# ODHADY PARAMETRŮ PRO VEKTOROVÁ GAUSS-MARKOVOVA NÁHODNÁ POLE

Antonín Otáhal, ÚTIA ČSAV, Praha

Pojem vektorového Gauss-Markovova náhodného pole je určitým zobecněním pojmu stacionární autoregresní posloupnosti - jeho spektrální hustota v jistém smyslu představuje vícerozměrnou verzi převrácené hodnoty trigonometrického polynomu, což je právě tvar spektrální hustoty stacionární autoregresní posloupnosti. Přesně je tento pojem zaveden v první části práce. V druhé části zobecňujeme postup pro odhad parametrů, který pro skalární Gauss-Markovova náhodná pole odvodil Janžura (1988), na vektorový případ. Třetí část se zabývá předpokladem separability kanálů, jehož přijetí umožňuje výrazně zjednodušit odhadování parametrů.

## 1. Základní pojmy

*Vektorové náhodné pole* X je systém n-rozměrných náhodných vektorů, které jsou indexovány d-ticemi celých čísel, $X = (X(t): t \in T)$, kde $T = Z^d$ a $X(t) = (X_1(t),\dots,X_n(T))^*$ ($^*$ značí transpozici). Budeme uvažovat *gaussovská* pole tohoto typu, tj. předpokládat, že každý konečný soubor náhodných veličin z X má sdružené normální rozdělení. Dále přijmeme předpoklad *stacionarity*, podle něhož všechny náhodné veličiny v X mají nulovou střední hodnotu a existuje maticová *kovarianční funkce* R typu $n \times n$ taková, že pro všechna $s, t \in T$ platí $EX(s)X(t)^* = R(s-t)$. Stacionární vektorové náhodné pole má (maticovou) *spektrální hustotu* f, jestliže pro všechna $t \in T$ platí

$$R(t) = \int_{I^d} e^{i\langle t,x\rangle} f(x)dx, \tag{1}$$

kde $I = [-\pi,\pi)$, integruje se "po složkách", tj. $\left(\int M(x)dx\right)_{jk} = \int M_{jk}(x)dx$, a $\langle .,.\rangle$ značí běžný skalární součin. Přitom matice $f(x)$ je pozitivně semidefinitní pro skoro všechna $x \in I^d$.

Jako *VGM-pole* (vektorové Gauss-Markovovo náhodné pole) budeme označovat náhodné pole, které vyhovuje všem výše uvedeným podmínkám a navíc pro všechna $x \in I^d$ je příslušnou hodnotou spektrální hustoty *reálná* pozitivně definitní matice $f_U(x)$ daná vztahem

$$f_U^{-1}(x) = \sum_{t\in A} U_t \cos\langle t,x\rangle, \tag{2}$$

kde $^{-1}$ značí maticovou inverzi, A je *konečná* množina, která obsahuje nulový index a všechny nenulové prvky množiny A jsou v lexikografickém uspořádání indexové množiny T větší než nulový index, $U = (U_t: t \in A)$ je soubor symetrických matic a matice $U_0$ je navíc pozitivně definitní.

Označme symbolem $\mathcal{U}(A)$ množinu všech souborů matic $U = (U_t: t \in A)$, které vyhovují všem těmto požadavkům.

**1. 1. Poznámka.** Předpoklad, že spektrální hustota je reálná, implicitně obsahuje předpoklad, že všechny matice $R(t)$, $t \in T$ jsou symetrické, tj. že uvažované náhodné pole je v tomto smyslu *symetrické*.

Konečný soubor matic $U = (U_t:\ t \in A)$ můžeme považovat za parametry, které určují pravděpodobnostní rozdělení daného pole; tyto parametry mají přirozenou interpretaci jako interakce mezi dvojicemi náhodných veličin daného pole. Taková interpretace úzce souvisí s pojetím z hlediska statistické fyziky, v jehož rámci se na tato pole pohlíží jako na gibbsovská náhodná pole příslušná párovým potenciálům U. Podrobněji o tomto pojetí viz např. Dobrušin (1980), Künsch (1981).

Důležitým a z hlediska aplikací zajímavým případem VGM-pole je pole s interakcemi omzenými na *nejbližší sousedy*, jehož množina A obsahuje pouze ty nenulové indexy, které mají jednotkovou euklidovskou normu. Pro případ rovinných polí, kdy $T = \mathbb{Z}^d$, to znamená spektrální hustotu tvaru

$$f^{-1}(x_1, x_2) = U_0 + U_1 \ .\ \cos x_1 + U_2 \ .\ \cos x_2, \tag{3}$$

kde $U = (U_0, U_1, U_2) = \left[U_{(0,0)}, U_{(0,1)}, U_{(1,0)}\right]$ je příslušný soubor parametrů.

## 2. Odhady parametrů

Nechť je dána konečná podmnožina M indexové množiny T. Uvažujme konečný úsek VGM-pole X určený touto množinou, tj. $X^M = (X(t): t \in M)$. Předpokládejme, že je známa množina A příslušná poli X podle vztahu (2) a že soubor parametrů $U \in \mathcal{U}(A)$ je neznámý. Naší snahou bude odhadnout U na základě $X^M$. K tomuto odhadu zobecníme postup, který pro skalární případ ($n = 1$) odvodil Janžura (1988). Vlastnosti odhadu budou studovány asymptoticky pro $M \to T$.

Pro $t \in T$ označme $M_t = \{s \in M;\ s+t \in M\}$. Pro pozdější asymptotické úvahy definujme konvergenci $M \to T$ v tom smyslu, že pro každé $t \in T$ platí $|M_t| \ /\ |M| \to 1$, kde $|.|$ značí počet prvků množiny.

Pro taková $t \in T$, pro která je $M_t \neq \emptyset$, a pro $j,k = 1,\dots,n$ označme

$$\hat{R}^M_{jk}(t) = \frac{1}{2.\ |M_t|} \sum_{s \in M_t} (X_j(s+t)X_k(s) + X_k(s+t)X_j(s)) \tag{4}$$

a dále

$$\hat{R}^M(t) = \left[\hat{R}^M_{jk}(t)\right]^n_{j,k=1} \tag{5}$$

**2. 1. Věta.** Vztahy (4) a (5) určují nestranný a pro $M \to T$ konzistentní odhad hodnoty kovarianční funkce $R(t)$.

**Důkaz** by bylo možné provést bezprostředním zobecněním standardních technik analýzy časových řad, viz např. Anderson (1971), 8.3.2. Zdůrazněme jen, že na pravé straně vztahu (4) se provádí symetrizace odhadnuté matice $R^M(t)$, která vychází ze symetrie uvažovaného pole ve smyslu poznámky 1.1. □

Označme symbolem $R^U$ kovarianční funkci, která podle vztahu (1) odpovídá spektrální hustotě $f_U$ dané souborem parametrů $U$ ve vztahu (2), tj. $R^U$ označuje kovarianční funkci VGM-pole, jehož parametry jsou $U$. Předpokládejme, že $M_t \neq \emptyset$ pro všechna $t \in A$ a definujme odhad parametrů $U$ jako ten soubor matic $\hat{U}(M) = (\hat{U}_t(M)$: $t \in A)$, pro který platí

$$R^{\hat{U}(M)}(t) = \hat{R}^M(t) \quad \text{pro všechna } t \in A \tag{6}$$

Každému odhadu $\hat{R}^M$ kovarianční funkce odpovídá nejvýše jeden odhad $\hat{U}(M)$, jak plyne z následujícího lemmatu.

**2.2. Lemma.** Nechť $U, V \in \mathcal{U}(A)$ a

$$R^U(t) = R^V(t) \quad \text{pro všechna } t \in A. \tag{7}$$

Potom platí $U = V$.

**Důkaz.** Označme $g_1(x) = f_U^{-1/2}(x) f_V(x) f_U^{-1/2}(x)$, $g_2(x) = f_V^{-1/2}(x) f_U(x) f_V^{-1/2}(x)$ a dále nechť $\lambda_1^1(x), \dots, \lambda_n^1(x)$ (resp. $\lambda_1^2(x), \dots, \lambda_n^2(x)$) jsou vlastní čísla matice $g_1(x)$ (resp. $g_2(x)$). Předpoklad $U, V \in \mathcal{U}(A)$ má za následek, že všechna tato vlastní čísla jsou kladná pro všechna $x \in I^d$. Má proto smysl definovat pro $j = 1,2$

$$h_j(x) = \mathrm{Tr}\, g_j(x) - n - \ln \det g_j(x) = \sum_{k=1}^{n} \left(\lambda_k^j - 1 - \ln \lambda_k^j\right). \tag{8}$$

Platí

$$\int_{I^d} (h_1(x) + h_2(x))dx = \mathrm{Tr} \int_{I^d} \left[f_U(x) - f_V(x)\right] \left[f_V^{-1}(x) - f_U^{-1}(x)\right] dx.$$

Integrál napravo je roven nule vzhledem k (7), protože podle (1) a (2) je lineární kombinací matic $\left[R_t^U - R_t^V\right]$, $t \in A$. Z (8) plyne $h_1(x) \geq 0$, $h_2(x) \geq 0$, a tedy platí $h_1(x) = h_2(x) = 0$. Opět z (8) dostáváme $\lambda_k^j = 1$ pro $j = 1,2$, $k = 1,\dots,n$, což vzhledem k symetrii matic $g_1(x)$, $g_2(x)$ znamená $f_U^{-1}(x) = f_V^{-1}(x)$ pro skoro všechna $x \in I^d$, proto podle (2) je $U = V$. □

**2.3. Lemma.** Odhad $\hat{U}(M)$ existuje s pravděpodobností, která pro $M \to T$ konverguje k 1.

**Důkaz** plyne z konzistence odhadů $\hat{R}^M(t)$, $t \in A$ - viz věta 2.1.

Obdobně jako Janžura (1988), věta 3.4, bychom mohli dokázat konzistenci a asymptotickou normalitu odhadu $\hat{U}(M)$.

Poznamenejme, že tento odhad jednak v jistém smyslu minimalizuje ve třídě VGM-polí vzdálenost od výběrové distribuce, jednak je na něj možné pohlížet jako na aproximaci (zpřesňující se pro $M \to T$) maximálně věrohodného odhadu parametrů $U$. Podrobněji tyto otázky zkoumá např. Künsch (1981).

## 3. Separabilita kanálů

Hledání odhadu $\hat{U}(M)$ na základě odhadnutých hodnot kovarianční funkce podle (5) představuje velmi obtížnou úlohu. Pro její zjednodušení budeme předpokládat jisté dodatečné vlastnosti závislostní struktury uvažovaného pole, které umožní převést tuto úlohu na vzájemně nezávislé skalární úlohy.

Konkrétně budeme předpokládat, že je možné *separovat kanály* uvažovaného VGM-pole X, tj. najít takovou lineární transformaci, představovanou regulární maticí L typu n x n, že skalární náhodná pole $Y_1, \dots, Y_n$ definovaná předpisem

$$Y(t) = L \,.\, X(t), \qquad T \in T \tag{9}$$

jsou vzájemně nekorelovaná. Ekvivalentně to znamená, že matice

$$L \,.\, R(t) \,.\, L^* \text{ je diagonální pro } t \in T. \tag{10}$$

Nyní se budeme snažit využít tohoto předpokladu pro odhady parametrů U. Nejprve ukážeme, že vlastnost (10) se přenáší na matice parametrů v U.

**3. 1. Lemma.** Nechť je pro VGM-pole X s parametry $U \in \mathcal{U}(A)$ splněna podmínka (10). Potom pro každé $t \in A$ je matice $V_t = L \,.\, U_t \,.\, L^*$ diagonální.

**Důkaz.** Vektorové náhodné pole Y definované v (9) je zřejmě VGM-pole s parametry $V = (V_t;\ T \in A) \in \mathcal{U}(A)$. Speciálně má, pro všechna $j = 1, \dots, n$, skalární náhodné pole $Y_j$ spektrální hustotu $f_j(x) = 1/\left[\sum_{t \in A} (V_t)_{jj} \cos\langle t, x\rangle\right]$, a tedy označíme-li $S_j(t) = (L.R(t).L^*)_{jj}$, platí

$$\begin{bmatrix} S_1(t) & & 0 \\ & \ddots & \\ 0 & & S_n(t) \end{bmatrix} = \int_{I^d} e^{i\langle t,x\rangle} \begin{bmatrix} f_1(x) & & 0 \\ & \ddots & \\ 0 & & f_n(x) \end{bmatrix} dx.$$

Podle předpokladu (10) stojí v poslední rovnosti nalevo přímo hodnota kovarianční funkce pole Y v indexu t, tj. diagonální matice $L \,.\, R(t) \,.\, L^*$, takže vzhledem k jednoznačnosti vzájemného vztahu mezi kovarianční funkcí a spektrální hustotou (viz např. Stein, Weiss (1971), věta VII.1.7) jsou všechny matice $V_t$, $t \in A$ diagonální. □

Je-li dáno VGM-pole X s kovarianční funkcí R, z pozitivní definitnosti spektrální hustoty plyne, že matice R(0) je pozitivně definitní. Má proto smysl definovat *korelační funkci*

$$\rho(t) = R(0)^{-1/2} \,.\, R(t) \,.\, R(0)^{-1/2}, \qquad t \in T. \tag{11}$$

Zřejmě je $\rho(0) = I$, kde I značí jednotkovou matici.

**3. 2. Lemma.** VGM-pole X s kovarianční funkcí R a korelační funkcí $\rho$ vyhovuje podmínce separability kanálů (10), právě když existuje ortogonální matice Q taková, že

$$Q \,.\, \rho(t) \,.\, Q^* \text{ je diagonální} \tag{12}$$

platí pro všechna $t \in T$.

Důkaz. 1. Platí-li (12), volbou $L = Q \cdot R(0)^{-1/2}$ plyne (10).
2. Ekvivalentním přepisem podmínky (10) plyne, že matice $L \cdot R(0)^{1/2} \cdot R(0)^{1/2} \cdot L^*$ je diagonální pro všechna $t \in T$, takže $Q = L \cdot R(0)^{1/2}$ vyhovuje podmínce (12). Speciální volbou $t = 0$ v (12) plyne $Q \cdot Q^* = I$, což znamená ortogonalitu matice $Q$. □

Uvažujme nyní situaci části 2., tj. odhady kovarianční funkce (4), (5) a definujme pro $t \in A$

$$\hat{\rho}^M(t) = (\hat{R}^M(0))^{-1/2} \cdot \hat{R}^M(t) \cdot (\hat{R}^M(0))^{-1/2}. \tag{13}$$

Tato definice má smysl, protože pro VGM-pole X je matice $\hat{R}^M(0)$ skoro jistě pozitivně definitní. Navíc z konzistence odhadů $\hat{R}^M$ plyne konzistence odhadů $\hat{\rho}^M$. Pro $\hat{\rho}^M$ však nemusí platit podmínka (12), a to ani při restrikci indexů na $t \in A$. Abychom mohli provést rozklad (12) na A alespoň přibližně, budeme hledat matici $\hat{Q}$ tak, aby se matice $\hat{Q} \cdot \hat{\rho}^M(t) \cdot \hat{Q}^*$ pro $t \in A$ co nejvíce blížily diagonálním.

Označme $\bar{A} = A \setminus \{0\}$ a symbolem $O_n$ označme grupu všech ortogonálních matic typu $n \times n$. Definujme funkci c na $O_n$ předpisem

$$c(P) = \sum_{t \in \bar{A}} \sum_{j=1}^{n} \sum_{\substack{k=1 \\ j \neq k}}^{n} \left[P . \hat{\rho}^M(t) . P^*\right]^2_{jk}, \tag{14}$$

tj. jako součet čtverců všech nediagonálních prvků všech matic, které jsou prostřednictvím matice P ortogonálně podobné všem maticím $\hat{\rho}^M(t)$, $t \in \bar{A}$. Jako odhad matice Q z (12) pak vezměme tu matici $\hat{Q} \in O_n$, pro kterou je minimalizována funkce c:

$$c(\hat{Q}) = \min_{P \in O_n} c(P). \tag{15}$$

Protože c je spojitá funkce definovaná na kompaktní grupě $O_n$, matice $\hat{Q}$ daná v (15) vždy existuje; nemusí však existovat jednoznačně. Způsob jejího nalezení je založen na tzv. *algoritmu redukcibilní aproximace*, viz Otáhal (1988).

Rozklad (12) přibližně provedeme tak, že "zapomeneme" nediagonální prvky matic $D(t) = \hat{Q} \cdot \hat{\rho}^M(t) \cdot \hat{Q}^*$, $t \in A$. To znamená, pro $j = 1,\dots,n$ položíme $r_j(0) = 1$, $r_j(t) = D_{jj}(t)$ pro $t \in \bar{A}$ a definujeme odhad $u_j = (u_j(t)\colon t \in A)$ soustavou rovnic

$$r_j(t) = \int_{I^d} \frac{\cos\langle t,x\rangle}{\sum_{s \in A} u_j(s)\cos\langle s,x\rangle}\, dx \quad \text{pro } t \in A, \tag{16}$$

tedy jako skalární verzi odhadu (6). Pro $u_j$ tedy speciálně platí lemma 2.2 a 2.3: odhad $u_j$ je definován s pravděpodobností, která konverguje k 1 pro $M \to T$ a pokud $u_j$ existuje, je určen jednoznačně pro dané hodnoty $(r_j(t)\colon t \in A)$.

Pro zpětný přechod k odhadu původních parametrů U položme $K = (R^M(0))^{1/2} \cdot \hat{Q}^*$ a

$$\tilde{U}(t) = K \cdot \begin{bmatrix} u_1(t) & & 0 \\ & \ddots & \\ 0 & & u_n(t) \end{bmatrix} K^* \quad \text{pro } t \in A. \tag{17}$$

**3.3. Věta.** Nechť VGM-pole X s parametry $\underset{\sim}{U} \in \underset{\sim}{U}(A)$ splňuje podmínku separability kanálů (10). Potom systém matic $\tilde{U} = (\tilde{U}(t):\ t \in A)$ definovaný v (17) je konzistentním odhadem parametrů U pro $M \to T$.

**Důkaz.** Kdybychom nalezení matice $\hat{Q}$ podle (14), (15) a výpočet $u_j$, $\tilde{U}$ v (16), (17) prováděli pro $\rho$ místo $\hat{\rho}^M$, matice $\hat{Q}$ by přesně diagonalizovala $\rho$ ve smyslu (12), a muselo by proto platit $\tilde{U} = U$. Protože odhady $\hat{\rho}^M$ jsou konzistentní, $\hat{\rho}^M$ konverguje k $\rho^M$ pro $M \to T$, plyne odtud $\tilde{U} \to U$ pro $M \to T$.

□

## Literatura

*Anderson T. W. (1971):* The Statistical Analysis of Time Series, Wiley & Sons, New York - London - Sydney - Toronto.

*Dubrušin, R. L. (1980):* Gaussian Random Fields - Gibbsian Point of View, in: Multicomponent Random Systems, eds. R. L. Dobrušin, Ya. G. Sinai, M. Dekker, New York.

*Janžura, M. (1988):* Asymptotic Theory of Parameter Estimation for Gauss-Markov Random Fields, Kybernetika 24, No. 3, 161-176.

*Künsch H. (1981):* Thermodynamics and Statistical Analysis of Gaussian Random Fields, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 55, 407-421.

*Otáhal A. (1988):* Algoritmus reducibilní aproximace symetrických matic, Výzkumná zpráva 1512, ÚTIA ČSAV Praha.

*Stein E. M., Weiss G. (1971):* Introduction to Fourier Analysis on Euclidean Spaces, Princeton Univ. Press.

Adresa: Antonín Otáhal, ÚTIA ČSAV, Pod vodárenskou věží 4, 182 08 Praha 8