dla danych, tzn.

$$d(k) = \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} X(t) e^{-ikt},$$

---

<sup>6</sup>Program jest dostępny pod adresem <http://www.statlab.uni-heidelberg.de/people/eichler/spectrum.html>

wówczas dla częstotliwości  $\omega_k = \frac{2\pi k}{T}$ ,  $\omega_k \in [0, 2\pi]$ , periodogram jest zdefiniowany jako

$$I(\omega_k) = \frac{1}{2\pi} d(k)d(k)^*.$$

Periodogram nie jest jednak zgodnym estymatorem gęstości spektralnej, nie jest również estymatorem dobrym (efektywnym i nieobciążonym), lecz może być, jeżeli zostanie odpowiednio wygładzony (patrz [BJ04]). Wygładzanie uzyskujemy licząc splot periodogramu z oknem wygładzającym  $W_r(j)$ , co prowadzi do następującego estymatora

$$\hat{f}^r(\omega_k) = \sum_{j=-\infty}^{\infty} W_r(j)I(\omega_{j+k}).$$

Wymagamy by okno wygładzające  $W_r(\cdot)$  było symetryczne i sumowało się do jedynki. Liczbę  $r$  będziemy nazywać szerokością okna. Przy odpowiednim doborze  $W_r(\cdot)$  jeśli  $T \rightarrow \infty$  oraz  $r(T)$  dąży do nieskończoności w taki sposób, że  $\frac{r(T)}{T}$  dąży do zera, to  $\hat{f}^r(\cdot)$  jest zgodnym estymatorem macierzy gęstości spektralnej (patrz [BD88]). W tym tekście nie będziemy się zajmować ani doбором optymalnego okna wygładzającego ani doбором jego szerokości odsyłając do artykułu Bacha i Jordana [BJ04].

Tak uzyskany estymator macierzy spektralnej  $\hat{f}$  odwracamy oraz odpowiednio skalujemy. Co jest podstawą do uzyskania estymatorów dla koherencji cząstkowych na podstawie twierdzenia 3.4. Aby testować istnienie krawędzi w grafie skonstruujemy obszar krytyczny dla statystyki  $|\hat{R}_{ij|\{i,j\}^c}(\omega)|^2$  pod warunkiem hipotezy, że  $R_{ij|\{i,j\}^c}(\omega) \equiv 0$ . Uznamy, że krawędź łącząca wierzchołek  $i$  i  $j$  istnieje jeśli choć dla jednej częstotliwości wykres tej statystyki przekroczy wyznaczoną granicę. Jako tą granicę będziemy brać  $1 - \alpha$ -kwantyl asymptotycznego rozkładu statystyki  $\max_{\omega} |\hat{R}_{ij|\{i,j\}^c}(\omega)|^2$  przy założeniu, że  $R_{ij|\{i,j\}^c} \equiv 0$  (patrz Dahlhaus i in. [DES97]).

## 4 Model dla gospodarki amerykańskiej

Jako przykład zastosowania opisywanych metod, przedstawimy graficzny model interakcji dla sześciowymiarowego szeregu czasowego złożonego z podstawowych miesięcznych zmiennych makroekonomicznych dla gospodarki amerykańskiej oraz trzech podstawowych wskaźników mających charakteryzować obecny i przyszły poziom aktywności gospodarczej oraz przyszły poziom inflacji.

Pierwotny 6-wymiarowy szereg zawierał komponenty niestacjonarne. Zastosowaliśmy powszechnie znaną procedurę pozbycia się trendu logarytmując pierwsze zmiany (miesięczne) w każdym z niestacjonarnych czynników. Standardowe testy statystyczne w postaci testów ADF potwierdziły, że uzyskany w ten sposób szereg złożony jest z komponentów stacjonarnych. Szereg ten obejmuje dane od lutego 1956 roku do stycznia 2006 roku, czyli łącznie 600 obserwacji miesięcznych. Poszczególne składniki analizowanego szeregu to:

**LII** Wyprzedzający Wskaźnik Inflacji - źródło: The Foundation For International Business And Economic Research, Inc.

**ci** Logarytm zmian Równoległego Wskaźnika Koniunktury - źródło: Economic Cycle Research Intitute.

**li** Logarytm zmian Wyprzedzającego Wskaźnika Koniunktury - źródło: Economic Cycle Research Intitute.

**INFL** Inflacja, względne zmiany roczne CPI, źródło: U.S. Department of Labor: Bureau of Labor Statistics.

**prod** Logarytm zmian wskaźnika produkcji przemysłowej - źródło: Board of Governors of the Federal Reserve System.

**UNEMP** Stopa bezrobocia w Stanach Zjednoczonych - źródło: U.S. Department of Labor: Bureau of Labor Statistics.

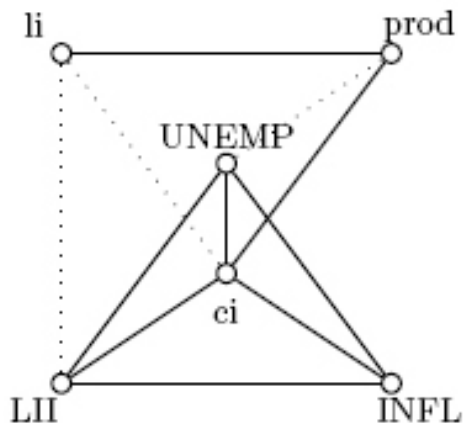
W analizie wskaźników gospodarczych, jako główną miarę jakości wskaźnika podaje się stopień w jakim wskaźnik odzwierciedla kształtowanie się swoich zmiennych referencyjnych. Jasne jest na przykład, że dobry wyprzedzający wskaźnik inflacji, to taki wskaźnik, który z dużym prawdopodobieństwem pozwala przewidzieć przyszłe kształtowanie się inflacji. W przypadku wskaźników koniunktury szereg referencyjny jest nie do końca zdefiniowany, gdyż stan aktywności gospodarczej to w istocie wypadkowa bardzo wielu składowych. W praktyce jednak oczekuje się, że wskaźniki koniunktury dobrze odzwierciedlają takie makrozmiennne jak produkt krajowy, produkcja, itd. Często w tym kontekście analizuje się po prostu funkcję korelacji pomiędzy parą szeregów. Zauważmy jednak, że w przypadku analizowania tak złożonego systemu jak gospodarka, może to prowadzić do zjawiska regresji pozornej. Rzeczywiście, współzależność dwóch zmiennych może wynikać z tego, że istnieje jakaś grupa zmiennych, które w analizie pominęliśmy, a które wpływają na obie te zmienne. W rezultacie powoduje to, że obserwujemy pozorny związek między tymi zmiennymi. Wszystko jest w porządku do póki nie chcemy tego związku interpretować jako związku bezpośredniego. Tak jest jednak na przykład w przypadku analizy jakości wskaźników gospodarczych. W tym wypadku będziemy stosować koherencje cząstkowe, które pozwalają wyodrębnić związek pomiędzy dwoma szeregami po eliminacji liniowych efektów wszystkich innych procesów w systemie. Metody modelowania graficznego dają nam wówczas możliwość dokładniejszego wejrzenia w naturę złożonego systemu, który badamy.

Na podstawie procedury opisanej w poprzedniej sekcji oraz z wykorzystaniem programu Spectrum, o którym już wcześniej wspomnieliśmy, oszacowany został model interakcji dla tych zmien-

nych. Uzyskaliśmy w ten sposób wykresy odpowiednich koherencji oraz koherencji cząstkowych. Wyniki przedstawione są na rysunku 1, gdzie powyżej diagonali pojawiają się koherencje, a poniżej koherencje cząstkowe. Przerywana linia odpowiada 95% przedziałowi ufności dla

$$\sup_{\omega} |\hat{R}_{ij|\{i,j\}^c}(\omega)|^2$$

wyliczonemu przy hipotezie, że  $R_{ij|\{i,j\}^c}(\cdot) \equiv 0$ , co zostało dokładniej opisane w poprzedniej sekcji. Na podstawie wykresów została zaproponowana poniższa postać grafu odpowiadającego temu modelowi.



Linia przerywaną zaznaczyliśmy te krawędzie, za których istnieniem dane przemawiają stosunkowo najslabiej. Z analizy wykresów 1 oraz powyższego grafu możemy wyciągnąć szereg niezwykle ciekawych wniosków. Trzeba być jednak ostrożnym. Zauważmy, że wskaźniki gospodarcze są złożonymi kompozytami, a zatem związek między nimi może wynikać z uwzględniania w nich podobnych komponentów. Jednak to co wydaje się najbardziej interesujące to bardzo wyraźny związek wyprzedzającego wskaźnika inflacji (FII) z inflacją (INFL). Odpowiedni wykres koherencji cząstkowej bardzo wyraźnie przekracza granicę odpowiedniego przedziału ufności. Równie wyraźny jest związek logarytmu zmian równoległego wskaźnika koniunktury (ci) oraz logarytmu wskaźnika produkcji (prod). Warto jednak podkreślić, że związek wskaźnika produkcji z wyprzedzającym wskaźnikiem koniunktury jest już znacznie słabszy (choć również istotny statystycznie). Innym ciekawym zaobserwowanym związkiem jest powszechni znany w makroekonomii (ujemny) związek pomiędzy bezrobociem a inflacją.

Należy jednak podkreślić kilka rzeczy. Po pierwsze w tak prostym przypadku nie wykorzystaliśmy wszystkich zalet modelowania graficznego. Przede wszystkim wyniki dla sześciu zmiennych jesteśmy w stanie analizować bezpośrednio i nie ma potrzeby dzielenia systemu na podsystemy. Po drugie ograniczając się do grafów nieskierowanych zrezygnowaliśmy z interpretacji krawędzi jako związków przyczynowych. Wydaje się, że to mogło być ciekawe w przypadku badania systemu

zmiennych makroekonomicznych (bez uwzględniania wskaźników). Celem tego badania było jednak przede wszystkim przedstawienie metody oraz zbadanie jakości wskaźników i ich użyteczności do opisu stanu faktycznego w gospodarce. Po drugie zawsze można się zastanawiać czy badany system jest wystarczająco kompletny. Założenie o kompletności jest niezwykle istotne jeśli koherencje cząstkowe chcemy interpretować jako statystykę odzwierciedlającą „czysty” związek pomiędzy parą szeregów. Oczywiście odpowiedź na pytanie, czy system jest kompletny nigdy nie może być pozytywna, jednak badanie odpowiednio dużej liczby szeregów znacznie obniża prawdopodobieństwo, że coś pominęliśmy.

## 5 Zakończenie

W pracy tej przedstawiliśmy podstawowe idee modelowania graficznego. Objęło to zarówno modele grafów skierowanych i nieskierowanych. Zaprezentowaliśmy również najprostszy model, który uogólnia modele graficzne grafów nieskierowanych na przypadek szeregów czasowych (wierzchołki grafu są procesami stochastycznymi). W tym wypadku analiza spektralna pozwala badać zależności cząstkowe pomiędzy grafami i sytuacja staje się analogiczna do klasycznej, jeśli zamiast macierzy korelacji rozpatrywać macierz koherencji spektralnych.

Dużo ciekawsze jest rozszerzenie analizy na grafy skierowane oraz mieszane. Pozwala to interpretować strzałki w grafie jako kierunek wpływu stochastycznego jednego procesu na drugi. Michael Eichler był pionierem w tych badaniach i bardzo pracowicie wyprowadził odpowiednie własności Markowa dla różnych grafów tego rodzaju. W pewnym sensie podsumowanie jego pracy z kilku ostatnich lat można znaleźć w [Eic06].

Drugim istotnym rozszerzeniem jest zastosowanie podejścia parametrycznego do estymacji. Ograniczamy się wówczas z reguły do modeli VAR, jednak pozwala to przy zachowaniu stosunkowo dużej ogólności znacznie sformalizować etap wnioskowania. Podejście parametryczne pozwala również wierzyć, że uda się tu z powodzeniem wykorzystać metodologię bayesowską, która od dawna wykorzystywana jest w klasycznych modelach graficznych. Pierwsze podejście do tego typu problemu zrobili Jukka Corander oraz Mattias Villani [CV05], jednak ich praca to jedynie początek.

Wydaje się, że właśnie ten kierunek jest szczególnie ważny w analizie szeregów czasowych za pomocą modelowania graficznego. Usystematyzowane wykorzystanie metodologii bayesowskiej pozwoliłoby efektywniej estymować te modele, a także (co ważne wobec wielkości przestrzeni możliwych modeli) uwzględnić elementy wiedzy a priori. Wykorzystanie MCMC mogłoby również pomóc w rozszerzeniu tych modeli o ukryte zmienne, co mogłoby mieć potencjalnie bardzo ważne zastosowania w ekonometrii i uogólniałyby na przykład filtry Kalmana i ukryte procesy Markowa.

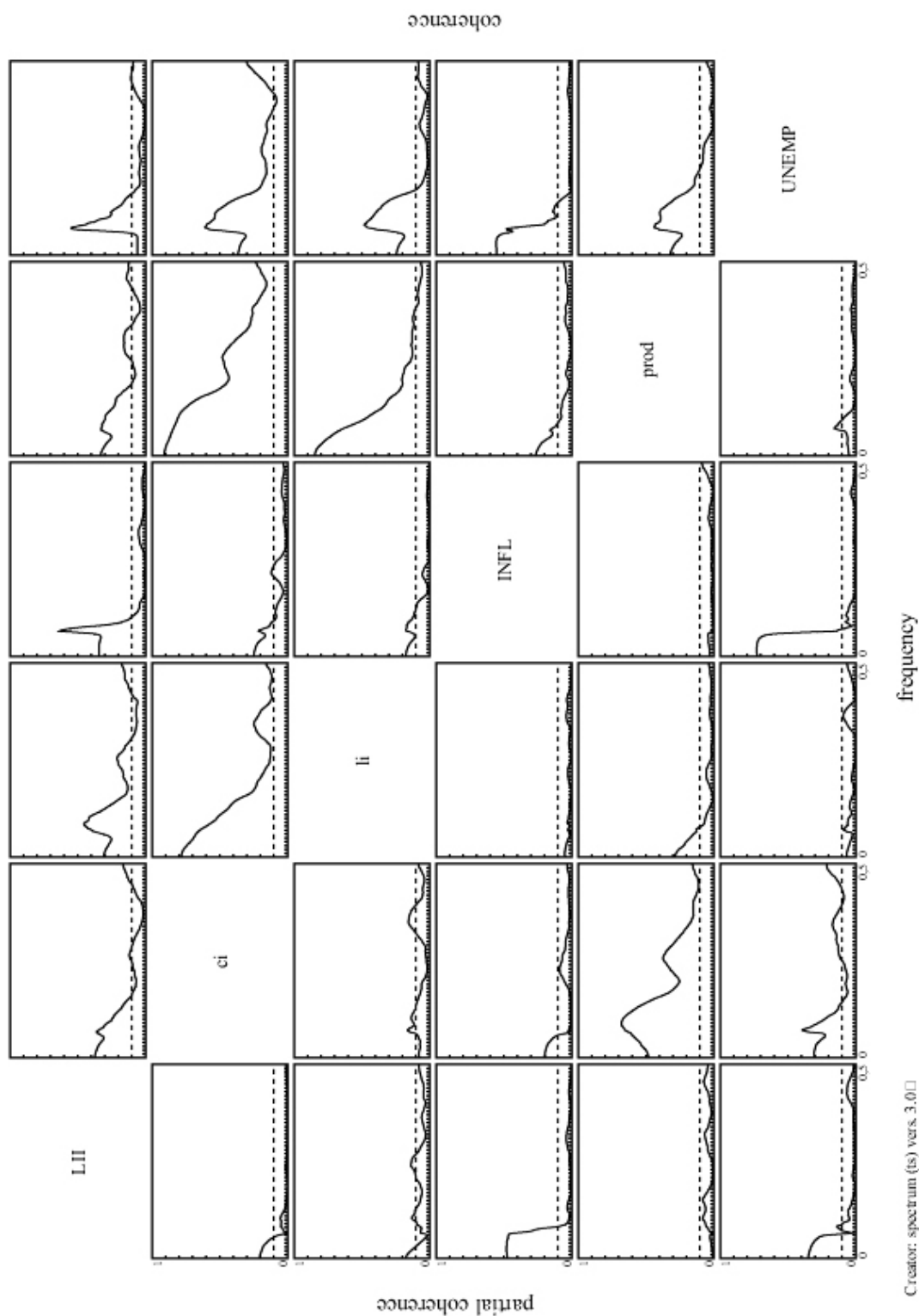
## Literatura

- [BD88] Peter J. Brockwell, Richard A. Davis. *Time Series: Theory and Methods*. Springer, New York, wydanie II, 1988.
- [BJ04] Francis R. Bach, Michael I. Jordan. Learning Graphical Models for Stationary Time Series. *IEEE Transactions on Signal Processing*, 52:2189–2199, 2004.
- [Bri81] David R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, Inc., San Francisco, 1981.
- [Bri96] David R. Brillinger. Remarks Concerning Graphical Models for Time Series and Point Processes. *Brazilian Review Econometrics*, 16:1–23, 1996.
- [CT91] Thomas M. Cover, Joy A. Thomas. *Elements of information theory*. Wiley New York, 1991.
- [CV05] Jukka Corander, Mattias Villani. A Bayesian Approach to Modelling Graphical Vector Autoregressions. *Journal of Time Series Analysis*, 27:141–156, 2005.
- [Dah00] Rainer Dahlhaus. Graphical Interaction Models for Multivariate Time Series. *Metrika*, 51:157–172, 2000.
- [DES97] Rainer Dahlhaus, Michael Eichler, Jürgen Sandkühler. Identification of Synaptic Connections in Neural Ensembles by Graphical Models. *Journal of Neuroscience Methods*, 77(1):93–107, 1997.
- [Edw00] David Edwards. *Introduction to Graphical Modelling*. Springer, wydanie second, 2000.
- [Eic99] Michael Eichler. *Graphical Models in Time Series Analysis*. Praca doktorska, University of Heidelberg, 1999.
- [Eic03] Michael Eichler. Graphical Interaction Models for Time Series: Parameter Estimation and Model Selection. <http://www.statlab.uni-heidelberg.de/people/eichler/modelselect.pdf>, 2003.
- [Eic06] Michael Eichler. Graphical Modelling of Multivariate Time Series. Raport instytutowy, University of Heidelberg, 2006. <http://www.statlab.uni-heidelberg.de/people/eichler/graphmodel.pdf>.



- [GP93] Dan Geiger, Judea Pearl. Logical and Algorithmic Properties of Conditional Independence and Graphical Models. *The Annals of Statistics*, 21(4):2001–2021, 1993.
- [Lau96] Steffen L. Lauritzen. *Graphical Models*. Oxford Stat. Sci. Series No. 17. Clarendon Press, Oxford, 1996.
- [Mon03] Alessio Moneta. Graphical Models for Structural Vector Autoregressions. LEM Papers Series 2003/07, Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy, 2003.
- [Mur02] Kevin P. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. Praca doktorska, University of California, Berkeley, 2002.
- [Pea00] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, NY, USA, 2000.
- [PP87] Judea Pearl, Azaria Paz. Graphoids: A Graph Based Logic for Reasoning about Relevancy Relations. B.D. Boulay, D. Hogg, L. Steel, redaktorzy, *Advances in Artificial Intelligence*, wolumen II, strony 357–363. North-Holland, Amsterdam, 1987.
- [Whi90] Joe Whittaker. *Graphical Models in Applied Multivariate Statistics*. Wiley New York, 1990.
- [Wil04] Robin J. Wilson. *Wprowadzenie do teorii grafów*. Państwowe Wydawnictwo Naukowe, wydanie II, 2004.

Rysunek 1: Graf interakcji dla szeregu wskaźników oraz zmiennych referencyjnych gospodarki amerykańskiej. Dane obejmują okres od lutego 1956 roku do stycznia 2006 roku.



Creator: spectrum (ts) vers. 3.0.0