POPIS STRUKTUR PODMÍNĚNÉ STOCHASTICKÉ NEZÁVISLOSTI POMOCÍ FORMULÍ SOUČINOVÉHO TYPU

Milan Studený, ÚTIA ČSAV
Pod vodárenskou věží 4, 182 08 Praha 8
tel: 815 2304, fax: 847452
e-mail: studeny@cspgas11.bitnet

V tomto příspěvku je krátce popsán alternativní přístup k matematickému popisu struktur podmíněné stochastické nezávislosti (pomocí tkz. imsetů). Tento přístup je vztažen ke klasickým metodám jejich popisu (pomocí tkz. závislostních modelů či pomocí grafů). Hlavní výsledek je tvrzení , že oba zmíněné přístupy (tj. pomocí imsetů a pomocí závislostních modelů) jsou ekvivalentní platnosti určité součinové formule odpovídající příslušnému modelu struktury podmíněné nezávislosti.

AMS klasifikace : 68T 30 , 62B 10.

Úvod

Zájem o studium vlastností podmíněné stochastické nezávislosti pramení z teorie pravděpodobnostních expertních systémů. Je to proto, že každý výrok o podmíněné nezávislosti lze interpretovat jakožto určitý kvalitativní vztah mezi symptomy či veličinami, které jsou popisovány. Tudíž pomocí takovýchto informací je možné určit vhodný strukturální model pravděpodobnostního expertního systému. Lze říci, že v teoretických základech různých přístupů ke kvalitativnímu popisu pravděpodobnostních modelů (influenční diagramy, Markovské sítě) je skryt pojem podmíněné nezávislosti (PN). Význam PN pro teorii pravděpodobnostních expertních systémů poprvé explicitně zdůraznil Pearl [1986], nicméně mnoho jiných autorů z této oblasti více či méně explicitně pracuje s tímto pojmem [Pearl & Geiger & Verma 1990], [Shachter 1988], [Smith 1989], [Spiegelhalter & Lauritzen 1990], [Ur & Paz 1991].

Pokud pomineme grafické přístupy, které neumožňují popsat všechny možné struktury PN, pak "klasický" přístup k popisu nezávislostních struktur je pomocí pojmu *závislostního modelu*, konkrétněji pomocí pojmu *semigrafoidu* [Pearl & Paz 1987]. Tento přístup motivoval pokusy "axiomatizovat" PN tj. charakterizovat vzájemné vztahy mezi výroky o PN jednoduchým syntaktickým způsobem. Hypotéza zněla, že modely PN lze charakterizovat jako závislostní modely uzavřené na konečný počet odvozovacích pravidel, které Pearl a jeho spolupracovníci nazývali axiomy. Nicméně, tato hypotéza byla vyvrácena v [Studený 1992], kde je dokázáno, modely PN nelze charakterizovat pomocí konečného počtu takovýchto odvozovacích pravidel. Tento výsledek nedávno rozšířili Geiger a Pearl [1993], kteří ukázali, že dokonce tkz. "disjunktivní" odvozovací pravidla nemohou pomoci. Tyto výsledky mne vedly k pokusu navrhnout alternativní přístup k popisu struktur PN [Studený 1993a], totiž pomocí tkz. *imsetů*. Tento přístup má naději odstranit výše zmíněnou nevýhodu klasického popisu struktur PN.

Cílem tohoto příspěvku je ukázat na další rys tohoto alternativního přístupu: platnost "imsetového" modelu struktury PN je ekvivalentní platnosti jisté formule součinového typu (pro příslušnou vícerozměrnou pravděpodobnostní míru). Tato ekvivalentní definice může usnadnit globální interpretaci takovýchto modelů. Tento text především přináší informaci; důkazy uvedených tvrzení a lemmat zde nejsou. Čtenář je může nalézt v rozšířené anglické versi [Studený 1993b].

Základní značení

V celém tomto textu se zabýváme následující situací:
Je dána konečná, alespoň dvouprvková množina N nazývaná *základní množina*.
Pro disjunktní množiny $A, B \subset N$ píšeme často AB namísto $A \cup B$. Třída všech podmnožin N bude označena $exp\, N$; třída minimálně dvouprvkových podmnožin $\mathcal{U}$ tj.
$\mathcal{U} = \{S \subset N,\ card\, S \geq 2\}$.
Každé množině $T \subset N$ přiřadíme funkci δ_T na $exp\, N$ (případně restringovanou na $\mathcal{U}$) následujícím způsobem:

$$\delta_T(S) = \begin{cases} 1 & \text{jestliže } S = T \\ 0 & \text{jestliže } S \neq T. \end{cases}$$

Je-li $\{X_i\,;\, i \in N\}$ soubor neprázdných konečných množin, $x \in \prod_{i \in N} X_i$ prvek příslušného kartézského součinu a $\emptyset \neq S \subset N$ množina, pak projekce x do $\prod_{i \in S} X_i$ bude označena x_S :
$x_S \in \prod_{i \in S} X_i$ je určeno vztahem $[\ \forall i \in S\ \ (x_S)_i = x_i\ ]$.
Je-li a, b dvojice reálných funkcí na konečné množině X jejich *skalární součin* bude označován $\langle a, b \rangle$:
$\langle a, b \rangle = \sum_{x \in X} a(x) \cdot b(x)$.
Množina celých čísel bude označena $\mathbb{Z}$, množina nezáporných celých čísel $\mathbb{Z}^+$ (tj. včetně nuly), množina přirozených čísel (bez nuly) $\mathbb{N}$ a množina reálných čísel $\mathbb{R}$.
Je-li Q pravděpodobnostní míra na Y (či funkce na $exp\, Y$) a $y \in Y$ budeme pro stručnost psát $Q(y)$ namísto $Q(\{y\})$.
Kladnou a zápornou část funkce $w : exp\, N \to \mathbb{Z}$ budeme označovat symboly w_+ a w_- :

$$w_+(S) = max\,\{w(S), 0\} \qquad \text{pro } S \subset N$$
$$w_-(S) = max\,\{-w(S), 0\} \qquad \text{pro } S \subset N.$$

§ 1 Formule součinového typu daná imsetem

Tento paragraf obsahuje pouze základní definice s příslušným komentářem. Zavádíme pojem imsetu, jeho přirozeného rozšíření a pojem formule součinového typu dané imsetem.

Def. 1 (imset)
Každá celočíselná funkce na $\mathcal{U}$ se nazývá *imset* (*na* $\mathcal{U}$). Třída všech imsetů na $\mathcal{U}$ bude označena $Z(\mathcal{U})$. Nezáporné imsety se nazývají *multisety*, jejich třída se značí $Z^+(\mathcal{U})$. Základní operace s imsety tj. sčítání, odčítání a násobení celým číslem jsou definovány po složkách. Imset $u \in Z(\mathcal{U})$ se nazývá *normalizovaný* jestliže soubor čísel $\{u(S); S \in \mathcal{U}\}$ je nesoudělný. Triviálními příklady imsetů jsou *nulový imset* (funkce přiřazující nulu každé množině z $\mathcal{U}$) značený symbolem 0 a funkce δ_T pro $T \in \mathcal{U}$.

Poznámka Termín "multiset" pochází z kombinatoriky [Aigner 1979] zatímco slovo imset je anglická zkratka pro integer-valued multiset.

Někdy je vhodné uvažovat imsety na $\mathcal{U}$ jakožto funkce na $exp\, N$. Korektnost definice příslušného rozšíření je založena na následujícím evidentním lemmatu.

Lemma Každý imset $u \in Z(\mathcal{U})$ má jednoznačně určené rozšíření $\overline{u} : exp\, N \to \mathbb{Z}$ splňující následující podmínky:
$(N.1)\ \ \sum\{\overline{u}(S); S \subset N\} = 0$
$(N.2)\ \ \forall r \in N \quad \sum\{\overline{u}(S);\ r \in S \subset N\} = 0$.
Toto přiřazení definuje vzájemně jednoznačnou korespondenci mezi $Z(\mathcal{U})$ a třídou celočíselných funkcí na $exp\, N$ splňujících $(N.1) - (N.2)$.

Def. 2 (přirozené rozšíření)
Je-li dán imset u na $\mathcal{U}$, pak jeho jednoznačně určené rozšíření $\overline{u}$ na $exp\,N$ splňující podmínky $(N.1)$ – $(N.2)$ se nazývá *přirozené rozšíření u*. Vždy bude označeno přidáním pruhu nad původní symbol.

Cílem našeho studia jsou struktury PN pro konečný počet náhodných veličin (tkz. náhodné systémy). Nicméně, v tomto textu se omezujeme na konečně-hodnotové náhodné veličiny. V literatuře je běžné zmiňovat se o náhodných veličinách, ale fakticky pracovat s pravděpodobnostními mírami (totiž s jejich rozděleními). Nicméně, v tomto textu se jak mluví, tak přímo pracuje s pravděpodobnostními mírami.

Vymezme tedy nejprve příslušnou třídu pravděpodobnostních měr. Neboť se jedná o rozdělení náhodných systémů indexovaných základní množinou N, jejich definiční obory budou kartézské součiny indexované N.

Def. 3 (pravděpodobnostní míra nad N)
Pravděpodobnostní míra nad N (s konečným definičním oborem) je vymezena souborem neprázdných konečných množin $\{X_i\ ;\ i \in N\}$ a pravděpodobnostní mírou na kartézském součinu $\prod_{i\in N} X_i$.
Jestliže máme $\emptyset \neq S \subsetneq N$ pak pod příslušnou *marginální mírou P* rozumíme pravděpodobnostní míru P^S na $\prod_{i\in S} X_i$ definovanou následovně:
$P^S(A) = P(A \times \prod_{i\in N\setminus S} X_i)$ kdykoliv $A \subset \prod_{i\in S} X_i$.
Budeme ji označovat symbolem původní míry doplněným horním indexem, který identifikuje marginální prostor. Marginální míra P na $\prod_{i\in N} X_i$ je míra P sama, tj. $P^N \equiv P$.

Jak bylo již naznačeno, každý předpoklad týkající se struktury PN pravděpodobnostní míry je ekvivalentní platnosti určité formule součinového typu. Následuje příslušná definice.

Def. 4 (formule součinového typu)
Nechť P je pravděpodobnostní míra nad N a u imset na $\mathcal{U}$. Řekneme, že P splňuje *formuli součinového typu* danou imsetem u jestliže platí:

$$(1.1)\quad \forall x \in \prod_{i\in N} X_i \quad \prod_{\emptyset\neq S\subset N} (P^S(x_S))^{\overline{u}_+(S)} = \prod_{\emptyset\neq S\subset N} (P^S(x_S))^{\overline{u}_-(S)}.$$

Podmínka (1.1) je zmíněná formule.

Poznámka Pokud přijmeme konvenci $P^\emptyset(x_\emptyset) = 1$ můžeme též psát:

$$(1.2)\quad \forall x \in \prod_{i\in N} X_i \quad \prod_{S\subset N} (P^S(x_S))^{\overline{u}_+(S)} = \prod_{S\subset N} (P^S(x_S))^{\overline{u}_-(S)}.$$

Následující jednoduchý příklad ilustruje zavedený pojem.

Příklad 1
Uvažujeme $S, T \in \mathcal{U}$ kde $S \cap T = \emptyset$ a $S \cup T = N$; položme $u = \delta_N - \delta_S - \delta_T$. Potom $\overline{u} = u + \delta_\emptyset$ a formule součinového typu daná u má následující tvar:

$$(1.3)\quad \forall x \in \prod_{i\in N} X_i \quad P(x) = P^S(x_S) \cdot P^T(x_T).$$

Jedná se tedy o ekvivalentní definici stochastické nezávislosti S a T.

§ 2 Popis struktur PN

V tomto paragrafu připomeneme "klasické" způsoby popisu struktur PN, zejména pomocí závislostních modelů.

Def. 5 (podmíněná nezávislost)
Nechť P je pravděpodobnostní míra nad N a (A, B, C) je trojice navzájem disjunktních podmnožin N kde A a B jsou neprázdné. Řekněme, že A je *podmíněně nezávislá na B za podmínky C* a píšeme $A \perp B|C\ (P)$ jestliže
$\forall x \in \prod_{i \in N} X_i \quad P^{ABC}(x_{ABC}) \cdot P^{C}(x_C) = P^{AC}(x_{AC}) \cdot P^{BC}(x_{BC})$.
Níže používáme frázi " *P podléhá trojici (A, B, C)* ".

Podmínka $A \perp B|C\ (P)$ má mnoho ekvivalentních vyjádření, například:

- $\forall a, \bar{a} \in \prod_{i \in A} X_i \quad b, \bar{b} \in \prod_{i \in B} X_i \quad c \in \prod_{i \in C} X_i$
  $$P^{ABC}([a,b,c]) \cdot P^{ABC}([\bar{a},\bar{b},c]) = P^{ABC}([a,\bar{b},c]) \cdot P^{ABC}([\bar{a},b,c])$$
- existují funkce $f : \prod_{i \in AC} X_i \rightarrow \mathbb{R}$ a $g : \prod_{i \in BC} X_i \rightarrow \mathbb{R}$ takové, že
  $$\forall x \in \prod_{i \in N} X_i \quad P^{ABC}(x_{ABC}) = f(x_{AC}) \cdot g(x_{BC})$$
- $\forall c \in \prod_{i \in C} X_i$ splňující $P^{C}(c) > 0$ je podmíněná pravděpodobnost $P_{AB|C}(\cdot|c)$ součinová míra na $(\prod_{i \in A} X_i) \times (\prod_{i \in B} X_i)$.

poslední uvedená ekvivalentní definice vede přímo k běžné interpretaci $A \perp B|C$: "poznáme-li hodnotu veličin z C, veličiny z A a B se stanou nezávislé tj. pravděpodobnostní informace o nich budou navzájem irelevantní". Zřejmě proto, aby se pojem PN přiblížil lidskému chápání, byly struktury PN v literatuře popisovány pomocí grafů. Můžeme odlišit dva směry: pomocí neorientovaných grafů (toto vychází z teorie Markovských polí a příslušný graf se nazývá Markovská *síť* viz [Darroch & Lauritzen & Speed 1980], [Moussouris 1984], [Lauritzen & Speed & Vijayan 1984]) a pomocí orientovaných acyklických grafů (zde dlouhá tradice počatá genetikem S. Wrightem [1921] vedla k pojmům influenčního diagramu [Howard & Matheson 1984], [Shachter 1988], [Smith 1989], [Pearl & Geiger & Verma 1990] a rekursivního modelu [Wermuth & Lauritzen 1983], [Kiiveri & Speed & Carlin 1984]). Nicméně oba grafické přístupy nemohou popsat všechny možné struktury PN.

Proto byl navržen jiný způsob: jednoduše popsat strukturu PN seznamem platných výroků o PN neboli trojic, kterým pravděpodobnostní míra podléhá. To vedlo k pojmu závislostního modelu zavedeného Pearlem a Pazem [1987]; jejich definice je zde nepatrně upravena.

Def. 6 (závislostní model, model nezávislostní struktury)
a) Nechť $T(N)$ označuje množinu všech trojic (A, B, C) navzájem disjunktních podmnožin N kde A a B jsou neprázdné. Každá podmnožina $T(N)$ se bude nazývat *závislostní model nad N*.
b) Buď P pravděpodobnostní míra nad N a I závislostní model nad N. Řekněme, že I je *submodel nezávislostní struktury P* jestliže P podléhá každé trojici z I.
Dále řekněme, že I je *model nezávislostní struktury P* jestliže I je přesně množina trojic, kterým P podléhá.

Díky známým vlastnostem PN, kterými se zabýval již Dawid [1979] (viz též [Smith 1989]) pouze některé závislostní modely jsou modely nezávislostní struktury pro nějakou míru. Proto Pearl a Paz [1987] zavedli pojem semigraphoidu jakožto závislostního modelu splňujícího tyto vlastnosti. O vyvrácení jejich původní hypotézy, že semigrafoidy se shodují s modely PN a dalším vývoji jsme se zmínili již v úvodu.

§ 3 Informačně–teoretický přístup

V tomto paragrafu připomeňme informačně-teoretický pojem multiinformační funkce, který se podobně jako pojem entropické funkce hodí na studium vlastností PN. Ukážeme, jak multiinformační funkce umožňuje popsat strukturu PN pomocí imsetů.

Def. 7 (multiinformační funkce)
Nechť P je pravděpodobnostní míra nad N. Její *multiinformační funkce* $M : exp\,N \rightarrow \mathbb{R}$ je definována následovně:

$$\begin{aligned} M(\emptyset) &= 0 \\ M(S) &= \sum_{x, P(x)>0} P(x) \cdot ln(P^S(x_S) / \prod_{i \in S} P^{(i)}(x_i)) \qquad \text{pro } \emptyset \neq S \subset N. \end{aligned}$$

Poznámka Multiinformace, definovaná jakožto relativní entropie míry vůči součinu jejich jednorozměrných marginál, zobecňuje známý informačně-teoretický pojem vzájemné informace a slouží jakožto kvantitativní charakteristika úrovně stochastické závislosti více (než dvou) náhodných veličin. Proto jsem v [Studený 1989] používal název "multiinformace". Jiný název "entaxy" použil Malvestuto [1983].

Následující vlastnoti, dokázané například v [Studený 1989] naznačují, proč je multiinformační funkce vhodný nástroj na studium PN.
(3.1) $M(ABC) + M(C) \geq M(AC) + M(BC)$ kdykoliv $\langle A, B, C\rangle \in T(N)$
(3.2) $M(ABC) + M(C) = M(AC) + M(BC) \Leftrightarrow A \perp B|C\ (P)$ pro $\langle A, B, C\rangle \in T(N)$.

Podobné vlastnosti má i tkz. entropická funkce (každé množině S je přiřazena entropie marginální míry P^S) , pouze nerovnost je opačná. Co se týče schopností popsat vlastnosti PN pro pravděpodobnostní míry s konečným definičním oborem jsou entropická a multiinformační funkce zcela zaměnitelné. Poněkud se liší jejich další možnosti: entropická funkce může být použita pro popis funkčních závislostí v diskrétních pravděpodobnostních mírách - viz [Matúš 1991], zatímco multiinformační funkce může být použita pro studium PN pro spojité či "smíšené" pravděpodobnostní míry - viz [Studený 1989]. Použití vlastností multiinformační funkce byl hlavní krok v důkazu neexistence konečné axiomatické charakterizace v [Studený 1992]. Nicméně, souvislost entropické a multiinformační funkce s PN byla známa dříve - viz [Nambiar 1980], [Malvestuto 1983].

Def. 8 (pravděpodobnostní míra vyhovuje imsetu)
Buď P pravděpodobnostní míra nad N a u imset na $\mathcal{U}$. Rekněme, že P *vyhovuje* u jestliže $\langle M, u\rangle = 0$ (tj. skalární součin u s restrikcí multiinformační funkce P na $\mathcal{U}$ je nulový).

Výše definovaný pojem je vztažen k platnosti formule součinového typu následovně:

Lemma 1
Nechť P je pravděpodobnostní míra nad N a u imset na $\mathcal{U}$. Uvažujme následující podmínky:
(a) P splňuje formuli součinového typu danou u
(b) $\forall\, x \in \prod_{i \in N} X_i$ splňující $P(x) > 0$ platí: $\prod_{S \subset N} P^S(x_S)^{u(S)} = 1$
(c) P vyhovuje u.
Potom (a) $\Rightarrow$ (b) $\Rightarrow$ (c).

§4 Strukturální imsety

V tomto paragrafu zavedeme třídu strukturálních imsetů a ke každému takovému imsetu přidáme závislostní model. Ukazuje se, že vyhovování pravděpodobnostní míry nějakému strukturálnímu imsetu může být interpretováno jakožto částečný popis její struktury PN (totiž pomocí přiřazeného závislostního modelu). Speciální pozornost je věnována otázce, jak rozpoznávat strukturální imsety. Poslední výsledek říká, že každá možná (stochastická) nezávislostní struktura může být popsána strukturálním imsetem.

Strukturální imsety budou zavedeny jakožto "kombinace" tkz. elementárních imsetů:

Def. 9 (elementární imset)
Imset $u \in Z(\mathcal{U})$ se nazývá *elementární* jestliže jeho přirozené rozšíření $\bar{u}$ má tvar:
$\bar{u} = \delta_{S \cup T} - \delta_S - \delta_T + \delta_{S \cap T}$ kde $S, T \subset N$ $card\, S \setminus T = card\, T \setminus S = 1$.
Množina elementárních imsetů je níže označena symbolem E.

V následujícím příkladě ilustrujeme tento pojem na speciálním případě $card\, N = 4$.

Příklad 2
Uvažujme $N = \{1,2,3,4\}$. Podle definice je každý elementární imset "vytvořen" dvojicí $[S,T]$, nutně $1 \leq card\, S = card\, T \leq card\, N - 1$.
Takže elementární imsety mohou být rozděleny do skupin podle kardinality "vytvořujících" množin. V uvažovaném případě odlišíme 3 skupiny:

$E_1 \ldots$ tj. $card\, S = card\, T = 1$
například $S = \{1\}$ a $T = \{2\}$ dává $\bar{u} = \delta_{\{1,2\}} - \delta_{\{1\}} - \delta_{\{2\}} + \delta_{\emptyset}$
a tudíž $u = \delta_{\{1,2\}}$.
Příslušný seznam následuje :
$\delta_{\{1,2\}}, \delta_{\{1,3\}}, \delta_{\{2,3\}}, \delta_{\{1,4\}}, \delta_{\{2,4\}}, \delta_{\{3,4\}}$.

$E_2 \ldots$ tj. $card\, S = card\, T = 2$
například $S = \{1,2\}$ a $T = \{2,3\}$ dává $\bar{u} = \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}} + \delta_{\{2\}}$
a tudíž $u = \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}}$.
Příslušný seznam následuje :
$\delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}}, \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{1,3\}}, \delta_{\{1,2,3\}} - \delta_{\{1,3\}} - \delta_{\{2,3\}},$
$\delta_{\{1,2,4\}} - \delta_{\{1,2\}} - \delta_{\{2,4\}}, \delta_{\{1,2,4\}} - \delta_{\{1,2\}} - \delta_{\{1,4\}}, \delta_{\{1,2,4\}} - \delta_{\{1,4\}} - \delta_{\{2,4\}},$
$\delta_{\{1,3,4\}} - \delta_{\{1,3\}} - \delta_{\{1,4\}}, \delta_{\{1,3,4\}} - \delta_{\{1,3\}} - \delta_{\{3,4\}}, \delta_{\{1,3,4\}} - \delta_{\{1,4\}} - \delta_{\{3,4\}},$
$\delta_{\{2,3,4\}} - \delta_{\{2,3\}} - \delta_{\{2,4\}}, \delta_{\{2,3,4\}} - \delta_{\{2,3\}} - \delta_{\{3,4\}}, \delta_{\{2,3,4\}} - \delta_{\{2,4\}} - \delta_{\{3,4\}}$.

$E_3 \ldots$ tj. $card\, S = card\, T = 3$
například $S = \{1,2,3\}$ a $T = \{1,2,4\}$ dává
$u = \bar{u} = \delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,2,4\}} + \delta_{\{1,2\}}$.
Příslušný seznam následuje :
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,2,4\}} + \delta_{\{1,2\}},$
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,3,4\}} + \delta_{\{1,3\}},$
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{2,3,4\}} + \delta_{\{2,3\}},$
$\delta_N - \delta_{\{1,2,4\}} - \delta_{\{1,3,4\}} + \delta_{\{1,4\}},$
$\delta_N - \delta_{\{1,2,4\}} - \delta_{\{2,3,4\}} + \delta_{\{2,4\}},$
$\delta_N - \delta_{\{1,3,4\}} - \delta_{\{2,3,4\}} + \delta_{\{3,4\}}$.

V uvedeném příkladě je tedy celkový počet elementárních imsetů 24. Není problém najít obecný vzorec pro počet elementárních imsetů : $card\, N \cdot (card\, N - 1) \cdot 2^{card\, N - 3}$.

Def. 10 (strukturální imset)
Imset $u \in Z(\mathcal{U})$ se nazývá *strukturální* jestliže platí:
(4.1) $\exists n \in \mathbb{N}\ \ k_v \in \mathbb{Z}^+$ (pro $v \in E$) $\quad n \cdot u = \sum_{v \in E} k_v \cdot v.$

Následující snadné lemma umožňuje ztotožnit výroky o PN se strukturálními imsety a zajistit korektnost další definice.

Lemma Kdykoliv $\langle A, B, C\rangle \in T(N)$ pak imset $u \in Z(\mathcal{U})$ jehož přirozené rozšíření má tvar $\overline{u} = \delta_{ABC} - \delta_{AC} - \delta_{BC} + \delta_C$ je strukturální.

Def. 11 (závislostní model odpovídající strukturálnímu imsetu)
a) Každé trojici $\langle A, B, C\rangle \in T(N)$ přiřaďme strukturální imset $i(\langle A, B, C\rangle)$ určený svým přirozeným rozšířením $\overline{u} = \delta_{ABC} - \delta_{AC} - \delta_{BC} + \delta_C$.
b) Nechť u je strukturální imset. *Závislostní model odpovídající* u označený I_u je definován následovně:
$\langle A, B, C\rangle \in I_u \Leftrightarrow [\ \exists n \in \mathbb{N}\ \ n \cdot u - i(\langle A, B, C\rangle)$ je strukturální imset].

Lemma 2
Buď P pravděpodobnostní míra nad N a u strukturální imset na $\mathcal{U}$. Pak následující podmínky jsou ekvivalentní:
(a) P vyhovuje u
(b) I_u je submodel nezávislostní struktury P.

V souvislosti s počítačovou realizací strukturálních imsetů vyvstává následující otázka: jak poznat, že nějaký daný imset je strukturální? Uvedená definice není vhodná pro tento záměr. Nicméně, strukturální imsety mohou být charakterizovány vhodnějším způsobem. K tomu účelu je potřeba zavést další pojem.

Def. 12 (úplně konvexní funkce)
Množinová funkce $c : \mathcal{U} \to \mathbb{R}$ se nazývá *úplně konvexní* jestliže její nulové rozšíření $\underline{c}$ (tj. $\underline{c}(T) = 0$ pro $T \in exp\, N \setminus \mathcal{U}$) splňuje podmínku konvexity:
$\underline{c}(K \cup L) + \underline{c}(K \cap L) \geq \underline{c}(K) + \underline{c}(L)$ kdykoliv $K, L \subset N$.

Poznámka Přídavné jméno "konvexní" pochází z teorie her - viz [Rosenmüller & Weidner 1974]

Tvrzení 1
a) Nechť C označuje třídu úplně konvexních množinových funkcí na $\mathcal{U}$. Pro každý imset $u \in Z(\mathcal{U})$ pak platí: [u je strukturální] $\Leftrightarrow$ [$\forall c \in C\ \ \langle c, u\rangle \geq 0$].
Navíc, C je největší třída splňující tuto podmínku.
b) Existuje nejmenší konečná třída normalizovaných imsetů A taková, že pro každý imset u na $\mathcal{U}$ platí: [u je strukturální] $\Leftrightarrow$ [$\forall a \in A\ \ \langle a, u\rangle \geq 0$].
(Podle první části nutně $A \subset C$.)

Takže, z teoretického hlediska je vyřešen problém rozpoznání strukturálních imsetů: jednoduše zkontrolovat platnost všech nerovností $\langle a, u\rangle \geq 0$ pro $a \in A$.

Tvrzení 2
Kdykoliv P je pravděpodobnostní míra nad N a I je model nezávislostní struktury P, pak existuje strukturální imset u, že I odpovídá u (tj. $I = I_u$).

§5 Výsledek o ekvivalentnosti

Hlavní věta tohoto příspěvku říká, že pravděpodobnostní míra vyhovuje nějakému strukturálnímu imsetu u právě tehdy, když splňuje formuli součinového typu danou u. Tento výsledek lze dokázat za určitého dodatečného formálního předpokladu na u, nazvaným regularitou.

Def. 13 (regulární strukturální imset)
Uvažujme strukturální imset $u \in Z(\mathcal{U})$ a položme:
$\mathcal{A}_u = \{S \subset N;\ S \subset T$ pro nějaké $T \subset N$ splňující $\overline{u}(T) < 0\}$
$\mathcal{B}_u = \{S \subset N;\ S \subset T$ pro nějaké $T \subset N$ splňující $\overline{u}(T) > 0\}$.
(Lze ověřit, že $\mathcal{A}_u \subset \mathcal{B}_u$) .
Řekněme, že u je *regulární* jestliže jediné $\mathcal{E} \subset \mathcal{B}_u$ splňující následující tři podmínky:
[a] $\mathcal{E}$ je dědičné (tj. $K \subset L \in \mathcal{E} \Rightarrow K \in \mathcal{E}$)
[b] $\mathcal{A}_u \subset \mathcal{E}$
[c] kdykoliv $K, L \in \mathcal{E}$ splňuje $(K\setminus L, L\setminus K, K \cap L) \in I_u$ pak $K \cup L \in \mathcal{E}$
je $\mathcal{B}_u$ samo.

Příklad 3
Každé $u \in Z(\mathcal{U})$ tvaru $i(\langle A, B, C\rangle)$ pro $\langle A, B, C\rangle \in T(N)$ je regulární strukturální imset. Speciálně, každý elementární imset je regulární.

Tvrzení 3
V případě $card\ N \leq 4$ je každý strukturální imset regulární.

Dosud neznám příklad neregulárního strukturálního imsetu. Dokonce se domnívám, že každý strukturální imset je regulární (je to zatím nedokázaná hypotéza).

Nyní můžeme formulovat hlavní výsledek.

VĚTA
Nechť P je pravděpodobnostní míra nad N a u je regulární strukturální imset na $\mathcal{U}$. Potom následující podmínky jsou ekvivalentní:

(a) P splňuje formuli součinového typu danou u

(b) $\forall x \in \prod_{i \in N} X_i$ splňující $P(x) > 0$ platí $\prod_{S \subset N} P^S(x_S)^{\overline{u}(S)} = 1$

(c) P vyhovuje u

(d) I_u je submodel nezávislostní struktury P.

Poznámka Předchozí věta platí pro striktně kladné pravděpodobnostní míry bez předpokladu regularity imsetu.

Závěr

Výše uvedená věta dává do souvislosti 3 přístupy k popisu (podmíněných) nezávislostních struktur:

- pomocí závislostních modelů
- pomocí imsetů (informačně–teoretická definice)
- pomocí platnosti formule součinového typu

a ukazuje jejich ekvivalentnost.

Poznamenejme, že popis pomocí imsetů je systematicky popsán a ilustrován příklady v [Studený 93a]. Poskytuje dedukční mechanismus, který umožňuje odvozovat platné výroky o PN z vstupní informace o nezávislostní struktuře (z teoretického hlediska konečně realizovatelný).

Popis pomocí formulí součinového typu může být chápán jako krok ke globální interpretaci těchto nezávislostních struktur. Podle mého názoru mají popsané modely podobné důvody (či právo) být nazývány vysvětlitelné jako hierarchické loglineární modely. Vždyť obecný loglineární model je určen jistou "formulí" pro pravděpodobnostní míru vyjadřující ji jakožto součin marginálních faktorů. To je blízké k uvedeným formulím součinového typu a v některých speciálních případech (rozložitelné modely) dokonce ekvivalentní.

Poznamenejme, že jak bylo odvozeno v [Studený 93a] některé grafické popisy nezávislostních struktur mohou být "přeloženy" do imsetů a pak dále do formulí součinového typu. Pro informaci uveďme bez důkazu příslušné vyjádření imsetů.

Buď dán influenční diagram (= orientovaný acyklický graf); nechť $\pi(k)$ označuje množinu rodičů uzlů $k \in N$. Odpovídající imset je dán přímo svým přirozeným rozšířením:

$$\overline{u} = \delta_N - \delta_\emptyset + \sum_{k \in N} \left\{ \delta_{\pi(k)} - \delta_{\{k\} \cup \pi(k)} \right\}.$$

Je-li dán rozložitelný model určený nějakým triangulovaným (neorientovaným) grafem, nechť $\mathcal{C} \subset exp\, N$ označuje třídu jeho maximálních klik. Odpovídající imset je určen takto:

$$\overline{u} = \delta_{\cup \mathcal{C}} + \sum_{\emptyset \neq \mathcal{B} \subset \mathcal{C}} (-1)^{card\, \mathcal{B}} \cdot \delta_{\cap \mathcal{B}}$$

($\cup \mathcal{B}$ resp. $\cap \mathcal{B}$ označuje sjednocení resp. průnik množin z $\mathcal{B}$).

Poděkování

Jsem zavázán F. Malvestutovi; jedna jeho rada během jeho loňské návštěvy v Praze mi pomohla překonat jistou obtíž. Jsem též vděčen svému kolegovi F. Matúšovi za podnětnou diskusi.

Literatura

M. Aigner (1979) Combinatiorial Theory, Springer Verlag, Berlin-Heidelberg-New York.

J.N. Darroch, S.L. Lauritzen, T.P. Speed (1980) Markov fields and loglinear interaction models for contingency tables, Ann. Statist. 8, n.3, pp. 522-539.

A.P. Dawid (1979) Conditional independence in statistical theory (with discussion), J. Roy. Statist. Soc. ser. B 41, pp. 1-31.

D. Geiger, J. Pearl (1993) Logical and algorithmic properties of conditional independence and graphical models, in draft.

R.A. Howard, J.E. Matheson (1984) Influence diagrams, in Principles and Applications of Decision Analysis, Strategic Decision Group, Menlo Park, California.

H. Kiiveri, T.P. Speed, J.B. Carlin (1984) Recursive causal models, J. of Austral. Math. Soc. 36, pp. 30-52.

S.L. Lauritzen, T.P. Speed, K. Vijayan (1984) Decomposable graphs and hypergraphs, J. Austral. Math. Soc. 36, pp. 12-29.

F.M. Malvestuto (1983) Theory of random observables in relational data bases, Inform. Systems 8, n.4, pp. 281-289.

F. Matúš (1991) Probabilistic conditional independence structures and matroid theory: background, Proc. of WUPES (Workshop on Uncertainty Processing in Expert Systems), September 9-12, 1991, Alšovice, Czechoslovakia.

J. Moussouris (1984) Gibbs and Markov random systems with constraints, J. of Statist. Physics 10, n. 1, pp. 11-33.

K.K. Nambiar (1980) Some analytic tools for design of relational database systems, Proc. Conf. on Very Large Data Bases.

J. Rosenmüller, H.G. Weidner (1974) Extreme convex set functions with finite carrier: general theory, Discrete Math. 10, pp. 343-382.

J. Pearl, A. Paz (1987) Graphoids: a graph-based logic for reasoning about relevance relations, Advances in Artificial Intelligence II (B. Du. Boulay et al. Eds.) North-Holland, Amsterdam, pp. 357-363.

J. Pearl (1986) Markov and Bayes networks: a comparison of two graphical representations of probabilistic knowledge, technical report CSD-860024, R-46-I, UGLA, Los Angeles.

J. Pearl, D. Geiger, T. Verma (1990) The logic of influence diagrams, chapter 4 in Influence Diagrams, Belief Nets and Decision Analysis, (R.M. Oliver and J.Q. Smith Eds.), John Wiley.

R.D. Shachter (1988) Probabilistic inference and influence diagrams, Operation Research 36, pp. 589-604.

J.Q. Smith (1989) Influence diagrams for statistical modelling, Ann. of Statist. 17, n. 2, pp. 654-672.

D.J. Spiegelhalter, S.L. Lauritzen (1990) Sequential updating of conditional probabilities of directed graphical structures, Network 20, n. 5, pp. 579-605.

M. Studený (1989) Multiinformation and the problem of characterization of conditional independence relations, Problems of Control and Information Theory 18, n.1, pp. 3-16.

M. Studený (1992) Conditional independence relations have no finite complete characterization, Trans. of 11-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, vol.B, pp.377-396, Academia, Prague.

M. Studený (1993a) Description of structures of stochastic conditional independence by means of faces and imsets, submitted to International Journal of General Systems.

M. Studený (1993b) Description of conditional independence structures by means of imsets: a connection with product formula validity, submitted to Proceedings of IPMU'92 (Palma de Mallorca, July 6-10, 1992), Springer Verlag edited volume.

S.Ur, A. Paz (1991) The representation power of probabilistic knowledge by undirected graphs and directed acyclic graphs: a comparison. Proc. of WUPES (Workshop on Uncertainty Processing in Expert Systems), September 9-12, 1991, Alšovice, Czechoslovakia.

N. Wermuth, S.L. Lauritzen (1983) Graphical and recursive models for contingency tables, Biometrika 70, pp. 537-552.

S. Wright (1921) Correlation and causation, J. of Agricultural Research 20, pp. 557-585.