

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Jan Zouhar

Lineární rekursivní systémy a struktury podmíněné nezávislosti

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Milan Studený, DrSc.
(ÚTIA AV ČR)

Studijní program: Obecná matematika

Praha, 2010

Poděkování. Rád bych na tomto místě poděkoval vedoucímu bakalářské práce dr. Studenému za poskytnutí materiálů, důsledné přečtení několika předběžných verzí práce a četné cenné rady a komentáře. Další díky patří prof. Radimovi Jirouškovi, který mne před několika lety uvedl do problematiky grafických modelů struktur podmíněné nezávislosti, a též mojí manželce za její nesmírnou trpělivost a toleranci mé duševní nepřítomnosti v době psaní této práce.

Prohlášení. Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 11. prosince 2010

Jan Zouhar

Abstrakt

Název práce: Lineární rekursivní systémy a struktury podmíněné nezávislosti
Autor: Jan Zouhar
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí práce: RNDr. Milan Studený, DrSc. (ÚTIA AV ČR, studenym@utia.cas.cz)

Lineární rekursivní systémy (LRS) popisují lineární funkční vztahy spojitých, zpravidla normálně rozdělených náhodných veličin. Pro kvalitativní popis těchto vztahů se využívá acyklických orientovaných grafů. Grafy se využívají i v jiné statistické disciplíně, a sice při popisu struktury podmíněné nezávislosti (PN) systému náhodných veličin. Jedním z cílů práce bylo ukázat, že v rámci regulárních gaussovských rozdělání oba uvedené přístupy splývají: je-li dán acyklický orientovaný graf, lze statistický model LRS vymezený tímto grafem ekvivalentně zavést jako třídu gaussovských distribucí, jejichž struktura PN odpovídá témuž grafu. Některé vztahy mezi grafem LRS a jeho strukturou PN jsme dále zobecnili i mimo rámec gaussovských distribucí. Dalším tématem je popis vztahu mezi grafem LRS a kovariancemi jeho veličin. Zde jsme odvodili vztah, který je jistou analogií metody koeficientů na cestách, kterou zavedl ve 20. letech minulého století americký genetik Sewall Wright.

Klíčová slova: lineární rekursivní systémy, struktury podmíněné nezávislosti, grafické modely, gaussovská rozdělání.

Abstract

Title: Linear Recursive Systems and Conditional Independence Structures
Author: Jan Zouhar
Department: Department of Probability and Mathematical Statistics
Supervisor: RNDr. Milan Studený, DrSc. (ÚTIA AV ČR, studenym@utia.cas.cz)

Linear recursive systems (LRS) describe linear relationships among continuous random variables (typically, normally distributed ones). Acyclic oriented graphs are used to provide a qualitative description of these relationships. In a different branch of statistics, graphs serve as a means to describe conditional independence (CI) structures in systems of random variables. One of the aims of the thesis is to show that within the class of regular Gaussian distributions, both approaches coincide: for a given acyclic oriented graph, the statistical model of LRS specified by the graph is equivalent to a class of Gaussian distributions with CI structures that accord with the same graph. Furthermore, we generalized some of the relations between a graph of LRS and its CI structure outside the scope of Gaussian distributions. Another focus of the thesis is the relation between the graph of a LRS and the covariances among its variables. We derived a relationship that is analogous to the method of path coefficients which was introduced in the 1920s by the American geneticist Sewall Wright.

Keywords: linear recursive systems, conditional independence structures, graphical models, Gaussian distributions.

Obsah

Úvod	1
Použité značení	3
1 Podmíněná nezávislost a grafické modely její struktury	5
1.1 Podmíněná pravděpodobnost a podmíněná hustota	5
1.2 Podmíněná nezávislost	8
1.3 Podmíněná střední hodnota, rozptyl a kovariance	10
1.4 Grafické modely struktury podmíněné nezávislosti	12
2 Lineární rekursivní systémy	19
2.1 Podmíněná nezávislost v lineárních rekursivních systémech	20
2.2 Podmíněná nezávislost v zobecněných lineárních rekursivních systémech	26
2.3 Grafický model lineárního rekursivního systému	33
3 Metoda koeficientů na cestách a lineární rekursivní systémy	35
3.1 Metoda koeficientů na cestách	35
3.2 Cesty, sledy a varianční matice v grafickém modelu LRS	39
Závěr	49
A Některé věty o maticích	51
B Některé věty o $\mathcal{N}(\mu, \Sigma)$	55
Literatura	57

Úvod

Lineární rekursivní systémy (LRS) patří mezi nejjednodušší modely v oblasti vícerozměrné statistické analýzy. Jde o speciální případ tzv. *modelů strukturálních rovnic* (*structural equation models*, SEM), které nacházejí časté praktické uplatnění při analýze vztahů mezi socioekonomickými a biologickými veličinami (viz např. [Gol73] nebo [Pug03]). SEM popisují soubor náhodných veličin (zpravidla spojitých, nejčastěji normálně rozdělených), mezi kterými existují lineární funkční závislosti, které bývají při praktických aplikacích často interpretovány coby kauzální vztahy. Pro kvalitativní popis těchto vztahů se v SEM tradičně používá grafického znázornění, označovaného obecně jako diagram. V případě LRS má tento diagram podobu acyklického orientovaného grafu, jehož vrcholy odpovídají jednotlivým náhodným veličinám. Daný graf potom představuje v rámci LRS jistá omezení na koeficienty funkčních vztahů, čímž vymezuje příslušný statistický model.

Podobné grafy se využívají i v poněkud odlišné statistické disciplíně, a sice při popisu struktur *podmíněné nezávislosti* (PN) v systému náhodných veličin. Studium struktur PN má poměrně krátkou historii – první výsledky pocházejí ze 70. let minulého století – a kromě vícerozměrné statistické analýzy je využíváno v teorii kontingenčních tabulek a v pravděpodobnostních expertních systémech (v tzv. *bayesovských sítích*, viz např. [Cow07]). Grafické modely zde slouží jako přehledný způsob zachycení struktury PN prostřednictvím tzv. *markovských vlastností* (přehled grafických modelů struktur PN uvádí [Lau96]). Je-li dán pevný distribuční rámec, můžeme zavést statistický model pomocí požadavků na strukturu PN, určených v podobě markovských vlastností vzhledem k nějakému grafu.

Jedním z hlavních cílů práce je ukázat, že v rámci regulárních gaussovských distribucí oba uvedené přístupy pro zavedení statistického modelu splývají: uvažujeme-li nějaký acyklický orientovaný graf, lze statistický model LRS vymezený tímto grafem ekvivalentně zavést jako třídu distribucí, jejichž struktura PN odpovídá témuž grafu (ve smyslu příslušných markovských vlastností). Některé vztahy mezi grafem LRS a jeho strukturou PN se pokusíme dále zobecnit i mimo rámec gaussovských distribucí.

Dalším cílem práce je nalezení vztahu mezi grafem LRS a kovariancemi mezi jeho veličinami. Půjde o jistou analogii k *metodě koeficientů na cestách* (*method of path coefficients*), kterou popsal ve 20. letech minulého století americký genetik Sewall Wright. Wrightova série prací o této metodě (zejména [Wri21] a [Wri34]) bývá někdy označována za počátek modelů strukturálních rovnic: systémy náhodných veličin, kterými se Wright zabýval, jsou totiž v zásadě shodné jako SEM a Wright byl prvním, kdo využil diagramy pro zachycení kvalitativních vztahů mezi veličinami a pro následné praktické výpočty.

Metoda koeficientů na cestách popisuje vztah mezi korelačními koeficienty veličin v uvažovaném systému a diagramem, ke kterému přísluší tzv. *koeficienty na cestách* (*path coefficients*). Naším cílem je nalézt vztah poněkud odlišný: namísto korelačních koeficientů uvažujeme kovariance náhodných veličin a místo koeficientů na cestách budeme pracovat s koeficienty, které přímo zachycují lineární funkční vztahy mezi veličinami; úměrně těmto rozdílům se budou lišit i získané výsledky.

Struktura práce je následující. V první kapitole zavedeme pojmy týkající se podmíněné nezávislosti, struktury PN a jejich grafických modelů; uvedeme zde rovněž některá tvrzení z této oblasti, která budou podstatná pro studium struktury PN v LRS. Všechny výsledky týkající se struktury PN v LRS obsahuje pak kapitola druhá. Ve třetí, poslední kapitole stručně popíšeme metodu koeficientů na cestách a následně odvodíme způsob, jak vyčíst varianční matici LRS z jeho grafu.

Pro studium struktury PN v LRS i jeho kovariancí využijeme řadu tvrzení z oblasti maticového počtu. Abychom zpřehlednili text v hlavní části práce, zařadili jsme všechna tato tvrzení do přílohy, označené jako příloha A. Podobně příloha B obsahuje věty týkající se vícerozměrných gaussovských rozdělení.

Použité značení

Typografické rozlišení názvů podle typu objektu

i, σ, x, φ	(malá písmena)	skaláry (přirozené nebo reálné), reálné proměnné, funkce
X, I, G	(velká písmena)	náhodné veličiny (jednorozměrné), množiny (čísel), grafy
$\mathbf{x}, \boldsymbol{\mu}$	(malá tučná písmena)	reálné vektory (vždy sloupcové)
$\mathbf{X}, \mathbf{A}, \boldsymbol{\Sigma}$	(velká tučná písmena)	náhodné vektory (sloupcové), reálné matice
\mathcal{I}, \mathcal{M}	(kaligrafická písmena)	množinové systémy, pravděpodobnostní rozdělení

Měřitelné prostory, σ -algebry, míry

$(\mathbb{R}, \mathcal{B})$	měřitelný prostor reálných čísel (\mathbb{R}) s borelovskou σ -algebrou (\mathcal{B})
$(\mathbb{R}_n, \mathcal{B}_n)$	měřitelný prostor reálných n -tic (\mathbb{R}_n) s borelovskou σ -algebrou (\mathcal{B}_n)
$(\Omega, \mathcal{A}, \text{Pr})$	pravděpodobnostní prostor („univerzální“, tj. pravd. míra vždy značena Pr)
$\text{Pr}(X \in A)$	zkrácený zápis pro $\text{Pr}(\{\omega \in \Omega : X(\omega) \in A\})$
$\mu \times \nu$	součin měr μ a ν (součinnová míra)
s.v. $[\mu]$	skoro všude vzhledem k míře μ
s.j.	skoro jistě (neboli s.v. [Pr])

Náhodné veličiny

$X \perp\!\!\!\perp Y$	X a Y jsou nezávislé náhodné veličiny
EX	střední hodnota náhodné veličiny X
$\text{var} X$	rozptyl náhodné veličiny X
$\text{cov}(X, Y)$	kovariance náhodných veličin X a Y
$\text{corr}(X, Y)$	korelační koeficient náhodných veličin X a Y
$E\mathbf{X}$	střední hodnota náhodného vektoru \mathbf{X}
$\text{var} \mathbf{X}$	varianční matice náhodného vektoru \mathbf{X}
$\mathcal{N}(\mu, \sigma^2)$	jednorozměrné normální rozdělení se střední hodnotou μ a rozptylem σ^2
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	vícerozměrné normální rozdělení se střední hodnotou $\boldsymbol{\mu}$ a varianční maticí $\boldsymbol{\Sigma}$

Množiny

$ A $	kardinalita (počet prvků) množiny A (půjde vždy o konečné množiny)
2^A	potenční množina množiny A
$A \times B$	kartézský součin množin A a B

Matice, vektory

$\mathbf{A}_{m \times n}$	matice \mathbf{A} typu $m \times n$ (použito v případě, že je třeba zdůraznit typ matice)
---------------------------	---

$\mathbf{0}_{m \times n}$	nulová matice typu $m \times n$ (tj. matice tvořená samými nulami)
\mathbf{I}_n	jednotková matice typu $n \times n$
$\text{diag } \mathbf{b}$	diagonální matice, jejíž diagonálu tvoří prvky vektoru \mathbf{b}
$\det \mathbf{A}$	determinant matice \mathbf{A}
\mathbf{A}^{-1}	inverze matice \mathbf{A}
\mathbf{A}^\top	transpozice matice \mathbf{A}

Kromě výše uvedených maticových zápisů se v textu často vyskytuje značení pro výběr podmatic nebo podvektorů pomocí indexových množin; toto značení nyní podrobně popíšeme.

Máme-li matici $\mathbf{A}_{m \times n} = (a_{ij})$ a neprázdné indexové množiny I a J splňující

$$I = \{i_1, i_2, \dots, i_r\} \subseteq \{1, 2, \dots, m\}, \quad i_1 < i_2 < \dots < i_r, \\ J = \{j_1, j_2, \dots, j_s\} \subseteq \{1, 2, \dots, n\}, \quad j_1 < j_2 < \dots < j_s,$$

budeme zápisem $\mathbf{A}_{I,J}$ značit matici

$$\mathbf{A}_{I,J} = (a_{ij})_{i \in I, j \in J},$$

neboli podmatici \mathbf{A} , která je tvořena řádky \mathbf{A} s pořadovými čísly v I a sloupci \mathbf{A} s pořadovými čísly v J . Přesněji zapsáno,

$$\mathbf{A}_{I,J} = (a_{i_k j_l})_{\substack{k=1, \dots, r, \\ l=1, \dots, s}} = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_r j_1} & a_{i_r j_2} & \dots & a_{i_r j_s} \end{pmatrix}.$$

Chceme-li při použití tohoto značení zohlednit, že jedna z indexových množin (tj. I nebo J) je sjednocením dvou nebo více dílčích disjunktních množin, např. $I = I_1 \cup I_2$, vynecháme v příslušném zápisu pro lepší čitelnost znak pro sjednocení, tedy

$$\mathbf{A}_{I,J} = \mathbf{A}_{I_1 I_2, J}.$$

Pokud potřebujeme zapsat některou z indexových množin výčtem jejích prvků, vynecháme složené závorky a čárky mezi prvky; např. pro $I = \{i_1, i_2\}$ zapíšeme

$$\mathbf{A}_{I,J} = \mathbf{A}_{i_1 i_2, J}.$$

Analogické značení, jako jsme zavedli pro matice, zavedeme i pro vektory. Máme-li vektor $\mathbf{b} = (b_1, b_2, \dots, b_n)^\top$ a množinu $I = \{i_1, i_2, \dots, i_k\} \subseteq \{1, 2, \dots, n\}$, kde $i_1 < i_2 < \dots < i_k$, potom

$$\mathbf{b}_I = (b_{i_1}, b_{i_2}, \dots, b_{i_k})^\top.$$

Podobně jako v případě matic, i u vektorů budeme při tomto značení vynechávat na příslušných místech znaky pro sjednocení, složené závorky a čárky.

Ukažme použití právě zavedeného značení pro vektory při zápisu integrálů z funkcí více proměnných. Mějme $\mathbf{x} = (x_1, x_2, \dots, x_5)^\top$, $f : \mathbb{R}_5 \rightarrow \mathbb{R}$. Chceme-li integrovat f podle proměnných x_2 a x_4 , definujeme $I = \{2, 4\}$ a zapíšeme

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x_1, x_2, \dots, x_5) dx_2 dx_4 = \int_{\mathbb{R}_2} f(\mathbf{x}) d\mathbf{x}_I.$$

Kapitola 1

Podmíněná nezávislost a grafické modely její struktury

1.1 Podmíněná pravděpodobnost a podmíněná hustota

V elementárním počtu pravděpodobnosti se podmíněná pravděpodobnost zavádí následujícím způsobem: máme-li pravděpodobnostní prostor $(\Omega, \mathcal{A}, \Pr)$ a jevy $A, B \in \mathcal{A}$, potom podmíněnou pravděpodobnost jevu A za podmínky, že nastal jev B , definujeme jako

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \quad (1.1)$$

pokud $\Pr(B) > 0$. (Pro $\Pr(B) = 0$ se případně dodefinuje $\Pr(A|B) = 0$.) S takovou definicí lze ovšem vystačit pouze v případě diskrétních náhodných veličin. Obecnější způsob, jak definovat podmíněnou pravděpodobnost, se opírá o následující definici podmíněné hustoty.

Definice 1.1 Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ a disjunktní indexové množiny I a J takové, že $I \cup J = \{1, 2, \dots, n\}$. Předpokládejme, že \mathbf{X} má hustotu f vzhledem k součinnové míře $\mu = \mu_I \times \mu_J$, kde μ_I je σ -konečná míra na $(\mathbb{R}_{|I|}, \mathcal{B}_{|I|})$ a μ_J je σ -konečná míra na $(\mathbb{R}_{|J|}, \mathcal{B}_{|J|})$, a že f_J je marginální hustota vektoru \mathbf{X}_J . *Podmíněnou hustotou* náhodného vektoru \mathbf{X}_I při daném \mathbf{X}_J nazveme takovou nezápornou měřitelnou funkci $f_{I|J}(\mathbf{x}_I|\mathbf{x}_J)$, která pro libovolné množiny $A \in \mathcal{B}_{|I|}$ a $B \in \mathcal{B}_{|J|}$ splňuje vztah

$$\Pr(\mathbf{X}_I \in A, \mathbf{X}_J \in B) = \int_B \left[\int_A f_{I|J}(\mathbf{x}_I|\mathbf{x}_J) d\mu_I(\mathbf{x}_I) \right] f_J(\mathbf{x}_J) d\mu_J(\mathbf{x}_J). \quad (1.2)$$

Platí přitom, že existují-li hustoty f a f_J z definice 1.1, existuje i podmíněná hustota $f_{I|J}$, je dokonce určena v jistém slova smyslu jednoznačně a lze ji vypočítat jako podíl sdružené hustoty f a hustoty podmínky f_J . Přesněji toto tvrzení formuluje následující věta¹.

Věta 1.1 *Mějme náhodný vektor \mathbf{X} , množiny I a J , součinnovou míru $\mu = \mu_I \times \mu_J$ a hustoty f a f_J jako v definici 1.1. Označme dále ν_J takovou míru na $(\mathbb{R}_{|J|}, \mathcal{B}_{|J|})$, která má hustotu f_J vzhledem k μ_J . Potom platí:*

¹ Znění věty i její důkaz jsou upraveny z [And07], věty 3.18, 3.19 a 3.20. Důkaz je zde uveden ze dvou důvodů: jednak jde o tvrzení zcela zásadní pro další práci, jednak průběh důkazu dobře ilustruje použití zavedeného značení (tj. výběr složek vektoru pomocí indexových množin).

(i) Bud' $q(\mathbf{x}_I|\mathbf{x}_J)$ funkce daná předpisem

$$q(\mathbf{x}_I|\mathbf{x}_J) = \begin{cases} \frac{f(\mathbf{x})}{f_J(\mathbf{x}_J)}, & \text{pokud } \mathbf{x}_J \notin N, \\ 0, & \text{pokud } \mathbf{x}_J \in N, \end{cases} \quad (1.3)$$

kde $N = \{\mathbf{x} \in \mathbb{R}_{|J|}: f_J(\mathbf{x}) = 0\}$. Pak q je podmíněná hustota \mathbf{X}_I při daném \mathbf{X}_J .

(ii) Podmíněná hustota je podmínkami nezápornosti a měřitelnosti a vztahem (1.2) určena jednoznačně až na ekvivalenci vzhledem k míře $(\mu_I \times \nu_J)$.

(iii) Je-li funkce $f_{I|J}$ podmíněná hustota \mathbf{X}_I při daném \mathbf{X}_J , platí

$$f(\mathbf{x}) = f_{I|J}(\mathbf{x}_I|\mathbf{x}_J)f_J(\mathbf{x}_J) \quad \text{s.v. } [\mu_I \times \mu_J]. \quad (1.4)$$

Důkaz. Začneme s bodem (i). Volme libovolně $A \in \mathcal{B}_{|I|}$ a $B \in \mathcal{B}_{|J|}$. Budeme nadále pro stručnost značit $\Pr(A \times B) = \Pr(\mathbf{X}_I \in A, \mathbf{X}_J \in B)$. Označme dále N^c doplněk N v $\mathbb{R}_{|J|}$. Potom můžeme psát

$$\Pr(A \times B) = \Pr[A \times (B \cap N)] + \Pr[A \times (B \cap N^c)].$$

Ukážeme nejprve, že $\Pr[A \times (B \cap N)] = 0$. Užijeme postupně Fubiniovu větu [Rud03, věta 8.8], nezápornost f a nakonec vztah pro výpočet marginální hustoty:

$$\begin{aligned} \Pr[A \times (B \cap N)] &= \int_{A \times (B \cap N)} f(\mathbf{x}) \, d(\mu_I \times \mu_J)(\mathbf{x}) = \\ &= \int_{B \cap N} \left[\int_A f(\mathbf{x}) \, d\mu_I(\mathbf{x}_I) \right] d\mu_J(\mathbf{x}_J) \leq \\ &\leq \int_N f_J(\mathbf{x}_J) \, d\mu_J(\mathbf{x}_J) = 0. \end{aligned}$$

Máme tedy celkem $\Pr(A \times B) = \Pr[A \times (B \cap N^c)]$, odkud můžeme za použití Fubiniovy věty a vztahu (1.3) vyjádřit:

$$\begin{aligned} \Pr(A \times B) &= \int_{A \times (B \cap N^c)} f(\mathbf{x}) \, d(\mu_I \times \mu_J)(\mathbf{x}) = \\ &= \iint_{A \times (B \cap N^c)} \frac{f(\mathbf{x})}{f_J(\mathbf{x}_J)} f_J(\mathbf{x}_J) \, d(\mu_I \times \mu_J)(\mathbf{x}) = \\ &= \int_{B \cap N^c} \left[\int_A q(\mathbf{x}_I|\mathbf{x}_J) \, d\mu_I(\mathbf{x}_I) \right] f_J(\mathbf{x}_J) \, d\mu_J(\mathbf{x}_J) = \\ &= \int_B \left[\int_A q(\mathbf{x}_I|\mathbf{x}_J) \, d\mu_I(\mathbf{x}_I) \right] f_J(\mathbf{x}_J) \, d\mu_J(\mathbf{x}_J), \end{aligned}$$

čili q splňuje (1.2), a tedy je podmíněnou hustotou \mathbf{X}_I při daném \mathbf{X}_J .

Ukažme nyní (ii) a (iii). Nechť $f_{I|J}$ je podmíněná hustota při daném \mathbf{X}_I při daném \mathbf{X}_J . Použitím Fubiniovy věty na vztah (1.2) dostaneme

$$\Pr(A \times B) = \int_{A \times B} f_{I|J}(\mathbf{x}_I|\mathbf{x}_J) \, d(\mu_I \times \nu_J)(\mathbf{x}).$$

Z věty o jednoznačném rozšíření součinnové míry [Luk02, věta 11.7] však musí takový vztah platit nejen pro množiny typu $A \times B$, ale pro všechny množiny z \mathcal{B}_n . Můžeme proto pro libovolnou $C \in \mathcal{B}_n$ psát:

$$\Pr(C) = \int_C f_{I|J}(\mathbf{x}_I|\mathbf{x}_J) d(\mu_I \times \nu_J)(\mathbf{x}), \quad (1.5)$$

Z (1.5) plyne, že funkce $f_{I|J}$ je hustotou pravděpodobnosti vektoru \mathbf{X} vzhledem k míře $(\mu_I \times \nu_J)$, a tvrzení (ii) o jednoznačnosti tedy plyne z Radonovynikodymovy věty [Rud03, věta 6.10]. Vztah (1.5) dále snadno upravíme na

$$\Pr(C) = \int_C f_{I|J}(\mathbf{x}_I|\mathbf{x}_J) f_J(\mathbf{x}_J) d(\mu_I \times \mu_J)(\mathbf{x}),$$

ze kterého zřejmě plyne (iii). □

V následujícím textu budeme pod označením $f_{I|J}$ rozumět vždy tu verzi podmíněné hustoty, která vyhovuje předpisu (1.3). Nyní můžeme přistoupit k poněkud odlišné definici podmíněné pravděpodobnosti, než kterou popisuje vztah (1.1).

Definice 1.2 Nechtě $A \in \mathcal{B}_{|I|}$ a $\mathbf{x}_J \in \mathbb{R}_{|J|}$. Pak *podmíněnou pravděpodobnost* jevu $\mathbf{X}_I \in A$ za podmínky $\mathbf{X}_J = \mathbf{x}_J$ definujeme vzorcem

$$\Pr(\mathbf{X}_I \in A | \mathbf{X}_J = \mathbf{x}_J) = \int_A f_{I|J}(\mathbf{x}_I|\mathbf{x}_J) d\mu_I(\mathbf{x}_I). \quad (1.6)$$

Oba uvedené přístupy k definici podmíněné pravděpodobnosti nejsou navzájem v konfliktním vztahu. Snadno se např. ukáže, že je-li μ_J čítací míra a $\Pr(\mathbf{X}_J = \mathbf{x}_J) > 0$, pak jsou závěry z použití obou zmíněných vzorců pro výpočet $\Pr(\mathbf{X}_I \in A | \mathbf{X}_J = \mathbf{x}_J)$ stejné, neboť:

$$\begin{aligned} \Pr(\mathbf{X}_I \in A | \mathbf{X}_J = \mathbf{x}_J) &= \int_A \frac{f(\mathbf{x})}{f_J(\mathbf{x}_J)} d\mu_I(\mathbf{x}_I) = \\ &= \frac{1}{f_J(\mathbf{x}_J)} \int_{\{\mathbf{x}_J\}} \left[\int_A f(\mathbf{x}) d\mu_I(\mathbf{x}_I) \right] d\mu_J(\mathbf{x}_J) = \\ &= \frac{\Pr(\mathbf{X}_I \in A, \mathbf{X}_J = \mathbf{x}_J)}{\Pr(\mathbf{X}_J = \mathbf{x}_J)}. \end{aligned}$$

V případě, že podmiňujeme spojitou náhodnou veličinou, nelze ovšem vztah (1.1) použít a je třeba pracovat s podmíněnou hustotou.

V dalším textu se budeme zabývat výhradně spojitými náhodnými veličinami. Termínem *spojitá náhodná veličina* přitom označujeme takové veličiny, pro které existuje hustota vzhledem *Lebesgueově míře* na $(\mathbb{R}, \mathcal{B})$. Lebesgueovu míru budeme označovat symbolem λ . Při zápisech integrálů však budeme tento symbol zpravidla pro stručnost vynechávat, tedy ztotožníme $\int f(x) d\lambda(x)$ a $\int f(x) dx$.

V případě, že budeme pracovat s náhodným vektorem $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$, budeme vždy vyžadovat, aby existovala sdružená hustota všech jeho složek vzhledem k Lebesgueově míře na $(\mathbb{R}_n, \mathcal{B}_n)$; tuto míru budeme pro $n \geq 2$ značit λ_n . Podotkněme, že podle terminologie zavedené v [Stu02] je potom rozdělení tohoto vektoru *marginálně spojitě*. Pojem marginální spojitosti však v této práci nebudeme zavádět.

1.2 Podmíněná nezávislost

Připomeňme nejprve prostou *nezávislost náhodných vektorů*, bez přívlastku *podmíněná*.

Definice 1.3 Řekneme, že náhodné vektory $\mathbf{X} = (X_1, X_2, \dots, X_m)^\top$ a $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^\top$ jsou *nezávislé*, pokud pro libovolné množiny $A \in \mathcal{B}_m$ a $B \in \mathcal{B}_n$ platí:

$$\Pr(\{\omega \in \Omega : \mathbf{X}(\omega) \in A, \mathbf{Y}(\omega) \in B\}) = \Pr(\{\omega \in \Omega : \mathbf{X}(\omega) \in A\}) \cdot \Pr(\{\omega \in \Omega : \mathbf{Y}(\omega) \in B\}),$$

neboli stručněji

$$\Pr(\mathbf{X} \in A, \mathbf{Y} \in B) = \Pr(\mathbf{X} \in A) \cdot \Pr(\mathbf{Y} \in B).$$

Skutečnost, že vektory \mathbf{X} a \mathbf{Y} jsou nezávislé, značíme $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$.

Předchozí definici je možné analogicky rozšířit z dvojice vektorů na obecně n -tici vektorů. Dodejme, že definice zahrnuje i případ jednorozměrných náhodných veličin (pak $m = n = 1$). Následující lemma je známé tvrzení, které za jistých předpokladů charakterizuje nezávislost náhodných vektorů pomocí jejich hustot.

Lemma 1.2 *Uvažujme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s hustotou f vzhledem k λ_n a dvě neprázdné disjunktní indexové množiny I a J takové, že $I \cup J = \{1, 2, \dots, n\}$. Označme marginální hustoty vektorů \mathbf{X}_I a \mathbf{X}_J vzhledem k $\lambda_{|I|}$ a $\lambda_{|J|}$ jako f_I a f_J . Potom vektory \mathbf{X}_I a \mathbf{X}_J jsou nezávislé právě tehdy, když pro skoro všechna $\mathbf{x} \in \mathbb{R}_n$ vzhledem k λ_n platí*

$$f(\mathbf{x}) = f_I(\mathbf{x}_I) \cdot f_J(\mathbf{x}_J). \quad (1.7)$$

Důkaz. Jde o bezprostřední důsledek věty 2.9 z [Eat07]. □

Při značení marginálních a podmíněných hustot se budeme nadále držet konvence, kterou jsme použili v předchozím lemmatu a ve větě 1.1. Máme-li náhodný vektor \mathbf{X} s hustotou f a neprázdné disjunktní podmnožiny indexů jeho složek I a J , budeme zápisem f_I vždy značit marginální hustotu vektoru \mathbf{X}_I a pod označením $f_{I|J}$ budeme rozumět podmíněnou hustotu \mathbf{X}_I při daném \mathbf{X}_J ve tvaru (1.3). Tuto konvenci dále rozšíříme i na případ, kdy množina I je prázdná: je-li $I = \emptyset$, definujeme $f_I = f_{I|J} = 1$.

Přistoupíme nyní k definici podmíněné nezávislosti. V kontextu zaměření této práce se přitom omezíme pouze na případ takových n -rozměrných náhodných vektorů, pro které existuje hustota vzhledem k Lebesgueově míře na $(\mathbb{R}_n, \mathcal{B}_n)$. Obecnější rámec pro zkoumání podmíněné nezávislosti lze najít např. v [Lau96] nebo [Stu02].

Definice 1.4 Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s rozdělením \mathcal{L} a hustotou f vzhledem k λ_n , a tři disjunktní indexové množiny $I, J, K \subseteq \{1, 2, \dots, n\}$. Řekneme, že vektory \mathbf{X}_I a \mathbf{X}_J jsou *nezávislé při daném \mathbf{X}_K vzhledem k \mathcal{L}* , platí-li pro skoro všechna $\mathbf{x} \in \mathbb{R}_n$ vzhledem k λ_n

$$f_{IJK}(\mathbf{x}_{IJK}) \cdot f_K(\mathbf{x}_K) = f_{IK}(\mathbf{x}_{IK}) \cdot f_{JK}(\mathbf{x}_{JK}), \quad (1.8)$$

kde výrazy f_{IJK} , f_K , f_{IK} a f_{JK} značí příslušné marginální hustoty, přičemž tuto podmíněnou nezávislost pak značíme zápisem $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K [\mathcal{L}]$. Vzhledem k tomu, že rozdělení zkoumaného vektoru bývá zpravidla pevně dáno, budeme jej zpravidla ze zápisu vynechávat a psát stručněji $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K$.

Poznamenejme, že pokud $K = \emptyset$, dostaneme při použití výše uvedené konvence $f_K = 1$, a (1.8) pak odpovídá vztahu pro (nepodmíněnou) nezávislost (1.7). V takovém případě můžeme tedy chápat podmíněnou nezávislost při daném \mathbf{X}_K jako závislost nepodmíněnou a ztotožnit výrazy $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K$ a $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J$. Je-li prázdná množina I nebo množina J , platí (1.8) všude na \mathbb{R}_n .

Zcela analogicky je možné zavést i podmíněnou nezávislost pro dříve zmíněný obecnější rámec marginálně spojitých rozdělání s tím, že Lebesgueovu míru na \mathbb{R}_n potom vystřídá obecnější součinnová míra, podrobněji viz [Stu02].

Při diskuzi o podmíněné nezávislosti je užitečné následující pozorování, díky kterému budeme moci zavést některá alternativní vyjádření definičního vztahu (1.8).

Lemma 1.3 *Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$, který má hustotu f vzhledem k λ_n , a dvě indexové množiny A a B takové, že $A \subset B \subseteq \{1, 2, \dots, n\}$. Potom*

$$f_A(\mathbf{x}_A) = 0 \implies f_B(\mathbf{x}_B) = 0 \quad \text{s.v. } [\lambda_n]. \quad (1.9)$$

Důkaz. Je-li $A = \emptyset$, je podle zavedené konvence $f_A = 1$, a tvrzení věty platí triviálně. Nechť tedy $A \neq \emptyset$. Bez újmy na obecnosti budeme předpokládat, že $A = \{1, 2, \dots, k\}$ a $B = \{1, 2, \dots, m\}$ při $k < m \leq n$, což zjednoduší značení. Zavedeme množiny N , N_B a N_A následujícím předpisem:

$$\begin{aligned} N &= \{\mathbf{x} \in \mathbb{R}_n : f_A(\mathbf{x}_A) = 0, f_B(\mathbf{x}_B) > 0\}, \\ N_B &= \{\mathbf{y} \in \mathbb{R}_{|B|} : f_A(\mathbf{y}_A) = 0, f_B(\mathbf{y}) > 0\}, \\ N_A &= \{\mathbf{z} \in \mathbb{R}_{|A|} : f_A(\mathbf{z}) = 0\}. \end{aligned}$$

Vztah (1.9) je nyní ekvivalentní s tvrzením, že $\lambda_n(N) = 0$. Zřejmě platí

$$N_B \subseteq N_A \times \mathbb{R}_{|B|-|A|}, \quad N = N_B \times \mathbb{R}_{n-|B|} \subseteq N_A \times \mathbb{R}_{n-|A|}. \quad (1.10)$$

Je-li proto $\lambda_{|A|}(N_A) = 0$, pak $\lambda_n(N) = 0$ a (1.9) platí. Nechť nyní $\lambda_{|A|}(N_A) > 0$. Z (1.10), nezápornosti hustot, Fubiniovy věty a z věty o marginální hustotě [And07, věta 3.10] postupně dostaneme

$$\begin{aligned} 0 &\leq \int_{N_B} f_B(\mathbf{y}) \, d\mathbf{y} \leq \int_{N_A \times \mathbb{R}_{|B|-|A|}} f_B(\mathbf{y}) \, d\mathbf{y} = \\ &= \int_{N_A} \left[\int_{\mathbb{R}_{|B|-|A|}} f_B(\mathbf{y}) \, d\mathbf{y}_{B \setminus A} \right] d\mathbf{y}_A = \\ &= \int_{N_A} f_A(\mathbf{y}_A) \, d\mathbf{y}_A = 0. \end{aligned}$$

Jelikož $f_B(\mathbf{y}) > 0$ na N_B , je nutně $\lambda_{|B|}(N_B) = 0$ a podle (1.10) i $\lambda_n(N) = 0$. \square

Díky předchozímu lemmatu máme zaručeno, že rovnost v (1.8) platí pro libovolnou hustotu f skoro všude tam, kde je některá z funkcí f_K , f_{IK} nebo f_{JK} nulová. V ostatních případech můžeme těmito funkcemi celou rovnost dělit. Přihlédneme-li ke tvaru podmíněné hustoty (1.3), můžeme vztah (1.8) vyjádřit ekvivalentně jako

$$f_{I|J|K}(\mathbf{x}_{IJ} \mid \mathbf{x}_K) = f_{I|K}(\mathbf{x}_I \mid \mathbf{x}_K) \cdot f_{J|K}(\mathbf{x}_J \mid \mathbf{x}_K), \quad (1.11)$$

případně

$$f_{I|JK}(\mathbf{x}_I \mid \mathbf{x}_{JK}) = f_{I|K}(\mathbf{x}_I \mid \mathbf{x}_K). \quad (1.12)$$

Analogie mezi nezávislostí \mathbf{X}_I a \mathbf{X}_J a podmíněnou nezávislostí těchto dvou vektorů při daném \mathbf{X}_K je patrná z porovnání (1.11) a (1.7). Tvar (1.12) naopak bývá využíván pro verbální popis vztahu podmíněné nezávislosti. Plyne z něj, že pokud jsou \mathbf{X}_I a \mathbf{X}_J podmíněně nezávislé při daném \mathbf{X}_K , potom známe-li \mathbf{X}_K , nemá pro nás dodatečná informace o \mathbf{X}_J význam při zkoumání \mathbf{X}_I . Jinými slovy, při známém \mathbf{X}_K již nemá \mathbf{X}_J na \mathbf{X}_I vliv.

Jak uvádí např. [Stu02], podmíněné nezávislosti v náhodném vektoru \mathbf{X} splňují řadu vlastností, které jsou souhrnně označovány jako *semi-grafoidové vlastnosti* a které zavedeme v následující definici; pro jednotlivé vlastnosti jsme ponechali původní anglické názvy.

Definice 1.5 Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s rozdělením \mathcal{L} a hustotou f vzhledem k λ_n a neprázdné disjunktní indexové množiny $I, J, K, L \subseteq \{1, 2, \dots, n\}$. Pak *semi-grafoidovými* vlastnostmi podmíněných nezávislostí vektoru rozumíme soubor následujících vlastností; při jejich zápisu budeme pro snazší čitelnost zkracovat zápisy typu $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K [\mathcal{L}]$ jako $I \perp\!\!\!\perp J \mid K$:

- (i) *triviality*: $I \perp\!\!\!\perp \emptyset \mid K$.
- (ii) *symmetry*: $I \perp\!\!\!\perp J \mid K \implies J \perp\!\!\!\perp I \mid K$.
- (iii) *decomposition*: $I \perp\!\!\!\perp JL \mid K \implies I \perp\!\!\!\perp J \mid K$.
- (iv) *weak union*: $I \perp\!\!\!\perp JL \mid K \implies I \perp\!\!\!\perp J \mid KL$.
- (v) *contraction*: $\{I \perp\!\!\!\perp J \mid KL \ \& \ I \perp\!\!\!\perp L \mid K\} \implies I \perp\!\!\!\perp JL \mid K$.

Výše uvedené vlastnosti podmíněné nezávislosti budou podstatné zejména při studiu struktur podmíněné nezávislosti v oddílu 1.4.

1.3 Podmíněná střední hodnota, rozptyl a kovariance

V následující definici zavedeme pojmy uvedené v názvu tohoto oddílu. Opět se pro jednoduchost omezíme na práci s n -rozměrnými náhodnými vektory, které mají sdruženou hustotu vůči λ_n . Uvedená definice se opírá o pojem podmíněné hustoty a je upravena z [And07, oddíl 3.6]. Dodejme, že podmíněná střední hodnota se často zavádí obecnějším způsobem, viz např. [Ště87, oddíl VI.1]; pro naše účely to ovšem nebude potřeba.

Definice 1.6 Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s hustotou f vzhledem k λ_n a neprázdné disjunktní indexové množiny I a J takové, že $I \cup J = \{1, 2, \dots, n\}$. Nechť $S(\mathbf{x})$ je měřitelná funkce n proměnných, takže $S = S(\mathbf{X})$ je náhodná veličina. Předpokládejme, že existuje konečná střední hodnota $\mathbf{E}S$. *Podmíněnou střední hodnotou* veličiny S za podmínky, že vektor \mathbf{X}_J nabyl hodnoty \mathbf{x}_J , definujeme vzorcem

$$\mathbf{E}(S \mid \mathbf{X}_J = \mathbf{x}_J) = \int_{\mathbb{R}_{|I|}} S(\mathbf{x}_I, \mathbf{x}_J) f_{I \mid J}(\mathbf{x}_I \mid \mathbf{x}_J) d\mathbf{x}_I, \quad (1.13)$$

pokud integrál na pravé straně existuje. Položme nyní $g(\mathbf{x}_J) = \mathbf{E}(S \mid \mathbf{X}_J = \mathbf{x}_J)$ a zaveďme symbol $\mathbf{E}(S \mid \mathbf{X}_J)$ předpisem $\mathbf{E}(S \mid \mathbf{X}_J) = g(\mathbf{X}_J)$. Pak náhodnou veličinu $\mathbf{E}(S \mid \mathbf{X}_J)$ nazveme *podmíněnou střední hodnotou* S při daném \mathbf{X}_J . Nechť navíc $\mathbf{E}S^2 < \infty$. Potom definujeme *podmíněný rozptyl* S při daném \mathbf{X}_J jako

$$\text{var}(S \mid \mathbf{X}_J) = \mathbf{E}\{(S - \mathbf{E}[S \mid \mathbf{X}_J])^2 \mid \mathbf{X}_J\}. \quad (1.14)$$

Mějme dále měřitelnou funkci n proměnných $T(\mathbf{x})$, označme $T = T(\mathbf{X})$ a předpokládejme, že $\mathbf{E}T^2 < \infty$. *Podmíněnou kovarianci* S a T při daném \mathbf{X}_J zavedeme předpisem

$$\text{cov}(S, T \mid \mathbf{X}_J) = \mathbf{E}\{[S - \mathbf{E}(S \mid \mathbf{X}_J)] \cdot [T - \mathbf{E}(T \mid \mathbf{X}_J)] \mid \mathbf{X}_J\}. \quad (1.15)$$

Z věty 1.1 plyne, že podmíněná střední hodnota $E(S | \mathbf{X}_J = \mathbf{x}_J)$ je coby funkce proměnné \mathbf{x}_J určena jednoznačně až na ekvivalenci vzhledem k míře ν , kde ν má hustotu f_J vzhledem k $\lambda_{|J|}$. Proto dvě různé verze podmíněné střední hodnoty se mohou lišit pouze na množině, která má nulovou pravděpodobnost (podrobněji viz [And07], str. 58). V následující větě shrneme některé další vlastnosti podmíněné střední hodnoty. V této větě (jako i v dalším textu) budeme funkci pro výpočet absolutní hodnoty značit „abs“. Připomeňme ještě, že zápis „skoro jistě“, příp. „s.j.“ interpretujeme jako „s.v. [Pr]“.

Věta 1.4 (Vlastnosti podmíněné střední hodnoty.) *Mějme náhodný vektor \mathbf{X} a indexové množiny I a J jako v definici 1.6. Uvažujme náhodné veličiny $S = S(\mathbf{X})$, $T = T(\mathbf{X})$ a reálná čísla a, b . Pak platí:*

- (i) $ES = E[E(S | \mathbf{X}_J)]$.
- (ii) $E(a | \mathbf{X}_J) = a$ skoro jistě.
- (iii) $E(aS + bT | \mathbf{X}_J) = aE(S | \mathbf{X}_J) + bE(T | \mathbf{X}_J)$.
- (iv) *Nechť $H = H(\mathbf{X}_J)$ je náhodná veličina, která nezávisí na \mathbf{X}_I , a platí $E[abs(S)] < \infty$, $E[abs(H \cdot S)] < \infty$. Potom $E(H \cdot S | \mathbf{X}_J) = H \cdot E(S | \mathbf{X}_J)$ skoro jistě.*

Důkaz. Viz [And07], věty 3.22, 3.23, 3.24 a 3.25.

Jak známo (viz např. [And07, věta 2.9]), nezávislé veličiny jsou nutně nekorelované, tj. pro náhodné veličiny X_i a X_j s konečnými druhými momenty platí

$$X_i \perp\!\!\!\perp X_j \implies \text{cov}(X_i, X_j) = 0.$$

V následující větě ukážeme, že analogické tvrzení platí i o podmíněných nezávislostech a podmíněných kovariancích.

Věta 1.5 *Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s hustotou f vzhledem k λ_n , čísla i a j splňující $1 \leq j < i \leq n$ a neprázdnou indexovou množinu $K \subseteq \{1, 2, \dots, n\} \setminus \{i, j\}$. Nechť dále X_i a X_j mají konečné druhé momenty. Pak platí:*

$$X_i \perp\!\!\!\perp X_j | \mathbf{X}_K \implies \text{cov}(X_i, X_j | \mathbf{X}_K) = 0 \quad \text{s.j.} \quad (1.16)$$

Důkaz. Označme podmíněnou hustotu \mathbf{X}_{ij} při daném \mathbf{X}_K jako $f_{ij|K}$, analogický význam budou mít symboly $f_{i|K}$ a $f_{j|K}$. Označme dále $A = X_i - E(X_i | \mathbf{X}_K)$, $B = X_j - E(X_j | \mathbf{X}_K)$. Tedy A není funkcí X_j , B není funkcí X_i a podle věty 1.4 zřejmě platí

$$E(A | \mathbf{X}_K) = E(B | \mathbf{X}_K) = 0 \quad \text{s.j.} \quad (1.17)$$

V následujícím výpočtu uijeme postupně (1.11), Fubiniovu větu a nakonec (1.17):

$$\begin{aligned} \text{cov}(X_i, X_j | \mathbf{X}_K = \mathbf{x}_K) &= E(AB | \mathbf{X}_K = \mathbf{x}_K) = \\ &= \int_{\mathbb{R}^2} AB f_{ij|K} d\mathbf{x}_{ij} = \\ &\stackrel{(1.11)}{=} \int_{\mathbb{R}^2} AB f_{i|K} f_{j|K} d\mathbf{x}_{ij} = \\ &\stackrel{\text{Fub. v.}}{=} \int_{\mathbb{R}} \left[\int_{\mathbb{R}} A f_{i|K} dx_i \right] B f_{j|K} dx_j = \\ &= \left[\int_{\mathbb{R}} A f_{i|K} dx_i \right] \cdot \left[\int_{\mathbb{R}} B f_{j|K} dx_j \right] = \\ &= E(A | \mathbf{X}_K = \mathbf{x}_K) \cdot E(B | \mathbf{X}_K = \mathbf{x}_K) = \\ &\stackrel{(1.17)}{=} 0 \quad \text{s.j.} \quad \square \end{aligned}$$

1.4 Grafické modely struktury podmíněné nezávislosti

V úvodu tohoto oddílu shrneme základní pojmy z teorie grafů, které budeme v dalším textu využívat, a zavedeme odpovídající značení. Předem podotkneme, že si pro naše potřeby vystačíme s užším terminologickým rámcem, než jaký se obecně v teorii grafů zavádí. Budeme např. předpokládat, že všechny grafy jsou konečné a že neobsahují smyčky ani souběžné hrany; takové grafy proto následující definice vůbec nepřipouští.

Definice 1.7 *Grafem* rozumíme dvojici (V, E) , kde V je konečná množina a E je množina uspořádaných dvojic různých prvků z V , neboli $E \subseteq \{(\alpha, \beta) : \alpha, \beta \in V, \alpha \neq \beta\}$. Prvky množiny V nazýváme *vrcholy* (*vertices*), prvky z množiny E nazýváme *hrany* (*edges*). Řekneme, že hrana $(\alpha, \beta) \in E$ je *neorientovaná*, pokud je i $(\beta, \alpha) \in E$; v takovém případě značíme $\alpha \leftrightarrow \beta$. Pokud naopak $(\beta, \alpha) \notin E$, nazveme $(\alpha, \beta) \in E$ hranou *orientovanou* a značíme $\alpha \rightarrow \beta$, příp. $\beta \leftarrow \alpha$. Graf, ve kterém se vyskytují pouze orientované hrany, nazveme *orientovaným grafem*. *Sledem délky n z α do β* rozumíme posloupnost ne nutně různých vrcholů $(\alpha_0, \alpha_1, \dots, \alpha_n)$ takovou, že $\alpha = \alpha_0$, $\beta = \alpha_n$ a pro $i = 1, 2, \dots, n$ je $(\alpha_{i-1}, \alpha_i) \in E$ nebo $(\alpha_i, \alpha_{i-1}) \in E$. Platí-li navíc, že pro $i = 1, 2, \dots, n$ je $(\alpha_{i-1}, \alpha_i) \in E$, hovoříme o sledu *orientovaném*. Připouštíme i sledy délky nula, tj. sledy typu (α_0) . Orientovaný sled nenulové délky z α do α , tj. sled, který začíná a končí ve stejném vrcholu, nazveme *orientovaným cyklem*. Orientovaný graf, který neobsahuje žádný orientovaný cyklus, nazveme *acyklickým orientovaným grafem*, zkracujeme jako ADG (z anglického *acyclic directed graph*). Sled $(\alpha_0, \alpha_1, \dots, \alpha_n)$, jehož všechny vrcholy jsou různé, nazveme *cestou délky n z α_0 do α_n* . Analogicky jako pro sledy zavedeme též *orientované cesty*.

S grafem je neodmyslitelně spjato jeho grafické vyjádření. Vrcholy grafu zakreslujeme pomocí koleček, do kterých budeme zpravidla zapisovat název vrcholu. Orientovanou hranu (α, β) zakreslujeme pomocí šipky vedoucí od kolečka vrcholu α ke kolečku β , neorientovanou hranu (α, β) vyjádříme pomocí obousměrné šipky spojující kolečka vrcholů α a β .² Např. je-li $V = \{\alpha, \beta, \gamma\}$ a $E = \{(\alpha, \beta), (\beta, \alpha), (\beta, \gamma)\}$, můžeme zakreslit graf $G = (V, E)$ následovně:



Potřebujeme-li v textu zapsat nějaký sled nebo cestu, zpravidla je nezapisujeme ve tvaru posloupnosti, nýbrž používáme zápis, ve kterém je vyznačena orientace hran mezi sousedními vrcholy: například cestu vedoucí z α do γ v předchozím grafu zapíšeme jako $\alpha \leftrightarrow \beta \rightarrow \gamma$.

Množinu vrcholů V v definici 1.7 tvoří libovolná konečná množina. Abychom mohli snadněji zachytit některé vlastnosti vztahující se k danému grafu, budeme pro přehlednost vrcholy grafu „číslovat“ od 1 do $|V|$. Formálně vzato, *očíslováním uzlů* v grafu $G = (V, E)$ můžeme rozumět nějakou bijekci $\varphi : V \rightarrow \{1, 2, \dots, |V|\}$. Označíme-li $V' = \{1, 2, \dots, |V|\}$ a $E' = \{(\varphi(\alpha), \varphi(\beta)) : (\alpha, \beta) \in E\}$, potom jsou zřejmě grafy $G = (V, E)$ a $G' = (V', E')$ izomorfní. Při známém očíslování φ můžeme proto ztotožňovat vrchol α s jeho očíslováním $\varphi(\alpha)$, resp. graf G s grafem G' .

V dalším textu se budeme zabývat výhradně acyklickými orientovanými grafy; pro tyto grafy budeme používat jisté význačné očíslování, o kterém hovoří následující definice.

²Neorientované hrany se často zakreslují pomocí prostých čar (tj. bez šipek na koncích). Zde jsme se přiklonili k výše uvedenému značení, neboť se pak shoduje se značením v diagramech, které se používají v metodě koeficientů na cestách, o níž pojednává oddíl 3.1.

Definice 1.8 Necht $G = (V, E)$ je ADG. Potom *topologickým očíslováním* grafu G rozumíme bijekci $\varphi : V \rightarrow \{1, 2, \dots, |V|\}$ takovou, že $\varphi(\alpha) < \varphi(\beta)$, kdykoli $\alpha \rightarrow \beta$. Podobně, pokud pro nějaký graf platí, že $V = \{1, 2, \dots, |V|\}$ a navíc $j < i$, kdykoli $j \rightarrow i$, pak řekneme, že tento graf je *topologicky očíslovaný*.

V následující větě ukážeme, že každý ADG lze topologicky očíslovat. Díky tomu pak bude možné bez újmy na obecnosti předpokládat, že všechny ADG, se kterými budeme pracovat, jsou topologicky očíslované, což zjednoduší zápis některých vztahů.

Věta 1.6 (O existenci topologického očíslování.) *Necht $G = (V, E)$ je ADG. Potom existuje topologické očíslování grafu G .*

Důkaz. V dalším textu budeme říkat, že vrchol je *koncový*, pokud z něj nevede žádná hrana. V první řadě si uvědomme, že v každém ADG musí nutně existovat alespoň jeden koncový vrchol. To je snadný důsledek toho, že každý ADG je acyklický a konečný (tj. má konečný počet vrcholů): pokud není v grafu s n vrcholy koncový vrchol, můžeme zřejmě nalézt z libovolného vrcholu cestu délky n , na cestě délky n se ovšem musí nutně opakovat dva stejné vrcholy, tedy uvažovaný graf obsahuje cyklus.

Tvrzení věty nyní dokážeme indukcí podle počtu vrcholů v grafu, označme $|V| = n$. Je-li $n = 1$, je tvrzení zřejmé. Zbývá dokázat indukční krok. Necht tedy platí tvrzení věty pro $n - 1$. Mějme ADG $G = (V, E)$ s n vrcholy. Najdeme v G koncový vrchol a označme jej jako β . Uvažujme libovolné očíslování φ takové, že $\varphi(\beta) = n$. Potom zřejmě pro všechna $\alpha \in V$ platí, že $\varphi(\alpha) < \varphi(\beta)$, pokud $\alpha \rightarrow \beta$. Jinými slovy, hrany vedoucí do β vždy splňují požadavek na topologické očíslování, nehledě na to, jak očísloujeme ostatní vrcholy. K očíslování ostatních vrcholů uijeme indukčního předpokladu. Označme $\bar{V} = V \setminus \{\beta\}$, $\bar{E} = E \cap (\bar{V} \times \bar{V})$, $\bar{G} = (\bar{V}, \bar{E})$. Potom \bar{G} je ADG s $n - 1$ vrcholy, a z indukčního předpokladu pro něj tudíž existuje topologické očíslování, které označíme $\bar{\varphi}$. Definujeme-li zobrazení φ předpisem

$$\varphi(\alpha) = \begin{cases} \bar{\varphi}(\alpha), & \text{pokud } \alpha \in V \setminus \{\beta\}, \\ n, & \text{pokud } \alpha = \beta, \end{cases}$$

potom z předchozího plyne, že φ je topologické očíslování G . □

Předchozí důkaz dává zároveň návod, jak topologické očíslování sestavit: stačí najít koncový vrchol, přiřadit mu nejvyšší číslo, poté tento vrchol spolu se vstupujícími hranami z grafu pomyslně vypustit a celý proces opakovat. Analogicky lze postupovat i „z druhé strany“, tj. od vrcholů, do kterých nevstupuje žádná hrana. Vzhledem k tomu, že daný graf může mít několik koncových vrcholů, nemusí být topologické očíslování dáno jednoznačně.

Topologicky očíslované ADG budeme využívat k popisu struktury podmíněné nezávislosti vícerozměrných rozdělání. Přitom strukturou podmíněné nezávislosti rozumíme soubor údajů o podmíněných nezávislostech mezi složkami vícerozměrné náhodné veličiny; přesněji tento pojem zavádí následující definice.

Definice 1.9 Označme symbolem $\mathcal{T}(n)$ systém všech uspořádaných trojic $\langle I, J | K \rangle$ disjunktních podmnožin množiny $\{1, 2, \dots, n\}$. Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s rozdělením \mathcal{L} a hustotou f vzhledem k λ_n . Řekneme, že $\mathcal{S}(\mathcal{L}) \subset \mathcal{T}(n)$ je *strukturou podmíněné nezávislosti* indukovanou \mathcal{L} , platí-li

$$\mathcal{S}(\mathcal{L}) = \{ \langle I, J | K \rangle \in \mathcal{T}(n) : \mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K [\mathcal{L}] \}. \quad (1.18)$$

Struktury podmíněné nezávislosti musí splňovat některé formální vlastnosti, které se odvíjejí od semi-grafoidových vlastností podmíněné nezávislosti, které jsme zavedli v definici 1.5. Z vlastnosti *triviality* je např. zřejmé, že každá struktura podmíněné nezávislosti obsahuje všechny tzv. *triviální trojice*, tj. trojice typu $\langle I, \emptyset | K \rangle$, z vlastnosti *symmetry* vyplývá, že s každou trojicí $\langle I, J | K \rangle$ musí struktura podmíněné nezávislosti obsahovat i k ní *symmetrickou* trojici $\langle J, I | K \rangle$ atd.

Nelze však zaručit, že každá množina $\mathcal{S} \subset \mathcal{T}(n)$ splňující uvedené semi-grafoidové vlastnosti je strukturou podmíněné nezávislosti nějakého rozdělení. Jinými slovy, není pravda, že semi-grafoidové vlastnosti charakterizují struktury podmíněné nezávislosti (viz [Lau96, str. 32]). Ve [Stu98] je dokonce dokázáno, že pro diskrétní pravděpodobnostní rozdělení (kterými se v této práci nezabýváme), nelze obecně charakterizovat struktury podmíněných nezávislostí pomocí konečného počtu formálních vlastností tohoto typu.

Předpokládejme nyní, že je pevně určen nějaký distribuční rámec, který má podobu třídy Φ n -rozměrných pravděpodobnostních rozdělení – např. třída všech n -rozměrných regulárních gaussovských rozdělení \mathcal{N}_{reg} z definice B.1 nebo třída všech diskrétních rozdělení na \mathbb{R}_n . Uvažujme nějaký soubor \mathcal{I} nezávislostních údajů o n -rozměrném rozdělení, které jsou podobně jako u struktury podmíněné nezávislosti zakódovány pomocí trojic typu $\langle I, J | K \rangle$, tedy $\mathcal{I} \subset \mathcal{T}(n)$. Potom \mathcal{I} lze chápat jako *statistický model*, určuje totiž třídu pravděpodobnostních rozdělení z Φ , jejichž struktura podmíněné nezávislosti obsahuje všechna nezávislostní tvrzení z \mathcal{I} , neboli třídu

$$\{\mathcal{L} \in \Phi : \mathcal{S}(\mathcal{L}) \supseteq \mathcal{I}\}.$$

Lze ukázat, že některé modely používané ve statistice lze ekvivalentně formulovat jako právě zavedené statistické modely určené pomocí požadavků na strukturu podmíněné nezávislosti; v kapitole 2 ukážeme takovouto ekvivalenci pro případ lineárních rekursivních systémů, další příklady viz [Stu02].

Různým množinám \mathcal{I} ve výše uvedeném statistickém modelu mohou odpovídat stejné třídy pravděpodobnostních rozdělení. To je patrné například ze semi-grafoidových vlastností: z vlastnosti *triviality* plyne, že vynecháme-li z množiny \mathcal{I} všechny triviální trojice, výsledná třída pravděpodobnostních rozdělení se nezmění. Podobně můžeme na základě vlastnosti *symmetry* ponechat pouze jeden z každé dvojice symetrických prvků. Ze zbylých tří semi-grafoidových vlastností dále přímo plyne, že

$$I \perp\!\!\!\perp JL | K \iff \{I \perp\!\!\!\perp J | KL \ \& \ I \perp\!\!\!\perp L | K\},$$

(použili jsme zde zkráceného zápisu podmíněných nezávislostí stejně jako v definici semi-grafoidových vlastností). Odkud je vidět, že libovolnou množinu \mathcal{I} lze nahradit seznamem tzv. *elementárních trojic*, což jsou trojice typu $\langle i, j | K \rangle$, neboli první dvě množiny v dané trojici jsou jednoprvkové.

Grafické modely struktury podmíněné nezávislosti umožňují zadat požadavky na strukturu podmíněné nezávislosti vizuálně názorným způsobem pomocí grafu. Pro tento účel se používá celá řada různých tříd grafů (dobré shrnutí uvádí např. [Lau96] nebo [And99]), v této práci se zaměříme pouze na využití acyklických orientovaných grafů. Existuje rovněž více možností, jak interpretovat graf v podobě požadavků na závislostní strukturu pravděpodobnostního rozdělení; hovoří se o tzv. *markovských a faktorizačních podmínkách*. Zde se budeme zabývat podmínkami markovskými, pomocí nichž lze z daného grafu přímo určit výše zavedený statistický model, tj. zapsat množinu \mathcal{I} . Abychom mohli tyto podmínky přehledně formálně definovat, je třeba nejprve zavést některé pojmy týkající se ADG.

Definice 1.10 Nechť $G = (V, E)$ je ADG, $\alpha, \beta \in V$. Skutečnost, že graf G neobsahuje hranu $\alpha \rightarrow \beta$, značíme $\alpha \nrightarrow \beta$ (tento zápis nevylučuje existenci hrany $\alpha \rightarrow \beta$). Řekneme,

že β je *rodičem* α , je-li $\beta \rightarrow \alpha$. Množinu všech rodičů vrcholu α označíme jako $\text{pa}(\alpha)$ (z anglického *parents*). Řekneme, že β je *následníkem* α , pokud v G existuje orientovaná cesta z α do β . Množinu všech vrcholů z $V \setminus \{\alpha\}$, které nejsou následníky α , označíme jako $\text{nd}(\alpha)$ (z anglického *non-descendants*).

Přejdeme nyní k definici prvních dvou markovských podmínek. V této definici, jakož i ve zbylé části této kapitoly, budeme pro přehlednost užívat konvenci pro zkrácený zápis podmíněných nezávislostí, kterou jsme zavedli pro účely definice 1.5, tj. zapíšeme podmíněné nezávislosti typu $\mathbf{X}_I \perp\!\!\!\perp \mathbf{X}_J \mid \mathbf{X}_K [\mathcal{L}]$ zkráceně jako $I \perp\!\!\!\perp J \mid K$.

Definice 1.11 Nechť $G = (V, E)$ je topologicky očíslovaný acyklický orientovaný graf s n vrcholy a $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ je náhodný vektor s rozdělením \mathcal{L} , pro které existuje hustota vzhledem k λ_n . Řekneme, že \mathcal{L} splňuje *párovou markovskou podmínku* (*PM*) vzhledem ke G , pokud pro libovolné dva vrcholy $i, j \in V$ takové, že $j < i$ a $j \nrightarrow i$, platí

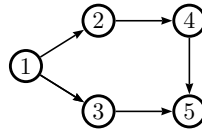
$$i \perp\!\!\!\perp j \mid \{1, 2, \dots, i-1\} \setminus \{j\}. \quad (PM)$$

Řekneme, že \mathcal{L} splňuje *lokální markovskou podmínku* (*LM*) vzhledem ke G , pokud pro libovolný uzel $i \in V \setminus \{1\}$ platí

$$i \perp\!\!\!\perp \{1, 2, \dots, i-1\} \setminus \text{pa}(i) \mid \text{pa}(i). \quad (LM)$$

Z předchozí definice je patrné, že každý vrchol daného grafu představuje jednu z náhodných veličin, jejichž sdružené rozdělení zkoumáme; to je společný znak všech grafických modelů struktury podmíněné nezávislosti. Rozdělení, které splňuje (*PM*), resp. (*LM*) vzhledem ke grafu G , budeme též označovat jako *párově*, resp. *lokálně markovské* vzhledem ke G . Z definice rovněž vidíme, že v případě (*PM*) bude zadaný graf reprezentovat soubor elementárních nezávislostních údajů (tj. údajů odpovídajících elementárním trojicím). Použitím vlastnosti *weak union* ihned dostaneme, že lokální markovská podmínka je silnější než podmínka párová, tj. že $(LM) \Rightarrow (PM)$, neboli lokálně markovské rozdělení je vždy i párově markovské (k danému grafu). Obě vlastnosti ilustruje následující příklad.

Příklad 1.1 Graf G má podobu:



Snadno se přesvědčíme, že G je topologicky očíslovaný ADG. Párová markovská podmínka vyžaduje následující soubor podmíněných nezávislostí:

$$\begin{aligned} 3 &\perp\!\!\!\perp 2 \mid 1, \\ 4 &\perp\!\!\!\perp 1 \mid \{2, 3\}, \\ 4 &\perp\!\!\!\perp 3 \mid \{1, 2\}, \\ 5 &\perp\!\!\!\perp 1 \mid \{2, 3, 4\}, \\ 5 &\perp\!\!\!\perp 2 \mid \{1, 3, 4\}. \end{aligned}$$

Lokální markovská podmínka vyžaduje podmíněné nezávislosti

$$\begin{aligned} 3 &\perp\!\!\!\perp 2 \mid 1, \\ 4 &\perp\!\!\!\perp \{1, 3\} \mid 2, \\ 5 &\perp\!\!\!\perp \{1, 2\} \mid \{3, 4\}. \end{aligned}$$

Pro úplnost dodejme, že párovou i lokální vlastnost je možné zavést i jiným způsobem, bez využití topologického očíslování (jako je tomu např. v [Lau96]). Namísto „množin vrcholů s nižším očíslováním“, které figurují v definici 1.11 jako množiny tvaru $\{1, 2, \dots, i-1\}$, se pak využívá množiny nenásledníků (*non-descendants*), kterou jsme zavedli v definici 1.10. Označíme-li pak alternativy k podmínkám (*PM*), resp. (*LM*) jako (*PN*), resp. (*LN*) (podle „neočíslované“), můžeme příslušné požadavky na podmíněné nezávislosti zapsat stručně ve tvaru:

$$\begin{aligned} \forall i, j \in V \text{ takové, že } j \in \text{nd}(i) \text{ a } j \not\rightarrow i, \text{ platí: } & i \perp\!\!\!\perp j \mid \text{nd}(i) \setminus \{j\}, & (PN) \\ \forall i \in V \text{ platí: } & i \perp\!\!\!\perp \text{nd}(i) \setminus \text{pa}(i) \mid \text{pa}(i). & (LN) \end{aligned}$$

Porovnáme-li tento zápis s definicí 1.11, zjistíme, že jsou v zásadě jen nahrazeny množiny typu $\{1, 2, \dots, i-1\}$ množinou $\text{nd}(i)$. Pro vrchol $i \neq 1$ v topologicky očíslovaném grafu zřejmě platí, že $\{1, 2, \dots, i-1\} \subseteq \text{nd}(i)$, což spolu s vlastností *weak union* ihned dává výsledek $(LN) \Rightarrow (LM)$. Podle [Cow07, věta 5.14] ovšem platí, že jsou podmínky (*LN*) a (*LM*) dokonce ekvivalentní. Podmínkami (*PN*) a (*LN*) se již dále nebudeme zabývat.

Kromě párové a lokální markovské podmínky se často hovoří ještě o třetí, obecně vzato nejsilnější, markovské podmínce, označované jako *globální*. Globální markovská podmínka vzhledem k ADG se v literatuře definuje dvěma formálně značně odlišnými, přesto však ve výsledku ekvivalentními způsoby: buď pomocí tzv. *morálních grafů*, nebo pomocí tzv. *d-separace*. Zde se budeme držet druhého z těchto postupů, pojem *d-separace* zavedeme v následující definici.

Definice 1.12 Nechť $G = (V, E)$ je orientovaný graf a $\sigma = (\alpha_0, \alpha_1, \dots, \alpha_n)$ je sled v G . Vrchol α_i nazveme *kolizním* ve sledu σ , pokud $1 \leq i < n$ a platí $\alpha_{i-1} \rightarrow \alpha_i$ a zároveň $\alpha_i \leftarrow \alpha_{i+1}$; v opačném případě je α *nekolizní* v σ . Řekneme, že množina $K \subset V$ *blokuje* sled σ , pokud K neobsahuje některý z kolizních vrcholů v σ nebo obsahuje některý nekolizní vrchol ze σ . Jsou-li I, J, K tři disjunktí podmnožiny V , řekneme, že množina K *d-separuje* množiny I a J , jestliže K blokuje každý sled $(\alpha_0, \alpha_1, \dots, \alpha_n)$ takový, že $\alpha_0 \in I$ a $\alpha_n \in J$.

Příklad 1.2 Uvažujme orientovaný graf



a sled $(\alpha, \beta, \gamma, \delta)$, neboli sled $\alpha \rightarrow \beta \rightarrow \gamma \leftarrow \delta$. V tomto případě je jediným kolizním vrcholem γ ; kolizní vrcholy jsou charakterizovány zápisem typu $\rightarrow \gamma \leftarrow$. V anglicky psané literatuře se kolizní vrcholy nazývají různě, např. *colliders* nebo *head-to-head vertices*; oba tyto názvy pramení z toho, že se v těchto vrcholech „srážejí“ vstupující hrany.

Sled $\alpha \rightarrow \beta \rightarrow \gamma \leftarrow \delta$ je blokován např. prázdnou množinou, neboť ta neobsahuje kolizní vrchol γ ; snadno dále ověříme, že tento sled není blokován množinou $\{\gamma\}$ a je blokován množinou $\{\beta, \gamma\}$. Jelikož každý sled mezi α a δ vede přes β a δ , můžeme říci, že jak prázdná množina, tak množina $\{\beta, \gamma\}$ *d-separují* vrcholy α a δ ; *d-separaci* dvou vrcholů přitom rozumíme *d-separaci* příslušných jednoprvkových množin, v tomto případě $\{\alpha\}, \{\delta\}$.

Můžeme nyní přistoupit k definici globální markovské podmínky. Abychom zachovali stejný rámec jako pro párovou a lokální markovskou podmínku, definujeme podmínku globální opět pro topologicky očíslované ADG. Jak bude ale z definice patrné, topologické očíslování nehraje v tomto případě v podstatě žádnou roli.

Definice 1.13 Necht $G = (V, E)$ je topologicky očíslovaný acyklický orientovaný graf s n vrcholy a $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ je náhodný vektor s rozdělením \mathcal{L} , pro které existuje hustota vzhledem k λ_n . Řekneme, že \mathcal{L} splňuje *globální markovskou podmínku* (GM) vzhledem ke G , pokud pro libovolnou trojici $\langle I, J|K \rangle \in \mathcal{T}(n)$ platí

$$K \text{ d-separuje } I \text{ a } J \text{ v grafu } G \implies I \perp\!\!\!\perp J|K.$$

Lze snadno ukázat, že požadavky na podmíněné nezávislosti v (GM) nutně obsahují i podmíněné nezávislosti, které vyžaduje (LM); stačí si rozmyslet, že pro vrcholy $i, j \in V$, $j \in \{1, 2, \dots, i-1\} \setminus \text{pa}(i)$ platí, že libovolný sled mezi i a j je blokován množinou $\text{pa}(i)$. Odtud hned vidíme, že platí implikace

$$(GM) \implies (LM) \implies (PM),$$

neboli že (GM) je nejsilnější z uvedených podmínek. Lze ovšem ukázat, že je ve skutečnosti (GM) ekvivalentní s (LM), viz [Cow07, věta 5.14], můžeme tedy psát

$$(GM) \iff (LM) \implies (PM).$$

Význam globální markovské podmínky spočívá m.j. v tom, že podává v jistém slova smyslu vyčerpávající seznam nezávislostních požadavků. Lze totiž ukázat (viz [Cow07], str. 73.), že pro každý topologicky očíslovaný ADG G s n vrcholy lze najít rozdělení \mathcal{L} s hustotou vzhledem k λ_n takové, že struktura podmíněné nezávislosti $\mathcal{S}(\mathcal{L})$ obsahuje trojici $\langle I, J|K \rangle$ právě tehdy, když K d-separuje I a J v G .

Vztah mezi (GM), (LM) a (PM) se dále zjednoduší, pokud si můžeme dovolit předpokládat, že hustota zkoumaného rozdělení je spojitá a striktně kladná, jak ukazuje následující věta.

Věta 1.7 Necht $G = (V, E)$ je topologicky očíslovaný acyklický orientovaný graf s n vrcholy a $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ je náhodný vektor s rozdělením \mathcal{L} , pro něž existuje hustota vzhledem k λ_n , která je spojitá a striktně kladná všude na \mathbb{R}_n . Potom platí

$$(GM) \iff (LM) \iff (PM).$$

Důkaz. Jde o přímý důsledek věty 3.34 z [Lau96]; tato věta je formulována pro obecnější rámec tzv. *řetězcových grafů*, které jsme zde nezaváděli, ale jejichž speciálním případem jsou právě ADG. \square

Tento výsledek bude podstatný při studiu lineárních rekursivních systémů v následující kapitole.

Kapitola 2

Lineární rekursivní systémy

Definice 2.1 Mějme náhodný vektor $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n)^\top$, jehož složky \mathcal{E}_i tvoří navzájem nezávislé spojité náhodné veličiny s nenulovým rozptylem, a matici $\mathbf{A}_{n \times n} = (a_{ij})$ takovou, že $a_{ij} = 0$, kdykoli $j \geq i$. Potom náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$, jehož složky vyhovují vztahům

$$\begin{aligned} X_1 &= \mathcal{E}_1, \\ X_2 &= a_{21}X_1 + \mathcal{E}_2, \\ &\vdots \\ X_i &= a_{i1}X_1 + a_{i2}X_2 + \dots + a_{i,i-1}X_{i-1} + \mathcal{E}_i, \\ &\vdots \\ X_n &= a_{n1}X_1 + a_{n2}X_2 + \dots + a_{n,n-1}X_{n-1} + \mathcal{E}_n, \end{aligned} \tag{2.1}$$

nazveme *zobecněným lineárním rekursivním systémem s maticí koeficientů \mathbf{A} a chybovým vektorem \mathcal{E}* . Pokud mají navíc všechny složky vektoru \mathcal{E} nedegenerované normální rozdělení, tj. platí-li

$$\mathcal{E}_i \sim \mathcal{N}(0, \gamma_i^2), \quad \gamma_i^2 > 0, \quad i = 1, 2, \dots, n,$$

můžeme souhrnně zapsat $\mathcal{E} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$, kde $\mathbf{\Gamma} = \text{diag}(\gamma_1^2, \gamma_2^2, \dots, \gamma_n^2)$ (viz příp. lemma B.1), a hovoříme potom o *lineárním rekursivním systému s maticí koeficientů \mathbf{A} a chybovou varianční maticí $\mathbf{\Gamma}$* . Jelikož název „lineární rekursivní systém“, resp. „zobecněný lineární rekursivní systém“ je dosti těžkopádný, budeme jej místo zkracovat jako LRS, resp. ZLRS.

Vzhledem k tomu, že matice \mathbf{A} má podle definice tvar

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{pmatrix},$$

můžeme stručně zapsat soustavu (2.1) ve tvaru

$$\mathbf{X} = \mathbf{A}\mathbf{X} + \mathcal{E}. \tag{2.2}$$

Soustava rovnic (2.1) je speciálním případem tzv. *modelů strukturálních rovnic (structural equation models, viz [Gol73])*, přičemž rovnice mají v tomto případě *rekursivní* tvar; jak naznačuje [Stu02], slovo „rekursivní“ zde odráží skutečnost, že při známých hodnotách realizace

náhodného vektoru \mathcal{E} můžeme vypočítat hodnoty realizace \mathbf{X} postupným dosazováním od první rovnice v (2.1). Rekursivní modely tohoto typu našly praktické uplatnění v nejrůznějších vědních disciplínách: např. v genetice (viz [Li75]), psychologii ([Hod89]) nebo v sociologii a ekonomii ([Gol73]); podrobnější shrnutí aplikačních oblastí uvádí [And99].

Zpravidla se v takových systémech vyžaduje normalita chybového vektoru (viz např. [Wer80], [Stu02]). Této konvence jsme se drželi i zde, definice LRS zahrnuje předpoklad normality; stejné pojetí lineárních rekursivních systémů obsahovalo i původní zadání této bakalářské práce. Není ovšem zcela bez zajímavosti, že řada výsledků ohledně struktury podmíněných nezávislostí v LRS zůstává v platnosti i bez předpokladu normality chybového vektoru. Z toho důvodu jsme výše definovali ještě ZLRS, které zahrnují LRS jako speciální případ.

Nutno přiznat, že z jazykového hlediska je takový postup poněkud neobratný – nebýt zmíněných konvencí, bylo by z hlediska užití patrně vhodnější zachovat termín „lineární rekursivní systém“ pro obecnější model a dodatečný předpoklad normality upřesnit vhodným přívlaskem. Např. samotný název kapitoly je nyní poněkud zavádějící, neboť jsou v této kapitole obsaženy LRS i ZLRS. Případný název „Zobecněné lineární rekursivní systémy“ by ovšem zase evokoval představu, že je zde kladen důraz právě na zmíněné zobecnění; přitom těžiště této kapitoly (a její prvotní cíl) spočívá ve zkoumání LRS.

Hlavní náplní této kapitoly bude studium souvislostí mezi maticí koeficientů LRS (příp. ZLRS) a jeho strukturou podmíněných nezávislostí. Začneme přitom s případem LRS, kde je díky normalitě chybového vektoru situace jednodušší; následně některé výsledky zobecníme na ZLRS. Na základě získaných vztahů navrhne grafický model pro LRS a ZLRS.

2.1 Podmíněná nezávislost v lineárních rekursivních systémech

V tomto oddílu se budeme zabývat pouze lineárními rekursivními systémy, tj. situací, kdy chybový vektor \mathcal{E} má regulární normální rozdělení. Nejprve ukážeme, že každý lineární rekursivní systém je náhodný vektor s regulárním normálním rozdělením.

Lemma 2.1 *Mějme lineární rekursivní systém \mathbf{X} s maticí koeficientů $\mathbf{A}_{n \times n}$ a varianční maticí $\mathbf{\Gamma}$. Potom \mathbf{X} má regulární normální rozdělení s nulovým vektorem středních hodnot.*

Důkaz. Z definice lineárního rekursivního systému a vztahu (2.2) ihned dostaneme $(\mathbf{I} - \mathbf{A})\mathbf{X} = \mathcal{E}$, kde \mathcal{E} je chybový vektor s regulárním rozdělením $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$. Z lemmatu A.1 je dále patrné, že $\det(\mathbf{I} - \mathbf{A}) = 1$, tedy matice $\mathbf{I} - \mathbf{A}$ je regulární. Můžeme proto vyjádřit $\mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1}\mathcal{E}$. Označíme-li $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$, má \mathbf{X} podle lemmatu B.2 n -rozměrné normální rozdělení s nulovou střední hodnotou a varianční maticí $\mathbf{B}\mathbf{\Gamma}\mathbf{B}^\top$, která je díky regularitě \mathbf{B} a $\mathbf{\Gamma}$ zřejmě regulární. \square

Díky předchozímu lemmatu se při zkoumání závislostí mezi složkami LRS můžeme opřít o znalost vlastností vícerozměrného normálního rozdělení, které jsou shrnuty v příloze B. Podotkněme, že pro zpřehlednění následující diskuze budeme vždy uvažovat pouze spojitě verze hustot vícerozměrného normálního rozdělení (viz též příloha B) a podmíněné hustoty budou vždy ve tvaru (1.3), tj. budou rovny podílu příslušných sdružených a marginálních hustot (a tím pádem rovněž spojitě). Veškeré hustoty i podmíněné hustoty tedy budou určeny zcela jednoznačně. V takovém případě se, neformálně řečeno, mění platnost řady tvrzení z kapitoly 1 ze „skoro jistě“ na „všude“. Této skutečnosti budeme automaticky využívat a dodatek „s.j.“ bude na příslušných místech vynechán.

Jednou z příjemných vlastností normálně rozdělených vektorů je skutečnost, že veškerá informace ohledně závislostní struktury mezi složkami těchto vektorů je obsažena v jejich varianční matici. Podle lematu B.4 např. platí, že dvě složky normálně rozděleného vektoru jsou nezávislé právě tehdy, když je jejich kovariance nulová. Jak níže ukážeme, podobný vztah platí i o podmíněných nezávislostech a kovariancích.

Formulujme toto tvrzení poněkud přesněji. Budeme uvažovat náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s regulárním normálním rozdělením $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, čísla $i, j \in \{1, 2, \dots, n\}$ a indexovou množinu $K \subseteq \{1, 2, \dots, n\} \setminus \{i, j\}$. Ukážeme, že platí

$$X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_K \iff \text{cov}(X_i, X_j \mid \mathbf{X}_K) = 0.$$

Podstatná je přitom především implikace zprava doleva, neboť v opačném směru nám platnost i pro jiná než normální rozdělení zaručuje věta 1.5. Při zkoumání implikace zprava doleva hraje klíčovou roli věta B.5, podle které odpovídá podmíněná hustota X_i a X_j při dané hodnotě $\mathbf{X}_K = \mathbf{x}_K$ hustotě regulárního normálního rozdělení se střední hodnotou

$$\boldsymbol{\mu}_{ij} + \boldsymbol{\Sigma}_{ij \cdot K} (\boldsymbol{\Sigma}_{K \cdot K})^{-1} (\mathbf{x}_K - \boldsymbol{\mu}_K)$$

a varianční maticí

$$\boldsymbol{\Sigma}_{ij \cdot ij} - \boldsymbol{\Sigma}_{ij \cdot K} (\boldsymbol{\Sigma}_{K \cdot K})^{-1} \boldsymbol{\Sigma}_{K \cdot ij}.$$

Podstatným rysem tohoto vztahu je fakt, že uvedená varianční matice nezávisí na hodnotě \mathbf{x}_K podmiňující veličiny \mathbf{X}_K . Náhodná veličina $\text{cov}(X_i, X_j \mid \mathbf{X}_K)$ je tedy v tomto případě konstantní. Podrobněji tento vztah popisuje následující lemma.

Lemma 2.2 *Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s regulárním normálním rozdělením $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, indexy $i, j \in \{1, 2, \dots, n\}$ a indexovou množinu $K \subseteq \{1, 2, \dots, n\} \setminus \{i, j\}$. Označme symbolem $\boldsymbol{\Sigma}_{ij|K}$ matici danou vztahem*

$$\boldsymbol{\Sigma}_{ij|K} = \boldsymbol{\Sigma}_{ij \cdot ij} - \boldsymbol{\Sigma}_{ij \cdot K} (\boldsymbol{\Sigma}_{K \cdot K})^{-1} \boldsymbol{\Sigma}_{K \cdot ij}, \quad (2.3)$$

prvky této matice označme následujícím způsobem:

$$\boldsymbol{\Sigma}_{ij|K} = \begin{pmatrix} s_{ii} & s_{ij} \\ s_{ji} & s_{jj} \end{pmatrix}.$$

Potom platí

- (i) $\text{cov}(X_i, X_j \mid \mathbf{X}_K) = s_{ij}$.
- (ii) $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_K \iff s_{ij} = 0$.

Důkaz. Mějme libovolné $\mathbf{x}_K \in \mathbb{R}_{|K|}$. Podle věty B.5 má podmíněná hustota \mathbf{X}_{ij} při $\mathbf{X}_K = \mathbf{x}_K$ tvar hustoty normálního rozdělení $\mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{ij|K})$, kde

$$\boldsymbol{\nu} = \boldsymbol{\mu}_{ij} + \boldsymbol{\Sigma}_{ij \cdot K} (\boldsymbol{\Sigma}_{K \cdot K})^{-1} (\mathbf{x}_K - \boldsymbol{\mu}_K), \quad (2.4)$$

a jednorozměrné podmíněné hustoty X_i , resp. X_j při $\mathbf{X}_K = \mathbf{x}_K$ odpovídají marginálním hustotám normálního vektoru s tímto rozdělením. Všimněme si, že podmíněné rozptyly a kovariance se vypočítají zcela analogicky jako jejich nepodmíněné protějšky s tím, že užíváme příslušných podmíněných hustot namísto hustot nepodmíněných. Uvažujme náhodný vektor $\mathbf{Y} = (Y_i, Y_j)^\top$ s rozdělením $\mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{ij|K})$; potom z předchozího platí, že

$$\text{cov}(X_i, X_j \mid \mathbf{X}_K = \mathbf{x}_K) = \text{cov}(Y_i, Y_j) = s_{ij},$$

přičemž poslední rovnost plyne z lemmatu B.2. Tím jsme dokázali bod (i).

Bod (ii) se dokáže podobně. Označme podmíněnou hustotu \mathbf{X}_{ij} při $\mathbf{X}_K = \mathbf{x}_K$ jako $f_{ij|K}$, analogicky definujeme i podmíněné hustoty $f_{i|K}$ a $f_{j|K}$. Potom $f_{ij|K}$ je zároveň sdružená hustota Y_i a Y_j a příslušné marginální hustoty jsou $f_{i|K}$ a $f_{j|K}$. Z lemmat B.2 a B.4 plyne:

$$s_{ij} = 0 \iff Y_i \perp\!\!\!\perp Y_j \iff f_{ij|K} = f_{i|K} \cdot f_{j|K}.$$

Jelikož tento vztah platí pro libovolné $\mathbf{x}_K \in \mathbb{R}_{|K|}$, plyne bod (ii) okamžitě z (1.11). \square

Předchozí lemma dává návod, jak vyčíst případnou přítomnost podmíněných nezávislostí v gaussovském vektoru z jeho varianční matice. Tímto způsobem lze zkoumat samozřejmě i lineární rekursivní systémy (které mají podle lemmatu 2.1 regulární normální rozdělení). V případě LRS by to ovšem bylo zbytečně těžkopádné – o některých podmíněných nezávislostech lze totiž rozhodnout na první pohled z matice koeficientů. Jak postupně ukážeme, pro LRS \mathbf{X} s maticí koeficientů $\mathbf{A}_{n \times n} = (a_{ij})$ platí následující tvrzení: máme-li $1 \leq j < i \leq n$ a $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$, potom platí

$$a_{ij} = 0 \iff X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}. \quad (2.5)$$

Ve zbytku této kapitoly se budeme zabývat výhradně podmíněnými nezávislostmi tohoto typu. Budeme-li zkoumat podmíněnou nezávislost složek X_i a X_j pro $j < i$, budeme za podmíňující veličinu brát složky vektoru \mathbf{X} s indexy v množině $\{1, 2, \dots, i-1\} \setminus \{j\}$, kterou budeme v takovém případě vždy označovat symbolem \bar{K} .

Všimněme si, že pokud zkoumáme podmíněné nezávislosti typu $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$, zajímáme se pouze o prvních i složek vektoru \mathbf{X} . Označíme-li $L = \{1, 2, \dots, i\}$, můžeme říci, že přítomnost podmíněné nezávislosti z (2.5) je dána marginálním rozdělením vektoru \mathbf{X}_L . Přitom vektor \mathbf{X}_L zřejmě představuje LRS s maticí $\mathbf{A}_{L,L}$. Jinými slovy, marginální rozdělení vektoru \mathbf{X}_L odpovídá rozdělení LRS s maticí $\mathbf{A}_{L,L}$. (Podrobnější důkaz tohoto tvrzení pro obecnější případ ZLRS je proveden ve větě 2.10.) V následujících lemmatech budeme proto rovnou předpokládat, že $i = n$, což zjednoduší značení.

Lemma 2.3 *Uvažujme LRS \mathbf{X} s maticí koeficientů $\mathbf{A}_{i \times i}$ a chybovým vektorem \mathcal{E} . Mějme dále $1 \leq j < i$ a neprázdnou indexovou množinu $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$.*

$$\text{cov}(\mathcal{E}_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0.$$

Důkaz. Nejprve si uvědomme, že chybová složka \mathcal{E}_i je nezávislá s prvními $i-1$ složkami vektoru \mathbf{X} (neboli $\mathcal{E}_i \perp\!\!\!\perp \mathbf{X}_{j\bar{K}}$), neboť $\mathbf{X}_{j\bar{K}}$ je funkcí $\mathcal{E}_{j\bar{K}}$ a z definice LRS je $\mathcal{E}_i \perp\!\!\!\perp \mathcal{E}_{j\bar{K}}$. Označme $\mathbf{Y} = (X_1, X_2, \dots, X_{i-1}, \mathcal{E}_i)^\top$, $\mathbf{\Lambda} = \text{var } \mathbf{Y}$, $\gamma_i^2 = \text{var } \mathcal{E}_i$. Potom z uvedených nezávislostí plyne, že matice $\mathbf{\Lambda}$ má tvar

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_{j\bar{K} \cdot j\bar{K}} & \mathbf{0}_{i-1 \times 1} \\ \mathbf{0}_{1 \times i-1} & \gamma_i^2 \end{pmatrix}$$

a zřejmě platí $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$. Z lemmatu 2.2 a tvaru matice $\mathbf{\Lambda}$ dostaneme, že

$$\begin{aligned} \text{cov}(\mathcal{E}_i, X_j \mid \mathbf{X}_{\bar{K}}) &= \mathbf{\Lambda}_{i \cdot j} - \mathbf{\Lambda}_{i \cdot \bar{K}} (\mathbf{\Lambda}_{\bar{K} \cdot \bar{K}})^{-1} \mathbf{\Lambda}_{\bar{K} \cdot j} = \\ &= 0 - \mathbf{0}_{i-2 \times 1} (\mathbf{\Lambda}_{\bar{K} \cdot \bar{K}})^{-1} \mathbf{\Lambda}_{\bar{K} \cdot j} = 0, \end{aligned}$$

což jsme měli dokázat. \square

Lemma 2.4 *Mějme LRS \mathbf{X} s maticí koeficientů $\mathbf{A}_{i \times i}$ a chybovým vektorem \mathcal{E} . Mějme dále $1 \leq j < i$, a neprázdnou indexovou množinu $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$. Potom platí*

$$\text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = a_{ij} \text{var}(X_j | \mathbf{X}_{\bar{K}}).$$

Důkaz. Označme výraz $[X_i - \mathbf{E}(X_i | \mathbf{X}_{\bar{K}})] \cdot [X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})]$ symbolem V . Platí tedy $\text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = \mathbf{E}(V | \mathbf{X}_{\bar{K}})$. Výraz V nyní postupně upravíme. Nejprve vyjádříme:

$$X_i = \mathbf{A}_{i, \bar{K}} \mathbf{X}_{\bar{K}} + a_{ij} X_j + \mathcal{E}_i,$$

odtud za použití bodů (iii) a (iv) z věty 1.4 dostaneme, že

$$\begin{aligned} \mathbf{E}(X_i | \mathbf{X}_{\bar{K}}) &\stackrel{1.4(iii)}{=} \sum_{k \in \bar{K}} a_{ik} \mathbf{E}(X_k | \mathbf{X}_{\bar{K}}) + a_{ij} \mathbf{E}(X_j | \mathbf{X}_{\bar{K}}) + \mathbf{E}(\mathcal{E}_i | \mathbf{X}_{\bar{K}}) = \\ &\stackrel{1.4(iv)}{=} \sum_{k \in \bar{K}} a_{ik} X_k + a_{ij} \mathbf{E}(X_j | \mathbf{X}_{\bar{K}}) + \mathbf{E}(\mathcal{E}_i | \mathbf{X}_{\bar{K}}) = \\ &= \mathbf{A}_{i, \bar{K}} \mathbf{X}_{\bar{K}} + a_{ij} \mathbf{E}(X_j | \mathbf{X}_{\bar{K}}) + \mathbf{E}(\mathcal{E}_i | \mathbf{X}_{\bar{K}}). \end{aligned}$$

Výraz V můžeme odtud upravit následujícím způsobem:

$$\begin{aligned} V &= [X_i - \mathbf{E}(X_i | \mathbf{X}_{\bar{K}})] \cdot [X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})] = \\ &= \{a_{ij}[X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})] + \mathcal{E}_i - \mathbf{E}(\mathcal{E}_i | \mathbf{X}_{\bar{K}})\} \cdot \{X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})\} = \\ &= a_{ij}[X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})]^2 + [\mathcal{E}_i - \mathbf{E}(\mathcal{E}_i | \mathbf{X}_{\bar{K}})] \cdot [X_j - \mathbf{E}(X_j | \mathbf{X}_{\bar{K}})]. \end{aligned}$$

Podmíněnou kovarianci X_i a X_j při daném $\mathbf{X}_{\bar{K}}$ můžeme nyní vyjádřit jako

$$\text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = \mathbf{E}(V | \mathbf{X}_{\bar{K}}) = a_{ij} \text{var}(X_j | \mathbf{X}_{\bar{K}}) + \text{cov}(\mathcal{E}_i, X_j | \mathbf{X}_{\bar{K}}),$$

podle lemmatu 2.3 je $\text{cov}(\mathcal{E}_i, X_j | \mathbf{X}_{\bar{K}}) = 0$, a důkaz je tedy hotov. \square

Z předchozího lemmatu je ihned vidět, že pro zkoumaný LRS platí implikace

$$a_{ij} = 0 \implies \text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = 0,$$

která spolu s lemmatem 2.2 již dává do souvislosti podmíněné nezávislosti a koeficienty v LRS (připomeňme, že právě tuto souvislost nyní zkoumáme). Pomocí následujícího lemmatu ukážeme, že platí i implikace opačná.

Lemma 2.5 *Mějme vektor $\mathbf{X} = (X_1, X_2, \dots, X_{i-1})^\top$ mající regulární rozdělení $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Nechť dále $1 \leq j \leq i-1$ a $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$. Potom $\text{var}(X_j | \mathbf{X}_{\bar{K}}) > 0$.*

Důkaz. Podle věty o podmíněném rozdělení gaussovského vektoru B.5 je podmíněný rozptyl $\text{var}(X_j | \mathbf{X}_{\bar{K}} = \mathbf{x}_{\bar{K}})$ roven výrazu

$$\sigma_{jj} - \boldsymbol{\Sigma}_{j, \bar{K}} (\boldsymbol{\Sigma}_{\bar{K}, \bar{K}})^{-1} \boldsymbol{\Sigma}_{\bar{K}, j}.$$

který nezávisí na hodnotě $\mathbf{x}_{\bar{K}}$ podmiňujícího vektoru $\mathbf{X}_{\bar{K}}$. Stačí tedy ukázat, že tento výraz je různý od nuly. Předně si uvědomme, že matice $\boldsymbol{\Sigma}_{jK, jK}$ a $\boldsymbol{\Sigma}_{K, K}$ jsou podle lemmatu A.7 pozitivně definitní, tedy nutně i regulární (viz příp. lemma A.3). Podle věty o determinantu pozitivně definitní matice A.8 platí

$$\det \boldsymbol{\Sigma}_{j\bar{K}, j\bar{K}} = \det \boldsymbol{\Sigma}_{\bar{K}, \bar{K}} \cdot \det(\sigma_{jj} - \boldsymbol{\Sigma}_{j, \bar{K}} (\boldsymbol{\Sigma}_{\bar{K}, \bar{K}})^{-1} \boldsymbol{\Sigma}_{\bar{K}, j}).$$

Vzhledem ke zmíněné regularitě matic platí $\det \boldsymbol{\Sigma}_{j\bar{K}, j\bar{K}} \neq 0 \neq \det \boldsymbol{\Sigma}_{\bar{K}, \bar{K}}$, tedy nutně též $\sigma_{jj} - \boldsymbol{\Sigma}_{j, \bar{K}} (\boldsymbol{\Sigma}_{\bar{K}, \bar{K}})^{-1} \boldsymbol{\Sigma}_{\bar{K}, j} \neq 0$. \square

Nyní můžeme konečně formulovat tvrzení o vztahu koeficientů LRS a jeho podmíněných nezávislostech, které jsme naznačili v (2.5), v podobě následující věty.

Věta 2.6 (O podmíněných nezávislostech v LRS.) *Nechť $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ je LRS s maticí koeficientů \mathbf{A} . Mějme dále přirozená čísla i a j splňující $1 \leq j < i \leq n$ a neprázdnou indexovou množinu $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$. Potom platí:*

$$a_{ij} = 0 \iff X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}.$$

Důkaz. Všechna potřebná tvrzení pro důkaz této věty jsou obsažená v předchozích lemmatech. Podle lemmatu 2.2 je

$$X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}} \iff \text{cov}(X_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0$$

a z lemmat 2.4 a 2.5 bezprostředně plyne, že

$$\text{cov}(X_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0 \iff a_{ij} = 0,$$

čímž je věta dokázána. \square

Dodejme, že použijeme-li konvenci zmíněnou v oddílu 1.2, podle které při $K = \emptyset$ ztotožníme výrazy $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$ a $X_i \perp\!\!\!\perp X_j$, můžeme rozšířit tvrzení předchozí věty i na případ, kdy $K = \emptyset$. Tento případ nastane jedině při $j = 1$ a $i = 2$, a věta potom říká, že

$$a_{21} = 0 \iff X_2 \perp\!\!\!\perp X_1,$$

což plyne ihned z normality LRS (viz příp. lemma B.4).

Z předchozí věty je patrné, že lineární rekursivní systémy, které mají nuly na stejných pozicích v matici koeficientů, vykazují podobnou strukturu podmíněných nezávislostí; takové LRS zahrneme v následující definici do stejné třídy. Pro popis postavení nul v matici koeficientů budeme používat acyklických orientovaných grafů, jejichž vrcholy odpovídají jednotlivým složkám LRS.

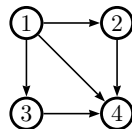
Definice 2.2 Uvažujme topologicky očíslovaný acyklický orientovaný graf $G = (V, E)$ s n vrcholy. Řekneme, že LRS \mathbf{X} s maticí koeficientů $\mathbf{A}_{n \times n} = (a_{ij})$ *vyhovuje grafu G* , pokud platí:

$$(j, i) \notin E \implies a_{ij} = 0, \quad i, j = 1, 2, \dots, n,$$

neboli koeficient a_{ij} je nulový, pokud graf G neobsahuje hranu $j \rightarrow i$. (Podotkněme, že pro $j \geq i$ platí vždy $a_{ij} = 0$ z definice LRS.) Skutečnost, že \mathbf{X} vyhovuje grafu G , budeme značit zápisem $\mathbf{X} \angle G$. Třídou lineárních rekursivních systémů omezených grafem G budeme potom rozumět systém pravděpodobnostních rozdělení $\mathcal{R}(G)$ daný předpisem

$$\mathcal{R}(G) = \{\mathcal{L} \in \mathcal{N}_{\text{reg}} : \exists \text{ LRS } \mathbf{X} \text{ takový, že } \mathbf{X} \angle G \text{ a } \mathbf{X} \sim \mathcal{L}\}.$$

Příklad 2.1 Mějme graf G ve tvaru



a LRS \mathbf{X} daný soustavou

$$\begin{aligned} X_1 &= \mathcal{E}_1, \\ X_2 &= X_1 + \mathcal{E}_2, \\ X_3 &= X_1 + \mathcal{E}_3, \\ X_4 &= X_1 + X_3 + \mathcal{E}_4. \end{aligned}$$

Můžeme se snadno přesvědčit, že $\mathbf{X} \angle G$; v grafu G totiž vede příslušně orientovaná hrana mezi každým párem vrcholů, vyjma vrcholů 2 a 3. Má-li být $\mathbf{X} \angle G$, nesmí se v rovnici pro X_3 (tj. ve třetí rovnici) vyskytovat na pravé straně X_2 , což v tomto případě platí. Libovolné rozdělení z třídy $\mathcal{R}(G)$ je potom rozdělení nějakého LRS, jehož matici koeficientů \mathbf{A} lze zapsat ve tvaru

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{pmatrix}.$$

Písmeno \mathcal{R} v zavedeném značení napovídá, že se jedná o třídu omezených (*restricted*) lineárních rekursivních systémů. Je-li graf G úplným topologicky očíslovaným ADG, tj. platí-li $j \rightarrow i$, kdykoli $j < i$, potom zřejmě vyhovuje grafu G libovolný LRS příslušného rozměru; v takovém případě někdy $\mathcal{R}(G)$ nazýváme *třídou neomezených lineárních rekursivních systémů*.

Podobný princip restrikce koeficientů lineárních rekursivních systémů používá Nanny Wermuth v práci [Wer80]. Namísto pojmů omezené, resp. neomezené systémy používá označení úplné (*complete*), resp. neúplné (*incomplete*), místo omezení grafem G používá tzv. *zero pattern*, neboli vzor postavení nul v matici koeficientů. Pro naše účely se však ukazuje jako výhodné popisovat postavení nul právě pomocí grafů, neboť, jak dále ukážeme, vztah rozdělení z třídy $\mathcal{R}(G)$ a grafu G potom přesně odpovídá grafickým modelům, které jsme popsali v oddílu 1.4.

Podobně, jako jsme zavedli třídy $\mathcal{R}(G)$, můžeme zavést třídy regulárních gaussovských rozdělení, které jsou globálně markovské vzhledem k danému acyklickému orientovanému grafu.

Definice 2.3 Necht G je acyklický orientovaný graf. Pod označením $\mathcal{M}(G)$ rozumíme *třídou regulárních gaussovských rozdělení markovských vzhledem ke grafu G* danou předpisem

$$\mathcal{M}(G) = \{ \mathcal{L} \in \mathcal{N}_{\text{reg}} : \mathcal{L} \text{ splňuje } (GM) \text{ vzhledem ke grafu } G \}.$$

V zbylé části tohoto oddílu budeme dokazovat, že pro daný topologicky očíslovaný ADG G jsou třídy $\mathcal{R}(G)$ a $\mathcal{M}(G)$ totožné. Budeme k tomu m.j. potřebovat tvrzení, že každé regulární gaussovské rozdělení lze vyjádřit v podobě nějakého LRS. Toto tvrzení nyní vyslovíme samostatně jako následující větu.

Věta 2.7 Necht náhodný vektor \mathbf{X} má regulární rozdělení $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. Potom existují matice $\mathbf{A}_{n \times n}$ a $\mathbf{\Gamma}_{n \times n}$ takové, že \mathbf{X} je lineární rekursivní systém s maticí koeficientů \mathbf{A} a chybovou varianční maticí $\mathbf{\Gamma}$.

Důkaz. Podle věty A.6 o rozkladu pozitivně definitní matice existují dolní trojúhelníková matice s jednotkovou diagonálou \mathbf{B} a diagonální matice s kladnými prvky na diagonále $\mathbf{\Gamma}$ takové, že $\mathbf{\Sigma} = \mathbf{B}\mathbf{\Gamma}\mathbf{B}^\top$. Podle lemmatu A.1 je matice \mathbf{B} regulární a její inverzní matice \mathbf{B}^{-1} je dolní trojúhelníková s jednotkovými prvky na hlavní diagonále. Zavedeme-li matici \mathbf{A} předpisem $\mathbf{A} = \mathbf{I} - \mathbf{B}^{-1}$, je \mathbf{A}

dolní trojúhelníková s nulovou diagonálou, a může tedy vystupovat v roli matice koeficientů lineárního rekursivního systému. Označíme-li \mathbf{X} lineární rekursivní systém s maticí koeficientů \mathbf{A} a chybovou varianční maticí $\mathbf{\Gamma}$, můžeme analogicky jako v důkazu lemmatu 2.1 vyjádřit $\mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1} \mathcal{E} = \mathbf{B} \mathcal{E}$, odkud již zřejmě plyne, že $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. \square

Nyní můžeme přistoupit k větě o ekvivalenci $\mathcal{R}(G)$ a $\mathcal{M}(G)$, kterou jsme výše naznačili; hlavní myšlenky vedoucí k jejímu důkazu již máme k dispozici ve dvou předchozích větách, tj. ve větách 2.7 a 2.6.

Věta 2.8 (O ekvivalenci tříd $\mathcal{R}(G)$ a $\mathcal{M}(G)$.) *Nechť G je topologicky očíslovaný acyklický orientovaný graf. Potom*

$$\mathcal{R}(G) = \mathcal{M}(G).$$

Důkaz. Položme $n = |V|$, a mějme libovolné n -rozměrné rozdělení $\mathcal{L} \in \mathcal{N}_{\text{reg}}$. Nechť $\mathbf{X} \sim \mathcal{L}$, označme hustotu pravděpodobnosti náhodného vektoru \mathbf{X} vzhledem k λ_n jako f . Z tvaru hustoty vícerozměrného normálního rozdělení (B.3) plyne, že $f(\mathbf{x}) > 0$ pro všechna $\mathbf{x} \in \mathbb{R}_n$ (připomeňme, že jsme se omezili na spojitě verze hustot). Podle věty 1.7 je v takovém případě párová markovská podmínka (PM) ekvivalentní s globální markovskou podmínkou (GM). Proto třída $\mathcal{M}(G)$ je totožná se systémem distribucí

$$\{\mathcal{L} \in \mathcal{N}_{\text{reg}} : \mathcal{L} \text{ splňuje (PM) vzhledem ke grafu } G\}.$$

Volme nyní libovolně $\mathcal{L} \in \mathcal{R}(G)$. Nechť $\mathbf{X} \sim \mathcal{L}$. Potom z definice třídy $\mathcal{R}(G)$ lze \mathbf{X} zapsat jako lineární rekursivní systém s nějakou maticí koeficientů \mathbf{A} , který vyhovuje grafu G . Uvažujme i a j taková, že $1 \leq j < i \leq n$ a G neobsahuje hranu $j \rightarrow i$. Jelikož $\mathbf{X} \angle G$, je nutně $a_{ij} = 0$. Podle věty 2.6 je pak $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$, kde $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$. To ovšem znamená, že \mathcal{L} splňuje (PM) vzhledem ke grafu G , neboli $\mathcal{L} \in \mathcal{M}(G)$. Celkem odtud dostáváme, že $\mathcal{R}(G) \subseteq \mathcal{M}(G)$.

Nechť naopak $\mathcal{L} \in \mathcal{M}(G)$, $\mathbf{X} \sim \mathcal{L}$. Potom podle věty 2.7 lze \mathbf{X} opět zapsat jako lineární rekursivní systém s nějakou maticí koeficientů \mathbf{A} . Pokud nyní pro nějaké $1 \leq j < i \leq n$ neexistuje v grafu G hrana $j \rightarrow i$, z (PM) platí $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$. Užitím věty 2.6 dostaneme pak nutně $a_{ij} = 0$, tedy $\mathbf{X} \angle \mathbf{B}$ a $\mathcal{L} \in \mathcal{R}(G)$. Je tedy rovněž $\mathcal{R}(G) \supseteq \mathcal{M}(G)$ a věta je dokázána. \square

2.2 Podmíněná nezávislost v zobecněných lineárních rekursivních systémech

V tomto oddílu zobecníme některé poznatky z oddílu minulého, který se týkal lineárních rekursivních systémů (jež jsou speciálním případem ZLRS). Ačkoli jsme se při studiu LRS opírali o vlastnosti normálního rozdělení, lze k některým výsledkům, které popisují vztah mezi koeficienty LRS a jeho strukturou podmíněných nezávislostí, dospět i při obecnějších podmínkách. Konkrétně ukážeme, že i pro ZLRS platí obdoba věty 2.6 o podmíněných nezávislostech LRS, tj. že za jistých podmínek platí ekvivalence

$$a_{ij} = 0 \iff X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}.$$

Pro přehlednost rozdělíme tento vztah do dvou implikací,

$$a_{ij} = 0 \implies X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}, \quad (2.6)$$

$$X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}} \implies a_{ij} = 0, \quad (2.7)$$

a na důkaz každé z nich (včetně přípravných tvrzení) vyhradíme samostatný pododdíl.

2.2.1 Důkaz implikace (2.6)

V minulém oddílu jsme zkoumali vztah mezi koeficienty LRS a podmíněnými nezávislostmi typu $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$ výhradně prostřednictvím podmíněných kovariancí, jejichž vynulování bylo díky normalitě LRS nutnou a postačující podmínkou pro přítomnost podmíněné nezávislosti. Pro důkaz implikace (2.6) se ovšem v obecnějším případě ZLRS o podmíněnou kovarianci nelze opřít, její vynulování (skoro jistě) je zde pouze podmínkou nutnou, nikoli však postačující. Výskyt podmíněné nezávislosti budeme v tomto případě ověřovat přímo pomocí sdružených a marginálních hustot ZLRS.

V následujících tvrzeních budeme zpravidla vycházet z identických předpokladů a používat stejného značení. Abychom zjednodušili a zpřehlednili zápisy, tyto předpoklady nyní formulujeme a označíme je (α) , (β) .

(α) $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ je zobecněný lineární rekursivní systém s maticí koeficientů \mathbf{A} a chybovým vektorem $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n)^\top$. Připomeňme, že z definice ZLRS plyne, že složky vektoru \mathcal{E} jsou navzájem nezávislé spojité náhodné veličiny s nenulovým rozptylem.

(β) Složka \mathcal{E}_k vektoru \mathcal{E} má hustotu φ_k vzhledem k λ pro $k = 1, 2, \dots, n$.

Často budeme vybírat z n -složkových vektorů několik prvních složek. Zavedeme proto následující značení. Pro $k \geq 1$ označíme symbolem L_k indexovou množinu danou předpisem

$$L_k = \{1, 2, \dots, k\}$$

a symbol L_0 ztotožníme s prázdnou množinou. Máme-li např. vektor $\mathbf{v} = (v_1, v_2, \dots, v_n)^\top$, pak je vektor \mathbf{v}_{L_k} pro $1 \leq k \leq n$ tvořen prvními k složkami vektoru \mathbf{v} .

Je-li \mathbf{A} matice nějakého zobecněného lineárního rekursivního systému \mathbf{X} , můžeme pomocí právě zavedeného značení vyjádřit její k -tý řádek při $k \geq 2$ blokovým zápisem

$$(\mathbf{A}_{k \cdot L_{k-1}}, \mathbf{0}_{1 \times n-k+1}),$$

neboť v k -tém řádku může být podle definice 2.1 pouze prvních $k - 1$ složek nenulových. První řádek matice \mathbf{A} je vždy nulový. Výraz $\mathbf{A}\mathbf{X}$ z pravé strany (2.2) můžeme potom vyjádřit ve tvaru

$$\mathbf{A}\mathbf{X} = \begin{pmatrix} 0 \\ \mathbf{A}_{2 \cdot L_1} \mathbf{X}_{L_1} \\ \mathbf{A}_{3 \cdot L_2} \mathbf{X}_{L_2} \\ \vdots \\ \mathbf{A}_{n \cdot L_{n-1}} \mathbf{X}_{L_{n-1}} \end{pmatrix}.$$

Zápis některých tvrzení se zjednoduší, zavedeme-li ještě konvenci, že pro libovolný řádkový vektor \mathbf{b} a sloupcový vektor \mathbf{c} je „prázdný“ skalární součin nulový: $\mathbf{b}_\emptyset \mathbf{c}_\emptyset = \sum_{i \in \emptyset} b_i c_i = 0$.

Potom můžeme psát $\mathbf{A}_{1:L_0}\mathbf{X}_{L_0} = \mathbf{A}_{1:\emptyset}\mathbf{X}_{\emptyset} = 0$ a celý vektor $\mathbf{A}\mathbf{X}$ můžeme vyjádřit ve tvaru

$$\mathbf{A}\mathbf{X} = \begin{pmatrix} \mathbf{A}_{1:L_0}\mathbf{X}_{L_0} \\ \mathbf{A}_{2:L_1}\mathbf{X}_{L_1} \\ \vdots \\ \mathbf{A}_{n:L_{n-1}}\mathbf{X}_{L_{n-1}} \end{pmatrix}. \quad (2.8)$$

Vztah (2.8) bude užitečný v nadcházející diskuzi.

Věta 2.9 *Nechť jsou splněny předpoklady (α) , (β) . Potom existuje hustota pravděpodobnosti f vektoru \mathbf{X} vzhledem k λ_n a platí*

$$f(\mathbf{x}) = \prod_{k=1}^n \varphi_k(x_k - \mathbf{A}_{k:L_{k-1}}\mathbf{x}_{L_{k-1}}). \quad (2.9)$$

Důkaz. Nejprve si uvědomme, že vzhledem k nezávislosti složek \mathcal{E}_k můžeme hustotu φ vektoru \mathcal{E} vzhledem k míře λ_n vyjádřit jako

$$\varphi(\boldsymbol{\varepsilon}) = \prod_{k=1}^n \varphi_k(\varepsilon_k). \quad (2.10)$$

Vztah (2.2) z definice ZLRS můžeme přepsat jako $(\mathbf{I} - \mathbf{A})\mathbf{X} = \mathcal{E}$. Z lemmatu A.1 plyne, že $\det(\mathbf{I} - \mathbf{A}) = 1$, tedy $\mathbf{I} - \mathbf{A}$ je regulární. Můžeme tedy definovat zobrazení

$$\begin{aligned} g: \mathbb{R}_n &\rightarrow \mathbb{R}_n, \\ g: \mathbf{y} &\mapsto (\mathbf{I} - \mathbf{A})^{-1}\mathbf{y}. \end{aligned}$$

Zobrazení g je zřejmě regulární a prosté na \mathbb{R}_n a platí $\mathbf{X} = g(\mathcal{E})$. Můžeme tedy použít větu o transformaci náhodného vektoru [And07, věta 3.7] pro výpočet hledané hustoty f . V následujícím výpočtu označíme zobrazení inverzní ke g jako g^{-1} , jeho jakobián označíme $J_{g^{-1}}$ a absolutní hodnotu z tohoto jakobiánu jako $\text{abs}(J_{g^{-1}})$. Použijeme postupně větu o transformaci a rovnost $\det(\mathbf{I} - \mathbf{A}) = 1$ a dostaneme

$$\begin{aligned} f(\mathbf{x}) &= \varphi[g^{-1}(\mathbf{x})] \cdot \text{abs}(J_{g^{-1}}) = \\ &= \varphi[(\mathbf{I} - \mathbf{A})\mathbf{x}] \cdot \text{abs}[\det(\mathbf{I} - \mathbf{A})] = \\ &= \varphi[(\mathbf{I} - \mathbf{A})\mathbf{x}]. \end{aligned} \quad (2.11)$$

Výraz $(\mathbf{I} - \mathbf{A})\mathbf{x}$ je sloupcový vektor, jehož složky můžeme vzhledem k (2.8) rozepsat následovně:

$$(\mathbf{I} - \mathbf{A})\mathbf{x} = \begin{pmatrix} x_1 - \mathbf{A}_{1:L_0}\mathbf{x}_{L_0} \\ x_2 - \mathbf{A}_{2:L_1}\mathbf{x}_{L_1} \\ \vdots \\ x_n - \mathbf{A}_{n:L_{n-1}}\mathbf{x}_{L_{n-1}} \end{pmatrix}.$$

Můžeme potom pokračovat ve výpočtu (2.11) s tím, že nakonec užijeme vztah (2.10):

$$\begin{aligned} f(\mathbf{x}) &= \varphi \left[\begin{pmatrix} x_1 - \mathbf{A}_{1:L_0}\mathbf{x}_{L_0} \\ x_2 - \mathbf{A}_{2:L_1}\mathbf{x}_{L_1} \\ \vdots \\ x_n - \mathbf{A}_{n:L_{n-1}}\mathbf{x}_{L_{n-1}} \end{pmatrix} \right] = \\ &= \prod_{k=1}^n \varphi_k(x_k - \mathbf{A}_{k:L_{k-1}}\mathbf{x}_{L_{k-1}}). \quad \square \end{aligned}$$

V dalším textu budeme dodržovat konvenci, že f představuje vždy hustotu pravděpodobnosti ZLRS \mathbf{X} vzhledem k λ_n , a je-li $I \subset \{1, 2, \dots, n\}$, pak f_I představuje marginální hustotu vektoru X_I vzhledem k $\lambda_{|I|}$. V následující větě odvodíme marginální rozdělení pro prvních i složek zobecněného lineárního rekursivního systému.

Věta 2.10 *Nechť platí (α) , (β) a nechť $i \in \{1, 2, \dots, n\}$. Potom marginální hustota pravděpodobnosti f_{L_i} vektoru \mathbf{X}_{L_i} vzhledem k λ_i je rovna*

$$f_{L_i}(\mathbf{x}_{L_i}) = \prod_{k=1}^i \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}). \quad (2.12)$$

Důkaz. Připomeňme, že vektor \mathbf{X}_{L_i} má tvar $(X_1, X_2, \dots, X_i)^\top$, neboli prvních i složek vektoru \mathbf{X} . Prvních i složek vektoru \mathbf{X} je ovšem funkcí pouze prvních i složek vektoru \mathcal{E} . Přesněji řečeno, \mathbf{X}_{L_i} tvoří ZLRS s maticí koeficientů $\mathbf{A}_{L_i \cdot L_i}$ a chybovým vektorem \mathcal{E}_{L_i} . Tvrzení potom plyne přímo z věty 2.9.

Alternativně lze postupovat tak, že vypočítáme marginální hustotu f_{L_i} přímo ze sdružené hustoty (2.9). Platnost vztahu (2.12) lze pak asi nejpréhledněji ukázat úplnou zpětnou matematickou indukci podle i při libovolném n . Pro $i = n$ je tvrzení dáno přímo vztahem (2.9). Předpokládejme nyní, že tvrzení platí pro nějaké i , kde $2 \leq i \leq n$, a dokažme jej pro $i - 1$. Z věty o marginální hustotě [And07, věta 3.10] platí

$$\begin{aligned} f_{L_{i-1}}(\mathbf{x}_{L_{i-1}}) &= \int_{\mathbb{R}} f_{L_i}(\mathbf{x}_{L_i}) dx_i = \\ &= \int_{\mathbb{R}} \prod_{k=1}^i \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}) dx_i = \\ &= \prod_{k=1}^{i-1} \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}) \int_{\mathbb{R}} \varphi_i(x_i - \mathbf{A}_{i \cdot L_{i-1}} \mathbf{x}_{L_{i-1}}) dx_i = \\ &= \prod_{k=1}^{i-1} \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}). \end{aligned}$$

Poslední rovnost platí díky tomu, že φ_i je hustota pravděpodobnosti složky \mathcal{E}_i vzhledem k λ , tedy $\int_{\mathbb{R}} \varphi_i dx_i = 1$. \square

Sdružených a marginálních hustot odvozených v předchozích větách můžeme nyní využít k tomu, abychom konečně dokázali platnost (2.6).

Věta 2.11 *Nechť jsou splněny předpoklady (α) , (β) . Mějme dále přirozená čísla i a j taková, že $1 \leq j < i \leq n$, a neprázdnou indexovou množinu $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$. Potom platí:*

$$a_{ij} = 0 \implies X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}.$$

Důkaz. Budeme postupně při $a_{ij} = 0$ dokazovat rovnost $f_{ij\bar{K}} \cdot f_{\bar{K}} = f_{i\bar{K}} \cdot f_{j\bar{K}}$ z definice podmíněné nezávislosti, kde symboly $f_{\bar{K}}$, $f_{i\bar{K}}$, $f_{j\bar{K}}$ a $f_{ij\bar{K}}$ označují příslušné marginální hustoty. Vzhledem k tomu, že vektor $\mathbf{X}_{j\bar{K}}$ obsahuje prvních $i-1$ složek vektoru \mathbf{X} a vektor $\mathbf{X}_{ij\bar{K}}$ prvních i složek, vyjádříme hustoty $f_{j\bar{K}}$

a $f_{ij\bar{K}}$ podle věty 2.9 jako

$$f_{j\bar{K}}(\mathbf{x}_{j\bar{K}}) = \prod_{k=1}^{i-1} \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}),$$

$$f_{ij\bar{K}}(\mathbf{x}_{ij\bar{K}}) = \prod_{k=1}^i \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}),$$

přičemž hustotu $f_{ij\bar{K}}$ můžeme za platnosti $a_{ij} = 0$ dále upravit na

$$f_{ij\bar{K}}(\mathbf{x}_{ij\bar{K}}) = \varphi_i(x_i - \mathbf{A}_{i \cdot \bar{K}} \mathbf{x}_{\bar{K}}) \cdot \prod_{k=1}^{i-1} \varphi_k(x_k - \mathbf{A}_{k \cdot L_{k-1}} \mathbf{x}_{L_{k-1}}) =$$

$$= \varphi_i(x_i - \mathbf{A}_{i \cdot \bar{K}} \mathbf{x}_{\bar{K}}) \cdot f_{j\bar{K}}(\mathbf{x}_{j\bar{K}}).$$

Hustoty $f_{\bar{K}}$ a $f_{i\bar{K}}$ vyjádříme pomocí hustot $f_{j\bar{K}}$ a $f_{ij\bar{K}}$ následujícím způsobem:

$$f_{\bar{K}}(\mathbf{x}_{\bar{K}}) = \int_{\mathbb{R}} f_{j\bar{K}}(\mathbf{x}_{j\bar{K}}) dx_j,$$

$$f_{i\bar{K}}(\mathbf{x}_{i\bar{K}}) = \int_{\mathbb{R}} f_{ij\bar{K}}(\mathbf{x}_{ij\bar{K}}) dx_j =$$

$$= \varphi_i(x_i - \mathbf{A}_{i \cdot \bar{K}} \mathbf{x}_{\bar{K}}) \cdot \int_{\mathbb{R}} f_{j\bar{K}}(\mathbf{x}_{j\bar{K}}) dx_j =$$

$$= \varphi_i(x_i - \mathbf{A}_{i \cdot \bar{K}} \mathbf{x}_{\bar{K}}) \cdot f_{\bar{K}}(\mathbf{x}_{\bar{K}}).$$

Odtud již snadno získáme rovnost (pro přehlednost vynecháme některé zřejmé argumenty funkcí)

$$f_{ij\bar{K}} \cdot f_{\bar{K}} = \varphi_i(x_i - \mathbf{A}_{i \cdot \bar{K}} \mathbf{x}_{\bar{K}}) \cdot f_{j\bar{K}} \cdot f_{\bar{K}} = f_{i\bar{K}} \cdot f_{j\bar{K}},$$

což je přesně definiční vztah (1.8) pro podmíněnou nezávislost. \square

2.2.2 Důkaz implikace (2.7)

Pro důkaz implikace (2.7) využijeme podobného přístupu jako v případě LRS, a sice vztahu mezi koeficienty v matici \mathbf{A} a příslušnými podmíněnými kovariancemi. Podle věty 1.5 totiž platí

$$X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}} \implies \text{cov}(X_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0 \quad \text{s.j.},$$

stačí tedy pro (2.7) pouze ukázat, že

$$\text{cov}(X_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0 \quad \text{s.j.} \implies a_{ij} = 0.$$

Vzhledem k tomu, že budeme pracovat s podmíněnými kovariancemi, potřebujeme vesměs předpokládat, že složky vektoru \mathcal{E} mají konečné druhé momenty (snadno se ověří, že v takovém případě mají konečné druhé momenty i složky příslušného ZLRS). Pro stručnější zápis tvrzení v tomto pododdílu rozšíříme stávající předpoklady (α) , (β) o předpoklad (γ) a konvenci (δ) :

(γ) $E \mathcal{E}_k^2 < \infty$ pro $k = 1, 2, \dots, n$.

(δ) Uvažujeme pevně daná přirozená čísla i a j splňující $1 \leq j < i \leq n$ a neprázdnou indexovou množinu $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$.

Lemma 2.12 *Nechť platí $(\alpha), (\beta), (\gamma), (\delta)$. Potom existuje sdružená hustota náhodného vektoru $(\mathbf{X}_{j\bar{K}}, \mathcal{E}_i)^\top$ vzhledem k λ_i a je rovna $f_{j\bar{K}} \cdot \varphi_i$.*

Důkaz. Připomeňme, že $\mathbf{X}_{j\bar{K}} = (X_1, X_2, \dots, X_{i-1})^\top$. Tedy $\mathbf{X}_{j\bar{K}}$ závisí pouze na prvních $i - 1$ složkách vektoru \mathcal{E} . Jinak řečeno, $\mathbf{X}_{j\bar{K}}$ je funkcí vektoru $\mathcal{E}_{j\bar{K}}$. Z definice ZLRS jsou ale složky chybového vektoru \mathcal{E} navzájem nezávislé, čili $\mathcal{E}_i \perp\!\!\!\perp \mathcal{E}_{j\bar{K}}$, a tedy i $\mathcal{E}_i \perp\!\!\!\perp \mathbf{X}_{j\bar{K}}$. Je tudíž sdružená hustota $\mathbf{X}_{j\bar{K}}$ a \mathcal{E}_i rovna součinu jejich marginálních hustot, což jsme měli dokázat. \square

Lemma 2.13 *Nechť platí $(\alpha), (\beta), (\gamma), (\delta)$. Potom je*

$$\text{cov}(X_j, \mathcal{E}_i | \mathbf{X}_{\bar{K}}) = 0 \quad \text{s.j.}$$

Důkaz. Stačí dokázat, že $X_j \perp\!\!\!\perp \mathcal{E}_i | \mathbf{X}_{\bar{K}}$, neboť pak již podle věty 1.5 nutně platí $\text{cov}(\mathcal{E}_i, \mathcal{E}_j | \mathbf{X}_{\bar{K}}) = 0$ skoro jistě. Podle lemmatu 2.12 existuje sdružená hustota $g_{ij\bar{K}}$ vektoru $\mathbf{Y} = (\mathbf{X}_{j\bar{K}}, \mathcal{E}_i)^\top$ vzhledem k λ_i a platí

$$g_{ij\bar{K}} = \varphi_i \cdot f_{j\bar{K}}. \quad (2.13)$$

Označme $g_{i\bar{K}}$, $g_{j\bar{K}}$ a $g_{\bar{K}}$ marginální hustoty vektorů $\mathbf{Y}_{i\bar{K}}$, $\mathbf{Y}_{j\bar{K}}$ a $\mathbf{Y}_{\bar{K}}$. Je tedy zřejmě $g_{j\bar{K}} = f_{j\bar{K}}$ a $g_{\bar{K}} = f_{\bar{K}}$. Použitím věty o marginální hustotě na vztah (2.13) (nebo analogickým postupem jako v lemmatu 2.12) ihned dostaneme, že

$$g_{i\bar{K}} = \varphi_i \cdot g_{\bar{K}}. \quad (2.14)$$

Celkem máme

$$g_{ij\bar{K}} \cdot g_{\bar{K}} \stackrel{(2.13)}{=} \varphi_i \cdot g_{j\bar{K}} \cdot g_{\bar{K}} \stackrel{(2.14)}{=} g_{i\bar{K}} \cdot g_{j\bar{K}},$$

neboli $Y_j \perp\!\!\!\perp Y_i | \mathbf{Y}_{\bar{K}}$, tj. $X_j \perp\!\!\!\perp \mathcal{E}_i | \mathbf{X}_{\bar{K}}$. \square

Lemma 2.14 *Nechť platí $(\alpha), (\beta), (\gamma), (\delta)$. Potom je*

$$\text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = a_{ij} \text{var}(X_i | \mathbf{X}_{\bar{K}}) \quad \text{s.j.}$$

Důkaz. Zcela analogicky jako v důkazu lemmatu 2.4 dostaneme vztah pro podmíněnou kovarianci X_i a X_j při daném $\mathbf{X}_{\bar{K}}$ jako

$$\text{cov}(X_i, X_j | \mathbf{X}_{\bar{K}}) = a_{ij} \text{var}(X_j | \mathbf{X}_{\bar{K}}) + \text{cov}(X_j, \mathcal{E}_i | \mathbf{X}_{\bar{K}}) \quad \text{s.j.}$$

(Podotkněme, že normalita chybového vektoru v LRS nebyla pro tuto úpravu v lemmatu 2.4 vůbec potřeba.) Podle lemmatu 2.13 je $\text{cov}(X_j, \mathcal{E}_i | \mathbf{X}_{\bar{K}}) = 0$ skoro jistě, a důkaz je tedy hotov. \square

Můžeme nyní přistoupit k formulaci a důkazu věty o platnosti implikace (2.7).

Věta 2.15 *Nechť platí $(\alpha), (\beta), (\gamma), (\delta)$. Potom*

$$X_i \perp\!\!\!\perp X_j | \mathbf{X}_{\bar{K}} \implies a_{ij} = 0.$$

Důkaz. Větu dokážeme sporem. Předpokládejme pro spor, že platí $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}}$ a zároveň $a_{ij} \neq 0$. Podle věty 1.5 plyne z podmíněné nezávislosti X_i a X_j , že $\text{cov}(X_i, X_j \mid \mathbf{X}_{\bar{K}}) = 0$ skoro jistě, použitím lemmatu 2.14 odtud dostaneme

$$a_{ij} \text{var}(X_j \mid \mathbf{X}_{\bar{K}}) = 0 \quad \text{s.j.},$$

tedy při $a_{ij} \neq 0$ máme

$$\text{var}(X_j \mid \mathbf{X}_{\bar{K}}) = 0 \quad \text{s.j.} \quad (2.15)$$

Označme pravděpodobnostní míru na \mathbb{R}_{i-1} indukovanou náhodným vektorem $\mathbf{X}_{j\bar{K}}$ jako μ . Ukážeme, že za platnosti (2.15) není μ absolutně spojitá vzhledem k λ_{i-1} . Zavedme funkci g předpisem

$$\begin{aligned} g &: \mathbb{R}_{i-2} \rightarrow \mathbb{R}, \\ g &: \mathbf{y} \mapsto \mathbf{E}(X_j \mid \mathbf{X}_{\bar{K}} = \mathbf{y}) \end{aligned}$$

a množinu M ve tvaru

$$M = \{\mathbf{y} \in \mathbb{R}_{i-1} : \mathbf{y}_j = g(\mathbf{y}_{\bar{K}})\}.$$

Z definice funkce g můžeme psát $\text{var}(X_j \mid \mathbf{X}_{\bar{K}}) = \mathbf{E}\{[X_j - g(\mathbf{X}_{\bar{K}})]^2 \mid \mathbf{X}_{\bar{K}}\}$. Za platnosti (2.15) a s přihlédnutím k větě 1.4 (i) pak dostaneme

$$\mathbf{E}[X_j - g(\mathbf{X}_{\bar{K}})]^2 \stackrel{\text{v.1.4}}{=} \mathbf{E}[\text{var}(X_j \mid \mathbf{X}_{\bar{K}})] \stackrel{(2.15)}{=} 0,$$

tudíž platí $X_j = g(\mathbf{X}_{\bar{K}})$ skoro jistě. Proto $\Pr(\mathbf{X}_{j\bar{K}} \in M) = \mu(M) = 1$. Přitom zřejmě $\lambda_{i-1}(M) = 0$. Tudíž μ není absolutně spojitá vzhledem k λ_{i-1} . Nemůže proto existovat hustota pravděpodobnosti $\mathbf{X}_{j\bar{K}}$ vzhledem k λ_{i-1} , což je spor s větou 2.10, která nám existenci této hustoty zaručuje. \square

Výsledky ohledně souvislosti mezi koeficienty ZLRS a jeho strukturou podmíněných nezávislostí shrnuje následující věta. Abychom mohli zapsat výsledky co nejstručněji, využijeme opět konvenci, že při $K = \emptyset$ ztotožníme výrazy $X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_K$ a $X_i \perp\!\!\!\perp X_j$.

Věta 2.16 *Nechť \mathbf{X} je ZLRS s maticí koeficientů $\mathbf{A}_{n \times n}$ a chybovým vektorem \mathcal{E} a nechť všechny složky vektoru \mathcal{E} mají konečné druhé momenty. Potom pro libovolná čísla i a j taková, že $1 \leq j < i \leq n$ platí*

$$a_{ij} = 0 \iff X_i \perp\!\!\!\perp X_j \mid \mathbf{X}_{\bar{K}},$$

kde $\bar{K} = \{1, 2, \dots, i-1\} \setminus \{j\}$.

Důkaz. Pro $\bar{K} \neq \emptyset$, plyne tvrzení dokazované věty bezprostředně z vět 2.11 a 2.15. Při $\bar{K} = \emptyset$ máme nutně $i = 2$ a $j = 1$ a v rámci výše zmíněné konvence má tvrzení dokazované věty podobu

$$a_{21} = 0 \iff X_2 \perp\!\!\!\perp X_1.$$

Pokud $a_{21} = 0$, je $X_2 = \mathcal{E}_2$ a $X_1 = \mathcal{E}_1$, z předpokladu o nezávislosti chybových složek ZLRS plyne $X_2 \perp\!\!\!\perp X_1$. Je-li naopak $a_{21} \neq 0$, vyjádříme

$$\text{cov}(X_1, X_2) = a_{ij} \text{var} X_1 + \text{cov}(\mathcal{E}_1, \mathcal{E}_2) = a_{ij} \text{var} X_1 \neq 0,$$

tedy X_1 a X_2 nemohou být nezávislé. \square

2.3 Grafický model lineárního rekursivního systému

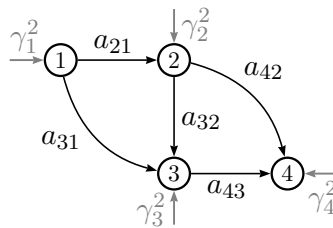
Výsledky předchozích dvou oddílů poukazují na velmi jednoduchý a přirozený způsob sestavení grafického modelu struktury podmíněných nezávislostí v LRS (příp. ZLRS) přímo z matice jeho koeficientů. Součástí grafického modelu, který zde zavedeme, jsou kromě informací ohledně struktury podmíněné nezávislosti také některé hodnoty z matice koeficientů a rozptyly složek chybového vektoru. Budeme tedy pracovat s *ohodnocenými grafy*, ve kterých je každé hraně a/nebo každému vrcholu přiřazeno reálné číslo; hovoříme potom o *hranově a/nebo vrcholově ohodnocených grafech*.

Definice 2.4 Mějme LRS $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ s maticí koeficientů $\mathbf{A} = (a_{ij})$ a chybovým vektorem \mathcal{E} , označme rozptyl i -té složky chybového vektoru jako γ_i^2 . *Grafickým modelem lineárního rekursivního systému \mathbf{X}* rozumíme orientovaný graf s množinou vrcholů $\{1, 2, \dots, n\}$, který obsahuje hranu $j \rightarrow i$, pokud $a_{ij} \neq 0$, a který je hranově i vrcholově ohodnocený tak, že ohodnocení hrany $j \rightarrow i$ je rovno a_{ij} a ohodnocení uzlu i je γ_i^2 . Grafický model LRS \mathbf{X} značíme $G(\mathbf{X})$. Zcela analogicky definujeme i *grafický model zobecněného lineárního rekursivního systému \mathbf{X}* , značíme jej shodně $G(\mathbf{X})$.

Příklad 2.2 Uvažujme například lineární rekursivní systém \mathbf{X} s maticí koeficientů

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ 0 & a_{42} & a_{43} & 0 \end{pmatrix},$$

jehož chybový vektor má složky s rozptyly $\gamma_1^2, \gamma_2^2, \gamma_3^2, \gamma_4^2$. Z tvaru matice \mathbf{A} je patrné, že $G(\mathbf{X})$ obsahuje hranu $j \rightarrow i$, kdykoli $j < i$, vyjma hrany $1 \rightarrow 4$, neboť $a_{41} = 0$. $G(\mathbf{X})$ má tedy následující podobu (ohodnocení vrcholů je pro přehlednost zachyceno šedou barvou pomocí „ohodnocených šipek“):



Z definice 2.4 je patrné, že vrchol i v grafu $G(\mathbf{X})$ odpovídá náhodné veličině X_i . Vzhledem k tomu, že matice koeficientů LRS (příp. ZLRS) je dolní trojúhelníková s nulovou diagonálou, je graf $G(\mathbf{X})$ automaticky topologicky očíslovaným ADG.

Podle věty 2.8 platí pro libovolný LRS \mathbf{X} , že rozdělení \mathbf{X} splňuje (GM) (je globálně markovské) vzhledem ke $G(\mathbf{X})$. Pokud je \mathbf{X} ZLRS, pak jeho rozdělení podle věty 2.16 splňuje (PM) , tj. je párově markovské, vzhledem ke $G(\mathbf{X})$. Mohli bychom pro ZLRS \mathbf{X} dále snadno ukázat, že je-li hustota všech chybových složek *kladná* všude na \mathbb{R} , je i hustota \mathbf{X} kladná všude na \mathbb{R}_n (to plyne ihned z věty 2.9), a podle věty 1.7 pak rozdělení \mathbf{X} splňuje dokonce (GM) vzhledem ke $G(\mathbf{X})$. Graf $G(\mathbf{X})$ tedy zachycuje strukturu podmíněné nezávislosti \mathbf{X} ve smyslu příslušných markovských vlastností.

Abychom z grafického modelu ZLRS nebo LRS vyčetli příslušné podmíněné nezávislosti, nepotřebujeme znát ohodnocení hran ani vrcholů v tomto grafu. Užitečnost tohoto ohodnocení ukážeme v následující kapitole.

Kapitola 3

Metoda koeficientů na cestách a lineární rekursivní systémy

Jak bylo zmíněno v úvodu práce, jednou z prvních aplikací grafických modelů ve statistice byla *metoda koeficientů na cestách* (*method of path coefficients*) amerického biologa a statistika Sewalla Wrighta, který ji publikoval ve dvacátých letech minulého století. V této kapitole metodu stručně popíšeme a poukážeme na souvislosti mezi systémy náhodných veličin, kterými se metoda zabývá, a lineárními rekursivními systémy, jež jsme zavedli a studovali v předchozí kapitole. Následně uvedeme postup, který se podobá metodě koeficientů na cestách a pomocí něhož lze vyčíst z grafického modelu LRS jeho varianční matici.

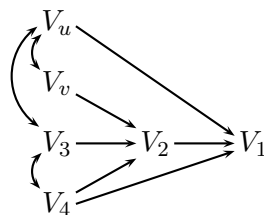
3.1 Metoda koeficientů na cestách

Úvodem tohoto oddílu si dovolíme malou poznámku. Sewall Wright byl původem biolog a při návrhu metody koeficientů na cestách sledoval především konkrétní aplikační cíle. Jeho pojednání o této metodě (např. [Wri21], [Wri34], [Wri60]) nejsou tudíž po formální stránce vedena jako matematický text. Jednotlivé pojmy nejsou rigorózně definovány a obsažená tvrzení nejsou formulována v podobě matematických vět, které by vymezily nezbytné předpoklady a byly nakonec korektním a ucelným způsobem dokázány. Tuto skutečnost nezmiňujeme coby kritiku; snažíme se pouze vysvětlit, proč se zpracování tohoto oddílu liší od zbylého textu. Cílem zde je totiž pouze představit metodu koeficientů na cestách v původním Wrightově pojetí pro účely pozdějšího srovnání se studiem (zobecněných) lineárních rekursivních systémů. Jediný odklon od původní Wrightovy terminologie bude představovat práce s grafickým modelem – v tomto ohledu využijeme pojmy a notaci zavedené v oddílu 1.4.

Důraz na využití grafického znázornění vztahů mezi veličinami v metodě koeficientů na cestách je patrný už z jejího názvu; účelové zaměření metody pak vytyčil Wright názvem článku [Wri21], *Correlation and causation*. Metoda koeficientů na cestách se totiž, stručně řečeno, zabývá vztahy mezi korelačními koeficienty v systému náhodných veličin, jejichž kauzální vztahy jsou zachyceny graficky ve formě diagramu.

Wrightovy diagramy se skládají ze symbolů představujících jednotlivé náhodné veličiny a z jednosměrných a obousměrných šipek, které tyto symboly spojují. Příklad takového diagramu je uveden na obrázku 3.1. Z obrázku je patrné, že na Wrightův diagram lze pohlížet jako na graf, jehož vrcholy představují náhodné veličiny a jehož hrany odpovídají znázorněným šipkám ve smyslu grafického značení, které jsme zavedli v oddílu 1.4: obousměrným šipkám odpovídají hrany neorientované, jednosměrným šipkám hrany orientované.

V dalším textu budeme tento grafický model i nadále nazývat diagramem, abychom jej odlišili od grafů použitých pro LRS. Pro práci s diagramy budeme nicméně využívat pojmů, které jsme zavedli pro grafy v první kapitole (jako jsou např. hrany, vrcholy, cesty apod.).



Obrázek 3.1: Příklad diagramu pro popis kauzálních vztahů mezi veličinami

Náhodné veličiny v systému zkoumaném metodou koeficientů na cestách lze rozdělit do dvou skupin:

- *exogenní veličiny*: jde o veličiny, jejichž hodnoty jsou determinovány mimo uvažovaný systém – z hlediska tohoto systému jsou tedy dány „shůry“. Jelikož vztahům mezi veličinami přisuzuje Wright kauzální charakter, představují tyto veličiny *prvotní příčiny*, příp. *prvotní faktory* (v článku [Wri60] jsou označovány jako *ultimate factors*). V příslušném diagramu se jedná o veličiny, do nichž nevstupuje žádná orientovaná hrana – na obrázku 3.1 jsou to V_u , V_v , V_3 a V_4 . Exogenní veličiny mohou být navzájem korelované, v takovém případě je spojuje neorientovaná hrana.
- *endogenní veličiny*: tyto veličiny jsou zcela determinovány ve zkoumaném systému, tj. jejich hodnoty jsou jednoznačně určeny veličinami exogenními. Přesněji řečeno, každá endogenní veličina je funkcí všech veličin, od kterých k ní v diagramu vedou orientované hrany – tj. je funkcí svých rodičů. Na obrázku 3.1 jsou endogenními veličinami V_1 a V_2 . Veličina V_2 je např. funkcí V_3 , V_4 a V_v . Při kauzální interpretaci tohoto vztahu bychom chápali V_3 , V_4 a V_v jako *bezprostřední faktory* (*immediate factors*) vyvolávající V_2 coby *důsledek* (*effect*).

Pracujeme-li s reálně existujícími (pozorovanými) náhodnými veličinami, často si nemůžeme dovolit předpokládat přesný funkční vztah mezi nimi. Z tohoto důvodu je třeba do uvažovaného systému zařadit kromě pozorovaných veličin také tzv. *residuální faktory* (*residual factors*), které představují hypotetický souhrn všech nepozorovaných nebo neměřitelných vlivů, které „dovysvětlují“ danou endogenní veličinu (residuální faktory jsou vždy brány jako exogenní). Ve formálním zápisu rozlišuje Wright tyto veličiny pomocí indexů: zatímco pro residuální faktory používá písmenné indexy, pozorované veličiny indexuje arabskými číslicemi. Na obrázku 3.1 tedy představují V_u a V_v residuální faktory pro pozorované veličiny V_1 a V_2 .

Automaticky se předpokládá, že residuální faktor pro danou endogenní veličinu je nezávislý na jejích ostatních bezprostředních faktorech (tj., s použitím grafického vyjádření, na ostatních rodičích endogenní veličiny); proto nejsou v diagramu na obrázku 3.1 přítomny hrany $V_v \leftrightarrow V_3$, $V_v \leftrightarrow V_4$ a $V_u \leftrightarrow V_4$. Obecně se však nevyklučuje, že jsou tyto residuální faktory závislé na „vzdálenějších předchůdcích“, což v diagramu zachycuje hrana $V_u \leftrightarrow V_3$.

Musí dále platit, že v systému neexistují obousměrné kauzální vztahy. To v grafu znamená, že neexistuje orientovaný cyklus vedoucí přes některou endogenní veličinu.

Posledním klíčovým předpokladem při použití metody koeficientů na cestách je lineární podoba funkční závislosti endogenních proměnných na jejich bezprostředních faktorech – rodičích (již bylo zmíněno, že každá endogenní veličina je funkcí svých rodičů; nyní jsme

pouze upřesnili, že jde o funkci lineární). Např. pro systém veličin z obrázku 3.1 tento předpoklad říká, že existují taková reálná čísla $c_{10}, c_{12}, c_{14}, c_{1u}$, že

$$V_1 = c_{10} + c_{12}V_2 + c_{14}V_4 + c_{1u}V_u,$$

a taková reálná čísla $c_{20}, c_{23}, c_{24}, c_{2v}$, že

$$V_2 = c_{20} + c_{23}V_3 + c_{24}V_4 + c_{2v}V_v.$$

Obecně, označíme-li veličiny v libovolném uvažovaném systému pomocí indexovaného písmena V , můžeme endogenní veličinu V_i vyjádřit jako

$$V_i = c_{i0} + \sum_{j: V_j \in \text{pa}(V_i)} c_{ij}V_j \quad (3.1)$$

(připomeňme, že zápisem $\text{pa}(V_i)$ rozumíme množinu všech rodičů vrcholu V_i , viz definice 1.10).

Označme nyní pro libovolné k směrodatnou odchylku veličiny V_k jako σ_k a definujme veličinu X_k předpisem

$$X_k = \frac{V_k - \mathbb{E}V_k}{\sigma_k}, \quad (3.2)$$

tedy veličina X_k představuje *normovanou* (*standardized*) verzi veličiny V_k . Můžeme pak (3.1) vyjádřit v podobě

$$X_i = \sum_{j: V_j \in \text{pa}(V_i)} p_{ij}X_j, \quad (3.3)$$

kde $p_{ij} = c_{ij} \frac{\sigma_j}{\sigma_i}$. Takto zavedené koeficienty p_{ij} jsou nazývány *koeficienty na cestách* (*path coefficients*). Můžeme tedy říci, že koeficienty na cestách zachycují parametry lineárních funkčních vztahů mezi normovanými verzemi veličin z uvažovaného systému. Další způsoby, jak koeficienty na cestách interpretovat, uvádí např. [Wri34].

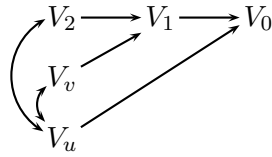
Hlavním Wrightovým výsledkem bylo nalezení vztahu mezi koeficienty na cestách a párovými korelačními koeficienty veličin v uvažovaném systému. Tento vztah Wright vyjadřuje verbálním způsobem, který není příliš přehledný (viz např. [Wri34, str. 163] nebo [Wri60, str. 193]). Použijeme-li však terminologii, kterou jsme zavedli pro práci s grafy, můžeme tento vztah shrnout do následujících dvou bodů:

- (i) *Souhrnné ohodnocení cesty* (*compound path coefficient*) v diagramu zkoumaného systému je dáno součinem ohodnocení všech hran podél této cesty; přitom ohodnocením orientované hrany $V_i \rightarrow V_j$ rozumíme koeficient p_{ji} a ohodnocením neorientované hrany $V_i \leftrightarrow V_j$ rozumíme korelační koeficient $\text{corr}(V_i, V_j)$.
- (ii) Korelační koeficient $\text{corr}(V_i, V_j)$ je roven součtu souhrnných ohodnocení všech takových cest nenulové délky, které vedou z V_i do V_j , nemají žádný kolizní vrchol a vedou přes nanejvýš jednu neorientovanou hranu.

Podotkněme, že bod (ii) je zajímavý pouze v případě, že alespoň jedna z veličin V_i, V_j je veličinou endogenní. V případě, že jde o dvě exogenní veličiny, říká bod (ii) pouze, že $\text{corr}(V_i, V_j)$ je roven souhrnnému ohodnocení cesty (resp. ohodnocení hrany) $V_i \leftrightarrow V_j$ – snadno se totiž ověří, že jakákoli jiná cesta mezi dvěma různými exogenními veličinami obsahuje buď kolizní vrchol, nebo alespoň dvě neorientované hrany.

Použití uvedeného vztahu mezi koeficienty na cestách a párovými korelačními koeficienty ilustruje následující příklad. Příklad je převzat z [Wri60, str. 198].

Příklad 3.1 Uvažujme systém pozorovaných náhodných veličin V_0, V_1, V_2 s residuálními faktory V_u, V_v , jejichž vztahy jsou dány následujícím diagramem:



Označme pro stručnost korelační koeficient typu $\text{corr}(V_i, V_j)$ symbolem r_{ij} . Zajímá-li nás např. korelace mezi V_0 a V_2 , můžeme podle bodu (ii) vyjádřit r_{02} jako součet souhrnného ohodnocení cest $V_0 \leftarrow V_1 \leftarrow V_2$ a $V_0 \leftarrow V_u \leftrightarrow V_2$; jiné cesty z V_0 do V_2 obsahují buď kolizní vrchol, nebo alespoň dvě neorientované hrany. Vyjádříme-li souhrnné ohodnocení uvedených cest podle bodu (i), můžeme psát

$$r_{02} = p_{01}p_{12} + p_{0u}r_{u2}.$$

Z podmínek kladených na residuální faktory plyne, že veličiny V_u a V_1 jsou nekorelované. Tedy musí platit

$$0 = r_{1u} = p_{1v}r_{vu} + p_{12}r_{2u}. \quad (3.4)$$

Vztah mezi korelačními koeficienty a koeficienty na cestách lze uplatnit i na korelační koeficienty typu $\text{corr}(V_i, V_i)$, které mají zřejmě vždy hodnotu 1. Např. koeficient r_{00} je podle bodu (ii) roven součtu souhrnných ohodnocení cest (v tomto případě cyklů)

$$\begin{aligned} V_0 &\leftarrow V_1 \rightarrow V_0, \\ V_0 &\leftarrow V_u \rightarrow V_0, \\ V_0 &\leftarrow V_1 \leftarrow V_2 \leftrightarrow V_u \rightarrow V_0, \\ V_0 &\leftarrow V_1 \leftarrow V_v \leftrightarrow V_u \rightarrow V_0, \end{aligned}$$

tedy platí

$$1 = r_{00} = p_{01}^2 + p_{0u}^2 + p_{01}p_{12}r_{2u}p_{0u} + p_{01}p_{1v}r_{vu}p_{0u} \stackrel{(3.4)}{=} p_{01}^2 + p_{0u}^2.$$

Podobně můžeme vyjádřit i ostatní párové korelační koeficienty, celkem dostaneme:

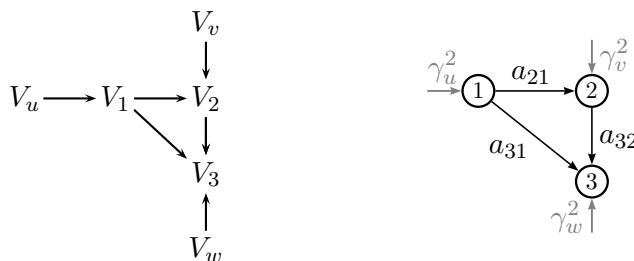
$$\begin{aligned} r_{00} &= 1 = p_{01}^2 + p_{0u}^2, & r_{0v} &= p_{01}p_{1v} + p_{0u}r_{uv}, \\ r_{01} &\stackrel{(3.4)}{=} p_{01}, & r_{11} &= 1 = p_{12}^2 + p_{1v}^2, \\ r_{02} &= p_{01}p_{12} + p_{0u}r_{u2}, & r_{12} &= p_{12}, \\ r_{0u} &\stackrel{(3.4)}{=} p_{0u}, & r_{1u} &= 0 = p_{1v}r_{vu} + p_{12}r_{2u}. \end{aligned}$$

Využití metody koeficientů na cestách v praktických aplikacích může mít v zásadě dvě různé podoby:

- 1) Známe-li korelační koeficienty mezi sledovanými veličinami, můžeme pomocí uvedených vztahů dopočítat koeficienty na cestách; přínos koeficientů na cestách pro další analýzu pak spočívá v jejich interpretaci. Wright uvádí [Wri20, str. 329], že koeficienty na cestách lze chápat např. jako relativní důležitost jednotlivých faktorů pro danou endogenní proměnnou. Autoři [Den06] se domnívají, že právě toto využití bylo zřejmě původním Wrightovým motivem pro vybudování celé metody.
- 2) Známe-li hodnoty koeficientů na cestách, potažmo tvar některých funkčních vztahů mezi uvažovanými veličinami, můžeme z nich odvodit vztahy mezi jejich korelačními koeficienty. Wright uvedl několik aplikací tohoto typu v oblasti populační genetiky, viz [Wri34]; v tomto případě slouží pro určení funkčních vztahů mezi veličinami poznatky ohledně pravidel pro křížení genotypů.

Ačkoli Wright kladl ve svých raných textech o metodě koeficientů na cestách důraz na kauzální interpretaci vztahů mezi zkoumanými veličinami, z popisu metody je zřejmé, že kauzalita jako taková se ve fungování metody nijak neprojevuje; slouží pouze coby verbální odůvodnění pro přítomnost funkční závislosti endogenní veličiny na jejích rodičích. Jak je patrné z jeho pozdějších prací, Wright si tento fakt sám dobře uvědomoval a např. v [Wri60, str. 191] zmiňuje, že použití metody koeficientů na cestách „není v žádném případě omezeno na vztahy, které lze označit za vztahy mezi příčinou a důsledkem. Lze ji použít na čistě matematické systémy lineárních vztahů.“

Příkladem takových systémů s lineárními vztahy jsou lineární rekursivní systémy. Roli residuálních faktorů a současně jediných exogenních veličin zde hrají složky chybového vektoru. Uvažujme např. lineární rekursivní systém $\mathbf{V} = (V_1, V_2, V_3)^\top$ s maticí koeficientů $\mathbf{A} = (a_{ij})$ a chybovým vektorem $(V_u, V_v, V_w)^\top$, jehož složky mají po řadě rozptyl γ_u^2, γ_v^2 a γ_w^2 . Předpokládejme dále, že koeficienty a_{21}, a_{31} a a_{32} jsou nenulové. Diagram pro metodu koeficientů na cestách má potom podobu, kterou znázorňuje obrázek 3.2 vlevo. V diagramu se nevyskytují žádné neorientované hrany, neboť složky chybového vektoru jsou z definice LRS nezávislé, tudíž nekorelované. Připomeňme, že při použití metody koeficientů na cestách chápeme všechny hrany jako ohodnocené, přičemž ohodnocením hrany $V_i \rightarrow V_j$ rozumíme koeficient p_{ji} .



Obrázek 3.2: Srovnání diagramu pro metodu koeficientů na cestách a grafického modelu LRS

Srovnajme tento diagram s grafickým modelem uvažovaného LRS (tj. s $G(\mathbf{V})$), který je uveden na obrázku 3.2 vpravo; najdeme zde několik odlišností:

- $G(\mathbf{V})$ neobsahuje vrcholy představující chybové složky (tj. residuální faktory).
- Ohodnocení hran v $G(\mathbf{V})$ představují koeficienty a_{ij} , které se obecně liší koeficientů na cestách p_{ij} .
- V grafickém modelu $G(\mathbf{V})$ jsou ohodnocené uzly.

V následujícím oddílu ukážeme, že z grafického modelu LRS lze vyčíst celou jeho varianční matici, a to způsobem, který se v jádru podobá metodě koeficientů na cestách; díky odlišnému cíli (popis varianční matice namísto korelačních koeficientů) a odlišným grafickým modelům se však oba postupy v několika ohledech liší.

3.2 Cesty, sledy a varianční matice v grafickém modelu lineárního rekursivního systému

Připomeňme nejprve, že máme-li LRS \mathbf{X} s maticí koeficientů \mathbf{A} a chybovou varianční maticí $\mathbf{\Gamma}$, můžeme na základě (2.2) vyjádřit \mathbf{X} jako $\mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{\mathcal{E}}$. Označíme-li pro přehlednost $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$, můžeme psát

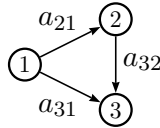
$$\text{var } \mathbf{X} = \mathbf{B} \mathbf{\Gamma} \mathbf{B}^\top. \quad (3.5)$$

Postupně ukážeme, jak lze hodnotu výrazu na pravé straně (3.5) určit z grafického modelu LRS. Začneme tím, že vyjádříme prvky matice $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ pomocí *hodnot cest* v grafu; hodnoty cest jsou analogií k souhrnným ohodnocení cest v metodě koeficientů na cestách a zavedeme je v následující definici.

Definice 3.1 Necht' $\gamma = (\alpha_0, \alpha_1, \dots, \alpha_n)$ je cesta v hranově ohodnoceném ADG. Označme pro $i = 1, 2, \dots, n$ ohodnocení hrany mezi vrcholy α_{i-1} a α_i jako v_i . Potom *hodnotou cesty* γ rozumíme hodnotu $\prod_{i=1}^n v_i$. Budeme přitom uvažovat i cesty délky 0, tj. cesty typu $\gamma = (\alpha_0)$, jejich hodnotu definujeme jako 1 (což odpovídá výsledku prázdného součinu $\prod_{i \in \emptyset} v_i = 1$). Zcela analogicky zavedeme i *hodnotu sledu*.

Vztah mezi hodnotou cest v grafickém modelu LRS a maticí $(\mathbf{I} - \mathbf{A})^{-1}$ ilustruje následující příklad.

Příklad 3.2 Mějme LRS \mathbf{X} s grafickým modelem (vynecháváme pro přehlednost ohodnocení vrcholů)



Matici koeficientů \mathbf{A} a matici $(\mathbf{I} - \mathbf{A})^{-1}$ můžeme vyjádřit jako

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}, \quad (\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} + a_{32}a_{21} & a_{32} & 1 \end{pmatrix}.$$

Můžeme se snadno přesvědčit, že prvek v i -tém řádku a j -tém sloupci matice $(\mathbf{I} - \mathbf{A})^{-1}$ lze vyjádřit jako součet hodnot všech orientovaných cest z j do i . Např. pro $i = 3$ a $j = 1$ bychom sčítali hodnoty cest $1 \rightarrow 3$ a $1 \rightarrow 2 \rightarrow 3$, což dává $a_{31} + a_{32}a_{21}$. Pro $i < j$ neexistuje v grafu žádná orientovaná cesta z j do i a součet hodnot cest je tedy nulový, pro $i = j$ existuje z j do i pouze cesta délky 0, jejíž hodnota je 1.

Jak později dokážeme, tento výsledek je možné zobecnit: je-li \mathbf{A} matice koeficientů LRS \mathbf{X} a $(\mathbf{I} - \mathbf{A})^{-1}$, lze prvek i -tém řádku a j -tém sloupci matice $(\mathbf{I} - \mathbf{A})^{-1}$ vyjádřit jako součet hodnot všech orientovaných cest, které vedou v $G(\mathbf{X})$ z j do i . Nejprve zavedeme značení, které nám umožní zapsat součet hodnot všech orientovaných cest z j do i pomocí prvků matice \mathbf{A} .

Definice 3.2 Necht' $\mathbf{A}_{n \times n}$ je matice koeficientů lineárního rekursivního systému. Zavedeme funkci

$$\pi_{\mathbf{A}} : 2^{\{1,2,\dots,n\}} \longrightarrow \mathbb{R}$$

následujícím způsobem. Necht' I je libovolná alespoň dvouprvková indexová množina, $I \subseteq \{1, 2, \dots, n\}$. Položme $m = |I|$ a prvky množiny I označme jako i_1, i_2, \dots, i_m tak, aby platilo $i_1 < i_2 < \dots < i_m$. Potom funkční hodnotu $\pi_{\mathbf{A}}(I)$ definujeme předpisem

$$\pi_{\mathbf{A}}(I) = a_{i_m i_{m-1}} a_{i_{m-1} i_{m-2}} \cdots a_{i_2 i_1}. \quad (3.6)$$

Pro jednoprvkové množiny z $2^{\{1,2,\dots,n\}}$ a pro množinu prázdnou nebudeme funkci $\pi_{\mathbf{A}}$ využívat, příslušné funkční hodnoty je možné dodefinovat jako nulu.

Snadno se přesvědčíme, že hodnota funkce $\pi_{\mathbf{A}}(I)$ odpovídá hodnotě orientované cesty, která vede v grafu LRS s maticí koeficientů \mathbf{A} přes vrcholy z I . Je-li $I = \{i_1, i_2, \dots, i_m\}$, kde $i_1 < i_2 < \dots < i_m$, vrací funkce $\pi_{\mathbf{A}}$ hodnotu cesty

$$i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_m,$$

příčemž požadavek na orientaci cesty plyne ze seřazení prvků I podle velikosti a topologického očíslování grafu LRS. Pokud taková cesta v grafu neexistuje, tj. pokud např. hrana $i_k \rightarrow i_{k+1}$ v grafu chybí, je z definice grafického modelu LRS nutně $a_{i_{k+1}i_k} = 0$, a tedy $\pi_{\mathbf{A}}(I) = 0$. Dodejme, že označení funkce pramení z anglického *path*.

Příklad 3.3 Uvažujme opět LRS z příkladu 3.2. Souhrnné ohodnocení cesty $1 \rightarrow 2 \rightarrow 3$ můžeme pomocí funkce $\pi_{\mathbf{A}}$ vyjádřit jako $\pi_{\mathbf{A}}(\{1, 2, 3\})$ a matici $(\mathbf{I} - \mathbf{A})^{-1}$ můžeme zapsat ve tvaru

$$(\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \pi_{\mathbf{A}}(\{1, 2\}) & 1 & 0 \\ \pi_{\mathbf{A}}(\{1, 3\}) + \pi_{\mathbf{A}}(\{1, 2, 3\}) & \pi_{\mathbf{A}}(\{2, 3\}) & 1 \end{pmatrix}.$$

Zavedeme nyní značení, které nám umožní zapsat pomocí funkce $\pi_{\mathbf{A}}$ součet hodnot všech orientovaných cest z j do i . Mějme přirozená čísla i a j taková, že $i > j$. Symbolem \mathcal{I}_{ij} budeme označovat množinový systém daný předpisem

$$\mathcal{I}_{ij} = \{I : \{j, i\} \subseteq I \subseteq \{j, j+1, \dots, i\}\}, \quad (3.7)$$

tedy např. $\mathcal{I}_{41} = \{\{1, 4\}, \{1, 2, 4\}, \{1, 3, 4\}, \{1, 2, 3, 4\}\}$. Máme-li grafický model LRS s maticí koeficientů \mathbf{A} , můžeme součet hodnot všech orientovaných cest z j do i vyjádřit pro $i > j$ jako

$$\sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I), \quad (3.8)$$

neboť díky topologickému očíslování mohou vést orientované cesty z j do i pouze přes vrcholy $j+1, j+2, \dots, i-1$; např. vrcholy každé orientované cesty z 1 do 4 musí odpovídat jedné z množin systému \mathcal{I}_{41} . Připomeňme, že pokud pro nějaké $I \in \mathcal{I}_{ij}$ orientovaná cesta přes vrcholy z I neexistuje, je nutně $\pi_{\mathbf{A}}(I) = 0$; výraz (3.8) tedy představuje součet hodnot všech *existujících* orientovaných cest z j do i .

Příklad 3.4 Uvažujme opět LRS z příkladu 3.2. Pomocí právě zavedeného značení můžeme matici $(\mathbf{I} - \mathbf{A})^{-1}$ zapsat ve tvaru

$$(\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \sum_{I \in \mathcal{I}_{21}} \pi_{\mathbf{A}}(I) & 1 & 0 \\ \sum_{I \in \mathcal{I}_{31}} \pi_{\mathbf{A}}(I) & \sum_{I \in \mathcal{I}_{32}} \pi_{\mathbf{A}}(I) & 1 \end{pmatrix}, \quad (3.9)$$

neboť např. $\sum_{I \in \mathcal{I}_{31}} \pi_{\mathbf{A}}(I) = \pi_{\mathbf{A}}(\{1, 3\}) + \pi_{\mathbf{A}}(\{1, 2, 3\}) = a_{31} + a_{32}a_{21}$.

V následující větě ukážeme, že vztah (3.9) lze příslušným způsobem zobecnit, tj. že zcela analogicky lze vyjádřit matici $(\mathbf{I} - \mathbf{A})^{-1}$ v libovolném LRS s maticí koeficientů \mathbf{A} .

Věta 3.1 *Nechť $\mathbf{A}_{n \times n}$ je matice koeficientů lineárního rekursivního systému. Zavedme matici $\mathbf{B}_{n \times n} = (b_{ij})$ předpisem*

$$b_{ij} = \begin{cases} 1, & \text{pokud } i = j, \\ 0, & \text{pokud } i < j, \\ \sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I), & \text{pokud } i > j. \end{cases} \quad (3.10)$$

Potom $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$.

Větu dokážeme později, nejprve uvedeme některá pomocná lemmata.

Lemma 3.2 *Nechť $\mathbf{A}_{n \times n}$ je matice koeficientů lineárního rekursivního systému a I je indexová množina splňující $I \subset \{1, 2, \dots, n\}$ a $|I| \geq 2$. Označme $m = \max\{i : i \in I\}$. Nechť dále $m < k \leq n$. Potom platí:*

$$a_{km} \pi_{\mathbf{A}}(I) = \pi_{\mathbf{A}}(I \cup \{k\}). \quad (3.11)$$

Důkaz. Plyne bezprostředně z definice 3.2. \square

V kontextu cest v grafu LRS hovoří předchozí lemma o prodloužení orientované cesty vedoucí přes vzestupně očíslované vrcholy z I do dalšího vrcholu k , přičemž musí být opět respektováno topologické očíslování, tedy vrchol k musí mít vyšší číslo než všechny vrcholy z I .

Lemma 3.3 *Nechť $\mathbf{A}_{n \times n}$ je matice koeficientů LRS. Mějme dále přirozená čísla j a k taková, že $1 \leq j < k \leq n$. Potom platí:*

$$\sum_{I \in \mathcal{I}_{kj}} \pi_{\mathbf{A}}(I) = \pi_{\mathbf{A}}(\{j, k\}) + \sum_{i=j+1}^{k-1} \left(\sum_{J \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(J \cup \{k\}) \right). \quad (3.12)$$

Důkaz. Nejprve si uvědomme, že jak na levé, tak na pravé straně se pouze sčítají hodnoty funkce $\pi_{\mathbf{A}}$ pro různé hodnoty vstupních argumentů. Stačí tedy ukázat, že množiny, které vystupují na levé straně coby argumenty funkce $\pi_{\mathbf{A}}$, tvoří právě systém \mathcal{I}_{kj} , přes který se sčítá na straně pravé.

Volme libovolně $I \in \mathcal{I}_{kj}$. Nechť nejprve $|I| = 2$. Potom nutně $I = \{j, k\}$, což je argument u prvního sčítance na pravé straně. Nechť nyní $|I| > 2$. Označme $i = \max\{h : h \in I \setminus \{k\}\}$; zřejmě je $j < i < k$. Pak ale existuje indexová množina $J \in \mathcal{I}_{ij}$ taková, že $I = J \cup \{k\}$, a I se tedy vyskytuje jako argument funkce $\pi_{\mathbf{A}}$ u nějakého sčítance v sumě na pravé straně (3.12).

Vezmeme-li libovolný argument funkce $\pi_{\mathbf{A}}$ na pravé straně (3.12), snadno se přesvědčíme, že patří do \mathcal{I}_{kj} . Zbývá už jen ověřit, že se na pravé straně nevyskytuje žádný sčítanec opakovaně. V první řadě si uvědomme, že všechny argumenty v druhém výrazu na pravé straně mají alespoň 3 různé prvky: j , i a k ; tedy $\{j, k\}$ se vyskytuje na pravé straně jen jednou. Je-li $j < i_1 < i_2 < k$, jsou libovolné dvě množiny $J_1 \in \mathcal{I}_{i_1j}$ a $J_2 \in \mathcal{I}_{i_2j}$ nutně různé, tedy všechny argumenty v sumě na pravé straně jsou různé a důkaz je hotov. \square

Dodejme, že lemma 3.3 lze opět snadno graficky interpretovat: výraz na levé straně (3.12) popisuje součet hodnot všech orientovaných cest jdoucích z j do k a ve výrazech na levé straně jsou hodnoty těchto cest seskupeny podle toho, který vrchol je na cestě předposlední. První výraz na pravé straně (3.12) představuje hodnotu cesty $j \rightarrow k$ a v druhém výrazu se sčítají hodnoty všech cest typu $j \rightarrow \dots \rightarrow i \rightarrow k$ pro i jdoucí od $j+1$ do $k-1$.

Před dalším lemmatem připomeňme, že podle značení zavedeného v oddílu 2.2 je $L_k = \{1, 2, \dots, k\}$ a $\mathbf{A}_{L_k \cdot L_k}$ představuje podmatici \mathbf{A} tvořenou výběrem prvních k řádků a sloupců.

Lemma 3.4 *Mějme matice \mathbf{A} a \mathbf{B} jako ve větě 3.1 a přirozené číslo k splňující $1 \leq k \leq n$. Nechť dále $\bar{\mathbf{A}} = \mathbf{A}_{L_k \cdot L_k}$ a $\bar{\mathbf{B}}_{k \times k} = (\bar{b}_{ij})$ je matice daná předpisem*

$$\bar{b}_{ij} = \begin{cases} 1, & \text{pokud } i = j, \\ 0, & \text{pokud } i < j, \\ \sum_{I \in \mathcal{I}_{ij}} \pi_{\bar{\mathbf{A}}}(I), & \text{pokud } i > j. \end{cases} \quad (3.13)$$

Potom $\bar{\mathbf{B}} = \mathbf{B}_{L_k \cdot L_k}$.

Důkaz. Nejprve poznamenejme, že (3.13) odpovídá přesně (3.10), nahradíme-li b v (3.10) symbolem \bar{b} a \mathbf{A} symbolem $\bar{\mathbf{A}}$. Matice \mathbf{B} i $\bar{\mathbf{B}}$ jsou obě dolní trojúhelníkové s jednotkovou diagonálou, stačí proto dokázat rovnost prvků v oblasti pod hlavní diagonálou. Chceme tedy pro všechna i, j splňující $1 \leq j < i \leq k$ ověřit, že $b_{ij} = \bar{b}_{ij}$. Pro taková i, j je zřejmě $\mathcal{I}_{ij} \subseteq 2^{\{1, 2, \dots, k\}}$, a tedy výraz $\sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I)$ obsahuje pouze prvky z prvních k řádků a sloupců matice \mathbf{A} , neboli pouze prvky z matice $\mathbf{A}_{L_k \cdot L_k} = \bar{\mathbf{A}}$. Odtud ihned dostaneme

$$b_{ij} = \sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I) = \sum_{I \in \mathcal{I}_{ij}} \pi_{\bar{\mathbf{A}}}(I) = \bar{b}_{ij}, \quad 1 \leq j < i \leq k$$

a lemma je dokázáno. \square

Lemmat 3.2, 3.4 a 3.3 využijeme při důkazu věty 3.1, který nyní provedeme.

Důkaz věty 3.1. Dokazovat budeme matematickou indukcí podle rozměru matice $\mathbf{A}_{n \times n}$. Mějme nejprve $n = 1$. Potom nutně $\mathbf{A} = 0$, čili $(\mathbf{I} - \mathbf{A})^{-1} = 1$, a rovněž $\mathbf{B} = b_{11} = 1$, věta 3.1 tedy v tomto případě platí. Připomeňme rovněž, že platnost věty pro $n = 3$ jsme ukázali v příkladu 3.4.

Dokažme nyní indukční krok, tj. předpokládejme, že tvrzení věty platí pro rozměr $n = k$, a ukažme, že odtud plyne jeho platnost pro $n = k + 1$. Nechť tedy $\mathbf{A}_{(k+1) \times (k+1)}$ je libovolná matice koeficientů LRS a \mathbf{B} je matice vypočtená z \mathbf{A} podle (3.10). Matice \mathbf{A} a \mathbf{B} můžeme díky jejich specifickému tvaru vyjádřit blokově jako

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{L_k \cdot L_k} & \mathbf{0}_{k \times 1} \\ \mathbf{A}_{(k+1) \cdot L_k} & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{L_k \cdot L_k} & \mathbf{0}_{k \times 1} \\ \mathbf{B}_{(k+1) \cdot L_k} & 1 \end{pmatrix}. \quad (3.14)$$

Odtud ihned máme

$$\mathbf{I} - \mathbf{A} = \begin{pmatrix} \mathbf{I}_k - \mathbf{A}_{L_k \cdot L_k} & \mathbf{0}_{k \times 1} \\ -\mathbf{A}_{(k+1) \cdot L_k} & 1 \end{pmatrix},$$

a podle lemmatu o inverzi blokové matice A.2 můžeme vyjádřit matici $(\mathbf{I} - \mathbf{A})^{-1}$ blokově jako

$$(\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} (\mathbf{I}_k - \mathbf{A}_{L_k \cdot L_k})^{-1} & \mathbf{0}_{k \times 1} \\ \mathbf{A}_{(k+1) \cdot L_k} (\mathbf{I}_k - \mathbf{A}_{L_k \cdot L_k})^{-1} & 1 \end{pmatrix}. \quad (3.15)$$

Připomeňme, že naším cílem je ukázat rovnost matic \mathbf{B} a $(\mathbf{I} - \mathbf{A})^{-1}$, tuto rovnost ověříme po jednotlivých blocích. Nejprve ukážeme, že $\mathbf{B}_{L_k \cdot L_k} = (\mathbf{I}_k - \mathbf{A}_{L_k \cdot L_k})^{-1}$. Označme $\bar{\mathbf{A}} = \mathbf{A}_{L_k \cdot L_k}$. Potom zřejmě $\bar{\mathbf{A}}$ má tvar matice koeficientů LRS s k složkami, podle indukčního předpokladu je matice $\bar{\mathbf{B}}$ vypočtená podle (3.13) shodná s maticí $(\mathbf{I}_k - \bar{\mathbf{A}})^{-1}$ a podle lemmatu 3.4 je $\bar{\mathbf{B}} = \mathbf{B}_{L_k \cdot L_k}$. Celkem tedy máme

$$\mathbf{B}_{L_k \cdot L_k} = \bar{\mathbf{B}} = (\mathbf{I}_k - \bar{\mathbf{A}})^{-1} = (\mathbf{I}_k - \mathbf{A}_{L_k \cdot L_k})^{-1}.$$

Díky této rovnosti přepíšeme (3.15) jako

$$(\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} \mathbf{B}_{L_k \cdot L_k} & \mathbf{0}_{k \times 1} \\ \mathbf{A}_{(k+1) \cdot L_k} \mathbf{B}_{L_k \cdot L_k} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{B}_{L_k \cdot L_k} & \mathbf{0}_{k \times 1} \\ \mathbf{c} & 1 \end{pmatrix}, \quad (3.16)$$

přičemž jsme pro přehlednost zavedli vektor \mathbf{c} předpisem $\mathbf{c} = \mathbf{A}_{(k+1) \cdot L_k} \mathbf{B}_{L_k \cdot L_k}$. Tedy $\mathbf{c} = (c_1, c_2, \dots, c_k)$ je řádkový vektor, jehož j -tou složku lze zapsat jako

$$c_j = \sum_{i=1}^k a_{k+1,i} b_{ij}, \quad j = 1, 2, \dots, k. \quad (3.17)$$

Z porovnání (3.16) s blokovým zápisem matice \mathbf{B} v (3.14) plyne, že dokončení důkazu stačí ověřit, že $\mathbf{c} = \mathbf{B}_{(k+1) \cdot L_k}$, neboli že

$$c_j = b_{k+1,j}, \quad j = 1, 2, \dots, k.$$

Volme libovolné $j \in \{1, 2, \dots, k\}$. V následujících úpravách uijeme postupně vztah (3.17), definici prvků b_{ij} (3.10) (nejprve pro $i \leq j$, posléze pro $i > j$) a nakonec lemmata 3.2 a 3.3:

$$\begin{aligned} c_j &\stackrel{(3.17)}{=} \sum_{i=1}^k a_{k+1,i} b_{ij} = \\ &\stackrel{(3.10)}{=} a_{k+1,j} + \sum_{i=j+1}^k a_{k+1,i} b_{ij} = \\ &\stackrel{(3.10)}{=} \pi_{\mathbf{A}}(\{k+1, j\}) + \sum_{i=j+1}^k \left(a_{k+1,i} \cdot \sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I) \right) = \\ &= \pi_{\mathbf{A}}(\{k+1, j\}) + \sum_{i=j+1}^k \left(\sum_{I \in \mathcal{I}_{ij}} a_{k+1,i} \cdot \pi_{\mathbf{A}}(I) \right) = \\ &\stackrel{\text{lemma 3.2}}{=} \pi_{\mathbf{A}}(\{k+1, j\}) + \sum_{i=j+1}^k \left(\sum_{I \in \mathcal{I}_{ij}} \pi_{\mathbf{A}}(I \cup \{k+1\}) \right) = \\ &\stackrel{\text{lemma 3.3}}{=} \sum_{I \in \mathcal{I}_{k+1,j}} \pi_{\mathbf{A}}(I) = \\ &\stackrel{(3.10)}{=} b_{k+1,j}. \end{aligned}$$

Máme tedy $\mathbf{c} = \mathbf{B}_{(k+1) \cdot L_k}$, tudíž i $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ a indukční krok a s ním i celá věta jsou dokázány. \square

Věta 3.1 dává návod, jak vyjádřit prvky matice $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ pomocí hodnot cest v grafu LRS. Tvar matice \mathbf{B} nyní využijeme pro popis varianční matice LRS.

Věta 3.5 *Mějme lineární rekursivní systém \mathbf{X} s maticí koeficientů \mathbf{A} , chybovou varianční maticí $\mathbf{\Gamma} = \text{diag}(\gamma_1^2, \gamma_2^2, \dots, \gamma_n^2)$. Označme varianční matici vektoru \mathbf{X} jako $\mathbf{\Sigma} = (\sigma_{ij})$ a matici $\mathbf{B} = (b_{ij})$ zavedme podle (3.10). Potom platí:*

$$\sigma_{ij} = \sum_{k=1}^{\min(i,j)} \gamma_k^2 b_{ik} b_{jk}. \quad (3.18)$$

Důkaz. Věta 3.1 dává rovnost $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$. Použijeme-li (3.5), dostaneme

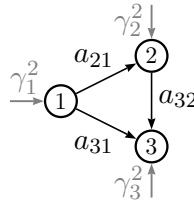
$$\mathbf{\Sigma} = \text{var} \mathbf{X} = \mathbf{B} \mathbf{\Gamma} \mathbf{B}^\top.$$

Označíme-li prvky matice $\mathbf{B}\mathbf{\Gamma}$ jako c_{ik} , pak díky diagonálnímu tvaru $\mathbf{\Gamma}$ platí $c_{ik} = \gamma_k^2 \cdot b_{ik}$. Odtud ihned dostaneme

$$\sigma_{ij} = \sum_{k=1}^n c_{ik} b_{jk} = \sum_{k=1}^n \gamma_k^2 b_{ik} b_{jk}.$$

Jelikož pro $k > i$ je $b_{ik} = 0$, můžeme upravit horní mez v přechodí sumě z n na $\min(i, j)$ a jsme hotovi. \square

Příklad 3.5 Uvažujme opět LRS \mathbf{X} z příkladu 3.2, jeho grafický model (tentokrát včetně ohodnocení vrcholů) vypadá následujícím způsobem:



Označme stejně jako v předchozí větě $\text{var } \mathbf{X} = \mathbf{\Sigma} = (\sigma_{ij})$ a $(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{B} = (b_{ij})$. Pomocí věty 3.5 spočítáme σ_{23} , tj. kovarianci veličin X_2 a X_3 . Podle (3.18) je

$$\begin{aligned} \sigma_{23} &= \sum_{k=1}^2 \gamma_k^2 b_{2k} b_{3k} = \\ &= \gamma_1^2 b_{21} b_{31} + \gamma_2^2 b_{22} b_{32} = \\ &= \gamma_1^2 a_{21} (a_{31} + a_{32} a_{21}) + \gamma_2^2 a_{32} = \\ &= \gamma_1^2 (a_{21} a_{31} + a_{21}^2 a_{32}) + \gamma_2^2 a_{32}. \end{aligned} \quad (3.19)$$

Součin $b_{ik} b_{jk}$ ze vzorce (3.18) má vcelku přehlednou grafickou interpretaci, kterou nyní popíšeme.

Lemma 3.6 Uvažujme LRS \mathbf{X} a matici $\mathbf{B} = (b_{ij})$ jako ve větě 3.5. Mějme dána čísla $i, j, k \in \{1, 2, \dots, n\}$ taková, že $k \leq \min(i, j)$. Potom výraz $b_{ik} b_{jk}$ představuje součet hodnot všech takových sledů, které vedou z i do j , neobsahují žádný kolizní vrchol a jejich vrchol s nejnižším očíslováním je k .

Důkaz. Uvažujme nejprve dvě orientované cesty: první vede z k do i a druhá z k do j . Symbolicky lze tyto cesty zapsat jako $k \rightarrow \dots \rightarrow i$ a $k \rightarrow \dots \rightarrow j$. Vynásobíme-li spolu hodnoty těchto cest, získáme vlastně hodnotu sledu, který vznikne spojením obou cest v jejich výchozím vrcholu, čili sledu $i \leftarrow \dots \leftarrow k \rightarrow \dots \rightarrow j$. Dodejme, že tato úvaha funguje i v případě, že některá z cest je nulové délky, např. když $i = k$; cesta nulové délky má totiž podle definice 3.1 hodnotu 1.

Vraťme se nyní k výrazu $b_{ik} b_{jk}$ z dokazovaného lemmatu. Připomeňme, že b_{ik} představuje součet hodnot všech orientovaných cest typu $k \rightarrow \dots \rightarrow i$, podobně b_{jk} sčítá hodnoty orientovaných cest typu $k \rightarrow \dots \rightarrow j$. Z předchozí diskuze plyne, že součin $b_{ik} b_{jk}$ pak vyjadřuje součet hodnot všech sledů typu

$$i \leftarrow \dots \leftarrow k \rightarrow \dots \rightarrow j. \quad (3.20)$$

Z topologického očíslování grafu $G(\mathbf{X})$ plyne, že vrchol k musí mít v takových sledech vždy nejnižší očíslování. Zřejmě žádný sled typu (3.20) neobsahuje kolizní vrchol, podobně každý sled z i do j bez kolizního vrcholu je nutně typu (3.20) (připomeňme, že jsme nevyloučili možnost, že $i = k$ nebo $j = k$). Všechny sledy typu (3.20) lze tedy charakterizovat jako sledy z i do j , které neobsahují kolizní vrchol a jejichž vrchol s nejnižším očíslováním je k . \square

Uvedené vztahy ilustruje výpočet v (3.19). Z porovnání druhého a posledního řádku výpočtu je patrné, že výraz $b_{21}b_{31}$ sčítá hodnoty sledů $2 \leftarrow 1 \rightarrow 3$ a $2 \leftarrow 1 \rightarrow 2 \rightarrow 3$, zatímco výraz $b_{22}b_{32}$ je hodnotou sledu $2 \rightarrow 3$.

Ukázali jsme, jak lze graficky interpretovat výraz $b_{ik}b_{jk}$. Sčítance v (3.18) jsou ovšem tvaru $\gamma_k^2 b_{ik}b_{jk}$. To vede k následující definici.

Definice 3.3 Nechť σ je sled v grafickém modelu LRS, označme vrchol sledu σ s nejnižším očíslováním jako k . *Rozšířenou hodnotou sledu σ* rozumíme součin hodnoty sledu σ a ohodnocení vrcholu k .

Pomocí rozšířených hodnot sledů můžeme jednoduchým způsobem formulovat pravidlo pro výpočet kovariancí v LRS z jeho grafického modelu.

Věta 3.7 Uvažujme LRS $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ a přirozená čísla $i, j \in \{1, 2, \dots, n\}$. Potom hodnota $\text{cov}(X_i, X_j)$ je rovna součtu rozšířených hodnot všech sledů, které vedou v $G(\mathbf{X})$ z i do j a neobsahují žádný kolizní vrchol.

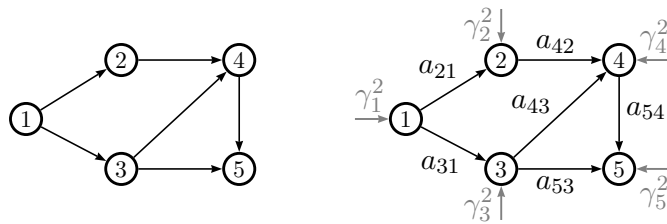
Důkaz. Podle věty 3.5 lze spočítat kovarianci X_i a X_j jako

$$\text{cov}(X_i, X_j) = \sum_{k=1}^{\min(i,j)} \gamma_k^2 b_{ik}b_{jk}. \quad (3.21)$$

Podle lemmatu 3.6 lze každý ze sčítanců na pravé straně interpretovat jako součet rozšířených hodnot všech takových sledů, které vedou v $G(\mathbf{X})$ z i do j , neobsahují žádný kolizní vrchol a jejich vrchol s nejnižším očíslováním je k . Zřejmě platí, že sledy z i do j mají hodnotu nejnižšího očíslování nižší nebo rovnu $\min(i, j)$, a suma v (3.21) tedy pokrývá všechny sledy mezi i a j bez kolizního vrcholu. \square

Podotkněme, že jsme v předchozí větě nevyloučili možnost, že $i = j$; v takovém případě věta popisuje vztah mezi $\text{var}X_i$ a sledy z i do i . Porovnáváme-li sledy z i do i , záleží na pořadí vrcholů ve sledu, tj. např. sledy $i \leftarrow l \leftarrow k \rightarrow i$ a $i \leftarrow k \rightarrow l \rightarrow i$ chápeme jako různé a při vyjádření hodnoty $\text{var}X_i$ započítáme rozšířené hodnoty obou z nich. Použití věty 3.7 ilustruje následující příklad.

Příklad 3.6 Uvažujme LRS \mathbf{X} s grafickým modelem (uvádíme pro přehlednost variantu grafu bez hodnocení i s ním)



Vypočítejme pomocí věty 3.7 kovarianci X_3 a X_4 a rozptyl X_4 . Sledy vedoucí z vrcholu 3 do vrcholu 4 bez kolizního vrcholu jsou

$$\begin{aligned} &3 \rightarrow 4, \\ &3 \leftarrow 1 \rightarrow 2 \rightarrow 4, \\ &3 \leftarrow 1 \rightarrow 3 \rightarrow 4, \end{aligned}$$

můžeme tedy psát

$$\text{cov}(X_3, X_4) = \gamma_3^2 a_{43} + \gamma_1^2 a_{31} a_{21} a_{42} + \gamma_1^2 a_{31}^2 a_{43}.$$

Sledy vedoucí z vrcholu 4 do vrcholu 4 bez kolizního vrcholu jsou

$$\begin{aligned}
 &4, \\
 &4 \leftarrow 2 \rightarrow 4, \\
 &4 \leftarrow 3 \rightarrow 4, \\
 &4 \leftarrow 2 \leftarrow 1 \rightarrow 2 \rightarrow 4, \\
 &4 \leftarrow 2 \leftarrow 1 \rightarrow 3 \rightarrow 4, \\
 &4 \leftarrow 3 \leftarrow 1 \rightarrow 2 \rightarrow 4, \\
 &4 \leftarrow 3 \leftarrow 1 \rightarrow 3 \rightarrow 4,
 \end{aligned}$$

což dává (po menší algebraické úpravě)

$$\text{var} X_4 = \gamma_4^2 + \gamma_3^2 a_{43}^2 + \gamma_2^2 a_{42}^2 + \gamma_1^2 (a_{21}^2 a_{42}^2 + 2a_{42} a_{21} a_{31} a_{43} + a_{31}^2 a_{43}^2).$$

Závěrem této kapitoly porovnáme postup pro nalezení vztahů mezi sledy a kovariancemi, který popisuje věta 3.7, s metodou koeficientů na cestách. Obě metody pracují podobným způsobem a mají podobný cíl: slouží k vyjádření vztahu mezi grafickým modelem funkčních vztahů v daném systému na jedné straně a příslušnými ukazateli míry lineární závislosti jeho veličin na straně druhé. Přesto se oba oba postupy v mnoha ohledech liší. Jejich rozdíly uvádí souhrnným způsobem tabulka 3.1.

Tabulka 3.1: Porovnání metody koeficientů na cestách (**A**) a postupu z věty 3.7 (**B**)

	A	B
Popisuje vztah mezi grafem a ...	korelačními koeficienty	kovariancemi
V grafu se pracuje s ...	cestami	sledy
Délka uvažovaných cest/sledů je ...	nenulová	libovolná
Graf má ohodnocené ...	hrany	hrany i vrcholy
Násobí se ohodnocení ...	hran	hran a 1 vrcholu

Závěr

Připomeňme ještě jednou cíle, které jsme vytyčili v úvodu práce. Při použití značení zavedeného později v textu je lze shrnout do následujících tří bodů:

1. Dokázat ekvivalenci $\mathcal{R}(G)$ a $\mathcal{M}(G)$ pro LRS.
2. Popsat souvislost mezi grafickým modelem ZLRS a jeho strukturou podmíněných nezávislostí.
3. Popsat způsob, jak z grafického modelu LRS vyčíst kovariance mezi jeho veličinami, a porovnat jej s metodou koeficientů na cestách.

Bod 1 byl v původním zadání bakalářské práce označen za její hlavní cíl. Ačkoli se tvrzení, které je předmětem důkazu v bodu 1, všeobecně považuje v oblasti lineárních rekursivních systémů za korektní, jeho kompletní rigorózní důkaz v literatuře doposud chybí. Hlavním smyslem první části práce bylo proto takový důkaz najít a v kompaktní podobě jej předložit. Tvrzení z bodu 1 jsme formulovali v podobě věty 2.8 a k jeho důkazu směřuje celý oddíl 2.1, jsou zde nicméně podstatně využity i výsledky z kapitoly 1 a příloh A a B.

Všechny důkazy v těchto i dalších částech práce jsme se snažili v maximální možné míře řešit odkazem na existující literaturu; je-li tedy důkaz v textu proveden, jde o důkaz vlastní (s výjimkou důkazu věty 1.1; důvod pro zařazení tohoto důkazu je v textu vysvětlen). Podotkněme, že v případě některých důkazů z kapitoly 1 a přílohy A jde často o jednoduchá a/nebo známá tvrzení, která se nám ovšem nepodařilo najít v potřebném znění (a s důkazem) v existujících pramenech.

Bod 2 se na rozdíl od bodu 1 v původním zadání práce nevyskytoval a představuje zobecnění některých výsledků, které jsme dokázali pro (gaussovsky rozdělené) LRS, i mimo rámec gaussovských distribucí. Přínos této části práce lze tedy označit za „invenční“. Podařilo se nám ukázat, že zachováme-li spojitost chybových složek, můžeme i při opuštění předpokladu jejich normality konstatovat, že výsledná struktura podmíněné nezávislosti splňuje párovou markovskou vlastnost (PM) vzhledem ke grafickému modelu daného rekursivního systému. Toto tvrzení je v zásadě obsahem věty 2.16, která završila oddíl 2.2; věta je navíc formulována ve tvaru ekvivalence, což dále zajišťuje, že se v daném ZLRS nevyskytují jiné podmíněné nezávislosti párově markovského typu¹, než které lze vyčíst z jeho grafu. Jak jsme dále poznamenali, za dodatečných předpokladů o striktně kladné hustotě chybových složek lze na základě věty 1.7 ztotožnit globální a párovou markovskou vlastnost, a ZLRS pak automaticky splňují i (GM) vzhledem ke svému grafickému modelu.

Bodem 3 jsme se zabývali ve třetí kapitole. Výpočet kovariancí v LRS na základě grafického modelu popisuje věta 3.7. Uvedený postup se v mnoha ohledech podobá metodě koeficientů na cestách, kterou jsme stručně popsali v původním Wrightově pojetí v oddílu 3.1. Obě metody mají nicméně řadu odlišností, které jsme shrnuli v závěru třetí kapitoly (jejich stručný výčet uvádí tabulka 3.1).

¹ Tj. nezávislosti typu $i \perp\!\!\!\perp j \mid \{1, 2, \dots, i-1\} \setminus \{j\}$.

Ačkoli jsme se ve třetí kapitole zabývali vztahem mezi grafem a kovariancemi pouze pro případ LRS, je nasnadě, že zcela identickým způsobem lze postupovat i v případě ZLRS – předpoklad normality nebyl v tvrzeních oddílu 3.2 nijak využit. Uvedený vztah by bylo dále možné zkoumat i pro širší třídu systémů náhodných veličin, nežli jsou LRS nebo ZLRS; mohli bychom například připustit, aby složky chybového vektoru byly korelované, podobně jako je tomu v případě metody koeficientů na cestách. Takové rozšíření je však již nad rámec této bakalářské práce.

Příloha A

Některé věty o maticích

V této příloze jsou uvedena některá tvrzení o maticích, která jsou využita v hlavní části práce. Připomeňme, že v úvodu práce (v kapitole *Použitá značení*) je definováno značení pro výběr podmatic (příp. podvektorů) pomocí indexových množin, a zápis některých matic s prvky pouze nula nebo jedna. Dodejme, že všechny matice, kterými se zde budeme zabývat, jsou výhradně matice reálné.

Lemma A.1 *Nechť $\mathbf{A}_{n \times n} = (a_{ij})$ je dolní trojúhelníková matice. Potom platí:*

- (i) $\det \mathbf{A} = \prod_{i=1}^n a_{ii}$,
- (ii) \mathbf{A} je regulární právě tehdy, když jsou všechny prvky na její hlavní diagonále nenulové,
- (iii) je-li \mathbf{A} regulární, potom její inverzní matice \mathbf{A}^{-1} je dolní trojúhelníková.
- (iv) má-li \mathbf{A} jednotkovou diagonálu (tj. všechny prvky na hlavní diagonále jsou rovny jedné), potom její inverzní matice \mathbf{A}^{-1} je dolní trojúhelníková s jednotkovou diagonálou.

Důkaz. Tvrzení (i) je patrné z výpočtu determinantu rozvojem podle řádku (vždy podle prvního řádku aktuální matice), (ii) je přímým důsledkem (i), neboť \mathbf{A} je regulární právě tehdy, když $\det \mathbf{A} \neq 0$. Bod (iii) plyne např. z (i) a ze známého pravidla pro výpočet inverzní matice pomocí determinantů (viz např. [Bic02, věta 8.12]): označíme-li $\mathbf{A}^{-1} = (b_{ij})$, můžeme prvky b_{ij} inverzní matice spočítat jako

$$b_{ij} = \frac{(-1)^{i+j} \det \mathbf{A}_{\setminus ji}}{\det \mathbf{A}}, \quad (\text{A.1})$$

kde $\mathbf{A}_{\setminus ji}$ označuje matici, která vznikne z \mathbf{A} vynecháním j -tého řádku a i -tého sloupce. Lze snadno ukázat, že pro $j > i$ je $\mathbf{A}_{\setminus ji}$ dolní trojúhelníková, jejíž i -tý diagonální prvek je nulový, podle (ii) je pak $\det \mathbf{A}_{\setminus ji} = 0$; celkem tedy $b_{ij} = 0$ pro $j > i$. Bod (iv) ukážeme podobným způsobem. Předně podle (ii) je \mathbf{A} regulární a podle (iii) je \mathbf{A}^{-1} dolní trojúhelníková. Označme opět $\mathbf{A}^{-1} = (b_{ij})$. Jelikož \mathbf{A} má jednotkovou diagonálu, podle (i) platí $\det \mathbf{A} = 1$ a $\det \mathbf{A}_{\setminus ii} = 1$ pro všechna i . Tvrzení o jednotkové diagonále \mathbf{A}^{-1} potom plyne ihned z (A.1). \square

Lemma A.2 (O inverzi blokové matice.) *Nechť \mathbf{B} je regulární matice typu $m \times m$ a \mathbf{C} je matice typu $n \times m$. Definueme matici \mathbf{A} blokovým zápisem*

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{0}_{m \times n} \\ \mathbf{C} & \mathbf{I}_n \end{pmatrix}.$$

Potom \mathbf{A} je regulární a její inverzní matice \mathbf{A}^{-1} má tvar

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0}_{m \times n} \\ -\mathbf{CB}^{-1} & \mathbf{I}_n \end{pmatrix}. \quad (\text{A.2})$$

Důkaz. Vynásobením ověříme, že matice ze vztahu (A.2) je inverzní k \mathbf{A} :

$$\mathbf{AA}^{-1} = \begin{pmatrix} \mathbf{B} & \mathbf{0}_{m \times n} \\ \mathbf{C} & \mathbf{I}_n \end{pmatrix} \cdot \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0}_{m \times n} \\ -\mathbf{CB}^{-1} & \mathbf{I}_n \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{I}_n \end{pmatrix} = \mathbf{I}_{m+n},$$

matice \mathbf{A} je tudíž invertovatelná, tedy i regulární. \square

Lemma A.3 *Nechť symetrická matice $\mathbf{A}_{n \times n}$ je pozitivně definitní. Pak \mathbf{A} je regulární.*

Důkaz. Je-li \mathbf{A} singulární, potom jsou její řádky lineárně závislé, tedy existuje existuje vektor $\mathbf{0} \neq \mathbf{c} \in \mathbb{R}_n$ takový, že $\mathbf{c}^\top \mathbf{A} = \mathbf{0}$. Potom rovněž $\mathbf{c}^\top \mathbf{A} \mathbf{c} = 0$, takže \mathbf{A} není pozitivně definitní. \square

Poznamenejme, že obdobným způsobem bychom mohli dokázat, že regulární pozitivně semidefinitní matice je již nutně pozitivně definitní, neboli že regulární varianční matice jsou pozitivně definitní. Právě poznatky o regulárních variančních maticích jsou klíčové pro studium některých vlastností lineárních rekursivních systémů v kapitole 2. Následují dvě věty o rozkladu pozitivně semidefinitních a definitních matic: první je větou o skeletním rozkladu pozitivně semidefinitní matice, druhá v pořadí je známá věta o *Choleského rozkladu* pozitivně definitní matice; běžně se tato věta formuluje pro hermitovské komplexní matice, pro naše účely postačí její „reálná“ verze.

Věta A.4 (O skeletním rozkladu pozitivně definitní matice.) *Nechť $\mathbf{A}_{n \times n}$ je symetrická pozitivně semidefinitní matice, $h(\mathbf{A}) = r \geq 1$. Potom existuje matice $\mathbf{B}_{n \times r}$ taková, že $h(\mathbf{A}) = r$ a platí $\mathbf{A} = \mathbf{BB}^\top$.*

Důkaz. Viz [And07, věta A.7].

Věta A.5 (Choleského rozklad.) *Nechť $\mathbf{A}_{n \times n}$ je symetrická pozitivně definitní matice. Potom existuje dolní trojúhelníková matice $\mathbf{B}_{n \times n}$ taková, že $\mathbf{A} = \mathbf{BB}^\top$.*

Důkaz. Viz [Wat02, věta 1.4.7]. \square

Věta A.6 (O rozkladu pozitivně definitní matice.) *Nechť $\mathbf{A}_{n \times n}$ je symetrická pozitivně definitní matice. Potom existují matice $\mathbf{C}_{n \times n}$ a $\mathbf{D}_{n \times n}$ takové, že*

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{21} & 1 & 0 & \dots & 0 \\ c_{31} & c_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n2} & \dots & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix},$$

kde $d_i > 0$ pro $i = 1, 2, \dots, n$, a přitom platí $\mathbf{A} = \mathbf{CDC}^\top$.

Důkaz. Věta o Choleského rozkladu zaručuje existenci dolní trojúhelníkové matice $\mathbf{B}_{n \times n} = (b_{ij})$ splňující $\mathbf{A} = \mathbf{B}\mathbf{B}^\top$. Pro hodnotu matice \mathbf{B} zřejmě platí $h(\mathbf{B}) \geq h(\mathbf{B}\mathbf{B}^\top) = h(\mathbf{A})$. Podle lemmatu A.3 je matice \mathbf{A} regulární, tedy $h(\mathbf{A}) = n$, tudíž je regulární i matice \mathbf{B} . Podle lemmatu A.1 jsou její diagonální prvky nenulové. Můžeme proto definovat matice \mathbf{M} a \mathbf{D} předpisem

$$\mathbf{M} = \begin{pmatrix} \frac{1}{b_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{b_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{b_{nn}} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} b_{11}^2 & 0 & \dots & 0 \\ 0 & b_{22}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_{nn}^2 \end{pmatrix},$$

diagonální prvky matice \mathbf{D} budou zřejmě kladné. Dále platí $\mathbf{D} = \mathbf{M}^{-1}(\mathbf{M}^{-1})^\top$, a tedy $\mathbf{B}\mathbf{M}\mathbf{D}\mathbf{M}^\top\mathbf{B}^\top = \mathbf{A}$. Označíme-li $\mathbf{C} = \mathbf{B}\mathbf{M}$, dostaneme vztah $\mathbf{C}\mathbf{D}\mathbf{C}^\top = \mathbf{A}$ a snadno se přesvědčíme, že matice \mathbf{C} má tvar, který vyžaduje od stejnojmenné matice dokazovaná věta. \square

Slevíme-li z požadavku, aby prvky matice \mathbf{D} byly nezáporné, je možné dokázat předchozí větu i pro širší třídu symetrických matic, nežli jsou matice pozitivně definitní, viz [Wat02, věta 1.7.30]. Pak ale pro důkaz samozřejmě nelze využít Choleského rozklad. Vzhledem k tomu, že budeme používat větu A.6 pouze k rozkladu regulárních variančních matic, není zde třeba její obecnější formu uvádět.

Lemma A.7 *Uvažujme symetrickou pozitivně definitní matici $\mathbf{A}_{n \times n}$ a neprázdnou indexovou množinu $J \subseteq \{1, 2, \dots, n\}$. Potom $\mathbf{A}_{J,J}$ je pozitivně definitní.*

Důkaz. Označme \mathbf{I} jednotkovou matici typu $n \times n$ a matici \mathbf{M} zavedme předpisem $\mathbf{M} = \mathbf{I}_{J \cdot \{1, 2, \dots, n\}}$. Vynásobením lze snadno ověřit, že $\mathbf{A}_{J,J} = \mathbf{M}\mathbf{A}\mathbf{M}^\top$. Uvažujme libovolný vektor $\mathbf{b} \in \mathbb{R}_{|J|}$ a označme $\mathbf{c} = \mathbf{M}^\top\mathbf{b}$. Potom platí:

$$\mathbf{b}^\top\mathbf{A}_{J,J}\mathbf{b} = \mathbf{b}^\top\mathbf{M}\mathbf{A}\mathbf{M}^\top\mathbf{b} = \mathbf{c}^\top\mathbf{A}\mathbf{c} > 0,$$

tedy matice $\mathbf{A}_{J,J}$ je pozitivně definitní. \square

Věta A.8 (O determinantu pozitivně definitní matice.) *Uvažujme symetrickou pozitivně definitní matici $\mathbf{A}_{n \times n}$ a neprázdné disjunktní indexové množiny I a J takové, že $I \cup J = \{1, 2, \dots, n\}$. Pak platí:*

$$\det \mathbf{A} = \det \mathbf{A}_{J,J} \cdot \det(\mathbf{A}_{I,I} - \mathbf{A}_{I,J}\mathbf{A}_{J,J}^{-1}\mathbf{A}_{J,I}).$$

Důkaz. Vzhledem k tomu, že při symetrické permutaci řádků a sloupců se determinant nezmění, lze bez újmy na obecnosti předpokládat, že $I = \{1, 2, \dots, k\}$, $J = \{k+1, k+2, \dots, n\}$. Matice \mathbf{A} má potom tvar

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{I,I} & \mathbf{A}_{I,J} \\ \mathbf{A}_{J,I} & \mathbf{A}_{J,J} \end{pmatrix}.$$

Podle lemmatu A.7 je matice $\mathbf{A}_{J,J}$ pozitivně definitní, tedy i invertovatelná. Definujeme-li matice \mathbf{B} a \mathbf{C} jako

$$\mathbf{B} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{I,J}\mathbf{A}_{J,J}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{A}_{I,I} - \mathbf{A}_{I,J}\mathbf{A}_{J,J}^{-1}\mathbf{A}_{J,I} & \mathbf{0} \\ \mathbf{A}_{J,I} & \mathbf{A}_{J,J} \end{pmatrix},$$

můžeme se snadno přesvědčit, že $\mathbf{BA} = \mathbf{C}$. Podle věty o rozvoji determinantu platí

$$\det \mathbf{B} = 1, \quad \det \mathbf{C} = \det \mathbf{A}_{J,J} \cdot \det(\mathbf{A}_{I,I} - \mathbf{A}_{I,J} \mathbf{A}_{J,J}^{-1} \mathbf{A}_{J,I}).$$

Je tedy $\det \mathbf{C} = \det(\mathbf{BA}) = \det \mathbf{B} \cdot \det \mathbf{A} = \det \mathbf{A}$, odkud již tvrzení věty zřejmě plyne. \square

Poznamenejme, že pro důkaz předchozí věty nebylo podstatné, že matice \mathbf{A} je pozitivně definitní. Jediné, k čemu byla tato vlastnost potřeba, byl její důsledek ohledně regularity matice $\mathbf{A}_{J,J}$. Tvrzení by šlo tedy příslušným způsobem zobecnit, pro naše potřeby to ovšem nemá význam.

Příloha B

Některé věty o $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

V této příloze uvedeme některá tvrzení o vícerozměrném normálním rozdělení, která jsou základními stavebními kameny pro studium lineárních rekursivních systémů.

Definice B.1 Řekneme, že náhodná veličina X má *nedegenerované normální rozdělení* s parametry $\mu \in \mathbb{R}, \sigma^2 > 0$, pokud je její hustota f vzhledem k Lebesgueově míře λ dána vztahem

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad \text{s.v. } [\lambda]. \quad (\text{B.1})$$

Píšeme potom $X \sim \mathcal{N}(\mu, \sigma^2)$. V celé práci se přitom omezíme pouze na spojité verze hustot z (B.1). Mezi normální rozdělení lze počítat i situace, kdy veličina X nabývá hodnoty μ s pravděpodobností jedna, potom hovoříme o *degenerovaném normálním rozdělení* a značíme $X \sim \mathcal{N}(\mu, 0)$. Mějme dále náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$. Nechť $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^\top$ a $\boldsymbol{\Sigma}_{n \times n}$ je symetrická pozitivně semidefinitní matice. Řekneme, že vektor \mathbf{X} má *n -rozměrné normální rozdělení* s parametry $\boldsymbol{\mu}$ a $\boldsymbol{\Sigma}$, jestliže pro libovolný vektor $\mathbf{c} \in \mathbb{R}_n$ platí

$$\mathbf{c}^\top \mathbf{X} \sim \mathcal{N}(\mathbf{c}^\top \boldsymbol{\mu}, \mathbf{c}^\top \boldsymbol{\Sigma} \mathbf{c}). \quad (\text{B.2})$$

Značíme potom $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Řekneme dále, že toto rozdělení je *regulární*, je-li matice $\boldsymbol{\Sigma}$ regulární, v opačném případě hovoříme o rozdělení *singulárním*. Množinu všech regulárních normálních rozdělení budeme značit \mathcal{N}_{reg} .

Lemma B.1 Mějme pro $i = 1, 2, \dots, n$ je $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Nechť dále náhodné veličiny X_i a X_j jsou nezávislé pro $i \neq j$ a je dáno $a, b \in \mathbb{R}$. Potom platí:

- (i) $a + bX_1 \sim \mathcal{N}(a + b\mu_1, b^2\sigma_1^2)$.
- (ii) $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- (iii) $\sum_{i=1}^n X_i \sim \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$.
- (iv) Vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ má rozdělení $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^\top$ a $\boldsymbol{\Sigma} = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2\}$.

Důkaz. Viz [And07], věta 4.1 a lemmata 4.6, 4.7 a 4.8. □

Lemma B.2 Mějme $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{n \times n})$. Potom platí:

- (i) $\mathbf{E}\mathbf{X} = \boldsymbol{\mu}, \text{var}\mathbf{X} = \boldsymbol{\Sigma}$.
- (ii) Je-li $\mathbf{Y} = \mathbf{a}_{m \times 1} + \mathbf{B}_{m \times n}\mathbf{X}$, pak $\mathbf{Y} \sim \mathcal{N}(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top)$.
- (iii) Nechť $J \subset \{1, 2, \dots, n\}$. Pak marginální rozdělení vektoru \mathbf{X}_J je $\mathcal{N}(\boldsymbol{\mu}_J, \boldsymbol{\Sigma}_{J,J})$.

(iv) Nechť $h(\boldsymbol{\Sigma}) = r$. Pak existuje matice $\mathbf{B}_{n \times r}$ taková, že $\mathbf{X} = \boldsymbol{\mu} + \mathbf{B}\mathbf{U}$, kde $\mathbf{U} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_r)$.

Důkaz. Body (i), (ii) a (iii) odpovídají větám 4.2, 4.4 a 4.5 z [And07]. Ukažme nyní bod (iv). Z věty A.4 o skeletním rozkladu pozitivně definitní matice existuje matice $\mathbf{B}_{n \times r}$ splňující $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}^\top$. Potom podle (iii) platí

$$\boldsymbol{\mu} + \mathbf{B}\mathbf{U} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{B}\mathbf{I}\mathbf{B}^\top) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \quad \square$$

Věta B.3 (O hustotě vícerozměrného normálního rozdělení.) Nechť náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ má regulární normální rozdělení $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Potom jeho hustota pravděpodobnosti vzhledem k Lebesgueově míře na \mathbb{R}_n má tvar

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \boldsymbol{\Sigma})^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}. \quad (\text{B.3})$$

Důkaz. Věta je dokázána v [And07], věta 4.10.

Lemma B.4 Mějme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ takový, že $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, a disjunktní indexové množiny I a J , kde $I \cup J \subseteq \{1, 2, \dots, n\}$. Potom platí: $\boldsymbol{\Sigma}_{I,J} = \mathbf{0}$ právě tehdy, když jsou náhodné vektory \mathbf{X}_I a \mathbf{X}_J nezávislé.

Důkaz. Jde o bezprostřední důsledek vět 4.5 a 4.11 z [And07]. □

Věta B.5 (O podmíněném rozdělení gaussovského vektoru.) Nechť \mathbf{X} má regulární rozdělení $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, I a J nechť jsou opět disjunktní indexové množiny splňující $I \cup J \subseteq \{1, 2, \dots, n\}$. Potom podmíněné rozdělení vektoru \mathbf{X}_I při dané hodnotě vektoru $\mathbf{X}_J = \mathbf{x}_J$ je regulární rozdělení

$$\mathcal{N}(\boldsymbol{\mu}_I + \boldsymbol{\Sigma}_{I,J} \boldsymbol{\Sigma}_{J,J}^{-1} (\mathbf{x}_J - \boldsymbol{\mu}_J), \boldsymbol{\Sigma}_{I,I} - \boldsymbol{\Sigma}_{I,J} \boldsymbol{\Sigma}_{J,J}^{-1} \boldsymbol{\Sigma}_{J,I}).$$

Důkaz. Jedná se o větu analogickou, jako je věta 4.12 z [And07]. □

Tvrzení předchozí věty o *podmíněném rozdělení* (což je pojem, který jsme v této práci nedefinovali) lze chápat tak, že podmíněnou hustotu $f_{I|J}(\mathbf{x}_I | \mathbf{x}_J)$ lze vyjádřit ve tvaru hustoty vícerozměrného normálního rozdělení (B.3), pokud dosadíme v (B.3) za $\boldsymbol{\mu}$ výraz

$$\boldsymbol{\mu}_I + \boldsymbol{\Sigma}_{I,J} \boldsymbol{\Sigma}_{J,J}^{-1} (\mathbf{x}_J - \boldsymbol{\mu}_J)$$

a za $\boldsymbol{\Sigma}$ výraz

$$\boldsymbol{\Sigma}_{I,I} - \boldsymbol{\Sigma}_{I,J} \boldsymbol{\Sigma}_{J,J}^{-1} \boldsymbol{\Sigma}_{J,I}. \quad (\text{B.4})$$

Literatura

- [And99] ANDERSSON S., MADIGAN, D., PERLMAN, M. D., RICHARDSON, T. S. (1999): Graphical Markov models in multivariate analysis. A chapter in *Multivariate Analysis, Design of Experiments, and Survey Sampling* (S. Ghosh ed.). Marcel Dekker, New York. Kapitola 8, 187–230.
- [And74] ANDERSON, J., EVANS, F. (1974): Causal models in educational research: recursive models. *American Educational Research Journal*, **11**(1), 29–39.
- [And07] ANDĚL, J. (2007): *Základy matematické statistiky*. MatfyzPress, Praha.
- [Bic02] BICAN, L. (2002): *Lineární algebra a geometrie*. Academia, Praha.
- [Cow07] COWELL, R. G., DAWID A. P., LAURITZEN, S. L., SPIEGELHALTER, D. J. (2007): *Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks*. Springer, New York.
- [Den06] DENIS, D. J., LEGERSKI, J. (2006): Causal modeling and the origins of path analysis. *Theory & Science* [online], **7**(1), [cit. 2010-11-20]. Dostupný z WWW: <<http://theoryandscience.icaap.org/content/vol7.1/denis.html>>
- [Eat07] EATON, M. L. (2007): *Multivariate Statistics: A Vector Space Approach*. Institute of Mathematical Statistics, Beachwood.
- [Gol73] GOLDBERGER, A., DUNCAN, O. (1973): *Structural Equation Models in the Social Sciences*. Seminar Press, New York.
- [Hod89] HODAPP, V. (1989): Anxiety, fear of failure, and achievement: Two path-analytical models. *Anxiety, Stress & Coping*, **1**(4), 301–312.
- [Li75] LI, C. (1975): *Path Analysis: A Primer*. The Boxwood Press, Pacific Grove.
- [Luk02] LUKEŠ, J., MALÝ, J. (2005): *Measure and Integral*. MatfyzPress, Praha.
- [Lau96] LAURITZEN, S. L. (1996): *Graphical Models*. Clarendon Press, Oxford.
- [Pug03] PUGESEK, B. H., TOMER, A., VON EYE, A. (2003): *Structural Equation Modeling: Applications in Ecological and Evolutionary Biology*. Cambridge University Press, Cambridge.
- [Rud03] RUDIN, W. (2003): *Analýza v reálném a komplexním oboru*. Academia, Praha.
- [Stu02] STUDENÝ, M. (2002): O použití řetězcových grafů pro popis struktur podmíněné nezávislosti. In *ROBUST'2002, sborník prací dvanácté zimní školy JČMF* (J. Antoch, G. Dohnal, J. Klaschka ed.). JČMF, Praha, 292–314.

- [Stu98] STUDENÝ, M., VEJNAROVÁ, J. (1998): The multiinformation function as a tool for measuring stochastic dependence. In *Learning in Graphical Models* (M. I. Jordan ed.). Kluwer Academic Publishers, Boston, 261–298.
- [Ště87] ŠTĚPÁN, J. (1987): *Teorie pravděpodobnosti*. Academia, Praha.
- [Wat02] WATKINS, D. S. (2002): *Fundamentals of Matrix Computations*. John Wiley & Sons, New York.
- [Wer80] WERMUTH, N. (1980): Linear recursive equations, covariance selection, and path analysis. *Journal of the American Statistical Association*, **75**(372), 963–972.
- [Wri18] WRIGHT, S. (1918): On the nature of size factors. *Genetics*, **3**, 367–374.
- [Wri20] WRIGHT, S. (1920): The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proceedings of the National Academy of Sciences*, **6**, 320–332.
- [Wri21] WRIGHT, S. (1921): Correlation and causation. *Journal of Agricultural Research*, **20**, 557–585.
- [Wri34] WRIGHT, S. (1934): The method of path coefficients. *Annals of Mathematical Statistics*, **5**, 161–215.
- [Wri60] WRIGHT, S. (1960): Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, **16**(2), 189–202.