Obsah

Příloha A

Pokus o „historický“ přehled

A.1 Komentovaný přehled literatury o PN

Následuje jakýsi pokus o přehled literatury o PN a grafických modelech ve spojení s historickým kontextem.

A.1.1 Podmíněná nezávislost

(Loève 1955) Již v 50. letech minulého století Loève z francouzské školy pravděpodobnostníků ve své knize definuje PN pro σ-algebry.

(Dawid 1979) Angličan Phil Dawid byl patrně prvním statistikem, který explicitně formuloval určité formální vlastnosti stochastické PN. Hlavním sdělením jeho článku z roku 1979 bylo pozorování, že mnohé statistické pojmy, jako například pojem *postačující statistiky* (anglicky *sufficient statistic*), či pojem tzv. *podružné statistiky* (anglicky *anciliary statistic*), lze ekvivalentně definovat v termínech zobecněné PN. To pak umožňuje odvodit některé tradiční výsledky mnohem elegantněji, právě s využitím těchto formálních vlastností PN.

(Spohn 1980) Nezávisle na Dawidovi formuloval nedlouho po něm tyto formální vlastnosti PN také Němec Spohn, kterého zajímaly z hlediska filosofické logiky. Spohn se zabýval spíše interpretací pojmu PN a jeho souvislostí s kauzalitou.

(Mouchart, Rolin 1984) Tytéž vlastnosti, formulované ovšem v termínech σ-algeber, použili později Belgičané Mouchart a Rolin. Jedná se o statistika a pravděpodobnostníka z frankofonní části Belgie; přirozeně tedy byli ovlivněni francouzskou pravděpodobnostní školou. Jak jsem se dozvěděl později, Dawid s Mouchartem se setkali koncem 70. let dvacátého století a podle Moucharta spolu tenkrát vymysleli symbol $\perp\!\!\!\perp$ pro PN, který se později ujal.

(Florens, Mouchart, Rolin 1990) Většinu obsahu onoho článku z roku 1984 pak zařadili do pozdější knihy o bayesovské statistice.

(van Putten, van Shuppen 1985) Nedlouho po Mouchartovi a Rolinovi publikovali v *Annals of Probability* článek van Putten a van Shuppen z Holandska.

Výsledky byly velmi podobné těm z článku Moucharta a Rolina, kteří v tomtéž roce 1985 publikovali noticku ve formě dopisu editorovi, kde tvrdí, že až na dvě tvrzení je vše, co je v článku van Puttena a van Shuppena, analogické tomu, co vyšlo v jejich vlastním článku z roku 1984. Podle toho, co mi naznačil Mouchart při rozhovoru, se totiž domnívali, že van Putten a van Shuppen přepsali obsah jejich článku – prý jim dva roky předtím poskytli preprint.

(Pearl, Paz 1985) Význam pojmu stochastické PN pro pravděpodobnostní rozhodování rozpoznali a zdůraznili v půli 80. let dvacáteho století Američan židovského původu Judea Pearl a Azaria Paz z Izraele. Rovněž explicitně formulovali stejné formální vlastnosti PN jako Dawid a zároveň si povšimli, že tytéž vlastnosti platí pro určité ternární relace indukované neorientovanými grafy pomocí jistého separačního kritéria. To je vedlo k myšlence popisovat tyto ternární relace s pomocí grafů.

Pearl a Paz tak zavedli pojem *semigrafoidu* jakožto formální ternární relaci axiomaticky splňující dříve zmíněné formální vlastnosti stochastické PN. V původním článku ale používali termín grafoid. Rovněž formulovali hypotézu, že tyto formální vlastnosti charakterizují struktury pravděpodobnostní PN, tedy že semigrafoidy se shodují s třídou ternárních relací indukovaných pravděpodobnostními rozděleními přes pojem PN.

A.1.2 Grafy

Myšlenka použít pro popis kvalitativních vztahů mezi veličinami grafy, jejichž uzly odpovídají náhodným veličinám, je však mnohem starší. Pearl a Paz na ni pouze v jistém smyslu navázali. Tradičně lze rozlišit dva směry:

- použití neorientovaných grafů,
- použití acyklických orientovaných grafů.

Dodatečně se ukázalo, že mnohé statistické modely popisované takto grafy lze chápat jako modely struktur PN.

Neorientované grafy

Neorientované grafy se objevily ve statistické fyzice, konkrétně v teorii markovských polí, coby nástroj pro popis vztahů mezi diskrétními náhodnými veličinami.

(Moussouris 1974) Moussouris zavedl několik markovských podmínek vůči neorientovanému grafu a dokázal jejich ekvivalenci s faktorizační podmínkou nazývanou gibbsovskou podmínkou. Tento výsledek, známý jako Hammersleyova-Cliffordova věta, byl již předtím dokázán v nepublikovaném rukopise z roku 1971 zmíněnými dvěma pány.

(Speed 1979) Další, kdo v této oblasti bádal, byl Australan Terry Speed, který později ovlivnil Dána Steffena Lauritzena v tom smyslu, že ten si uvědomil, že grafy lze využít k popisu statistických modelů používaných v teorii kontingenčních tabulek.

(Goodman 1973) (Haberman 1974) K rozvoji teorie kontingenčních tabulek, včetně grafických přístupů, přispěl mnohými publikacemi Goodman. Základy této teorie shrnul ve své disertaci Haberman.

(Darroch, Lauritzen, Speed 1980) Darroch, Lauritzen a Speed přišli se speciálními grafickými log-lineárními modely.

(Lauritzen 1982) Vhodným učebním textem, který tento přístup zpopularizoval a vyzdvihnul roli PN, se stal přepis Lauritzenových přednášek na univerzitě v Aalborgu.

(Dempster 1972) (Wermuth 1976) Neorientované grafy se také paralelně objevily i v mnohorozměrné statistické analýze. Američan kanadského původu Artur Dempster zavedl takzvané *covariance selection models*. Jeho studentka Nanny Wermuth(ová), původem z Německa, pak tyto modely interpretovala v termínech PN.

(Whittaker 1990) Angličan Joe Whittaker ve své knize pak o těchto statistických modelech ukázal, že se dají dobře popsat s pomocí neorientovaných grafů.

Orientované grafy

Orientované grafy se v literatuře v souvislosti s popisem vztahů mezi náhodnými veličinami objevily dokonce dříve než neorientované grafy.

(Wright 1934) Genetik Sewall Wright již ve třicátých letech dvacátého století navrhl určitou metodu výpočtu koeficientů korelace mezi gaussovskými náhodnými veličinami, které splňují nějaký systém strukturálních rovnic. Tato metoda, údajně prý zpočátku ignorovaná statistiky, byla později proslavena díky ekonomům, psychologům a sociologům coby *metoda koeficientů na cestách* (anglicky *method of path coefficients*). Pro reprezentaci zmíněného systému strukturálních rovnic použil Wright grafy, které měly jak jednosměrné, tak obousměrné šipky.

Použití acyklických orientovaných grafů přišlo později v souvislosti s *influenčními diagramy* v teorii rozhodování.

(Howard, Matheson 1984) Influenční diagramy zavedl Howard a Matheson coby nástroje na grafický popis rozhodovacích situací, tedy v podstatě vztahů mezi náhodnými veličinami.

(Shachter 1986) To inspirovalo Shachtera, který se v 80. letech minulého století systematicky věnoval influenčním diagramům, aby přeformuloval své algoritmy pro grafickou manipulaci s influenčními diagramy v termínech PN.

(Smith 1989) Angličan Jim Smith, mladší kolega Phila Dawida, si povšiml možnosti využití dříve zmíněných semigrafoidových vlastností PN pro snadnější důkaz korektnosti jistých operací s influenčními diagramy.

(Pearl 1988) Nicméně, podstatný vliv na propagaci grafických metod v umělé inteligenci a popularizaci pojmu PN měla Pearlova kniha o pravděpodobnostním rozhodování. Pearl zrekapituloval metody používající neorientované grafy, které nazýval *markovské sítě*. Více se však věnoval acyklickým orientovaným grafům, které byly hlavním objektem jeho výzkumných snah v té době. Tyto grafy začal nazývat *bayesovské sítě* a po něm tuto terminologii převzala většina výzkumníků zajímajících se o pravděpodobnostní rozhodování. Alternativní název jsou *belief networks*, anebo zkratka DAG z *directed acyclic graphs*. Kromě toho Pearl v této knize vypíchl úzkou souvislost obou grafických směrů s pojmem PN a zavedl grafické kritérium pro identifikaci nezávislostních údajů na základě příslušného acyklického orientovaného grafu, tzv. *d-separační kritérium* (d značí *directional*).

A.1.3 Metoda lokálních výpočtů

V souvislosti s pravděpodobnostním rozhodováním je vhodné se zmínit o každoroční konferenci *Uncertainty in Artificial Intelligence* (UAI). Jedná se o mezinárodní konferenci s již více než dvacetiletou tradicí, převážně pořádanou na americkém kontinentu, povětšinou v USA. Začala jako neformální workshop někdy v druhé půli 80. let dvacátého století a postupně se v průběhu 90. let proměnila v konferenci, na níž se setkávali odborníci z oblasti počítačových věd zajímající se o grafické modely. Jednou z vůdčích osobností, která tehdy ovlivnila zaměření konference, byl výše zmíněný Judea Pearl.

Údajně prý v průběhu prvních workshopů UAI probíhala diskuse mezi zastánci dvou různých přístupů: pravděpodobnostního a logicko-algebraického. Pravděpodobnostní přístup byl kritizován celkem oprávněnou námitkou, že příslušné výpočty nejsou prakticky realizovatelné na počítači. Tuto nedůvěru se podařilo posléze překonat tím, že byly navrženy algoritmy, které umožnily provádět *lokální výpočty*. O vývoj této metody se hodně zasloužili výzkumníci z Evropy, zejména lidé okolo Dána Steffena Lauritzena.

(Lauritzen, Spiegelhalter 1988) Výsledkem spolupráce Dána Lauritzena s Angličanem Spiegelhalterem byl metodický postup, jak při zadané bayesovské síti konkrétně provádět požadované výpočty. Tento postup později proslul pod názvem *metoda lokálních výpočtů* (anglicky *local computation method*). Jejich článek se do jisté míry stal kuchařkou pro následovatele.

(Spiegelhalter, Lauritzen 1990) Jiný článek těchže autorů, tentokrát s jejich obráceným pořadím, položil základy bayesovského přístupu k učení bayesovských sítí. Přesněji řečeno, jedná se o učení parametrů daného statistického modelu.

Lauritzen spolu s Angličany Dawidem, Cowellem a Spiegelhalterem vytvořili pracovní skupinu nazývanou *BAIES group*. Je to zkratka za *Bayesian Analysis In*

Expert Systems. Na přelomu 80. a 90. let dvacátého století úzce spolupracovali právě v oblasti pravděpodobnostních expertních systémů.

(Lauritzen 1996) Lauritzen později shrnul matematické základy těchto grafických metod v knize, která se stala jakýmsi zdrojovým textem pro ty, kteří toužili po větší matematické přesnosti.

(Cowell, Dawid, Lauritzen, Spiegelhalter 1999) Celá skupina pak později napsala knížku, která se naopak více věnovala algoritmickým hlediskům. Největší podíl na ní měl asi Robert Cowell.

Lauritzen, než se přestěhoval do Oxfordu, byl v Aalborgu vůdčí osobností skupiny ODIN, která se zaměřila na vytvoření konkrétního pravděpodobnostního expertního systému HUGIN.

(Jensen 2001) Šéfem skupiny vyvíjející tento, dalo by se říci komerčně úspěšný, expertní systém byl Dán Finn Jensen (Finn je křestní jméno), který napsal učebnici vhodnou zejména pro studenty počítačové vědy.

A.1.4 Vývoj v Čechách

Také v Čechách vznikla skupina, která se zabývala podobnou problematikou.

(Perez 1977) Byla to skupina z Ústavu teorie informace a automatizace (ÚTIA) Československé akademie věd (ČSAV) okolo Alberta Pereze, který již koncem 70. let minulého století přišel s myšlenkou aproximace mnohorozměrného diskrétního rozdělení pomocí rozdělení získaného násobením jeho málorozměrných marginál, respektive podmíněných pravděpodobností. Vlastně za touto ideou byla implicitně myšlenka přijetí jistého modelu struktury PN.

Pozoruhodná skutečnost je, že v Praze probíhala v druhé půli 80. let dvacátého století podobná výměna názorů mezi zastánci pravděpodobnostního a logicko-algebraického přístupu jako na americkém kontinentě mezi účastníky konference UAI a v západní Evropě mezi statistiky okolo skupiny BAIES. Vlastně to byl nezávislý a souběžný vývoj pravděpodobnostních metod v umělé inteligenci.

Dnes to asi nemusí být každému jasné, ale faktem je, že jakožto součást socialistického tábora byli vědci z Československa přeci jen poněkud izolováni a omezováni, co se týče účasti na zahraničních konferencích.

Výše zmíněná diskuse probíhala na pravidelném středečním semináři doktora Petra Hájka na matematickém ústavu ČSAV. Hájek byl považován za představitele logicko-algebraického směru, zatímco Perez za představitele pravděpodobnostního směru. Takovým arbitrem, nebo lépe řečeno překladatelem a možná občas i usmiřovatelem, byl doktor Tomáš Havránek, který byl odborníkem na grafické log-lineární modely. Havránek se později, začátkem 90. let, stal prvním ředitelem nově zřízeného ústavu informatiky (ÚI) Akademie věd. Po jeho předčasné smrti se stal (dlouholetým) ředitelem tohoto ústavu Hájek, který z tohoto důvodu odešel z matematického ústavu.

(Perez, Jiroušek 1985) Perez spolu se svými dvěma kolegy, Radimem Jirouškem a Otakarem Křížem, položili základy tzv. intensionálního přístupu k pravděpodobnostním expertním systémům, který kromě myšlenky aproximace mnohorozměrného rozdělení z dříve zmíněného Perezova článku zahrnuje *metody teorie informace* a grafické nástroje. Pro výzkumné účely vytvořil Ota Kříž i několik pracovních verzí expertního systému INES, což je zkratka z anglického *INtensional Expert System*.

(Hájek, Havránek, Jiroušek 1992) Později, po odchodu Pereze do důchodu, napsali Hájek, Havránek a Jiroušek společně knížku, která se věnovala základům jak logicko-algebraických, tak pravděpodobnostních expertních systémů.

Další společnou aktivitou výzkumníků z Čech byla organizace mezinárodního workshopu WUPES. Původně se jednalo o zkratku z anglického *Workshop on Uncertainty in Expert Systems*, nyní je zkratka interpretována jako *Workshop on Uncertainty ProcESsing*. První workshop se konal v roce 1988 v Alšovicích, kde tehdy Akademie věd měla školící středisko, a pomohl tak trochu prolomit isolaci české skupiny a navázat kontakty se Západem. Další workshopy se konaly pravidelně každé tři roky někde v Čechách a hlavním organizátorem byl Jiroušek, později ve spolupráci s kolegyní z ÚTIA Jiřinou Vejnarovou. V září 2012 se již devátý ročník konal v Mariánských Lázních.

(Studený 1992) Já, coby Perezův aspirant, což byla tehdejší obdoba dnešního PhD studenta, se stal členem této rozrůstající se skupiny díky svému zájmu o PN. Pomocí informačně-teoretických metod, na které mne Perez navedl, se mi podařilo vyvrátit dříve zmíněnou Pearlovu hypotézu o charakterizaci struktur PN pomocí semigrafoidových vlastností. O PN jsem se začal poté zajímat systematicky a v této souvislosti, pod vlivem zahraničních kolegů, i o novější trendy v oblasti grafických modelů, například o řetězcové grafy.

(Matúš, Studený 1995), (Matúš 1995, 1999) Zájem o téma PN projevil i kolega z ÚTIA František Matúš, s kterým jsme začali spolupracovat na problému charakterizace struktur PN indukovaných čtyřmi diskrétními náhodnými veličinami. Po delší době tento problém vyřešil sám Matúš.

Možná pod vlivem Petra Hájka a dřívějšího kolegy z ÚTIA doktora Ivana Kramosila, který později přešel za Hájkem do ÚI, se členové skupiny v ÚTIA začali zajímat i o PN v alternativních kalkulech pro popis nejistoty v umělé inteligenci.

(Vejnarová 1999) Například studiem PN v rámci Zadehovy teorie možností (anglicky *possibility theory*) se začala speciálně zabývat Jiřina Vejnarová.

A.2 Význačné konference a zahraniční skupiny

A.2.1 Konference UAI

Jak jsem se již zmínil, v průběhu 90. let dvacátého století se konference UAI přeměnila v hojně navštěvovanou konferenci. Navíc, zájem účastníků této konference se posunul více ke studiu bayesovských sítí, takže tuto konferenci bylo možno tehdy s určitou nadsázkou nazvat konferencí o bayesovských sítích. Nicméně koncem 90. let v souvislosti s nárůstem účastníků nastala asi nevyhnutelná diversifikace témat příspěvků, a dnes už asi nelze tvrdit, že hlavním tématem příspěvků na této konferenci jsou grafické modely.

V následujícím přehledu jsou zmíněni někteří tehdejší pravidelní účastníci této konference a některé jejich významné výsledky z 90. let dvacátého století.

(Geiger, Pearl 1990) Pearlův PhD student Dan Geiger dokázal, že každá ternární relace indukovaná bayesovskou sítí představuje skutečně strukturu PN. Tento výsledek vlastně říká, že je oprávněné používat acyklické orientované grafy pro popis struktur PN.

(Verma, Pearl 1991) Jiný Pearlův PhD student Tom Verma charakterizoval v grafických termínech ekvivalentní bayesovské sítě, tj. acyklické orientované grafy indukující stejnou strukturu PN.

(Chickering 1995) Chickering byl pozdější Pearlův PhD student, který přišel s transformační charakterizací ekvivalentních bayesovských sítí.

(Spirtes, Glymour, Scheines 1993) Spirtes a jeho kolegové z Pittsburgu přišli s alternativní kauzální interpretací acyklických orientovaných grafů, která později začala zajímat i Pearla. Tato interpretace se liší od interpretace v termínech PN: šipky v grafu odpovídají určitým funkcionálním závislostem.

(Meek 1995) Spirtesův žák Meek navrhl tzv. rekonstrukční algoritmus (anglicky *recovery algorithm*) na učení bayesovských sítí.

(Richardson 1996) Spirtes přišel s myšlenkou použít i obecné orientované grafy s možnými cykly a dal tomuto přístupu určitou interpretaci. Další jeho PhD student Thomas Richardson charakterizoval ekvivalentní orientované grafy.

(Heckerman 1995) Heckerman byl jeden z prvních PhD studentů Pearla. Svého času se zabýval bayesovskými metodami učení bayesovských sítí. Nechal se zaměstnat u Billa Gatese a založil v Microsoftu v Seattlu skupinu, do které posléze přilákal Meeka, Chickeringa a jiné.

(Neapolitan 1990) Neapolitan napsal knihu, relativně přesnější než Pearl, v níž shrnul matematické základy bayesovských sítí.

A.2.2 Program HSSS a nové grafické přístupy

Kontakty s ostatním světem naší skupině z ÚTIA usnadnil výzkumný program *Highly Structured Stochastic Systems* (HSSS), který proběhl v letech 1997–2000. Určitá skupina evropských výzkumníků, v níž vůdčí roli hrál Lauritzen, získala od *European Science Foundation* (ESF) nějaké prostředky, které použila na organizaci série speciálních konferencí a workshopů. Některé z nich se týkaly i pravděpodobnostních metod v umělé inteligenci a zprostředkovaně PN.

(Lauritzen, Wermuth 1984) Tak jsme měli možnosti potkat Nanny Wermuth(ovou), která v 80. letech minulého století přišla spolu s Lauritzenem s myšlenkou řetězcových grafů.

(Frydenberg 1990) Jednou jsme potkali i Frydenberga. To byl Lauritzenův PhD student, který podstatně přispěl k prohloubení teoretických základů řetězcových grafů.

(Andersson, Perlman 1993) Setkali jsme se se statistiky Anderssonem a Perlmanem, kteří již v roce 1993 přišli s negrafickou metodou popisu některých struktur PN pomocí konečných svazů. Steen Andersson je Dán, který se usídlil v Americe, Michael Perlman je Američan, žijící v Seattlu.

(Andersson, Madigan, Perlman 2001) Andersson, Madigan, Perlman a později i Levitz rozvinuli též teorii alternativních řetězcových grafů. David Madigan byl mladší kolega Perlmana, Michael Levitz byl Perlmanův PhD student.

(Cox, Wermuth 1996) Wermuth(ová) spolu s Coxem zavedli velmi širokou třídu řetězcových grafů, které zahrnují jak klasické, tak alternativní řetězcové grafy.

Dodal bych, že David Cox je ten proslulý statistik, po němž je pojmenován standardně používaný model v tzv. analýze přežití.

(Koster 1996) Jan Koster z Holandska přišel s třídou reciprokých grafů, které zahrnují jak klasické řetězcové grafy, tak orientované grafy s cykly. Tyto grafy použil k reprezentaci obecných modelů strukturálních rovnic (tzv. SEM modely).

(Kauerman 1996) Němec Kauerman publikoval článek o alternativně interpretovaných neorientovaných grafech.

Další příležitostí k setkání výzkumníků zajímajích se o PN byl speciální program Fieldsova ústavu na univerzitě v Torontu v Kanadě, který se konal na podzim 1999 pod odbornou garancí Kanaďanky francouzského původu Helène Massam a Steffena Lauritzena. V rámci tohoto programu jsme spolu s Matúšem zorganizovali workshop o PN a grafických modelech. Zmíněný workshop měl přes 60 účastníků, takže spíše to byla malá konference.

Existuje i pravidelná evropská konference zaměřená na metody popisu nejistoty v umělé inteligenci. Nazývá se ECSQARU (zkratka za *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*) a koná se zhruba každé dva roky. Na rozdíl od UAI je její záběr širší; je více věnována alternativním přístupům k popisu nejistoty v umělé inteligenci. Na této konferenci můžete potkat jak výzkumníky zaměřené na speciální kalkuly nejistoty v umělé inteligenci, tak ty, kteří se zajímají o grafické modely.

(de Campos, Huete, Moral 1993) Tak například španělská skupina z Granady okolo Serafína Morala se zabývala pojmem PN v rámci kalkulu pravděpodobnostních intervalů.

(Bouckaert 1995) Na ECSQARU jsem se seznámil s Holanďanem Bouckaertem, s kterým jsem později napsal nějaké články. Byl to on, kdo mne navedl na problematiku řetězcových grafů.

Nicméně s počátkem novém století snad někteří evropští výzkumníci v oblasti umělé inteligence začali pociťovat, že chybí nějaké pravidelné evropské setkání zaměřené pouze na grafické modely. Proto skupina španělských výzkumníků, povětšinou bývalých studentů Morala a de Campose, přišla s iniciativou organizovat každé dva roky workshop či konferenci s tímto zaměřením. Tak vznikl workshop či spíše malá konference *Probabilistic Graphical Models* (PGM). První PGM se konalo v roce 2002 ve Španělsku v městě Cuenca, druhé v roce 2004 v Holandsku v Leidenu a třetí jsme na výzvu zorganizovali v roce 2006 v Praze. Já jsem posloužil coby předseda programového výboru. Předsedou organizačního výboru byl můj starší kolega Radim Jiroušek, motorem organizačního výboru pak mladší kolega u ÚTIA, někdejší Jiroušků PhD student, Jiří Vomlel. Pražská PGM konference měla nakonec okolo 60 účastníků. PGM v roce 2008 byla v Aalborgu v Dánsku, v roce 2010 v Helsinkách a v roce 2012 se konala opět ve Španělsku, tentokrát v Granadě. V roce 2014 by se s desetiletým odstupem měla opět konat v Holandsku, v Leidenu.

A.2.3 Alternativní kalkuly v umělé inteligenci

Ještě se sluší zmínit nepravděpodobnostní PN.

(Sagiv, Walecka 1982) Například v *teorii relačních databází* byla známa analogie pojmu PN nazývaná EMVD (anglická zkratka za *embedded multivalued dependency*). Sagiv a Walecká dokázali, že neexistuje konečná axiomatická charakterizace příslušných indukovaných modelů.

(Malvestuto 1992) Této analogie si povšiml počátkem 90. let dvacátého století i Ital Franco Malvestuto.

(Wong, Wang 1994) A ještě později si té analogie povšiml i Kanaďan čínského původu Wong.

(Shenoy 1994) Američan indického původu Prakash Shenoy, jehož specializací je rozhodování za nejistoty, si povšiml možnosti zavést pojem PN i v rámci různých nepravděpodobnostních kalkulů pro popis nejistoty v umělé inteligenci. Dokázal, že odvozený pojem PN splňuje známé formální vlastnosti pravděpodobnostní PN (semigrafoidové vlastnosti).

(Zadeh 1978) (Dubois, Prade 1988) Jednou z oblastí, na které Shenoy upozornil, byla Zadehova teorie možnosti (anglicky *possibility theory*). Tento přístup rozpracovala zejména francouzská skupina z Tolouse.

(Spohn 1988) Zvláštní filosofickou motivaci má kalkul tzv. ordinálních podmíněných funkcí navržený Spohnem, později nazývaný v UAI komunitě *κ-kalkulus*.

(Dempster 1967) (Shafer 1976) Motivován článkem Dempstera z konce 60. let dvacátého století, Američan Glenn Shafer navrhl určitou metodu popisu nejisté znalosti, která může být chápána jako zobecnění pravděpodobnostního kalkulu. Je to zobecnění v tom smyslu, že příslušnou důvěrovou funkci (anglicky *belief function*) lze chápat jako popis jisté konvexní množiny pravděpodobnostních rozdělení. Jeho teorie se stala populární pod názvem *teorie evidence* či Dempsterova-Shaferova teorie.

(Shafer 1996) Shafer se v jedné ze svých pozdějších knih věnoval i otázce, kdy lze v rámci nějakého alternativního kalkulu použít zobecněnou metodu lokálních výpočtů.

Existuje i pravidelná konference zaměřená na zobecněné pravděpodobnostní přístupy. Nazývá se ISIPTA (zkratka za anglické *International Symposium on Imprecise Probabilities and Their Applications*) a koná se zhruba každé dva roky. V roce 2007 se konala v Praze již sedmá konference této řady a její hlavní organizátorkou byla kolegyně Jiřina Vejnarová.

A.2.4 Geometrický přístup a metody moderní algebry

Snad bych se měl zmínit i o trendech posledního desetiletí, včetně výzkumného směru, kterému v současného době věnuji nejvíce.

(Studený 2005) Pozorování, že pomocí grafů není z principiálních důvodů možné popsat všechny možné (diskrétní) struktury PN, mne motivovalo k návrhu popisovat tyto struktury pomocí jistých speciálních celočíselných vektorů, nazývaných *strukturální imsety*. Tento přístup, částečně inspirován informačně-teoretickými metodami, zároveň umožňuje v této oblasti aplikovat pokročilé geometrické metody. Například v posledních letech se (ve spolupráci s odborníky na optimalizaci) věnuji záměru využít metody lineárního a celočíselného lineárního programování v oblasti učení grafických modelů.

(Hemmecke, Morton, Shiu, Sturmfels, Wienand 2008) Výše zmíněný přístup však zároveň vedl k celé řadě otevřených otázek, které jsem shrnul v jedné kapitole knížky, která se této metodě věnovala. Některé z těchto otevřených problémů zaujaly skupinu algebraiků a geometrů okolo Bernda Sturmfelse z University of Berkeley. Ve spolupráci s odborníkem na kombinatorickou optimalizaci Raymondem Hemmeckem (nyní na TU Mnichov) se jim některé z nich podařilo vyřešit.

(Drton, Sturmfels, Sullivant 2009) Důvodem jejich zájmu o téma podmíněné nezávislosti bylo to, že jsem nebyl zdaleka jediný, kdo přišel s myšlenkou použít algebraické a geometrické metody na řešení problémů ve statistice a jejich aplikacích. Vlastně jsem se mimoděk trefil do trendu, který vedl ke vzniku nového oboru, tzv. *algebraické statistiky*. Ta se dá zhruba charakterizovat jako oblast, v níž se kombinují metody kombinatoriky, geometrie polytopů, výpočetní algebry a algebraické geometrie při (formulaci, interpretaci a) řešení statistických úloh. Jedním z propagátorů tohoto přístupu je právě Sturmfels a lidé z jeho skupiny. Konalo se již několik konferencí na toto téma (v USA i Japonsku), nedávno byl dokonce založen časopis *Journal of Algebraic Statistics*. Sturmfels spolu s dvěma jinými kolegy napsal přehledovou knížku na toto téma, jejíž jedna kapitola je věnována PN a požití metod komutativní algebry v této oblasti.

Avšak jedná se o pokročilejší (a hlubší) témata a metody dosud vyvíjené; proto tato skripta, jejichž účelem je spíše čtenáři vysvětlit matematické základy, tyto metody nepopisují podrobněji. Nicméně, spolu s kolegou Matúšem se snažíme sledovat moderní trendy a proto ve spolupráci se zahraničními kolegy, konkrétně Caroline Uhler z IST Austria (což je byvší PhD studentka Sturmfelse) a Raymondem Hemmeckem z TU Munich (toho času), plánujeme v roce 2014 zorganizovat v rámci setkání *Prague Stochastics 2014* specializovaný workshop o algebraické statistice.

Příloha B

Přehled elementárních pojmů

V tomto dodatku jsou zopakovány některé základní matematické pojmy a pozorování. Studentům MFF UK by sice povětšinou měly být známy, ale protože skripta jsou zamýšlena i jako doplňkový učební text pro studenty jiných vysokých škol, kteří s těmito matematickými pojmy nemusí být obeznámeni, rozhodl jsem se připojit tento přehled. Pro rychlé vyhledání příslušného hesla doporučuji využít rejstřík.

B.1 Pojmy z teorie míry

- *Měřitelný prostor* je dvojice $(X, \mathcal{X})$, kde $X \neq \emptyset$ je neprázdná množina a $\mathcal{X} \subseteq \mathcal{P}(X) \equiv \{Y : Y \subseteq X\}$ je *σ-algebra* jejích podmnožin, což značí, že splňuje následující podmínky:
  - (i) $X \in \mathcal{X}$,
  - (ii) $\mathbb{A} \in \mathcal{X} \;\Rightarrow\; X \setminus \mathbb{A} \in \mathcal{X}$,
  - (iii) $\mathbb{A}_n \in \mathcal{X}$, $n = 1, 2, \ldots \;\Rightarrow\; \bigcup_n \mathbb{A}_n \in \mathcal{X}$.

  Libovolný průnik σ-algeber na X je σ-algebra na X. Speciálně, $\forall\, \mathcal{B} \subseteq \mathcal{P}(X)$ existuje nejmenší σ-algebra na X obsahující $\mathcal{B}$, která se nazývá σ-algebra *generovaná* $\mathcal{B}$ a označuje $\sigma(\mathcal{B})$.
- *$\mathcal{X}$-měřitelná funkce* je funkce $f : X \to [-\infty, +\infty]$ taková, že

$$\forall\, a \in \mathbb{R} \qquad f_{-1}[-\infty, a] \equiv \{x \in X : f(x) \in [-\infty, a]\} \in \mathcal{X}\,.$$

- *Součin měřitelných prostorů*. Buď $(\mathsf{X}_i, \mathcal{X}_i)$, $i \in N$ konečný systém měřitelných prostorů. Potom položíme $X \equiv \prod_{i \in N} \mathsf{X}_i$ a $\mathcal{X} \equiv$ σ-algebra generovaná měřitelnými obdélníky tvaru $\prod_{i \in N} \mathbb{A}_i$, $\mathbb{A}_i \in \mathcal{X}_i$. Součinovou σ-algebru $\mathcal{X}$ budeme označovat symbolem $\prod_{i \in N} \mathcal{X}_i$.
- *Nezáporná míra* na $(X, \mathcal{X})$ je funkce $\mu : \mathcal{X} \to [0, +\infty]$ splňující
  - (i) $\mu(\emptyset) = 0$,
  - (ii) $\mathbb{A}_n \in \mathcal{X}$, $n = 1, 2, \ldots$ po dvou disjunktní

$$\mu\Big(\bigcup_{n=1}^{\infty} \mathbb{A}_n\Big) = \sum_{n=1}^{\infty} \mu(\mathbb{A}_n)\,.$$

Podmínka (ii) se nazývá spočetná aditivita. Nezáporná míra je

– *konečná*, jestliže $\mu(X) < \infty$,
– *σ-konečná*, jestliže $\exists\, \mathbb{A}_n \in \mathcal{X}, n = 1, 2, \ldots$ takové, že platí $\mu(\mathbb{A}_n) < \infty$ a $X = \bigcup_n \mathbb{A}_n$,
– *pravděpodobnostní*, jestliže $\mu(X) = 1$.

Příklady.

1. Elementárním příkladem je *aritmetická míra* na konečné neprázdné množině X. V tomto diskrétním případě se implicitně předpokládá $\mathcal{X} = \mathcal{P}(X)$, tedy σ-algebrou je potenční množina X. Aritmetická míra pak každé podmnožině $\mathbb{A} \subseteq X$ přiřadí počet jejích prvků: $\nu(\mathbb{A}) = |\mathbb{A}|$.
2. Jiným základním příkladem je *jednorozměrná Lebesgueova míra*. V tomto spojitém případě je $X = \mathbb{R}$ a $\mathcal{X}$ tzv. *borelovská σ-algebra*, tj. σ-algebra generovaná systémem polouzavřených intervalů: $\mathcal{X} = \sigma(\{(a, b] \,:\, a < b\})$. Potom na $\mathcal{X}$ existuje jediná nezáporná míra λ taková, že míra intervalu je jeho délka: $\forall\, a < b \quad \lambda((a, b]) = b - a$. Tato míra se nazývá Lebesgueova.

- *Rovnost skoro všude* vzhledem k nezáporné míře. Buď $f, g : X \to \mathbb{R}$ funkce, obě $\mathcal{X}$-měřitelné, a μ míra na $(X, \mathcal{X})$. Pak definujeme

$$f = g \ \mu\text{-skoro všude} \quad \Leftrightarrow \quad \mu(\,\{x \in X \,:\, f(x) \neq g(x)\}\,) = 0\,.$$

Připomínám, že $\{x \in X \,:\, f(x) \neq g(x)\} \in \mathcal{X}$ neboť rozdíl $f - g$ je $\mathcal{X}$-měřitelná funkce, což plyne z $\mathcal{X}$-měřitelnosti funkcí f a g. Při značení se často používá zkratka: μ-s.v.

- Poznamenejme, že v diskrétním případě, tj. když X je konečná množina, je rovnost funkcí f a g skoro všude vzhledem k μ ekvivalentní požadavku, že funkce f a g se rovnají alespoň na nosiči míry μ, tedy na $\{x \in X \,:\, \mu(\{x\}) > 0\}$. To ale již není pravda obecně. Tím mám na mysli, že rovnost skoro všude nelze vždy charakterizovat jako rovnost funkcí na nějaké podmnožině (množiny X). Jinými slovy, množina tzv. plné míry, na níž se funkce f a g mají rovnat se může s dvojicí funkcí f a g měnit a nemusí existovat žádná „univerzální" množina jako v diskrétním případě. Jako příklad bych uvedl jednorozměrnou Lebesgueovu míru. Buď f funkce na $\mathbb{R}$ identicky rovná nule $f \equiv 0$, a pro každé $a \in \mathbb{R}$, buď g_a indikátorová funkce singletonu $\{a\}$: $g_a(x) = 1$ pro $x = a$, $g_a(x) = 0$ pro $x \neq a$. Pak pro každé $a \in \mathbb{R}$ je $f = g_a$ λ-skoro všude, ale množina bodů x kde $f(x) = g_a(x)$ pro všechna a je prázdná.
- *Lebesgueův integrál*. Buď $(X, \mathcal{X})$ měřitelný prostor a μ nezáporná míra na $(X, \mathcal{X})$. Každé $\mathcal{X}$-měřitelné funkci $f : X \to [0, +\infty]$ a každé množině $\mathbb{A} \in \mathcal{X}$ konstrukce, popsaná například v učebnici (Rudin 2003),

$$\text{přiřadí číslo} \quad \int_{\mathbb{A}} f(x)\, \mathrm{d}\mu(x) \in [0, +\infty]\,,$$

které se nazývá *integrál* z funkce f přes množinu $\mathbb{A}$ podle míry μ. Integrál zobecňuje pojem váženého součtu: pro X konečnou platí $\int_{\mathbb{A}} f(x)\, \mathrm{d}\mu(x) = \sum_{x \in \mathbb{A}} f(x) \cdot \mu(\{x\})$.

- *Integrovatelná funkce* je $\mathcal{X}$-měřitelná funkce $f : X \to \mathbb{R}$ taková, že
$$\int_X |f(x)| \, \mathrm{d}\mu(x) < \infty \, .$$
Buď $f^+(x) \equiv \max\{f(x), 0\}$, respektive $f^-(x) \equiv \max\{-f(x), 0\}$, pro $x \in X$, *kladná*, respektive *záporná, část* funkce f. Evidentně $f = f^+ - f^-$ a $|f| = f^+ + f^-$. Jestliže aspoň jeden z integrálů $\int_X f^+(x) \, \mathrm{d}\mu(x)$, $\int_X f^-(x) \, \mathrm{d}\mu(x)$ je konečný, pak se funkce f nazývá *kvazi-integrovatelná* a jejím integrálem je výraz
$$\int_X f(x) \, \mathrm{d}\mu(x) \equiv \int_X f^+(x) \, \mathrm{d}\mu(x) - \int_X f^-(x) \, \mathrm{d}\mu(x) \, .$$
Pro f, g kvazi-integrovatelné
$$f = g \quad \mu\text{-s.v.} \quad \Leftrightarrow \quad \forall\, \mathbb{A} \in \mathcal{X} \quad \int_{\mathbb{A}} f(x) \, \mathrm{d}\mu(x) = \int_{\mathbb{A}} g(x) \, \mathrm{d}\mu(x) \, .$$

- *Absolutní spojitost.* Nechť $(X, \mathcal{X})$ je měřitelný prostor, μ, ν nezáporné míry na $(X, \mathcal{X})$. Řekneme, že ν je absolutně spojitá vůči μ, zápis $\nu \ll \mu$, jestliže
$$\forall\, \mathbb{A} \in \mathcal{X} \quad \mu(\mathbb{A}) = 0 \;\Rightarrow\; \nu(\mathbb{A}) = 0 \, .$$
Alternativní termín je že *μ dominuje ν*.

Fakt (Radonova-Nikodymova věta). Buď $(X, \mathcal{X})$ měřitelný prostor, μ σ-konečná míra na $(X, \mathcal{X})$, ν konečná míra na $(X, \mathcal{X})$ a $\nu \ll \mu$. Potom existuje (nezáporná) $\mathcal{X}$-měřitelná funkce f nazývaná *Radonovou-Nikodymovou derivací ν* vzhledem k μ, taková, že
$$\forall\, \mathbb{A} \in \mathcal{X} \quad \nu(\mathbb{A}) = \int_{\mathbb{A}} f(x) \, \mathrm{d}\mu(x) \, .$$
Tato funkce je určena jednoznačně v rámci rovnosti μ-skoro všude. Proto každou jednotlivou funkci splňující výše zmíněnou rovnost budu nazývat *verzí Radonovy-Nikodymovy derivace*; standardně značená symbolem $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ či $\mathrm{d}\nu/\mathrm{d}\mu$. Alternativní název je *hustota* ν vzhledem k (dominující) míře μ.

- *Jensenova nerovnost.* Funkce $\phi : I \to \mathbb{R}$ (kde $I \subseteq \mathbb{R}$ je interval) je *konvexní* pokud $\forall\, \alpha \in [0, 1]$, $\forall\, x, y \in I$ platí
$$\phi(\alpha \cdot x + (1 - \alpha) \cdot y) \le \alpha \cdot \phi(x) + (1 - \alpha) \cdot \phi(y) \, .$$
Pokud, pro každé $x \neq y$ a $\alpha \in (0, 1)$, je nerovnost ostrá, pak ϕ je *ryze konvexní*. Pro pravděpodobnostní míru μ na měřitelném prostoru $(X, \mathcal{X})$, integrovatelnou $\mathcal{X}$-měřitelnou funkci $f : X \to I$ a výše zmíněnou konvexní funkci ϕ pak *Jensenova nerovnost* říká
$$\phi(\int_X f(x) \, \mathrm{d}\mu(x)) \le \int_X \phi(f(x)) \, \mathrm{d}\mu(x) \, .$$
V případě ryze konvexní funkce ϕ rovnost nastává právě tehdy, když existuje konstanta $k \in I$, že $f(x) = k$ pro μ-s.v. $x \in X$.

- *Součin σ-konečných měr.* Buď $(\mathsf{X}_i, \mathcal{X}_i)$, $i \in N$ měřitelné prostory a μ^i, $i \in N$ σ-konečné míry na $(\mathsf{X}_i, \mathcal{X}_i)$. Pak existuje právě jedna nezáporná míra μ na součinu měřitelných prostorů $(X, \mathcal{X}) = (\prod_{i \in N} \mathsf{X}_i, \prod_{i \in N} \mathcal{X}_i)$, taková, že na měřitelných obdélnících je určena následující hodnotou:

$$\mu(\prod_{i \in N} \mathbb{A}_i) = \prod_{i \in N} \mu^i(\mathbb{A}_i) \quad \text{pro } \mathbb{A}_i \in \mathcal{X}_i, \ i \in N\,.$$

Tato míra μ je rovněž σ-konečná, nazývá se *součinem* měr μ^i a označuje symbolem $\bigotimes_{i \in N} \mu^i$. V případě součinu dvou měr μ a ν píšeme $\mu \otimes \nu$.

Příklad. Příkladem je *vícerozměrná Lebesgueova míra*. V tomto případě je, pro každé $i \in N$, $\mathsf{X}_i = \mathbb{R}$, $\mathcal{X}_i$ borelovská σ-algebra na $\mathbb{R}$ a $\mu^i = \lambda$ jednorozměrná Lebesgueova míra.

Poznámka. Kartézský součin $\prod_{i \in N} \mathsf{X}_i$ v uvedeném případě $\mathsf{X}_i = \mathbb{R}$ budu označovat symbolem $\mathbb{R}^N$, kterýmžto značením chci zdůraznit, že množinu proměnných N apriorně nepovažujeme za uspořádanou. Naopak, v případě, že je účelné mít konečnou indexovou množinu N totálně uspořádanou, tj. když ji vlastně interpretujeme jako úsek celočíselné škály $N = [n] \equiv \{1, 2, \ldots, n\}$, $n \geq 1$, anebo pro to budou jiné důvody, pak namísto symbolu $\mathbb{R}^N$ budu používat běžnější značení pro n-rozměný reálný euklidovský prostor $\mathbb{R}^n$.

B.2 Pojmy z teorie svazů

- Uspořádaná množina, nebo *poset* $(L, \preceq)$, což je zkratka z anglického *partially ordered set*, je neprázdná množina L vybavená *částečným uspořádáním*, to jest binární relací $\preceq$, která je
  - (i) *reflexivní*: $\forall\, x \in L \quad x \preceq x\,,$
  - (ii) *antisymetrická*: $\forall\, x, y \in L \quad x \preceq y,\ y \preceq x \ \Rightarrow\ x = y\,,$
  - (iii) *transitivní*: $\forall\, x, y, z \in L \quad x \preceq y,\ y \preceq z \ \Rightarrow\ x \preceq z\,.$

  Uspořádání nazveme *totální*, jestliže navíc $\forall\, x, y \in L$ buď $x \preceq y$ nebo $y \preceq x$. Symbol $x \prec y$ značí $x \preceq y$ a $x \neq y$.
- Pokud $x \prec y$ a neexistuje $z \in L$ takové, že $x \prec z$ a $z \prec y$, pak se x nazývá *dolní soused* y, respektive y *horní soused* x.
- Konečný poset se zobrazuje pomocí *Hasseova diagramu*. V Hasseově diagramu je každý prvek konečného posetu zobrazen kolečkem (či oválem). Je-li x horním sousedem y, pak se udělá spojnice x a y, přičemž x je pak umístěno nad y.
- Jestliže $M \subseteq L$, pak prvek $x \in M$ je *minimální prvek* M vzhledem k $\preceq$, jestliže neexistuje $z \in M$ takové, že $z \prec x$. Analogicky, $y \in M$ je *maximální prvek* M vzhledem k $\preceq$, pokud neexistuje $z \in M$ splňující $z \succ y$.
- *Supremum* množiny $M \subseteq L$ v posetu $(L, \preceq)$, značené $\sup M$, a rovněž nazývané *nejmenší horní závora* M, je prvek $y \in L$ takový, že

(i) $z \preceq y$ pro každé $z \in M$, avšak

(ii) $y \preceq y'$ pro každé $y' \in L$ splňující $z \preceq y'$ pro $z \in M$.

Pokud supremum existuje, je jednoznačně určeno, díky antisymetrii uspořádání $\preceq$.

Spojení prvků $x, y \in L$, označované $x \vee y$ je supremum množiny $\{x\} \cup \{y\}$.

- Analogicky, *infimum* množiny $M \subseteq L$, značené $\inf M$, a nazývané také *největší dolní závora* M je prvek $x \in L$ takový, že

  (i) $x \preceq z$ pro každé $z \in M$, avšak

  (ii) $x' \preceq x$ pro každé $x' \in L$ splňující $x' \preceq z$ pro $z \in M$.

  Průsek prvků $x, y \in L$, označovaný $x \wedge y$, je infimum množiny $\{x\} \cup \{y\}$.
- *Spojový polovaz* je poset $(L, \preceq)$ takový, že pro každé $x, y \in L$ existuje supremum $x \vee y$ v L.
- *Svaz* je poset $(L, \preceq)$ takový, že pro každé $x, y \in L$ existuje jak supremum $x \vee y$ tak infimum $x \wedge y$ v L. Svaz je *distributivní* jestliže pro každé $x, y, z \in L$

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \quad \text{a} \quad x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)\,.$$

Příklad. Typický příklad distributivního svazu je *okruh podmnožin* konečné neprázdné množiny N, tj. soubor $\mathcal{R} \subseteq \mathcal{P}(N) \equiv \{Y : Y \subseteq N\}$ uzavřený na (konečné) průniky a sjednocení.

- *Úplný svaz* je poset $(L, \preceq)$ takový, že každá podmnožina $M \subseteq L$ má supremum i infimum v L.

Poznámka. Pro ověření, že poset L je úplný svaz stačí ověřit, že každá $M \subseteq L$ má infimum (nebo duálně, že každá podmnožina L má supremum). Každý konečný svaz je příkladem úplného svazu.

- *Nulovým prvkem* úplného svazu L se rozumí *nejmenší prvek* L, tj. $x_0 \in L$ takový, že $x_0 \preceq z$ pro každé $z \in L$. Není to nic jiného než supremum prázdné množiny v $(L, \preceq)$.
- *Jednotkovým prvkem* se rozumí *největší prvek* L, tj. $y_1 \in L$ takové, že $z \preceq y_1$ pro každé $z \in L$.
- Prvek x úplného svazu je *$\vee$-nerozložitelný* (budu říkat *supremum nerozložitelný*), jestliže

$$x \neq \sup\{z \in L : z \prec x\}\,.$$

- Prvek x je *$\wedge$-nerozložitelný* (to jest *infimum nerozložitelný*), pokud

$$x \neq \inf\{z \in L : x \prec z\}\,.$$

Poznámka. Prvek konečného svazu je $\vee$-nerozložitelný právě tehdy, když má právě jednoho dolního souseda. Analogicky, prvek je $\wedge$-nerozložitelný právě tehdy, když má právě jednoho horního souseda.

- Množina $\vee$-nerozložitelných prvků v konečném svazu $(L, \preceq)$ je nejmenší množina $M \subseteq L$ která je *supremum-hustá*, což značí, že pro každé $x \in L$ existuje $M' \subseteq M$ takové, že $x = \sup M'$.
- Analogicky, množina $\wedge$-nerozložitelných prvků v L je nejmenší množina $M \subseteq L$ která je *infimum-hustá*, tj. pro každé $y \in L$ existuje $M' \subseteq M$ splňující $y = \inf M'$.
- Standardní příklad $\vee$-nerozložitelného prvku úplného svazu L je jeho *atom*, což je horní soused nulového prvku L.
- *Koatomem* L se rozumí dolní soused jednotkového prvku L.
- Úplný svaz L se nazývá *atomický*, pokud množina jeho atomů je supremum-hustá v L; ekvivalentně, pokud jedinými $\vee$-nerozložitelnými prvky jsou atomy.
- Svaz je *koatomický*, pokud množina koatomů je infimum-hustá v L, tj. jediné $\wedge$-rozložitelné prvky jsou koatomy.
- Buď $(L_1, \preceq_1)$, $(L_2, \preceq_2)$ úplné svazy. *Antiisomorfismem uspořádání* těchto posetů se bude rozumět zobrazení $\varsigma : L_1 \to L_2$ (na L_2) takové, že

$$\forall\, x, y \in L_1 \quad x \preceq_1 y \;\Leftrightarrow\; \varsigma(y) \preceq_2 \varsigma(x)\,.$$

Evidentně, takové zobrazení ς je prosté. Zobrazení se nazve *svazovým antiisomorfismem*, jestliže platí

$$\begin{aligned} \forall\, M \subseteq L_1 \qquad & \varsigma(\sup M) = \inf\{\,\varsigma(z) \;:\; z \in M\} \\ \text{a} \quad & \varsigma(\inf M) = \sup\{\,\varsigma(z) \;:\; z \in M\}\,. \end{aligned}$$

Evidentně, každý antiisomorfismus uspořádání je i svazový antiisomorfismus.

B.3 Opakování o maticích

- Buď M, N neprázdné konečné množiny. Reálnou *$M \times N$-maticí* se bude rozumět reálná funkce na $M \times N$. Zápis $\mathbf{\Sigma} = (\sigma_{ij})_{i \in M, j \in N}$.
Symbol $\mathbf{\Sigma}_{A \cdot B}$ bude označovat *$A \times B$-podmatici* matice $\mathbf{\Sigma}$, kde $A \subseteq M$, $B \subseteq N$.
- Reálnou symetrickou $N \times N$-matici nazveme *positivně semidefinitní*, pokud

$$\boldsymbol{x}^\top \cdot \mathbf{\Sigma} \cdot \boldsymbol{x} \geq 0 \quad \text{pro každé } \boldsymbol{x} \in \mathbb{R}^N,$$

respektive *positivně definitní*, pokud

$$\boldsymbol{x}^\top \cdot \mathbf{\Sigma} \cdot \boldsymbol{x} > 0 \quad \text{pro každé } \mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^N.$$

Matice je positivně definitní, právě když je regulární a positivně semidefinitní. Inverzní matice k positivně definitní je rovněž positivně definitní.
- *Zobecněnou inverzí* $N \times N$-matice $\mathbf{\Sigma}$ se bude rozumět každá matice $\mathbf{\Sigma}^-$ taková, že $\mathbf{\Sigma} \cdot \mathbf{\Sigma}^- \cdot \mathbf{\Sigma} = \mathbf{\Sigma}$. Někteří autoři, například (Anděl 1985, 2007), používají namísto toho termín *pseudo-inverzní matice*. Taková matice vždy existuje, ale není určena jednoznačně. V případě, že $\mathbf{\Sigma}$ je regulární, pak je její zobecněná inverze jednoznačně určena a shoduje se s inverzní maticí $\mathbf{\Sigma}^{-1}$.

- Jestliže $\mathbf{\Sigma}$ je realná $N \times N$ matice a $A, C \subseteq N$ jsou disjunktní, pak *Schurův doplněk* podmatice $\mathbf{\Sigma}_{C\cdot C}$ (v $\mathbf{\Sigma}_{AC\cdot AC}$) je $A \times A$-matice definovaná vztahem

$$\mathbf{\Sigma}_{A|C} = \mathbf{\Sigma}_{A\cdot A} - \mathbf{\Sigma}_{A\cdot C} \cdot (\mathbf{\Sigma}_{C\cdot C})^{-} \cdot \mathbf{\Sigma}_{C\cdot A}\,,$$

přičemž platí konvence $\mathbf{\Sigma}_{A|\emptyset} \equiv \mathbf{\Sigma}_{A\cdot A}$. Pro positivně semidefinitní matice tato matice nezávisí na volbě pseudo-inverzní matice $(\mathbf{\Sigma}_{C\cdot C})^{-}$.

Fakt. Jestliže je $\mathbf{\Sigma}_{C\cdot C}$ regularní, pak je $\mathbf{\Sigma}_{A|C}$ regulární právě když je regulární matice $\mathbf{\Sigma}_{AC\cdot AC}$, a v tomto případě platí

$$(\mathbf{\Sigma}_{A|C})^{-1} = ((\mathbf{\Sigma}_{AC\cdot AC})^{-1})_{A\cdot A}\,.$$

Pokud je $\mathbf{\Sigma}$ positivně definitní, respektive positivně semidefinitní, pak též $\mathbf{\Sigma}_{A|C}$ je positivně definitní, respektive positivně semidefinitní. Konečně, platí nasledující *princip transitivity*: jsou-li $A, B, C \subseteq N$ po dvou disjunktní, a matice $\mathbf{\Sigma}_{BC\cdot BC}$, $\mathbf{\Sigma}_{C\cdot C}$ regulární, pak platí

$$\mathbf{\Sigma}_{A|BC} = (\mathbf{\Sigma}_{AB|C})_{A|B}\,.$$

B.4 Pojmy z konvexní geometrie

V tomto oddíle je záměrně označován (reálný) euklidovský prostor symbolem $\mathbb{R}^n$, kde $n \geq 1$ a nikoliv $\mathbb{R}^N$ pro množinu proměnných N, což je nejčastější označení euklidovského prostoru v těchto skriptech (viz poznámka na straně 16 či § B.6). To proto, že níže uvedená fakta jsou fakticky (až na jednu výjimku) využita v kontextu speciálního prostoru $\mathbb{R}^{\mathcal{P}(N)}$, kde $\mathcal{P}(N) \equiv \{A\ :\ A \subseteq N\}$ je systém podmnožin neprázdné konečné množiny proměnných N (zejména v § 3.3.2). Důkazy většiny zde uvedených tvrzení lze nalézt například v knize (Brøndsted 1983).

- Množina $K \subseteq \mathbb{R}^n$, $n \geq 1$ se nazývá *konvexní*, jestliže

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in K \quad \forall\, \alpha, \beta \geq 0,\ \alpha + \beta = 1 \qquad \alpha \cdot \boldsymbol{x} + \beta \cdot \boldsymbol{y} \in K\,.$$

Konvexní množina K je *uzavřená*, pokud s každou posloupností vektorů obsahuje i její limitu: $\lim_{\ell\to\infty} \boldsymbol{x}_\ell = \boldsymbol{x} \in \mathbb{R}^n$ pro $\boldsymbol{x}_\ell \in K$ implikuje $\boldsymbol{x} \in K$.

- *Dimense* (uzavřené) konvexní množiny $K \subseteq \mathbb{R}^n$ je definována jako dimense *afinního obalu* K, tedy množiny

$$\left\{ \boldsymbol{x} \in \mathbb{R}^n\ :\ \boldsymbol{x} = \sum_{\boldsymbol{z}\in J} \alpha_{\boldsymbol{z}} \cdot \boldsymbol{z} \quad \text{pro konečnou } J \subseteq K\,,\ \sum_{\boldsymbol{z}\in J} \alpha_{\boldsymbol{z}} = 1\,,\ \alpha_{\boldsymbol{z}} \in \mathbb{R} \right\},$$

což je (vždy) posunutý lineární podprostor, tedy množina $\{\boldsymbol{y} + \boldsymbol{z}\ :\ \boldsymbol{z} \in L\}$, kde L je lineární podprostor $\mathbb{R}^n$ a $\boldsymbol{y} \in \mathbb{R}^n$. Pro neprázdnou (konvexní) množinu K je dimense K (= dimense L) celé číslo mezi 0 a n.

Příklady.

1. Klasickým příkladem uzavřené konvexní množiny je *polytop*, definovaný jako *konvexní obal* nějaké konečné množiny $B \subseteq \mathbb{R}^n$:

$$conv\,(B) \equiv \{\, \boldsymbol{x} \in \mathbb{R}^n \ :\ \boldsymbol{x} = \sum_{\boldsymbol{z}\in B} \alpha_{\boldsymbol{z}} \cdot \boldsymbol{z} \quad \text{pro} \quad \alpha_{\boldsymbol{z}} \geq 0\,,\ \sum_{\boldsymbol{z}\in B} \alpha_{\boldsymbol{z}} = 1\,\}\,.$$

2. Obecnějším příkladem je *polyedr*, někdy též nazývaný *mnohostěn*, definovaný jako průnik konečného počtu uzavřených poloprostorů, tj. množin tvaru

$$\{\boldsymbol{x} \in \mathbb{R}^n \ :\ \langle \boldsymbol{x}, \boldsymbol{y}\rangle \geq c\}, \quad \text{kde} \quad \boldsymbol{y} \in \mathbb{R}^n,\, c \in \mathbb{R} \quad \text{a} \quad \langle \boldsymbol{x}, \boldsymbol{y}\rangle \equiv \sum_{i=1}^{n} x_i \cdot y_i\,.$$

   Anglický termín je *polyhedron*.

- Množina $K \subseteq \mathbb{R}^n$, $n \geq 1$ se nazývá *konvexním kuželem*, jestliže
  (i) $\mathbf{0} \equiv [0,\ldots,0] \in K$,
  (ii) $\forall\, \boldsymbol{x}, \boldsymbol{y} \in K \quad \forall\, \alpha, \beta \geq 0 \quad \alpha \cdot \boldsymbol{x} + \beta \cdot \boldsymbol{y} \in K$.

  Evidentně se jedná o konvexní množinu. Anglický termín je *convex cone*.

Příklady.

1. Klasickým příkladem uzavřeného konvexního kužele je *konický obal* neprázdné množiny $\emptyset \neq B \subseteq \mathbb{R}^n$:

$$con\,(B) \equiv \{\, \boldsymbol{x} \in \mathbb{R}^n \ :\ \boldsymbol{x} = \sum_{\boldsymbol{z}\in C} \alpha_{\boldsymbol{z}} \cdot \boldsymbol{z} \text{ pro konečnou } \emptyset \neq C \subseteq B \text{ a } \alpha_{\boldsymbol{z}} \geq 0\,\}\,.$$

   V případě prázdné množiny B přijímáme konvenci $con\,(\emptyset) = \{\mathbf{0}\}$.
2. Jiným příkladem uzavřeného konvexního kužele je *duální kužel* pro množinu $A \subseteq \mathbb{R}^n$:

$$A^* = \{\, \boldsymbol{x} \in \mathbb{R}^n \ :\ \forall\, \boldsymbol{y} \in A \quad \langle \boldsymbol{x}, \boldsymbol{y}\rangle \geq 0\,\}\,.$$

- Kužel nazveme *jehlan*, anebo též *polyedrický kužel*, jestliže existuje konečná $B \subseteq \mathbb{R}^n$ taková, že $K = con\,(B)$, respektive *racionální jehlan*, jestliže existuje konečná množina racionálních vektorů $B \subseteq \mathbb{Q}^n$, že $K = con\,(B)$.

  Poznámka. Pro konečnou $B \subseteq \mathbb{R}^n$ platí $con\,(B) = (B^*)^*$.

  Fakt. Množina $K \subseteq \mathbb{R}^n$, $n \geq 1$ je racionální jehlan právě tehdy, jestliže $K = A^*$ kde $A \subseteq \mathbb{Q}^n$ je konečná množina racionálních vektorů.

- *Stěnou* uzavřené konvexní množiny $K \subseteq \mathbb{R}^n$ se bude rozumět její konvexní podmnožina $F \subseteq K$ která má následující vlastnost: pokud pro $\boldsymbol{y}, \boldsymbol{z} \in K$

  otevřená úsečka $(\boldsymbol{y}, \boldsymbol{z}) \equiv \{\boldsymbol{x} \in \mathbb{R}^n \ :\ \boldsymbol{x} = \alpha \cdot \boldsymbol{y} + \beta \cdot \boldsymbol{z} \quad \alpha, \beta > 0,\ \alpha + \beta = 1\}$

protíná F, tj. pokud $F \cap (\boldsymbol{y}, \boldsymbol{z}) \neq \emptyset$, pak $\boldsymbol{y}, \boldsymbol{z} \in F$ (což fakticky značí, že celá uzavřená úsečka $[\boldsymbol{y}, \boldsymbol{z}]$ leží v F). Anglický termín pro stěnu je *face*.
Tato obecná definice pojmu „stěna“ uzavřené konvexní množiny je převzatá z knihy (Brøndsted 1983). Podle ní je prázdná množina $F = \emptyset$ (i celá množina $F = K$) stěnou K.

- Ekvivalentní definice stěny v případě polytopu $K \subseteq \mathbb{R}^n$ je průsek K s jeho *isolující nadrovinou*, tj. $F = K \cap \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = k\}$, kde $\boldsymbol{y} \in \mathbb{R}^N$ a $k \in \mathbb{R}$ jsou takové, že $K \subseteq \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \langle \boldsymbol{x}, \boldsymbol{y} \rangle \geq k\}$.
Tato definice je běžnější v učebnicích konvexní geometrie; v případě obecné uzavřené konvexní množiny K ale není ekvivalentní výše uvedené definici stěny nýbrž vede k speciálnějšímu pojmu, který Brøndsted (1983) nazývá anglickým termínem *exposed face*. Nicméně v případě polytopu K jsou obě definice ekvivalentní. Také podle této speciálnější definice jsou jak prázdná množina $F = \emptyset$ tak celý polytop $F = K$ stěnami, protože volbou $\boldsymbol{y} = \mathbf{0}$ a $k = -1$, respektive $k = 0$, dostaneme příslušnou isolující nadrovinu. Dodejme ale varování, že někteří autoři v této definici nepřipouští $\boldsymbol{y} = \mathbf{0}$ a tím tedy vylučují aby polytop $K \subseteq \mathbb{R}^n$ dimense n byl svou vlastní stěnou.
- Ekvivalentní definice neprázdné stěny v případě uzavřeného konvexního kužele $K \subseteq \mathbb{R}^n$ je množina $F \subseteq K$ splňující následující tři podmínky:
  (i) $\mathbf{0} \in F$,
  (ii) $\boldsymbol{x}, \boldsymbol{y} \in F,\ \alpha, \beta \geq 0 \quad \Rightarrow \quad \alpha \cdot \boldsymbol{x} + \beta \cdot \boldsymbol{y} \in F$,
  (iii) $\boldsymbol{x}, \boldsymbol{y} \in K,\ \boldsymbol{x} + \boldsymbol{y} \in F \quad \Rightarrow \quad \boldsymbol{x}, \boldsymbol{y} \in F$.

Poznámka. Stěna (racionálního) jehlanu je opět (racionální) jehlan.

- Neprázdné stěny uzavřené (neprázdné) konvexní množiny $\emptyset \neq K \subseteq \mathbb{R}^n$ jsou opět uzavřené konvexní množiny a lze je klasifikovat podle dimense, která nabývá hodnot mezi 0 a dimensí k množiny K (nutně $k \leq n$):
  – stěny dimense 0, neboli jednoprvkové množiny, se nazývají *vrcholy*, anebo též *extremální body* (anglický termín pro vrchol je *vertex* nebo *extreme point*),
  – stěny dimense 1 se nazývají *hrany*, anglicky *edges*,
  – stěny dimense 2 v tomto textu nazývám *rovinné stěny*,
  – stěny dimense $k-1$ se nazývají *fasety*, anglicky *facets*.

  Samozřejmě, jediná stěna K mající dimensi k je celá konvexní množina K.
- Uzavřený kužel $K \subseteq \mathbb{R}^n$ je *zaostřený*, jestliže $K \cap (-K) = \{\mathbf{0}\}$.

Fakt. Uzavřený kužel $K \subseteq \mathbb{R}^n$ je zaostřený, právě když

$$\exists\, \boldsymbol{y} \in \mathbb{R}^n \quad \forall\, \boldsymbol{x} \in K \setminus \{\mathbf{0}\} \qquad \langle \boldsymbol{x}, \boldsymbol{y} \rangle \equiv \sum_{i=1}^{n} x_i \cdot y_i > 0\,.$$

- *Paprsek* generovaný nenulovým vektorem $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{R}^n$ je množina $R = \{\alpha \cdot \boldsymbol{x} \,:\, \alpha \geq 0\}$.

- Je-li $K \subseteq \mathbb{R}^n$ zaostřený uzavřený konvexní kužel, pak jeho paprsek $R \subseteq K$ nazveme *extremální* (v K), jestliže

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in K \quad \boldsymbol{x} + \boldsymbol{y} \in R \ \text{ implikuje } \ \boldsymbol{x}, \boldsymbol{y} \in R\,.$$

Jinými slovy, R je stěnou K, konkrétně hranou K. Vlastně, v případě zaostřeného neprázdného polyedrického kužele (= jehlanu) jsou pojmy extremální paprsek a hrana synonyma.

Fakt (důsledek Krein-Milmanovy věty). Každý zaostřený uzavřený konvexní kužel K je konický obal svých extremálních paprsků. Dokonce K je konický obal jakékoli své podmnožiny, v níž je každý paprsek zastoupen alespoň jedním generujícím vektorem.

Fakt. Zaostřený uzavřený kužel je racionální jehlan právě tehdy, když má konečně mnoho extremálních paprsků, přičemž každý paprsek je generovaný racionálním vektorem.

Fakt. Třída stěn zaostřeného jehlanu uspořádaná inkluzí tvoří atomický svaz.

(i) Jednotkový prvek svazu je celé K.
(ii) Nulový prvek svazu je kužel $\{\mathbf{0}\}$.
(iii) Atomy tohoto svazu jsou extremální paprsky K.

B.5 Pojmy z teorie grafů

- *Hybridním grafem* se bude rozumět trojice $G = (N, \mathcal{L}, \mathcal{A})$, kde
  (i) N je neprázdná konečná množina *uzlů*,
  (ii) $\mathcal{L}$ je množina neorientovaných hran, neboli *linek*, tj. dvouprvkových podmnožin N: $\mathcal{L} \subseteq \{A \subseteq N \,:\, |A| = 2\}$.
  Zápis: $a — b$ v $G \quad \Leftrightarrow \quad \{a, b\} \in \mathcal{L}$,
  (iii) $\mathcal{A}$ je množina orientovaných hran, neboli *šipek*, tj. uspořádaných dvojic různých uzlů: $\mathcal{A} \subseteq N \times N \setminus \{(a,a) \,:\, a \in N\}$.
  Zápis: $a \rightarrow b$ v $G \quad \Leftrightarrow \quad (a, b) \in \mathcal{A}$,
  (iv) požaduje se *nenásobnost hran*:

$$\forall\, a, b \in N \ a \neq b \quad a \rightarrow b \ \text{ v } \ G \quad \Rightarrow \quad \neg\{b \rightarrow a \ \text{ v } \ G\} \ \wedge \ \neg\{b — a \ \text{ v } \ G\}\,.$$

Jestliže N je množina uzlů hybridního grafu G, pak říkáme, že G je *graf nad N*. Hybridní graf je *neorientovaný*, jestliže $\mathcal{A} = \emptyset$ a *orientovaný*, jestliže $\mathcal{L} = \emptyset$.

Poznámka. Způsob zavedení grafů vlastně zaručuje absenci smyček, tj. hran spojujících nějaký uzel se sebou samým.

- Dvojice uzlů $[a, b]$ je *hrana* v G jestliže nastane jedna ze situací:

(i) $a — b$ v G (což je totéž, co $b — a$ v G),
(ii) $a \to b$ v G,
(iii) $b \to a$ v G.

Poznamenejme, že formálně je hrana zavedena jako uspořádaná dvojice, ale fakticky je to neuspořádaná dvojice.

- Hybridní graf $\tilde{G} = (\tilde{N}, \tilde{\mathcal{L}}, \tilde{\mathcal{A}})$ nazveme *podgrafem* hybridního grafu $G = (N, \mathcal{L}, \mathcal{A})$, jestliže $\tilde{N} \subseteq N$, $\tilde{\mathcal{L}} \subseteq \mathcal{L}$ a $\tilde{\mathcal{A}} \subseteq \mathcal{A}$.
- *Indukovaný podgraf* grafu G pro neprázdnou množinu uzlů $T \subseteq N$ je graf

$$G_T = (T, \mathcal{L}_T, \mathcal{A}_T), \quad \text{kde } \mathcal{L}_T = \mathcal{L} \cap \mathcal{P}(T) \text{ a } \mathcal{A}_T = \mathcal{L} \cap (T \times T)\,.$$

- Je-li $a \in N$ uzel hybridního grafu, pak

$$\begin{aligned} ne_G(a) &= \{b \in N \,:\, a — b \ \text{ v } \ G\} && \text{je množina } \textit{sousedů} \text{ uzlu } a, \\ pa_G(a) &= \{b \in N \,:\, b \to a \ \text{ v } \ G\} && \text{je množina } \textit{rodičů} \text{ uzlu } a, \\ ch_G(a) &= \{b \in N \,:\, a \to b \ \text{ v } \ G\} && \text{je množina } \textit{dětí} \text{ uzlu } a. \end{aligned}$$

Značení je motivováno anglickou terminologií: soused = *neighbour*, rodič = *parent*, dítě = *child*.

- *Sledem* v hybridním grafu G se bude rozumět posloupnost uzlů tohoto grafu $\boldsymbol{\rho} : a_1, \dots, a_n$, $n \geq 1$ taková, že $\forall\, i = 1, \dots, n-1$ je $[a_i, a_{i+1}]$ hrana v G.

  Krajní uzly sledu $\boldsymbol{\rho}$ jsou a_1 a a_n, *vnitřní* uzly (jen v případě $n \geq 3$) jsou $a_2, \dots, a_{n-1}$. Jestliže $A, B \subseteq N$ jsou množiny uzlů, $a_1 \in A$ a $a_n \in B$, pak řekneme, že sled *vede z A do B* (někdy též je to sled mezi A a B).

  Upozorňuji, že definice sledu zahrnuje i případ $n = 1$!
- Sled se nazývá *cesta*, jestliže navíc $a_1, \dots, a_n$ jsou *různé* uzly.
- Sled je
  - *neorientovaný* jestliže $\forall\, i = 1, \dots, n-1$ je $a_i — a_{i+1}$ v G,
  - *sestupný*, pokud $\forall\, i = 1, \dots, n-1$ platí, že buď $a_i — a_{i+1}$ v G nebo $a_i \to a_{i+1}$ v G,
  - *striktně sestupný*, jestliže $\forall\, i = 1, \dots, n-1$ je $a_i \to a_{i+1}$ v G.

Neorientované grafy

- Jestliže G je neorientovaný graf nad N, pak množinu $A \subseteq N$ nazveme *úplnou* v G, jestliže

$$\forall\, a, b \in N,\ a \neq b \ \text{ je } a — b \ \text{ v } G\,.$$

- *Klikou* grafu G se bude rozumět maximální úplná podmnožina N ve smyslu inkluze. Samotný neorientovaný graf G nad N se nazve *úplný*, jestliže N je úplná množina v G.
- Neorientovaný graf G nad N nazveme *bipartitní*, jestliže existuje rozklad $N = A \cup B$, $A \cap B = \emptyset$ takový, že pokud $a — b$ v G, pak $|\{a, b\} \cap A| = 1 = |\{a, b\} \cap B|$.
- *Neorientovaný cyklus* v hybridním grafu G nad N je posloupnost uzlů tohoto grafu $a_1, \dots, a_{n+1}$, $n \geq 3$, taková, že

(i) $a_{n+1} = a_1$,
(ii) $a_1, \ldots, a_n$ jsou různé,
(iii) $\forall\, i = 1, \ldots, n \quad a_i - a_{i+1}$ v G.

Číslo n se pak nazývá *délkou cyklu*. *Tětivou* cyklu $a_1, \ldots, a_{n+1}$, kde $n \geq 4$, se rozumí linka $a_i - a_j$ v G taková, že $1 \leq i, j \leq n$ a $1 < j - i < n - 1$.

- Hybridní graf G nad N je *souvislý*, jestliže $\forall\, a, b \in N$ existuje v G neorientovaná cesta z a do b.
- Neorientovaný graf se nazývá *strom*, jestliže je souvislý a nemá žádný neorientovaný cyklus. *List* stromu G je uzel, který má jediného souseda v G. Snadno lze nahlédnout, že každý strom má list, v případě alespoň dvou uzlů alespoň dva listy.

  Neorientovaný graf, který nemá neorientovaný cyklus, se nazývá *les*.

  Fakt. Neorientovaný graf G nad N je strom právě tehdy, jestliže $\forall\, a, b \in N$ v G existuje právě jedna cesta z a do b.

Orientované grafy

- *Orientovaný cyklus* v hybridním grafu G nad N je posloupnost uzlů v tomto grafu $a_1, \ldots, a_{n+1}$, $n \geq 3$ taková, že

  (i) $a_{n+1} = a_1$,
  (ii) $a_1, \ldots, a_n$ jsou různé,
  (iii) $\forall\, i = 1, \ldots, n \quad a_i \to a_{i+1}$ v G.

  Číslo n se nazývá délkou cyklu. Orientovaný graf se nazývá *acyklický*, jestliže v něm neexistuje orientovaný cyklus.

  Fakt. Orientovaný graf G nad N je acyklický, právě když existuje očíslování $a_1, \ldots, a_n$, $n = |N|$ jeho uzlů, které je *souhlasné s orientací šipek* v G, tj.

$$\forall\, 1 \leq i, j \leq n \quad a_i \to a_j \text{ v } G \quad \Rightarrow \quad i < j\,.$$

- Řekneme, že uzel a je *předek* uzlu b v hybridním grafu G, respektive že b je *potomek* a v G, jestliže v G existuje sestupný sled z a do b (ekvivalentně, z a vede v G sestupná cesta do b). Množina potomků uzlu a v G se bude značit symbolem $ds_G(a)$, množina předků množiny $A \subseteq N$ symbolem $an_G(A)$.

  Poznámka. V orientovaném grafu G znamená, že a je předek b, že buď $a = b$ nebo v G vede z a do b striktně sestupný sled, ekvivalentně striktně sestupná cesta. Značení je motivováno anglickou terminologií: potomek = *descendant*, předek = *ancestor*.

B.6 Statistická rozdělení

- Buď P rozdělení nějakého (reálného) náhodného vektoru, tedy pravděpodobnostní míra na euklidovském prostoru $\mathbb{R}^N$, přičemž $|N| = n \geq 1$. Jeho základní numerickou charakteristikou je *vektor středních hodnot* $\boldsymbol{e} = [e_i]_{i \in N}$, jehož složky jsou definovány následovně:

$$e_i = \int_{\mathbb{R}^N} x_i \, \mathrm{d}P(\boldsymbol{x}) \qquad \text{pro } i \in N,$$

kde x_i označuje i-tou složku vektoru $\boldsymbol{x} \in \mathbb{R}^N$. V praxi se téměř výhradně uvažují rozdělení, pro něž zmíněné integrály e_i existují a jsou konečné.

- Další významnou numerickou charakteristikou P je jeho *kovarianční matice* $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j \in N}$, jejíž složky jsou definovány následující formulí:

$$\sigma_{ij} = \int_{\mathbb{R}^N} (x_i - e_i) \cdot (x_j - e_j) \, \mathrm{d}P(\boldsymbol{x}),$$

za implicitního předpokladu, že P má vektor středních hodnot. V praxi uvažovaná rozdělení mají zpravidla konečné hodnoty σ_{ij}, nazývané *kovariance*.

- *Regulární gaussovské rozdělení* je vícerozměrné spojité rozdělení. Pro $n \in \mathbb{N}$ sídlí příslušné n-rozměrné rozdělení na (výběrovém) prostoru $\mathbb{R}^N$, přičemž $|N| = n$. Je parametrizováno vektorem $\boldsymbol{e} \in \mathbb{R}^N$ a positivně definitní $N \times N$-maticí $\boldsymbol{\Sigma}$. Formule

$$f_{\boldsymbol{e},\boldsymbol{\Sigma}}(x) = \frac{1}{\sqrt{(2\pi)^{|N|} \cdot \det(\boldsymbol{\Sigma})}} \cdot \exp\left(- \frac{(x - \boldsymbol{e})^\top \cdot \boldsymbol{\Sigma}^{-1} \cdot (x - \boldsymbol{e})}{2} \right) \qquad \text{pro } x \in \mathbb{R}^N,$$

definuje hustotu příslušné *regulární gaussovské míry* vzhledem k Lebesgueově míře na $\mathbb{R}^N$. Značí se symbolem $\mathcal{N}(\boldsymbol{e}, \boldsymbol{\Sigma})$. Vektorem středních hodnot je pak $\boldsymbol{e}$ a kovarianční maticí $\boldsymbol{\Sigma}$.

- Rozdělení $\mathcal{N}(\boldsymbol{e}, \boldsymbol{\Sigma})$ se zavádí, ovšem komplikovanějším způsobem, i v případě obecnější positivně semidefinitní matice $\boldsymbol{\Sigma}$. Ale když matice $\boldsymbol{\Sigma}$ není positivně definitní (= regulární), pak definovaná míra již nesídlí na celém prostoru $\mathbb{R}^N$, nýbrž je soustředěná na nějakém afinním (= posunutém lineárním) podprostoru $\mathbb{R}^N$. Potom hovoříme o *singulární gaussovské míře* či *rozdělení*. Nicméně i v tomto obecnějším případě platí, že $\boldsymbol{e}$ je vektorem středních hodnot a $\boldsymbol{\Sigma}$ kovariační maticí.

- *Multinomické rozdělení* je vícerozměrné diskrétní rozdělení (= rozdělení „tabulek četností“, speciálně kontingenčních tabulek). Pro $r \in \mathbb{N}$ má příslušné r-rozměrné rozdělení následující parametry:
  (i) $\theta_1, \ldots, \theta_r$ takové, že $\theta_k > 0$ a $\sum_{k=1}^{r} \theta_k = 1$,
  (ii) $d \in \mathbb{N}$.

  Sídlí na konečné množině (= výběrovém prostoru)

$$\left\{ [d_1, \ldots, d_r] \; : \; d_k \in \mathbb{N}, \; \textstyle\sum_{k=1}^{r} d_k = d \right\}.$$

Hustota (vůči aritmetické míře) je určena vzorcem:

$$p([d_1, \ldots, d_r]) = \frac{d!}{d_1! \cdot \ldots \cdot d_r!} \cdot \theta_1^{d_1} \cdot \ldots \cdot \theta_r^{d_r}.$$

Značení $\mathcal{M}(\theta_1, \ldots, \theta_k, d)$.

Poznámka. Jestliže $\xi_1, \ldots, \xi_d$ je náhodný výběr (viz Definice 7.5) z rozdělení na $X = \{1, \ldots, r\}$ s hustotou $p(k) = \theta_k$, $k = 1, \ldots, r$, pak vektor $[\rho_1, \ldots, \rho_r]$, kde $\rho_k = |\{\ell \,:\, \xi_\ell = k\}|$, má výše uvedené rozdělení. Tento vektor je vlastně „tabulka četností“ sestavená na základě databáze (tj. posloupnosti hodnot náhodného výběru).

- *Dirichletovo rozdělení* je vícerozměrné spojité rozdělení (= rozdělení na „diskrétních pravděpodobnostních mírách“). Pro $r \in \mathbb{N}$, $r \geq 2$ má příslušné rozdělení parametry $\alpha_1, \ldots, \alpha_r$ takové, že $\alpha_k > 0$ pro $k = 1, \ldots, r$. Upozorňuji, že mezi parametry $\alpha_1, \ldots, \alpha_r$ *není* vazba! Dirichletovo rozdělení sídlí na množině (= výběrovém prostoru)

$$\Theta = \left\{ [\theta_1, \ldots, \theta_r] \,:\, \theta_k > 0,\ \textstyle\sum_{k=1}^{r} \theta_k = 1 \right\}.$$

Je zadáno hustotou vůči „transformované“ Lebesgueově míře. Přesněji: uvažujme restrikci $(r-1)$-rozměrné Lebesgueovy míry na

$$\left\{ [\theta_1, \ldots, \theta_{r-1}] \,:\, \theta_k > 0,\ \textstyle\sum_{k=1}^{r-1} \theta_k < 1 \right\}$$

a poté transformaci $[\theta_k]_{k=1}^{r-1} \mapsto [\theta_k]_{k=1}^{r}$, kde $\theta_r = 1 - \sum_{k=1}^{r-1} \theta_k$.
Hustota Dirichletova rozdělění (vůči zmíněné dominující míře) je

$$f([\theta_1, \ldots, \theta_r]) = \frac{\Gamma(\alpha_1 + \ldots + \alpha_r)}{\Gamma(\alpha_1) \cdot \ldots \cdot \Gamma(\alpha_r)} \cdot \theta_1^{\alpha_1 - 1} \cdot \ldots \cdot \theta_r^{\alpha_r - 1},$$

kde $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} \cdot e^{-y}\, \mathrm{d}y$, $\alpha > 0$ je hodnota *Gamma funkce*. Značení Dirichletova rozdělení je $\mathcal{D}(\alpha_1, \ldots, \alpha_r)$.
Vektor středních hodnot a kovarianční matice jsou určeny těmito vztahy:

$$\begin{aligned} \mathrm{E}(\theta_k) &= \tfrac{\alpha_k}{\alpha}, && \text{kde } \ \alpha = \textstyle\sum_{k=1}^{r} \alpha_r, \\ \mathrm{cov}(\theta_k, \theta_l) &= \tfrac{-\alpha_k \cdot \alpha_l}{\alpha^2 \cdot (\alpha+1)}, && k \neq l, \\ \mathrm{var}(\theta_k) &= \tfrac{(\alpha - \alpha_k) \cdot \alpha_k}{\alpha^2 \cdot (\alpha+1)}. \end{aligned}$$

Následující pozorování je důležité z bayesovského hlediska, viz Definice 7.9 v § 7.2, kde je zavedena příslušná terminologie.

Fakt (základní fakt z hlediska bayesovského přístupu k učení). Nechť $\boldsymbol{\theta} = [\theta_k]_{k=1}^{r}$, $r \geq 1$ označuje vektor parametrů r-rozměrného multinomického rozdělení a $\boldsymbol{d} = [d_k]_{k=1}^{r}$, $d_k \in \mathbb{N}$, kde $\sum_{k=1}^{r} d_k = d$, odpovídající tabulku četností pro databázi délky $d \in \mathbb{N}$. Jestliže

(i) hustota výběrové pravděpodobnosti $h(\boldsymbol{d}|\boldsymbol{\theta})$ má multinomické rozdělení $\mathcal{M}(\theta_1, \ldots, \theta_k, d)$,

(ii) apriorní hustota $\pi(\boldsymbol{\theta})$ má Dirichletovo rozdělení $\mathcal{D}(\alpha_1, \ldots, \alpha_r)$,

pak aposteriorní hustota $\pi(\boldsymbol{\theta}|\boldsymbol{d})$ má rozdělení $\mathcal{D}(\alpha_1 + d_1, \ldots, \alpha_r + d_r)$.

Vskutku,

$$\pi(\boldsymbol{\theta}|\boldsymbol{d}) \sim \pi(\boldsymbol{\theta}) \cdot h(\boldsymbol{d}|\boldsymbol{\theta}) \sim \prod_{k=1}^{r} \theta_k^{\alpha_k - 1} \cdot \prod_{k=1}^{r} \theta_k^{d_k} = \prod_{k=1}^{r} \theta_k^{(\alpha_k + d_k) - 1},$$

kde symbol $\sim$ značí rovnost až na multiplikativní faktor. Říkáme, že Dirichletovo rozdělení je *konjugované* k multinomickému.

Literatura

Andersson SA, Madigan D, Perlman MD (2001): Alternative Markov properties for chain graphs. *Scandinavian Journal of Statistics*, **28**(1), 33–85.

Andersson SA, Perlman MD (1993): Lattice models for conditional independence in a multivariate normal distribution. *Annals of Statistics*, **21**(3), 1318–1358.

Anděl J (1985): *Matematická statistika.* SNTL (Alfa), Praha.

Anděl J (2007): *Základy matematické statistiky.* MatfyzPress, Praha.

Bouckaert RR (1995): *Bayesian belief networks – from construction to inference.* PhD thesis, University of Utrecht.

Brøndsted A (1983): *An Introduction to Convex Polytopes.* Springer, New York.

Cox DR, Wermuth N (1996): *Multivariate Dependencies – Models, Analysis and Interpretation.* Monographs on Statistics and Applied Probability 67, Chapman and Hall, London.

Cowell RG, Dawid AP, Lauritzen SL, Spiegelhalter DJ (1999): *Probabilistic Networks and Expert Systems.* Springer, New York.

Darroch JN, Lauritzen SL, Speed TP (1980): Markov fields and log-linear interaction models for contingency tables. *Annals of Statistics*, **8**(3), 522–539.

Dawid AP (1979): Conditional independence in statistical theory. *Journal of the Royal Statistical Society B*, **41**(1), 1–31.

de Campos LM, Huete JF, Moral S (1993): *Probability intervals, a tool for uncertain reasoning.* Technical report DECSAI-93206, University of Granada.

Dempster AP (1967): Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, **11**, 325–339.

Dempster AP (1972): Covariance selection. *Biometrics*, **28**, 157–175.

Drton M, Sturmfels B, Sullivant S (2009): *Lectures on Algebraic Statistics.* Oberwolfach Seminars 39, Birkhäuser, Basel.

Florens J-P, Mouchart M, Rolin J-M (1990): *Elements of Bayesian Statistics.* Marcel Dekker, New York.

Frydenberg M (1990): The chain graph Markov property. *Scandinavian Journal of Statistics*, **17**(4), 333–353.

Geiger D, Pearl J (1990): On the logic of causal models. In: *Uncertainty in Artificial Intelligence 4* (R. D. Shachter, T. S. Lewitt, L. N. Kanal, J. F. Lemmer eds.), North-Holland, Amsterdam, 3–14.

Goodman LA (1973): The analysis of multidimensional contingency tables when some variables are posterio to others – a modified path analysis approach. *Biometrika*, **60**, 179-192.

Haberman SJ (1974): *The Analysis of Frequency Data.* University of Chicago Press, Chicago.

Hájek P, Havránek T, Jiroušek R (1992): *Uncertain Information Processing in Expert Systems.* CRC Press, Boca Raton.

Heckerman D (1995): *A tutorial on learning Bayesian networks.* Technical report MSR-TR-95-06, Microsoft Research, Redmont.

Hemmecke R, Morton J, Shiu A, Sturmfels B, Wienand O (2008): Three counter-examples on semi-graphoids. *Combinatorics, Probability and Computing*, **17**, 239–257.

Howard RA, Matheson JE (1984): Influence diagrams. In: *Readings in the Principles and Applications of Decision Analysis* (R. A. Howard, J. E. Matheson eds.), Strategic Decision Group, Menlo Park, 21–762.

Chickering DM (1995): A transformational characterization of equivalent Bayesian network structures. In: *Uncertainty in Artificial Intelligence 11* (P. Besnard, S. Hanks eds.), Morgan Kaufmann, San Francisco, 87–98.

Jensen FV (2001): *Bayesian networks and Decision Graphs.* Springer, New York.

Kauermann G (1996): On a dualization of graphical Gaussian models. *Scandinavian Journal of Statistics*, **23**(1), 105–116.

Koster JTA (1996): Markov properties of nonrecursive causal models. *Annals of Statistics*, **24**(5), 2148–2177.

Lauritzen SL (1982): *Lectures on Contingency Tables*, 2nd edition. University of Aalborg Press, Aalborg.

Lauritzen SL (1996): *Graphical Models.* Clarendon Press, Oxford.

Lauritzen SL, Spiegelhalter DJ (1988): Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society B*, **50**(2), 157–224.

Lauritzen SL, Wermuth N (1984): *Mixed interaction models.* Research report R-84-8, Inst. Elec. Sys., University of Aalborg, Aalborg.

Loève M (1955): *Probability Theory, Foundations, Random Processes.* D. van Nostrand, Toronto.

Malvestuto FM (1992): A unique formal system for binary decomposition of database relations, probability distributions and graphs. *Information Sciences*, **59**, 21–52.

Matúš F (1995): Conditional independences among four random variables II. *Combinatorics, Probability and Computing*, **4**(4), 407–417.

Matúš F (1999): Conditional independences among four random variables III., final conclusion. *Combinatorics, Probability and Computing*, **8**(3), 269–276.

Matúš F, Studený M (1995): Conditional independences among four random variables I. *Combinatorics, Probability and Computing*, **4**(4), 269–278.

Meek C (1995): Causal inference and causal explanation with background knowledge. In: *Uncertainty in Artificial Intelligence 11* (P. Besnard, S. Hanks eds.), Morgan Kaufmann, San Francisco, 403–410.

Moussouris J (1974): Gibbs and Markov properties over undirected graphs. *Journal of Statistical Physics*, **10**(1), 11–31.

Mouchart M, Rolin J-M (1984): A note on conditional independence with statistical applications. *Statistica*, **44**(4), 557–584.

Neapolitan RE (1990): *Probabilistic Reasoning in Expert Systems – Theory and Algorithms.* John Wiley, New York.

Pearl J (1988): *Probabilistic Reasoning in Intelligent Systems – Networks of Plausible Inference.* Morgan Kaufmann, San Francisco.

Pearl J, Paz A (1987): Graphoids, graph-based logic for reasoning about relevance relations. In: *Advances in Artificial Intelligence II* (B. Du Boulay, D. Hogg, L. Steels eds.), North-Holland, Amsterdam, 357–363.

Perez A (1977): ε-admissible simplifications of the dependence structure of a set of random variables. *Kybernetika*, **13**(6), 439–449.

Perez A, Jiroušek R (1985): Constructing an intensional expert system (INES). In: *Medical Decision Making: Diagnostic Strategies and Expert Systems* (J. V. van Bemmel, F. Grémy, J. Zvárová eds.), North-Holland, Amsterdam, 307–315.

Richardson TS (1996): A polynomial-time algorithm for deciding Markov equivalence of directed cyclic graphical models. In: *Uncertainty in Artificial Intelligence 12* (E. Horvitz, F. Jensen eds.), Morgan Kaufmann, San Francisco, 462–469.

Rudin W (2003): *Analýza v reálném a komplexním oboru.* Academia, Praha; český překlad třetího anglického vydání *Real and Complex Analysis*, McGraw-Hill 1987.

Sagiv Y, Walecka SF (1982): Subset dependencies and completeness result for a subclass of embedded multivalued dependencies. *Journal of Association for Computing Machinery*, **29**(1), 103–117.

Shafer G (1976): *A Mathematical Theory of Evidence.* Princeton Univerity Press, Princeton.

Shafer G (1996): *Probabilistic Expert Systems.* CBMS-NSF Regional Conference Series in Applied Mathematics 67, SIAM, Philadelphia.

Shachter RD (1986): Evaluating influence diagrams. *Operation Research*, **34**, 871–882.

Shenoy PP (1994): Conditional independence in valuation-based systems. *International Journal of Approximate Reasoning*, **10**(3), 203–234.

Smith JQ (1989): Influence diagrams for statistical modelling. *Annals of Statistics*, **17**(2), 654–672.

Speed TP (1979): A note on nearest-neighbour Gibbs and Markov probabilities. *Sankhyā, the Indian Journal of Statistics A*, **41**, 184–197.

Spiegelhalter DJ, Lauritzen SL (1990): Sequential updating of conditional probabilities on directed graphical structures. *Networks*, **20**(5), 579–605.

Spirtes P, Glymour C, Scheines R (1993): *Causation, Prediction and Search.* Springer, New York.

Spohn W (1980): Stochastic independence, causal independence and shieldability. *Journal of Philosophical Logic*, **9**(1), 73–99.

Spohn W (1988): Ordinal conditional functions – a dynamic theory of epistemic states. In: *Causation in Decision, Belief Change, and Statistics II* (W. L. Harper, S. Skyrms eds.), Kluwer, Dordrecht, 105–134.

Studený M (1992): Conditional independence relations have no finite complete characterization. In: *Information Theory, Statistical Decision Functions and Random Processes, Transactions of 11th Prague Conference B* (S. Kubík, J. Á. Víšek eds.), Kluwer/Academia, Dordrecht/Prague, 377–396.

Studený M (2005): *Probabilistic Conditional Independence Structures.* Springer, London.

van Putten C, van Shuppen JH (1985): Invariance properties of conditional independence relation. *Annals of Probability*, **13**(3), 934–945.

Vejnarová J (1999): Conditional independence in possibility theory. In: *Proceedings of ISIPTA 99 (1st International Symposium on Imprecise Probabilities and their Applications)* (G. de Cooman, F. G. Cozman, S. Moral, P. Walley eds.), 343–351.

Verma T, Pearl J (1991): Equivalence and synthesis of causal models. In: *Uncertainty in Artificial Intelligence 6* (P. P. Bonissone, M. Henrion, L. N. Kanal, J. F. Lemmer eds.), Elsevier, Amsterdam, 220–227.

Wermuth N (1976): Analogies between multiplicative models for contingency tables and covariance selection. *Biometrics*, **32**, 95–108.

Whittaker J (1990): *Graphical Models in Applied Multivariate Statistics.* John Wiley, Chichester.

Wong SKM, Wang ZW (1994): On axiomazation of probabilistic conditional independencies. In: *Uncertainty in Artificial Intellingence 10* (R. L. Mantaras, D. Poole eds.), Morgan Kaufmann, San Francisco, 591–597.

Wright S (1934): The method of path coefficients. *Annals of Mathematical Statistics*, **5**, 161–215.

Zadeh LA (1978): Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, **1**(1), 3–28.

Rejstřík