

# Náhodná pole a jejich aplikace

Martin Janžura, Antonín Otáhal

Ústav teorie informace a automatizace AV ČR

Pod vodárenskou věží 4, 182 08 Praha 8

## Úvod

Teorie náhodných polí je významným zdrojem modelů pro statistickou analýzu náhodných jevů prostorové povahy. Existuje několik základních přístupů, z nichž oběma nejhlavnějším, spektrální teorii zobecňující přístupy známé z časových řad i teorii Gibbsových stavů úzce navazující na statistickou fyziku, se zde stručně a spíše přehledově zmiňujeme. V celé práci se omezujeme na diskrétní indexovou množinu.

První kapitola se zabývá obecnou spektrální analýzou stacionárních náhodných polí druhého řádu, druhá dalšími možnými zobecněními na jiné indexové množiny se složitější strukturou. Třetí kapitola naznačuje přístup vycházející ze statistické fyziky, včetně statistické analýzy. Ve čtvrté kapitole se pak prolínají přístupy všech třech předchozích při analýze markovských gaussovských náhodných polí.

## 1 Spektrální analýza náhodných polí

Prvním studovaným přístupem k náhodným polím bude snaha o zobecňování postupů známých z teorie časových řad. Do určité míry jde o standardní "překlad" známých pojmů a výsledků, ale vyskytují se zde i závažné odlišnosti od jednorozměrného případu. Jako literaturu pro další studium tohoto zde stručně zmíněného přístupu uvedme například práce Whittle (1954), Helson a Lowdenslager (1958, 1961), Brillinger (1974), Ripley (1981), Tjøstheim (1978, 1983) i některé níže uvedené, více specializované práce.

### 1.1 Spektrální rozklady

1.1.1. *Náhodné pole* je v této kapitole stochastický proces  $X = (X(t) : t \in Z^d)$  indexovaný  $d$ -rozměrnými vektory s celočíselnými souřadnicemi, jehož množinou stavů je reálná přímka  $R$ . Předpokládáme, že střední hodnota všech náhodných veličin z  $X$  je stejná, bez újmy obecnosti můžeme předpokládat, že je rovna nule, a že všechny náhodné veličiny z  $X$  mají konečné druhé momenty. Takové náhodné pole nazveme (*slabě*) *stacionárním*, jestliže existuje *kovarianční funkce*, označme ji  $R$ , definovaná na  $Z^d$  tak, že pro všechna  $s, t \in Z^d$  platí

$$R(s - t) = \text{cov}(X(s), X(t)).$$

Z definice bezprostředně plyne, že každá kovarianční funkce stacionárního náhodného pole je pozitivně semidefinitní.

1.1.2. *Spektrální rozklad kovarianční funkce* vychází z této pozitivní semidefinitnosti a ze zobecněné Bochnerovy věty - viz Weil (1950). Říká, že pro každou kovarianční funkci existuje právě jedna konečná (nezáporná) míra  $F$  na  $I^d$  taková, že pro všechna  $t \in Z^d$  platí

$$R(t) = \int_{I^d} \exp\{2\pi i \cdot \langle t, \lambda \rangle\} dF(\lambda),$$

kde  $I = [0, 1)$  a  $\langle \cdot, \cdot \rangle$  je běžný  $d$ -rozměrný skalární součin.

1.1.3. (*Spektrální hustota*). Jestliže je spektrální míra  $F$  v předchozím bodě absolutně spojitá vůči Lebesgueově míře, nazývá se příslušná Radon-Nikodymova derivace spektrální hustotou. Spektrální hustota existuje, jestliže – zhruba řečeno – absolutní velikost hodnot kovarianční funkce klesá dostatečně rychle s rostoucí (euklidovskou) normou svého argumentu. Typickým příkladem je *absolutně sumabilní* kovarianční funkce  $R$ , daná podmínkou

$$\sum_{t \in \mathbb{Z}^d} |R(t)| < \infty,$$

pro kterou je spektrální hustota  $f$  dána *inverzním vzorcem*

$$f(\lambda) = \sum_{t \in \mathbb{Z}^d} R(t) \cdot \exp\{-2\pi i \cdot \langle t, \lambda \rangle\}$$

pro všechna  $\lambda \in I^d$ .

1.1.4. Obdobně jako v jednorozměrném případě podle Karhunenovy věty ze spektrálního rozkladu 1.1.2 plyne *spektrální rozklad* samotného *náhodného pole*  $X$  ve tvaru

$$X(t) = \int_{I^d} \exp\{2\pi i \cdot \langle t, \lambda \rangle\} dZ(\lambda),$$

kde  $Z$  je náhodná míra definovaná na borelovských podmnožinách množiny  $I^d$  s nulovou střední hodnotou taková, že pro každé dvě borelovské podmnožiny  $B, C \subset I^d$  platí

$$\text{cov}(Z(B), Z(C)) = F(B \cap C),$$

(příslušná spektrální míře  $F$ ) a integrál podle náhodné míry je obdobně jako v jednorozměrném případě chápán ve smyslu konvergence podle středu stupně 2.

1.1.5. Funkce  $\Phi$  definovaná na  $\mathbb{Z}^d$  je (*lineární*) *filtr*, jestliže

$$\sum_{t \in \mathbb{Z}^d} |\Phi(t)|^2 < \infty.$$

Tato podmínka je zejména splněna pro tzv. *konečný filtr*, jehož hodnoty jsou nenulové jen na konečné množině indexů. *Spektrální charakteristikou* filtru  $\Phi$  je funkce  $\varphi$  definovaná pro každé  $\lambda \in I^d$  předpisem

$$\varphi(\lambda) = \sum_{t \in \mathbb{Z}^d} \Phi(t) \cdot \exp\{-2\pi i \cdot \langle t, \lambda \rangle\}$$

ve smyslu konvergence podle středu stupně 2.

1.1.6. *Lineární filtrace* je postup pro získávání nového náhodného pole na základě daného. Jestliže tedy je  $X$  stacionární náhodné pole, pak jeho filtrací získáme pole  $Y = \Phi(X)$  dané pro každé  $t \in \mathbb{Z}^d$  předpisem

$$Y(t) = \sum_{s \in \mathbb{Z}^d} \Phi(s) \cdot X(t + s).$$

Jestliže v této situaci má pole  $X$  spektrální hustotu  $f$ , pak pole  $Y$  má rovněž spektrální hustotu, označme ji třeba  $g$ , a platí  $g = |\varphi|^2 \cdot f$  (odtud název spektrální charakteristika).

### 1.1.7. Příklady náhodných polí

(1) *Bílý šum* je stacionární náhodné pole  $N$  po dvou nekorelovaných náhodných veličin (termín "náhodné pole" je zde vlastně trochu zavádějící, protože v tomto případě nemá struktura indexové množiny souvislost s pravděpodobnostními vlastnostmi). Je zřejmé, že spektrální hustota bílého šumu existuje a je (skoro všude) konstantní funkcí na  $I^d$ .

(2) *Pole klouzavých součtů* vzniká lineární filtrací bílého šumu  $N$ . Jestliže je  $\sigma^2$  rozptyl veličin bílého šumu  $N$  a  $\varphi$  je spektrální charakteristikou příslušného filtru, pak podle 1.1.6 je spektrální hustota pole klouzavých součtů dána ve tvaru  $f = \sigma^2 \cdot |\varphi|^2$ .

(3) *Autoregresní náhodné pole* je dáno podmínkou, že jeho lineární filtrace je rovna bílému šumu. Při značení obdobným předchozímu příkladu je jeho spektrální hustota dána ve tvaru  $f = \sigma^2 \cdot |\varphi|^{-2}$ .

## 1.2 Jiná zobecnění stacionarity

1.2.1. Všimněme si, že spektrální rozklad předchozí podkapitoly je vlastně zvláštním případem obecnější věty. Nechť  $A$  je lokálně kompaktní komutativní (aditivní) topologická grupa a  $X = (X(a) : a \in A)$  je homogenní náhodné pole na  $A$ , tj. pole s konstantní střední hodnotou a kovariancemi invariantními vzhledem k automorfismům této grupy, tj. pro všechna  $a, b, c \in A$  platí

$$\text{cov}(X(a), X(b)) = \text{cov}(X(a+c), X(b+c)).$$

Nechť  $\hat{A}$  je duální grupa ke grupě  $A$  a  $\chi(a, \hat{a})$  jsou příslušné charaktery (grupové homomorfismy do multiplikativní grupy všech komplexních jednotek). Potom kovarianční funkci  $R$  pole  $X$ , definovanou obdobně jako v 1.1.1, můžeme vyjádřit pro všechna  $a \in A$  vztahem

$$R(a) = \int_{\hat{A}} \chi(a, \hat{a}) dF(\hat{a}),$$

kde  $F$  je jednoznačně určená konečná borelovská míra na  $\hat{A}$ .

1.2.2. V pracích Jaglom (1957, 1960, 1961) se toto pojetí dále zobecňuje na náhodná pole indexovaná tranzitivním prostorem lokálně kompaktní topologické grupy určitého typu, pro která jsou kovariance invariantní vzhledem k akci této grupy. Pak se místo charakterů integrují lineární reprezentace příslušné grupy přes určitý duální objekt, který je nyní obecnější než duální grupa.

Důležitým konkrétním příkladem jsou stacionární *izotropní* náhodná pole na  $d$ -rozměrném euklidovském prostoru  $R^d$ , jejichž kovariance jsou invariantní vzhledem ke všem shodným pohybům, a tedy závislé jen na euklidovské vzdálenosti – viz monografii Jadrenko (1980).

Studiem *stacionárních slabě izotropních* náhodných polí na  $Z^d$  obdobných izotropním (například vzniklých diskretním výběrem) se zabývají práce Otáhal (1984, 1986, 1988).

## 2 Markovská náhodná pole

Náhodným polem nyní budeme rozumět stochastický proces  $\mathcal{X} = (\mathcal{X}_i)_{i \in Z^d}$  s (pro jednoduhost) konečnou množinou stavů  $X$ .

Jestliže budeme chtít zobecnit pojem markovského řetězce, jakožto nejjednoduššího netriviálního modelu stochastické závislosti, narazíme na problém způsobený pouze částečným uspořádáním množiny vícerozměrného "času"  $Z^d$ . Nelze tedy zobecnění provést bezprostředně, je třeba nejprve důkladně analyzovat a pochopit podstatu markovské vlastnosti. K tomu nám ovšem nejlépe poslouží klasický jednorozměrný případ a tím také začneme v následujícím odstavci.

Náhodnými poli se z hlediska přístupu uplatňovaného v této kapitole zabývají vedle množství časopiseckých prací i některé speciální monografie, mimo jiné Ruelle (1969), Preston (1976) a Georgii (1988).

## 2.1 Náhodné posloupnosti s "dvoustrannou" markovskou vlastností

Nechť  $(\mathcal{X}_t)_{t \in Z}$  je homogenní markovský řetězec s konečnou množinou stavů  $X$ ,  $|X| < \infty$ , a maticí pravděpodobností přechodu  $P(\cdot|\cdot)$  s kladnými elementy, tj.  $P(x|y) > 0$  pro všechna  $x, y \in X$ . To znamená, že pro každé  $t \in Z$  platí

$$P(\mathcal{X}_t | \mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots) \stackrel{s.j.}{=} P(\mathcal{X}_t | \mathcal{X}_{t-1}) > 0,$$

a jedná se tedy o ergodický řetězec se všemi příslušnými "dobrymi vlastnostmi".

Potom můžeme psát

$$\begin{aligned} P(\mathcal{X}_t | \dots, \mathcal{X}_{t-2}, \mathcal{X}_{t-1}, \mathcal{X}_{t+1}, \mathcal{X}_{t+2}, \dots) &\stackrel{s.j.}{=} \lim_{N \rightarrow \infty} P(\mathcal{X}_t | \mathcal{X}_{t-N}, \dots, \mathcal{X}_{t-1}, \mathcal{X}_{t+1}, \dots, \mathcal{X}_{t+N}) = \\ &= \lim_{N \rightarrow \infty} \frac{P(\mathcal{X}_{t-N}, \dots, \mathcal{X}_{t+N})}{\sum_{y \in X} P(\mathcal{X}_{t-N}, \dots, \mathcal{X}_t = y, \dots, \mathcal{X}_{t+N})} = \frac{P(\mathcal{X}_{t+1} | \mathcal{X}_t) P(\mathcal{X}_t | \mathcal{X}_{t-1})}{P^2(\mathcal{X}_{t+1} | \mathcal{X}_{t-1})} = \\ &= P(\mathcal{X}_t | \mathcal{X}_{t-1}, \mathcal{X}_{t+1}) > 0. \end{aligned}$$

Jestliže nazveme výslednou rovnost dvoustrannou markovskou vlastností, dokázali jsme vlastně následující tvrzení.

**Věta 2.1.1.** Kladný homogenní markovský řetězec má i dvoustrannou markovskou vlastnost. □

Vyvstává přirozená otázka, zda platí i opačné tvrzení. Nechť tedy  $P(\mathcal{X}_t | \dots, \mathcal{X}_{t-1}, \mathcal{X}_{t+1}, \dots) = P(\mathcal{X}_t | \mathcal{X}_{t-1}, \mathcal{X}_{t+1}) > 0$  je homogenní posloupnost s "dvoustrannou" markovskou vlastností, tj.

$$P(\mathcal{X}_t = x | \mathcal{X}_{t-1} = y, \mathcal{X}_{t+1} = z) = \Pi(x|y, z)$$

pro každé  $t \in Z$ ,  $x, y, z \in X$ , kde

$$\Pi(\cdot|\cdot, \cdot) : X^3 \rightarrow R$$

splňuje  $\sum_{x \in X} \Pi(x|\cdot, \cdot) = 1$  a budeme ji nazývat lokální charakteristikou.

**Lemma 2.1.2.** Nechť  $P$  je kladný dvoustranně markovský řetězec s lokální charakteristikou

$\Pi(\cdot|\cdot, \cdot) > 0$ . Potom existuje matice  $Q = (Q(x, y))_{x, y \in X}$  s kladnými elementy  $Q(x, y) > 0$  pro všechna  $x, y \in X$ , pro niž platí

$$\Pi(x|y, z) = \frac{Q(z, x)Q(x, y)}{Q^2(z, y)} \quad \text{pro všechna } x, y, z \in X.$$

**Důkaz.** Zvolme pevné  $b \in X$ . Snadno se ukáže

$$\begin{aligned} \frac{\Pi(x|y, z) \Pi(y|b, b)}{P(b|y, z) \Pi(b|b, b)} &= \frac{P(\mathcal{X}_{t-1} = y, \mathcal{X}_t = x | \mathcal{X}_{t-2} = b, \mathcal{X}_{t+1} = z)}{P(\mathcal{X}_{t-1} = b, \mathcal{X}_t = b | \mathcal{X}_{t-2} = b, \mathcal{X}_{t+1} = z)} = \\ &= \frac{\Pi(y|b, x) \Pi(x|b, z)}{\Pi(b|b, x) \Pi(b|b, z)}. \end{aligned}$$

Nyní stačí dosadit  $Q(x, y) = \frac{\Pi(y|b, x)}{\Pi(b|b, x)} > 0$  a máme výsledek.  $\square$

**Důsledek 2.1.3.** Nechť  $\lambda > 0$  je maximální vlastní číslo matice  $Q$ ,  $r > 0$  nechť je příslušný levý vlastní vektor. Potom  $\tilde{P}(\cdot|\cdot)$ , kde

$$\tilde{P}(x|y) = \frac{Q(x, y)r(x)}{\lambda r(y)} > 0 \quad \text{pro každé } x, y \in X,$$

je maticí pravděpodobností přechodu kladného homogenního markovského řetězce  $\tilde{P}$ , pro který  $\tilde{P}(\mathcal{X}_t | \mathcal{X}_{t-1}, \mathcal{X}_{t+1}) = P(\mathcal{X}_t | \mathcal{X}_{t-1}, \mathcal{X}_{t+1})$  pro každé  $t \in \mathbb{Z}$ , a následně i  $\tilde{P} = P$ .

**Důkaz:** Vlastnosti  $\tilde{P}(\cdot|\cdot)$  se snadno ověří. Rovnost  $\tilde{P} = P$  vyplývá z neexistence fázových přechodů v dimenzi 1 (viz např. (5.37), Preston (1976)).  $\square$

**Věta 2.1.4.** Kladný homogenní dvoustranně markovský řetězec má i standardní jednostrannou markovskou vlastnost.

**Důkaz:** Plyne přímo z Důsledku 2.1.3.  $\square$

Ukázali jsme tedy ekvivalenci jednostranné a dvoustranné markovské vlastnosti. Její význam spočívá v tom, že pojem "bezprostředně předcházející" nelze dost dobře zobecnit z lineárně uspořádané množiny  $Z$  do částečně uspořádané množiny  $Z^d$ ,  $d \geq 2$ . Naproti tomu pojem "nejbližší sousedé" nebo "okolí" má naprosto přirozené zobecnění do více dimenzí. Je tudíž zřejmé, že zobecňovat budeme právě dvoustrannou a nikoli jednostrannou markovskou vlastnost. Střetneme se však s komplikací danou tím, že zdaleka ne každá funkce  $\Pi(\cdot|\cdot, \cdot)$  může být příslušnou lokální charakteristikou, jak ukazuje Lemma 2.1.2. Speciální tvar funkce  $\Pi(\cdot|\cdot, \cdot)$  souvisí s podmínkou konzistentnosti celého systému všech konečněrozměrných podmíněných pravděpodobností, odkud vyplývají některé vazby, které musí funkce  $\Pi(\cdot|\cdot, \cdot)$  splňovat, aby mohla být lokální charakteristikou nějakého markovského řetězce. Zde se nám tyto vazby podařilo poměrně snadno nalézt, obecně to však nemusí být snadné, a proto budeme hledat posléze cestu, jak se i těmto problémům vyhnout.

## 2.2 Markovská gibbsovská náhodná pole

Označme  $\mathcal{K} = \{A \subset Z^d; |A| < \infty\}$  systém všech konečných podmnožin  $Z^d$ . Náhodné pole s konečnou množinou stavů  $X$  je markovské, jestliže jeho rozdělení  $P$  splňuje

$$P_{A|Z^d \setminus A}(x_A | x_{Z^d \setminus A}) = P_{A|\partial A}(x_A | x_{\partial A}) \quad \text{pro skoro každé } x \in X^{Z^d}.$$

Zde pro každé  $A \in \mathcal{K}$  je  $\partial A \subset Z^d \setminus A$ ,  $\partial A \in \mathcal{K}$ , a  $P_{A|Z^d \setminus A}$  respektive  $P_{A|\partial A}$  je podmíněné rozdělení v konečné oblasti  $A$  při dané konfiguraci v jejím doplňku  $Z^d \setminus A$  respektive v "okoli"  $\partial A$ .

Markovské pole je homogenní, jestliže  $P_{A|\partial A} \stackrel{s.j.}{=} P_{A+t|\partial(A+t)}$  pro každé  $A \in \mathcal{K}$ ,  $t \in Z^d$ .

Máme-li markovské pole zadat, měli bychom mít k dispozici právě systém podmíněných rozdělení, neboli "specifikaci"

$$\Pi = \{\Pi_A(\cdot|\cdot)\}_{A \in \mathcal{K}},$$

kde  $\Pi_A : X^A \times X^{Z^d \setminus A} \rightarrow R$  pro každé  $A \in \mathcal{K}$  splňuje  $\sum_{y_A \in X^A} \Pi_A(y_A|\cdot) = 1$ .

Navíc budeme požadovat, aby specifikace byla

- i) markovská:  $\Pi_A(x_A | x_{Z^d \setminus A}) = \bar{\Pi}_A(x_A | x_{\partial A})$  pro každé  $x \in X^{Z^d}$ ,  
kde  $\bar{\Pi}_A : X^A \times X^{\partial A} \rightarrow R$  pro každé  $A \in \mathcal{K}$ ;
- ii) homogenní:  $\Pi_A = \Pi_{A+t}$  pro každé  $A \in \mathcal{K}$ ,  $t \in Z^d$ ;
- iii) konzistentní:  $\Pi_B \Pi_A = \Pi_B$  pro každé  $B \supset A$ ,  $A, B \in \mathcal{K}$ .

(Pozn.: Podmínka iii) odpovídá přirozené vlastnosti podmíněných distribucí: "podmíníme-li nejprve větší a pak menší  $\sigma$ -algebrou, je to jako bychom podmínili přímo tou menší.")

Jestliže

$$P_{A|Z^d \setminus A} \stackrel{s.j.}{=} \Pi_A \quad \text{pro každé } A \in \mathcal{K},$$

píšeme  $P \in G(\Pi)$  a říkáme, že (markovské) náhodné pole  $P$  je gibbsovské (Gibbsův stav) vzhledem ke specifikaci  $\Pi$ . Způsob zápisu  $P \in G(\Pi)$  je předurčen možností fázových přechodů, kdy  $|G(\Pi)| > 1$ , neboli Gibbsův stav není specifikací dán jednoznačně. (Vzhledem k tomu, že  $G(\Pi)$  je zřejmě konvexní množina, z vlastnosti  $|G(\Pi)| > 1$  už přímo plyne  $|G(\Pi)| = \infty$ .) Možnost existence různých markovských náhodných polí se stejnými konečněrozměrnými podmíněnými distribucemi je úkaz, který se v dimenzi 1 nevyskytuje a je zdánlivě překvapivý, má však velmi dobrou právě fyzikální interpretaci, neboť odpovídá existenci fázových přechodů.

Díky konzistentnosti specifikace stačí ověřit pouze

$$P_{t|Z^d \setminus \{t\}} \stackrel{s.j.}{=} \Pi_t \quad \text{pro každé } t \in Z^d,$$

a díky homogenosti pak stačí dokonce pouze pro  $t = 0$ , čili pro lokální charakteristiku  $\Pi_0(\cdot|\cdot)$ .

Nicméně stále zůstává problém, jak poznat, zda  $\Pi_0(\cdot|\cdot)$  může patřit do nějaké konzistentní specifikace. Bylo by užitečné mít k dispozici co nejjednodušší nástroj, jak konzistentní specifikace konstruovat. Tím se dostáváme k pojmu potenciálu.

## 2.3 Potenciály

Potenciálem nazveme systém funkcí  $\Phi = \{\Phi_\Lambda\}_{\Lambda \in \mathcal{K}}$ , kde  $\Phi_\Lambda : X^\Lambda \rightarrow R$  pro každé  $\Lambda \in \mathcal{K}$ . Budeme pracovat s potenciály, které jsou

i) homogenní:  $\Phi_\Lambda = \Phi_{\Lambda+t}$  pro každé  $\Lambda \in K$ ,  $t \in Z^d$ , a mají

ii) konečný rozsah:  $\Phi_\Lambda \equiv 0$  pokud  $\text{diam}(\Lambda) > r \geq 0$ .

Potom píšeme  $\Phi \in \mathcal{U}_r$  a můžeme definovat

$$\Pi_A^\Phi(x_A | x_{Z^d \setminus A}) = \exp \left\{ \sum_{\Lambda \in K, \Lambda \cap A \neq \emptyset} \Phi_\Lambda(x_A) \right\} / W_A^\Phi(x_{Z^d \setminus A})$$

kde  $W_A^\Phi(x_{Z^d \setminus A})$  je pro každé  $A \in K$  a  $x \in X^{Z^d}$  příslušná normalizační konstanta.

**Věta 2.3.1.** Nechť  $\Phi \in \mathcal{U}_r$ , potom specifikace  $\Pi^\Phi = \{\Pi_A^\Phi\}_{A \in K}$  je markovská, homogenní a konzistentní.

**Důkaz:** Markovská vlastnost plyne z konečného rozsahu potenciálu, když

$$\bar{A} = A \cup \partial A \subset \bigcup_{\substack{\Lambda \cap A \neq \emptyset \\ \text{diam}(\Lambda) \leq r}} \Lambda,$$

homogennost z homogennosti potenciálu a konzistentnost lze ověřit přímým výpočtem.  $\square$

**Poznámka 2.3.2.** Máme-li kladnou konzistentní homogenní markovskou specifikaci, můžeme k ní nalézt potenciál následujícím způsobem. Položme

$$\Phi_\Lambda(x_\Lambda) = \sum_{B \subset A} (-1)^{|A \setminus B|} \log \Pi_C(x_B, b_{C \setminus B} | b_{Z^d \setminus C})$$

pro libovolné  $C \supset \Lambda$  (to umožňuje konzistentnost), kde  $b = (b_t)_{t \in Z^d} \in X^{Z^d}$  je pevně zvolená konfigurace. Tento vzorec se nazývá Möbiova formule a dříve uvedená definice specifikace pomocí potenciálu je pak inverzní Möbiova formule.  $\square$

Budeme tedy zadávat specifikace pomocí potenciálů, jejichž definice nevyžaduje ověřovat žádné složité podmínky typu konzistentnosti. Specifikace pak označují třídy gibbovských distribucí. Budeme psát přímo  $G(\Phi)$  místo  $G(\Pi^\Phi)$  pro náhodná pole gibbovská vzhledem k potenciálu  $\Phi$ . Navíc potenciál můžeme chápat jako v širším slova smyslu "parametr" a tím přejít k parametrickým modelům a následně i k parametrickým úlohám v aplikacích.

Obvykle předpokládáme, že známe bázecké potenciály  $\Phi^1, \dots, \Phi^K$  a neznámý potenciál je pak jejich kombinací s neznámými koeficienty, tedy

$$\Phi^\theta = \sum_{i=1}^K \theta_i \Phi^i, \quad \text{kde } (\theta_1, \dots, \theta_K)^T = \theta \in R^K,$$

a pracovat budeme s parametrickou rodinou

$$\{G(\Phi^\theta)\}_{\theta \in R^K}.$$

Musíme mít stále na vědomí, že jednotlivé fáze, tj. různá  $P \in G(\Phi^\theta)$ , nejsou z hlediska parametru  $\theta$  rozlišitelné, a je-li naším cílem parametr odhadovat, je lhostejné, která konkrétní fáze  $P \in G(\Phi^\theta)$  je tou aktuální neznámou distribucí. Tím je plně opodstatněn přístup pracující nikoli s parametrickou třídou distribucí, ale s parametrickou třídou množin Gibbovských stavů.

## 2.4 Příklad: Isingův model

Zvolme  $d = 2$ ,  $X = \{0, 1\}$  a položme  $h = \{1, 0\}$ ,  $v = \{0, 1\} \in Z^2$ .

Pro  $\theta = (\theta_0, \theta_h, \theta_v)^T \in R^3$  a každé  $t \in Z^2$  definujeme

$$\begin{aligned}\Phi_t^\theta(x_t) &= \theta_0 x_t \\ \Phi_{(t, t+h)}^\theta(x_t, x_{t+h}) &= \theta_h x_t x_{t+h} \\ \Phi_{(t, t+v)}^\theta(x_t, x_{t+v}) &= \theta_v x_t x_{t+v}\end{aligned}$$

a  $\Phi_\lambda^\theta \equiv 0$  jinak.

Vidíme tedy, že pro  $0 \in Z^2$  je  $\partial 0 = \{t \in Z^2; \|t\| = 1\} = \{h, -h, v, -v\}$  a platí např.

$$\Pi_0^{\Phi^\theta}(x_0 | x_{\partial 0}) = \frac{\exp\{x_0[\theta_0 + \theta_h(x_h + x_{-h}) + \theta_v(x_v + x_{-v})]\}}{1 + \exp\{\theta_0 + \theta_h(x_h + x_{-h}) + \theta_v(x_v + x_{-v})\}}$$

pro každé  $x \in X^{Z^2}$ .

Parametry  $\theta_h, \theta_v$  můžeme interpretovat jako horizontální respektive vertikální interakce; pokud jsou stejné, tj.  $\theta_h = \theta_v = \theta_1$ , je příslušný model isotropní. Jestliže  $\theta_h, \theta_v > 0$ , nazývají se interakce ve shodě s fyzikální terminologií atraktivní a typická konfigurace obsahuje větší stejnobarevné "skvrny"; pokud  $\theta_h, \theta_v < 0$ , nazývají se repulsivní a typická konfigurace připomíná "šachovnici", tj. "sousedé" nemají obvykle stejnou barvu.

**Poznámka 2.4.1.** V dvourozměrném Isingově modelu je dobře známý příklad fázových přechodů, popsany v již klasické práci – Onsager (1944). Nechť  $\theta_h = \theta_v = \theta_1 > 2 \ln(\sqrt{2}+1)$  a  $\theta_0 = -2\theta_1$ , potom  $G(\Phi^\theta) = \overline{\text{co}}(P^0, P^1)$  kde  $P^0 \perp P^1$  jsou ergodické. Množina Gibbsovských stavů tvoří tedy konvexní obal dvou "čistých" fází.

## 3 Některé aplikace markovských náhodných polí

Aplikacemi zde budeme rozumět řešení statistických úloh, ve kterých náhodná pole figurují jako generující pravděpodobnostní distribuce. Budou to zejména

- identifikace modelu (odhad parametrů),
- výběr modelu (testování hypotéz, klasifikace),
- identifikace stavu (filtrace, restaurace obrazu),
- kódování, přenos informace, atd.

### 3.1 Data, asymptotika a parametrická rodina

Ve statistických úlohách budou data v  $n$ -tém kroku chápána jako jedna pevná konfigurace  $\hat{x}_n \in X^{V_n}$  v konečné objemu  $V_n \in \mathcal{K}$ . Všechny asymptotické výsledky pak budou ve smyslu expandujícího objemu  $V_n \nearrow Z^d$  pro  $n \rightarrow \infty$  v takovém smyslu, aby  $|V_n|^{-1}|V_n \cap (V_n + t)| \rightarrow 1$  pro  $n \rightarrow \infty$  a každé pevné  $t \in Z^d$ .

Ve shodě s paragrafem 2.3 nechť  $\Phi^1, \dots, \Phi^K \in \mathcal{U}_r$ ,  $r > 0$ , jsou známé. Budeme psát přímo  $\Pi^\theta$  a  $G(\theta)$  místo  $\Pi^{\Phi^\theta}$  a  $G(\Phi^\theta)$  a za příslušnou parametrickou rodinu

$$\{G(\theta)\}_{\theta \in \Theta}, \quad \text{kde } \Theta \in R^K.$$

a předpokládáme, že data jsou generována nějakým náhodným polem  $P^0 \in G(\theta^0)$  s neznámým  $\theta^0 \in \Theta$ .



### 3.2 Pseudo-věrohodnost

Pro pevné (relativně "malé")  $A \in \mathcal{K}$  zavedme pseudo-věrohodnostní funkci

$$\ell_A^n(\theta, \hat{x}_{V_n}) = \sum_{i: i+(A \cup \theta A) \subset V_n} \log \bar{\Pi}_{A+i}^\theta(\hat{x}_{A+i} | \hat{x}_{\partial(A+i)}).$$

Pak můžeme definovat maximálně pseudo-věrohodný (MPL) odhad předpisem

$$\hat{\theta}_A^n = \operatorname{argmax}_{\theta \in R^K} \ell_A^n(\theta, \hat{x}_{V_n}).$$

Dá se ukázat, že tento odhad je konzistentní (např. Gidas (1991)) a asymptoticky normální (pokud  $P^\theta$  je navíc ergodické - Guyon a Künsch (1992)). V úloze výběru modelu nechť  $\theta^1, \dots, \theta^J$  reprezentují příslušné třídy (hypotézy) mezi nimiž máme rozhodnout. Řešení je obdobné

$$\hat{\theta}_A^n = \operatorname{argmax}_{\theta \in \{\theta^1, \dots, \theta^J\}} \ell_A^n(\theta, \hat{x}_{V_n}).$$

Zde je možno ukázat (např. Janžura (1993a)), že pravděpodobnost chybného rozhodnutí jde pro rostoucí objem do nuly exponenciálně rychle, tedy

$$P_e^n = e^{-|V_n| \beta + o(|V_n|)}, \quad \text{kde } \beta > 0.$$

### 3.3 Proč ne věrohodnost

Volba pseudo-věrohodnostní funkce namísto standardní věrohodnostní má přinejmenším dva dobré důvody. První je praktický, spočívající v tom, že pseudo-věrohodnostní funkci  $\ell_A^n(\theta, \hat{x}_{V_n})$  umíme velmi snadno vyčíslit, neboť vzorec pro  $\bar{\Pi}_A^\theta(\cdot | \cdot)$  je (pro "malá"  $A \in \mathcal{K}$ !) poměrně jednoduchý. Na druhou stranu pro věrohodnost danou v rigidním slova smyslu formulí  $\hat{\ell}^n(\theta, \hat{x}_{V_n}) = \log P_{V_n}^\theta(\hat{x}_{V_n})$ , kde  $P_{V_n}^\theta$  jako restrikce příslušného náhodného pole  $P^\theta \in G(\theta)$  není v případě fázových přechodů vlastně ani jednoznačně dáno. Mohli bychom samozřejmě  $P_{V_n}^\theta$  přiblížit pomocí  $\bar{\Pi}_{V_n}^\theta(\cdot | \bar{x}_{\partial V_n})$  s nějakou volbou okrajové podmínky  $\bar{x}_{\partial V_n}$ , ale i tady bychom při vyčíslení normalizační konstanty  $W_{V_n}^\theta(\bar{x}_{\partial V_n})$  byli postaveni před nutnost sčítat přes všechny konfigurace z  $X^{V_n}$ , což je pro jen trochu větší objemy  $V_n$  zcela nemožné. Existuje sice cesta stochastické aproximace (Younés (1989)), ale v každém případě je tato cesta velmi neschůdná.

Druhý důvod je spíše teoretický. Je zřejmé, že asymptotická normalita odhadu je založena na platnosti příslušné verze centrální limitní věty, jejíž platnost pro náhodná pole je obecně omezena pouze na tzv. oblast jednoznačnosti (tj. bez fázových přechodů). Odhad založený na pseudo-věrohodnosti však vede ke zcela speciálnímu případu, kdy lze platnost centrální limitní věty odvodit (Janžura a Lachout (1994)). Tam, kde lze zaručit asymptotickou normalitu maximálně věrohodného odhadu, vidíme u MPL odhadu úbytek asymptotické eficienty a robustnosti (Janžura (1993b)), což je však zcela přijatelná cena v porovnání s nespornými výhodami.

### 3.4 Identifikace stavu

Tato úloha se svojí podstatou poněkud liší od předcházejících, což budeme ve stručnosti demonstrovat na úloze restaurace degradovaného obrazu.

To je pak neznámou konfigurace  $x_V \in X^V$  (původní obraz) v nějakém pevném objemu  $V \in \mathcal{K}$ . Známý je model, tj. apriorní rozdělení  $P_V$  na  $X^V$ , která je restrikcí nějakého  $P \in G(\Phi)$ ,  $\Phi \in \mathcal{U}_r$ ,  $r > 0$ . Známá jsou také napozorovaná data  $y_V \in X^V$  (degradovaný obraz) a "kanál"  $P(y_V|x_V)$ , který popisuje degradaci. Obvykle bude kanál bezpaměťový, tj.

$$P(y_V|x_V) = \prod_{i \in V} P(y_i|x_i),$$

odpovídající degradaci "nezávisle po bodech".

Řešením by byl

$$\hat{x}_V = \operatorname{argmax}_{x_V \in X^V} Q^\Phi(x_V|y_V),$$

kde  $Q^\Phi(\cdot|\cdot)$  je aposteriorní distribuce. Pro řešení takové úlohy však není k dispozici žádný deterministický algoritmus, využívá se tedy pravděpodobnostní algoritmus nazývaný pro svou fyzikální konotaci "simulovým žíháním". Spočívá v interaktivním postupu, kdy opakovaně procházíme body množiny  $V$  a v každém bodu  $t \in V$  vždy nahradíme stávající stav  $x_t$  stavem  $\tilde{x}_t$  (při nezměněném  $x_{V \setminus \{t\}}$ ) s pravděpodobností, která závisí na poměru

$$\frac{\prod_i^\Phi(\tilde{x}_i|x_{\partial i}) P(y_i|\tilde{x}_i)}{\prod_i^\Phi(x_i|x_{\partial i}) P(y_i|x_i)}.$$

Podrobnosti viz např. Janžura (1990).

## 4 Gaussovská pole – gibbsovský přístup a spektrální analýza

Spektrální teorie náhodných polí rozvíjená v první kapitole naráží na určité problémy při praktických aplikacích. Zejména nemožnost efektivního přenesení mocné věty o faktorizaci spektrální hustoty z jednorozměrného případu na vícerozměrný způsobuje, že se vytrácí jedna z hlavních předností spektrálního přístupu v teorii časových řad – názornost a poměrná jednoduchost statistického zpracování, odhadů parametrů, predikcí atd.

Jednou z možností, jak tuto nesnáz překonat, je kombinace přístupů prvních třech kapitol. Základní literaturou pro tento přístup je Dobrušín (1980) a zejména Künsch (1981). Viz též Janžura (1987) a Otáhal (1991).

### 4.1 Gaussovská náhodná pole jako gibbsovské stavy

4.1.1. *Gaussovské náhodné pole* na  $Z^d$  je dáno podmínkou, že všechna jeho konečněrozměrná rozdělení jsou normální.

Předpokládejme, že je dáno stacionární gaussovské náhodné pole  $X$  s nulovou střední hodnotou a kovarianční funkcí  $R$  takové, že existuje spektrální hustota  $f$  a že pro  $d$ -rozměrné Fourierovy koeficienty  $(a(t) : t \in Z^d)$  příslušné převrácené hodnotě spektrální hustoty, tj.  $a(t) = \int_{I^d} \exp\{2\pi i \cdot \langle t, \lambda \rangle\} f(\lambda)^{-1} d\lambda$ , platí

$$\sum_{t \in Z^d} |a(t)| < \infty.$$

Potom "nekonečné matice"  $(a(s-t))_{s,t \in \mathbb{Z}^d}$  a  $(R(s-t))_{s,t \in \mathbb{Z}^d}$  jsou "vzájemně inverzní" a příslušné pole je gibbovské pro nejvýše párové kvadratické potenciály

$$U_0(x) = \frac{1}{2}a(0) \cdot x^2,$$

$$U_s(x, y) = a(s) \cdot x \cdot y \quad \text{pro } s \neq 0.$$

Pokud předpokládáme, že jen konečně mnoho koeficientů  $a$  je nenulových, dostáváme tak "algebraické" zobecnění autoregresních posloupností konečného řádu, *markovská* gaussovská náhodná pole (kde autoregresní pole dle bodu 1.1.7.(3) představují ne vždy snadno interpretovatelné "geometrické" zobecnění).

4.1.2. Asymptotická analýza provedená v práci Künsch (1981) ukazuje, že rozumnou aproximací maximálně věrohodných odhadů koeficientů  $(a(t) : t \in D)$ , kde  $D$  je konečná množina dosahu interakcí markovského gaussovského pole, jsou "momentové" odhady  $\hat{a}$  dané soustavou rovnic

$$\int_{I^d} \exp\{2\pi i \cdot \langle t, \lambda \rangle\} \hat{f}(\lambda) d\lambda = \hat{R}(t) \quad \text{pro } t \in D,$$

kde

$$\hat{f}(\lambda) = \left( \sum_{t \in D} \hat{a}(t) \cdot \exp\{2\pi i \cdot \langle t, \lambda \rangle\} \right)^{-1}$$

a hodnoty  $\hat{R}$  jsou přirozené odhady kovarianční funkce založené na konečném dosti velkém pozorování. Práce Janžura (1987) se zabývá konzistencí a asymptotickou normalitou těchto odhadů, práce Otáhal (1991) rozšiřuje použití na vektorová náhodná pole, důležitá zejména z hlediska praktických aplikací při počítačovém zpracování družicových snímků, které se zaznamenávají v několika frekvenčních pásmech.

Poznamenejme, že i pro gaussovská markovská pole se mohou projevit fázové přechody, komplikující statistickou analýzu – viz Kurien a Sethumaran (1993).

## Literatura

- Brillinger, D. R. (1974): Fourier analysis of stationary processes. Proc. IEEE 62, 1628–1643.
- Dobrušin, R. L. (1980): Gaussian random fields – Gibbsian point of view. In: Multicomponent Random Systems, M. Dekker, N. York.
- Gidas, B. (1991): Parameter estimation for Gibbs distributions. I: Fully observed data. In: Markov Random Fields: Theory and Applications (R. Chellappa, A. Jain, eds.) Academic, New York.
- Georgii, H. O. (1988): Gibbs Measures and Phase Transitions. de Gruyter, Berlin.
- Guyon, X. a Künsch, H. R. (1992): Asymptotic comparison of estimators in the Ising model. In: Stochastic Models, Statistical Methods and Algorithms in Image Analysis. (P. Barone, A. Frigessi, M. Piccioni, eds.). Lecture Notes in Statistics 74, Springer, Berlin 177–198.
- Helson, H. a Lowdenslager, D. (1958, 1961): Prediction theory and Fourier series in several variables. Part I (1958), Acta Math. 99, 165–1.702, Part II (1961), Acta Math. 106, 175–1.713.

- Jadrenko, M. I. (1980): Spektralnaja teorija slučajnych polej. Kijev.
- Jaglom, A. M. (1957): Nekotoryje klassy slučajnych polej v n-mernom propstranstve, dodstvennyje stacionarnym slučajnym processam. Teorije verojat. i prim. 2, No. 3, 292–338.
- Jaglom, A. M. (1960): Položitelnyje opredelenyje funkcii i odnorodnyje slučajnyje polja na gruppach i odnorodnych prostranstvach. DAN SSSR 135, No. 6, 1342–1345.
- Jaglom, A. M. (1961): Second-order homogeneous random fields. Proc. 4th Berkeley Symp., Univ. of California Press, Berkeley and Los Angeles.
- Janžura, M. (1987): Odhady parametrů gaussovských náhodných polí. Výzk. zpráva ÚTIA ČSAV č. 1437.
- Janžura, M. (1990): O jednom pravděpodobnostním algoritmu pro optimalizační úlohy. ROBUST 90, JČSMF, 84–88.
- Janžura, M. (1993a): Asymptotic behaviour of the error probabilities in the pseudo-likelihood ratio test for Gibbs–Markov distributions. To appear in Proceedings of the Fifth Prague Symposium on Asymptotic Statistics.
- Janžura, M. (1993b): Asymptotic results in parameter estimation for Gibbs random fields. ÚTIA Research Report No. 1763.
- Janžura, M. a Lachout, P. (1994): A central limit theorem for stationary random fields. Preprint.
- Kurien, T. V. a Sethuraman, J. (1993): Singularities in Gaussian random fields. J. Theor. Prob. 6, No. 1, 89–99.
- Künsch, H. (1981): Thermodynamics and statistical analysis of Gaussian random fields. Z. Wahrs. v. Geb. 58, 407–421.
- Onsager, L. (1944): Crystal Statistics I. a two-dimensional model with an order-disorder transition. Phys. Rev. 65, 117–149.
- Otáhal, A. (1984): Spectrum Decomposition for stationary weakly isotropic random fields with bounded range interactions. Kybernetika 20, No. 6, 434–444.
- Otáhal, A. (1986): Isotropy of stationary random fields on lattice. Kybernetika 22, No. 3, 256–1.767.
- Otáhal, A. (1988): Spectrum decomposition for stationary weakly isotropic random fields in the plane. Trans. 10th Prague Conference. 223–1.728, Academia, Praha.
- Otáhal, A. (1991): Parameter estimation for nearest neighbor Gaussian random fields in the Plane. Problems of Control and Inf. Theory 20 (6), 429–439.
- Ripley, B. D. (1981): Spatial Statistics. Wiley, New York.
- Ruelle, D. (1969): Statistical Mechanics. Rigorous Results. Benjamin, New York.
- Preston, C. (1976): Random Fields (Lecture Notes in Math. 534), Springer Berlin.
- Tjøstheim, D. (1978, 1983): Statistical spatial series modelling. Part I (1978), Adv. Appl. Prob. 10, 130–154, Part II (1983), Adv. Appl. Prob. 15, 562–584.
- Weil, A. (1950): Integrovanije v topologičeskich gruppach i jevo primenenija. Moskva (překlad z francouzštiny).
- Whittle, P. (1954): On stationary processes in the plane. Biometrika 41, 450–462.
- Younés, L. (1989): Parametric inference for imperfectly observed Gibbsian fields. Probab. Theory Related Fields 82, 625–645.