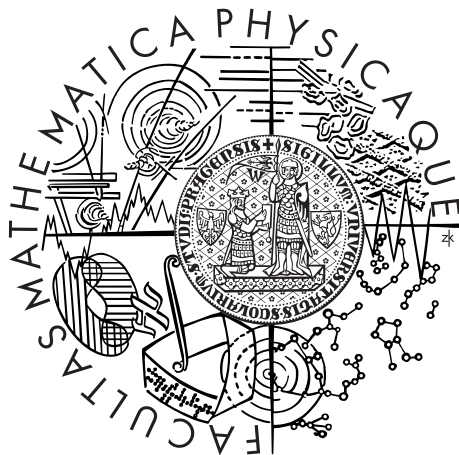


Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Karel Scheib

Bodový proces řízený Gaussovským polem

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Viktor Beneš, DrSc.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2012

Poděkování

Mé poděkování patří především vedoucímu mé bakalářské práce prof. RNDr. Viktoru Benešovi, DrSc., bez jehož trpělivého a neúnavného nasazení nemohla práce nikdy vzniknout. Dále děkuji všem, kteří mě podporovali během celého mého studia.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Název práce: Bodový proces řízený Gaussovským polem

Autor: Karel Scheib

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Viktor Beneš, DrSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce vyšetřuje hledání podprostoru redukce dimenze pro Poissonův bodový proces řízený Gaussovským náhodným polem. V práci je popsána metoda plátkové inverzní regrese, která je aplikována na prostředí bodového procesu řízeného náhodným polem. Dále je dokázána její funkčnost ve zmíněném kontextu. V prostředí programu R je metoda více způsoby implementována a testována na náhodných datech. Jednotlivé způsoby jsou popsány a výsledky vzájemně srovnány.

Klíčová slova: bodový proces, náhodné pole, plátková inverzní regrese, redukce dimenze

Title: A point process driven by a Gaussian field

Author: Karel Scheib

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Viktor Beneš, DrSc., Department of Probability and Mathematical Statistics

Abstract: The thesis investigates the search for dimension reduction subspace for the Poisson point process driven by a Gaussian random field. The work describes the method called sliced inverse regression, which is applied to a point process driven by random field. Its functionality in mentioned context is then proved. This method is in several ways implemented and tested in R software environment on random data. The individual implementations are described and results are then compared with each other.

Keywords: point process, random field, sliced inverse regression, dimension reduction

Obsah

Úvod	1
Značení	3
1 Teoretická část	4
1.1 Náhodná pole	4
1.2 Bodové procesy	4
1.3 Redukce dimenze	8
1.4 Redukce dimenze pro Coxův bodový proces řízený náhodným polem	10
1.5 SIR - Plátková inverzní regrese	12
1.6 Plátkování pomocí kót	13
1.7 Způsoby měření kvality nalezených směrů redukce dimenze	14
2 Praktická část	15
2.1 Volba modelu	15
2.2 Popis konstrukce simulace	16
2.2.1 Náhodné pole	16
2.2.2 Bodový proces	16
2.2.3 SIR	16
2.3 Výsledky	18
2.3.1 Model I.	18
2.3.2 Model II.	18

Závěr	19
Seznam použité literatury	21
Přílohy	22

Úvod

Chceme-li modelovat náhodná místa výskytu nějakého jevu v rovině (konkrétně na dvourozměrném Euklidovském prostoru) přirozeně zvolíme matematický model nazývající se náhodný bodový proces. Typickým a obvykle se vyskytujícím příkladem takového procesu je proces Poissonův. Mějme dále v každém bodě roviny vektor náhodných veličin s jistými vlastnostmi. Představíme-li si Poissonův bodový proces, který konkrétním způsobem závisí na tomto vektoru, získáme pojem Coxova procesu řízeného vícerozměrným náhodným polem. Pro účely práce se omezíme na stacionární Gaussovské náhodné pole. Toto náhodné pole pak určuje intenzitu zkoumaného Coxova bodového procesu. Často je však závislost bodového procesu různá pro jednotlivé složky řídicího náhodného pole. Pak bychom rádi věděli, které složky náhodného pole mají vliv na bodový proces a jak velký tento vliv je. Zkoumání této závislosti vystihuje hledání podprostoru redukce dimenze bodového procesu závislého na náhodném poli. Práce se dále zabývá hledáním centrálního podprostoru redukce dimenze, neboli podprostoru redukce dimenze s nejmenší dimenzí. Práce v podstatě následuje článek [3], vysvětluje jednotlivé jeho části při detailnější formulaci použitých pojmů. V druhé části obměněným způsobem aplikuje popsaný teoretický aparát na testovací data v prostředí programu R.

V teoretické části práce nejprve zavedeme pojmy náhodného pole, bodového procesu, podprostoru redukce dimenze a centrálního podprostoru redukce dimenze. Pojmy poté spojíme a definujeme centrální podprostor redukce dimenze pro první momentovou míru Coxova bodového procesu (neboli Poissonova bodového procesu závislého na vícerozměrném náhodném poli). Popíšeme metodu plátkové inverzní regrese (SIR), která obecně řeší problém hledání podprostoru redukce dimenze. Dále aplikujeme metodu plátkové inverzní regrese pro redukci dimenze na prostředí Coxova bodového procesu a odůvodníme její funkčnost. Pro správné použití důkazového aparátu dodatečně požadujeme ergodicitu náhodného pole a jeho zúžení na body bodového procesu. Protože metodu plátkové inverzní regrese lze v našem kontextu aplikovat více způsoby (způsob vytváření plátek může být různý) navrhujeme pro srovnání tři různé způsoby aplikace. Technicky použijeme prostředí tzv. kótovaného bodového procesu. Dále navrhujeme dvě statistické metody měřící účinnost získaného postupu.

Praktická část práce popisuje konkrétní naprogramování metod, popsaných v části teoretické. Pro dva různé modely se na náhodně vygenerovaných datech vyzkouší jednotlivé verze aplikace hledání podprostoru redukce dimenze. Tyto metody budou ve výsledku práce vzájemně srovnány a v závěru diskutovány.

Příloha práce obsahuje jednak histogramy uváděných výsledků, dále pak komentovaný zdrojový kód aplikace v programu R.

Značení

\mathbb{N}, \mathbb{R}	množiny přirozených, reálných čísel
\mathbb{N}_0	$\mathbb{N} \cup \{0\}$
\mathbb{R}^d	d -rozměrný Euklidovský prostor
\mathbb{R}_+	$\{x \in \mathbb{R}; x \geq 0\}$
$(\Omega, \mathcal{A}, \mathbb{P})$	pravděpodobnostní prostor
$\text{card}(\cdot)$	počet bodů množiny
\mathcal{B}^2	system borelovských podmnožin \mathbb{R}^2
\mathcal{B}_0^2	system omezených borelovských podmnožin \mathbb{R}^2
$\mathbf{1}_A$	charakteristická funkce množiny A
$ \cdot $	Lebesgueova míra množiny v \mathbb{R}^2
$\bar{0}$	p -rozměrný nulový vektor
\mathbf{I}_p	p -rozměrná jednotková matice
$\mathcal{S}(\mathbf{B})$	lineární podprostor generovaný sloupci matice \mathbf{B}
$\text{Cov}(\mathbb{X})$	kovarianční matice náhodného vektoru \mathbb{X}
$N(\bar{0}, \mathbf{I}_p)$	p -rozměrné normální rozdělení s nulovým vektorem středních hodnot a jednotkovou rozptylovou maticí

Kapitola 1

Teoretická část

1.1 Náhodná pole

Definice 1. Náhodné pole

Nechť $W \subset \mathbb{R}^2$, $|W| > 0$. *Náhodným polem* nazveme kolekci náhodných veličin $\{\mathbf{X}(s); s \in W\}$ definovaných na $(\Omega, \mathcal{A}, \mathbb{P})$.

Definice 2. Gaussovské náhodné pole

Náhodné pole $\{\mathbf{X}(s); s \in W\}$ nazveme *Gaussovským*, jestliže všechna jeho konečně rozměrná rozdělení jsou normální, tedy: $\forall n \in \mathbb{N}$ a $x_1, \dots, x_n \in W$ má vektor $(\mathbf{X}(x_1), \dots, \mathbf{X}(x_n))^T$ n -rozměrné normální rozdělení.

Definice 3. Stacionární náhodné pole

Náhodné pole $\{\mathbf{X}(s); s \in W\}$ nazveme (*striktně*) *stacionárním*, jestliže $\forall n \in \mathbb{N}$, $\forall x_1, \dots, x_n \in W$, a $\forall h \in \mathbb{R}^2$, $x_1 + h, \dots, x_n + h \in W$ jsou rozdělení $(\mathbf{X}(x_1), \dots, \mathbf{X}(x_n))^T$ a $(\mathbf{X}(x_1 + h), \dots, \mathbf{X}(x_n + h))^T$ stejná.

1.2 Bodové procesy

Definujme systém lokálně konečných podmnožin \mathbb{R}^2 jako

$$\mathcal{N} := \{\varphi \subseteq \mathbb{R}^2; \text{card}(\varphi \cap B) < \infty \forall B \subset \mathcal{B}_0^2\}.$$

Na \mathcal{N} zavedeme σ -algebru takto:

$$\mathfrak{N} = \sigma \{\mathcal{U}_{B,m} : m \in \mathbb{N}_0, B \in \mathcal{B}_0^2\}$$

kde $\mathcal{U}_{B,m} = \{\varphi \in \mathcal{N}; \text{card}(\varphi \cap B) = m\}$.

Definice 4. Bodový proces

Měřitelné zobrazení $\mathbf{N} : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathcal{N}, \mathfrak{N})$ nazveme *bodový proces*. Rozdělení bodového procesu \mathbf{N} je pravděpodobnostní míra $\mathbb{P}_{\mathbf{N}}$ na \mathfrak{N} , kde

$$\mathbb{P}_{\mathbf{N}}(F) = \mathbb{P}[\mathbf{N} \in F], \quad \forall F \in \mathfrak{N}.$$

Z **Definice 4.** vidíme, že pro bodový proces \mathbf{N} jakožto zobrazení $(\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathcal{N}, \mathfrak{N})$ je $\mathbf{N}(\omega)$ lokálně konečná (jistě tedy spočetná) množina pro jakékoliv $\omega \in \Omega$. Náhodnou veličinu počtu bodů množiny $\mathbf{N}(\omega)$ v dané množině $B \in \mathcal{B}_0^2$ označme symbolem $\mathbf{N}(B)$.

Definice 5. Kótovaný bodový proces

Buď $(\mathcal{M}, \mathfrak{M})$ měřitelný prostor. *Kótovaným bodovým procesem* Ψ v \mathbb{R}^2 rozumíme náhodnou posloupnost bodů $\{[x_n; m_n]\}$, kde $\{x_n\}$ tvoří realizaci bodového procesu \mathbf{N} v \mathbb{R}^2 , $\{m_n\} \in \mathcal{M}$, $n \in \mathbb{N}$.

Poznámky

[4, str. 105]: Vždy je možné interpretovat kótovaný bodový proces jako bodový proces nad $\mathbb{R}^2 \times \mathcal{M}$.

Struktura prostoru \mathcal{M} může být různá. Například v [4] se požaduje, aby \mathcal{M} byl polský prostor. V naší práci si vystačíme s případy: $\mathcal{M} = \{0, 1\}$, $\mathcal{M} = \mathbb{R}$ a $\mathcal{M} = \mathbb{R}^p$.

Definice 6. Náhodná míra

Zobrazení $\Gamma : \Omega \times \mathcal{B}^2 \rightarrow \mathbb{R}_+$ nazveme *náhodná míra*, jestliže platí:

- (i) $\forall \omega \in \Omega$ je $\Gamma(\omega, \cdot)$ míra na \mathcal{B}^2 ,
- (ii) $\forall B \in \mathcal{B}_0^2$ je $\Gamma(\cdot, B)$ náhodná veličina.

Definice 7. Stacionarita bodového procesu

Nechť \mathbf{N} je bodový proces. Řekneme, že \mathbf{N} je stacionární, je-li jeho rozdělení je invariantní vůči posunutím, tj.: rozdělení $\mathbf{N} + y = \{X + y; X \in \mathbf{N}\}$ je stejné jako rozdělení \mathbf{N} pro všechna $y \in \mathbb{R}^2$.

Kótovaný bodový proces Ψ je stacionární, jestliže rozdělení Ψ a Ψ_y je stejné pro všechna $y \in \mathbb{R}^2$, kde $\Psi_y = \{[x_1 + y, m_1], [x_2 + y, m_2], \dots\}$.

Definice 8. Lokálně konečná míra, difúzní míra

Nechť $(\mathbb{R}^2, \mathcal{B}_0^2, \mu)$ je prostor s mírou. Míru μ nazveme *lokálně konečnou* platí-li, že pro každou kompaktní množinu $K \subset \mathbb{R}^2$ je $\mu(K) < \infty$. Míru μ nazveme *difúzní* platí-li, že $\mu(\{x\}) = 0$ pro každé $x \in \mathbb{R}^2$.

Definice 9. Míra (funkce) intenzity bodového procesu

Pro bodový proces \mathbf{N} definujeme *míru intenzity* předpisem

$$\mu(B) = \mathbb{E}\mathbf{N}(B) = \int_{\mathcal{N}} \mathbf{N}(B) \mathbb{P}_{\mathbf{N}}(d\mathbf{N}), \quad B \in \mathcal{B}^2,$$

$\mu(B)$ je tedy střední počet bodů procesu \mathbf{N} v množině B . Existuje-li hustota λ vzhledem k Lebesgueově míře (tj. $\mu(B) = \int_B \lambda(x) dx$, $B \in \mathcal{B}^2$), potom se λ nazývá *funkce intenzity* bodového procesu \mathbf{N} .

Definice 10. Míra intenzity kótovaného bodového procesu, rozdělení kóty

Pro kótovaný bodový proces Ψ , $B \in \mathcal{B}^2$, $L \in \mathcal{M}$, definujeme jeho *míru intenzity* ν předpisem

$$\nu(B \times L) = \mathbb{E}\Psi(B \times L),$$

kde náhodná míra $\Psi(B \times L)$ je počet bodů Ψ v B s kótami v L .

Jistě $\forall L \in \mathcal{M}$ je $\nu(\cdot \times L)$ absolutně spojitá vzhledem k μ ¹. Proto existuje pravděpodobnostní míra $M_s(\cdot)$ na \mathcal{M} , $s \in \mathbb{R}^2$ tak, že

$$\nu(B \times L) = \int_B M_s(L) \mu(ds) = \int_B \int_L M_s(dx) \mu(ds) \quad (1.1)$$

která se nazývá *rozdělení kóty* v bodě s .

Campellova věta

- (i) Nechť \mathbf{N} je bodový proces, λ jeho funkce intenzity, $h: \mathbb{R}^2 \rightarrow \mathbb{R}_+$ měřitelná funkce. Pak platí:

$$\mathbb{E} \left[\sum_{\mathbf{x} \in \mathbf{N}} h(\mathbf{x}) \right] = \int_{\mathbb{R}^2} h(t) \lambda(t) dt \quad (1.2)$$

- (ii) Nechť Ψ je kótovaný bodový proces, ν jeho míra intenzity, $g: \mathbb{R}^2 \times \mathcal{M} \rightarrow \mathbb{R}_+$ měřitelná funkce. Pak platí:

$$\mathbb{E} \left[\sum_{[\mathbf{y}, m] \in \Psi} g(\mathbf{y}, m) \right] = \int_{\mathbb{R}^2} g(\mathbf{y}, m) \nu(d(\mathbf{y}, m)). \quad (1.3)$$

Důkaz: [4].

Definice 11. Momentová míra (funkce) bodového procesu

Pro bodový proces \mathbf{N} definujeme *momentovou míru k -tého řádu* předpisem

$$\mu_{\mathbf{N}}^{(k)}(B_1 \times \cdots \times B_k) = \mathbb{E} \{ \mathbf{N}(B_1) \cdots \mathbf{N}(B_k) \}, \quad B_1, \dots, B_k \in \mathcal{B}_0^2.$$

¹Pro $B \in \mathcal{B}^2$: $\mu(B) = 0 \Rightarrow \nu(B \times L) = 0$ pro jakoukoliv volbu $L \in \mathcal{M}$.

Definice 12. Faktoriální momentová míra (funkce) bodového procesu

Pro bodový proces N definujeme *faktoriální momentovou míru k -tého řádu* předpisem

$$\mu_N^{(k)!}(B_1 \times \cdots \times B_k) = \mathbb{E} \sum_{\xi_1, \dots, \xi_n \in N}^{\neq} \mathbf{1}_{\xi_1 \in B_1} \cdots \mathbf{1}_{\xi_n \in B_n},$$

kde $B_1, \dots, B_k \in \mathcal{B}_0^2$ a sčítá se přes všechny n -tice různých bodů N .

Je-li $\mu_N^{(k)!}$ lokálně konečná, a existuje-li hustota $\lambda_k(\cdot)$ vzhledem ke k -rozměrné Lebesgueově míře, potom $\lambda_k(\cdot)$ nazveme *funkcí intenzity k -tého řádu* bodového procesu N .

Speciálně: $\mu_N^{(1)!} = \mu$ z definice 9.

Definice 13. Poissonův bodový proces, homogenita procesu

Nechť μ je difúzní lokálně konečná míra na $(\mathbb{R}^2, \mathcal{B}_0^2)$. Bodový proces N splňující

- (i) $N(B)$ má Poissonovo rozdělení s parametrem $\mu(B)$ pro každé $B \in \mathcal{B}^2$,
- (ii) $N(B_1), \dots, N(B_n)$ jsou nezávislé pro každé $n \in \mathbb{N}$ a $B_1, \dots, B_n \in \mathcal{B}_0^2$ po dvou disjunktní,

nazveme *Poissonův bodový proces* s mírou intenzity μ . Existuje-li funkce intenzity, budeme ji značit λ . Poissonův bodový proces N nazýváme *homogenní* je-li λ konstantní.

Pozorování:

Homogenní Poissonův proces je stacionární.

Definice 14. Coxův (bodový) proces, rozdělení Coxova procesu

Nechť Λ je difúzní lokálně konečná náhodná míra na $\Omega \times \mathcal{B}^2$. Bodový proces N , který je podmíněně při $\Lambda = \mu$ Poissonův bodový proces s mírou intenzity μ nazveme *Coxův (bodový) proces*.

Dále, (viz. [4, str. 154]), je-li \mathbb{P}_Λ rozdělení Poissonova bodového procesu s mírou intenzity Λ , Q rozdělení Λ , pak *rozdělením Coxova procesu* N je pravděpodobnostní míra \mathbb{P}_N taková, že

$$\mathbb{P}_N(F) = \int \mathbb{P}_\Lambda(F) Q(d\Lambda), \quad \forall F \in \mathfrak{N}. \quad (1.4)$$

Poznámka

Λ se nazývá *řídící míra* Coxova procesu. Jeho míra intenzity je $\mathbb{E}\Lambda$. Existuje-li hustota λ řídící míry Λ vzhledem k Lebesgueově míře, nazývá se *řídící funkcí intenzity* Coxova procesu N . λ je pak náhodné pole na \mathbb{R}^2 .

Definice 15. Coxův bodový proces řízený náhodným polem

Nechť pro pevné $p \in \mathbb{N}$ je $\mathbb{X} = (\mathbf{X}_1(s), \mathbf{X}_2(s), \dots, \mathbf{X}_p(s))^T \in \mathbb{R}^p$; $s \in \mathbb{R}^2$ stacionární náhodné pole. Nechť $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ splňující:

$$\lambda(s) = f(\mathbf{B}^T \mathbb{X}(s))$$

kde f je měřitelná nezáporná funkce, \mathbf{B} reálná matice o rozměrech $p \times d$, $d \leq p$. Řekneme, že bodový proces s řídicí funkcí intenzity $\lambda(s)$ je *Coxův bodový proces řízený náhodným polem*.

Pozorování:

Coxův proces řízený stacionárním náhodným polem je stacionární.

Důkaz: Náhodné pole $\lambda(s) = f(\mathbf{B}^T \mathbb{X}(s))$ je stacionární jako transformace stacionárního náhodného pole. Dle definice 13. (i) je potom stacionarita λ postačující k stacionaritě Poissonova bodového procesu.

Důsledek 16. Ergodicita stacionárního bodového procesu

Definice ergodicity stacionárního bodového procesu je formálně zavedena v [4, str. 104]. Pro účely naší práce uvádíme důsledky ergodicity bodového a kótovaného bodového procesu uvedené rovněž v [4]:

Nechť \mathbf{N} je stacionární bodový proces s intenzitou $\rho > 0$ definovaný na \mathbb{R}^2 . Je-li \mathbf{N} ergodický, pak pro každou posloupnost $\{C_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}^2$ kompaktních konvexních množin splňujících $C_n \uparrow \mathbb{R}^2$ platí

$$\frac{\mathbf{N}(C_n)}{|C_n|} \rightarrow \rho \quad \text{v pravděpodobnosti.} \quad (1.5)$$

Nechť $\Psi = \{[s, m(s)]\}$ je stacionární kótovaný bodový proces takový, že kóta $m(s)$ je tvořena stacionárním náhodným polem $\mathbb{X}(s) \in \mathbb{R}^p$. Pak pro ergodický Ψ a C_n jako výše platí:

$$\frac{1}{|C_n|} \sum_{\substack{[s, \mathbb{X}(s)] \in \Psi \\ s \in C_n}} \mathbb{X}(s) \rightarrow \xi \quad \text{v pravděpodobnosti, } \xi \in \mathbb{R}^p. \quad (1.6)$$

1.3 Redukce dimenze

Nechť \mathbb{X} je p -rozměrné náhodné pole, \mathbf{N} Coxův bodový proces řízený náhodným polem \mathbb{X} . Buď \mathbf{B} je reálná matice o rozměrech $p \times d$, $d \leq p$ a $\mathcal{S}(\mathbf{B})$ lineární podprostor \mathbb{R}^p generovaný sloupci matice \mathbf{B} .

Definice 17. (Centrální) podprostor redukce dimenze

Řekněme, že $\mathcal{S}(\mathbf{B})$ je podprostor redukce dimenze, jestliže:

$$\forall_{k \in \mathbb{N}} \forall_{(C_1, \dots, C_k \in \mathcal{B}_0^2)} : \quad \{\mathbf{N}(C_1), \dots, \mathbf{N}(C_k)\} \perp\!\!\!\perp \{\mathbb{X}(C_1), \dots, \mathbb{X}(C_k)\} \mid \{\mathbf{B}^T \mathbb{X}(C_1), \dots, \mathbf{B}^T \mathbb{X}(C_k)\} \quad (1.7)$$

kde $\mathbb{X}(C) = \{(\mathbf{X}_1(s), \dots, \mathbf{X}_p(s))^T; s \in C\}$ a relace v (1.7) značí podmíněnou nezávislost² (*conditional independence*). Číslo d nazveme *dimenzí* podprostoru redukce dimenze.

Jinak řečeno, d -rozměrný vektor $\mathbf{B}^T\mathbb{X} = \{\mathbf{B}^T\mathbb{X}(s); s \in W\}$ obsahuje celou informaci z \mathbb{X} o \mathbf{N} . Poznamenejme, že takový prostor vždy existuje (např. \mathbb{R}^p).

Pro efektivní redukci dimenze hledáme však $\mathcal{S}(\mathbf{B})$ s nejmenší možnou dimenzí. Označme takový podprostor $\mathcal{S}_{\mathbf{N}|\mathbb{X}}$ a nazvěme ho *centrálním* podprostorem redukce dimenze. Stejně jako v [3] předpokládejme, že $\mathcal{S}_{\mathbf{N}|\mathbb{X}}$ existuje a má bázi $\mathbf{B}_0 \in \mathbb{R}^{p \times d_0}$, kde d_0 je dimenze $\mathcal{S}_{\mathbf{N}|\mathbb{X}}$.

Zaveďme pojem podprostoru redukce dimenze pro první momentovou míru bodového procesu a omezme se na případy, kdy $\mathcal{S}_{\mathbf{N}|\mathbb{X}}$ lze hledat jako podprostor redukce dimenze pro první momentovou míru.

Definice 18. (Centrální) podprostor redukce dimenze pro první momentovou míru

Řekněme, že $\mathcal{S}_{[1]}(\mathbf{B})$ je *podprostor redukce dimenze pro první momentovou míru*, jestliže

$$\lambda(s) = f(\mathbf{B}^T\mathbb{X}(s)), \quad (1.8)$$

pro nějakou nezápornou měřitelnou funkci $f(\cdot)$, kde $\lambda(\cdot)$ je řídicí funkce intenzity \mathbf{N} .

Jinak řečeno, d -rozměrný vektor $\mathbf{B}^T\mathbb{X} = \{\mathbf{B}^T\mathbb{X}(s); s \in W\}$ obsahuje celou informaci z \mathbb{X} o $\lambda(\cdot)$.

Konečně, označme

$$\mathcal{S}_1 := \bigcap_{\mathbf{B} \text{ splňující (1.8)}} \mathcal{S}_{[1]}(\mathbf{B})$$

Takový \mathcal{S}_1 nazveme *centrální podprostor redukce dimenze pro první momentovou míru*.

Tvrzení 1.

$$\mathcal{S}_1 \subseteq \mathcal{S}_{\mathbf{N}|\mathbb{X}}$$

Důkaz: $\mathcal{S}_{\mathbf{N}|\mathbb{X}}$ podle definice obsahuje veškerou informaci z \mathbb{X} o rozdělení \mathbf{N} . Zejména tedy informaci z \mathbb{X} o funkci intenzity λ bodového procesu \mathbf{N} . Tedy $\mathcal{S}_1 \subseteq \mathcal{S}_{\mathbf{N}|\mathbb{X}}$.

²Jevy A, B jsou podmíněně nezávislé za podmínky jevu Y tj. $A \perp\!\!\!\perp B|Y \Leftrightarrow \mathbb{P}(A \cap B|Y) = \mathbb{P}(A|Y)\mathbb{P}(B|Y)$

Tvrzení 2. (viz. [3, str. 371])

$$\mathcal{S}_{\mathbb{N}|\mathbb{X}} = \mathcal{S}_1,$$

pokud

$$\lambda_k(s_1, \dots, s_k) = \lambda_1(s_1) \dots \lambda_1(s_k) g_k(s_1, \dots, s_k), \quad (1.9)$$

kdykoliv $g_k(s_1, \dots, s_k)$ není pro $k \geq 2$ závislá na \mathbb{X} .

Všimněme si, že pro Coxův resp. Poissonův bodový proces je podmínka (1.9) z definice splněna.

1.4 Redukce dimenze pro Coxův bodový proces řízený náhodným polem

Nechť \mathbb{X} je p -rozměrné Gaussovské náhodné pole. Bez újmy na obecnosti předpokládejme, že \mathbb{X} je standardizované (tedy že $\mathbb{E}(\mathbb{X}) = \bar{0}$ a $Cov(\mathbb{X}) = \mathbf{I}_p$).

Buď \mathbf{N} Coxův bodový proces s řídicí funkcí intenzity

$$\lambda(s) = f(\mathbf{B}^T \mathbb{X}(s)), \quad s \in \mathbb{R}^2,$$

kde f je měřitelná nezáporná funkce a \mathbf{B} reálná matice o rozměrech $p \times d$, $d \leq p$. Pro realizaci \mathbf{N} v omezeném okně $W \subset \mathbb{R}^2$ položíme

$$\hat{\mathbf{B}}_{\text{SIR}} := \frac{1}{\mathbf{N}(W)} \sum_{s \in \mathbf{N} \cap W} \mathbb{X}(s). \quad (1.10)$$

Lemma o $\hat{\mathbf{B}}_{\text{SIR}}$.

Nechť \mathbf{N} a $\{\mathbb{X}(s); s \in \mathbf{N}\}$ jsou navíc ergodické, \mathbb{X} stacionární. Pak $\hat{\mathbf{B}}_{\text{SIR}}$ z (1.10) s $W = C_n$, $n \rightarrow \infty$, $\{C_n\}$ z důsledku 16. konverguje v pravděpodobnosti k \mathbf{B}_{SIR} , kde

$$\mathbf{B}_{\text{SIR}} = \frac{\mathbb{E}[\lambda(s)\mathbb{X}(s)]}{\mathbb{E}[\lambda(s)]}.$$

Důkaz: Z pozorování u definice 15. vidíme, že ze stacionarity \mathbb{X} plyne i stacionarita \mathbf{N} . Pro \mathbf{N} ergodické potom dle vzorce (1.5) máme

$$\frac{\mathbf{N}(C_n)}{|C_n|} \rightarrow \rho$$

pro posloupnost C_n z důsledku 16. Dále ze stacionarity \mathbf{N} máme $\rho = \mathbb{E}\lambda(U)$, $U = [0, 1] \times [0, 1]$. Ze stacionarity $\mathbb{X}(s)$ plyne současně i stacionarita náhodné posloupnosti $\{[s, \mathbb{X}(s)]; s \in \mathbf{N}\}$. Označme tuto posloupnost Ψ . Ψ je tedy kótovaný bodový

proces, který vznikl okótováním bodového procesu \mathbf{N} kótami $\mathbb{X}(s) \forall s \in \mathbf{N}$. Pro Ψ ergodický pak dle vzorce (1.6) platí

$$\frac{1}{|C_n|} \sum_{s \in \mathbf{N} \cap C_n} \mathbb{X}(s) \rightarrow \xi$$

pro libovolnou posloupnost C_n jako výše. Zde $\xi = \mathbb{E}[\lambda(s)\mathbb{X}(s)]$ ze stacionarity, neboť pro libovolnou omezenou $C \subset \mathbb{R}^2$ a kótovaný bodový proces Ψ máme:

$$\begin{aligned} \mathbb{E} \sum_{\substack{[s, \mathbb{X}(s)] \in \Psi \\ s \in C}} \mathbb{X}(s) &\stackrel{(i)}{=} \int \sum_{\substack{[s, \mathbb{X}(s)] \in \Psi \\ s \in C}} \mathbf{x}(s) \mathbb{P}_\Psi(d\Psi) \\ &\stackrel{(i)}{=} \int \int \sum_{\substack{[s, \mathbb{X}(s)] \in \Psi \\ s \in C}} \mathbf{x}(s) \mathbb{P}_\Lambda(d\Psi) \mathbb{Q}(d\Lambda) \\ &\stackrel{(ii)}{=} \int \int_C \int_{\mathbb{R}^p} \mathbf{x}(s) \Lambda(d(\mathbf{x}, s)) \mathbb{Q}(d\Lambda) \\ &\stackrel{(iii)}{=} \int_C \int \int_{\mathbb{R}^p} \mathbf{x}(s) \lambda(s) M_s(d\mathbf{x}) \mathbb{Q}(d\Lambda) ds \\ &\stackrel{(iv)}{=} \int_C \mathbb{E}[\mathbb{X}(s)\lambda(s)] ds = |C|\xi. \end{aligned}$$

Přičemž označené rovnosti zdůvodníme takto:

- (i) Ψ je Coxův kótovaný bodový proces (kóta v bodě s je $\mathbb{X}(s)$), \mathbb{P}_Ψ jeho rozdělení viz. (1.1) a (1.4).
- (ii) Použití Cambellovy věty pro kótovaný bodový proces, vzorec (1.3).
- (iii) Viz. (1.1).
- (iv) Po integraci přes \mathbf{x} již integrand nezávisí na Λ neboť $\lambda(s) = f(\mathbf{B}^T \mathbb{X}(s))$ a tedy λ je funkcí \mathbb{X} .

Dohromady, tedy za předpokladu ergodicity \mathbf{N} , $\{\mathbb{X}(s); s \in \mathbf{N}\}$ a stacionarity \mathbb{X} platí pro libovolnou posloupnost kompaktních konvexních $\{C_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}^2$ splňujících $C_n \uparrow \mathbb{R}^2$:

$$\frac{\sum_{s \in \mathbf{N} \cap C_n} \mathbb{X}(s)}{\mathbf{N}(C_n)} \rightarrow \frac{\xi}{\rho} \quad \text{v pravděpodobnosti.}^3$$

a

$$\frac{\xi}{\rho} = \frac{\mathbb{E}[\lambda(s)\mathbb{X}(s)]}{\mathbb{E}[\lambda(s)]}$$

což bylo tvrzení lemmatu.

³Podíl posloupností náhodných veličin konverguje k podílu jejich limit neboť ρ, ξ jsou konstanty, $\rho > 0$.

Věta 1. (viz. [3, str. 372])

Za předpokladů uvedených na začátku sekce 1.4 chceme dokázat, že

$$\mathcal{S}(\mathbf{B}_{\text{SIR}}) \subseteq \mathcal{S}_1.$$

Poznamenejme, že \mathbf{B}_{SIR} je vlastně odhadem $\mathbb{E}[\mathbb{X}(s); s \in \mathbb{N}]$.

Důkaz Věty 1:

Nechť \mathbf{B} je taková matice, že $\mathcal{S}(\mathbf{B}) = \mathcal{S}_1$ (neboli \mathbf{B} je báze \mathcal{S}_1). Ve smyslu Lemmatu o $\hat{\mathbf{B}}_{\text{SIR}}$ a definice 18. pak tedy stačí dokázat, že

$$\mathcal{S}(\mathbb{E}[\lambda(s)\mathbb{X}(s)]) = \mathcal{S}(\mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))\mathbb{X}(s)]) \subset \mathcal{S}_1.$$

Definujme projekční matice $\mathbf{P}_\mathbf{B} := \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ a $\mathbf{Q}_\mathbf{B} := \mathbf{I}_p - \mathbf{P}_\mathbf{B}$. Vidíme, že

$$\begin{aligned} \mathbf{P}_\mathbf{B}\mathbf{Q}_\mathbf{B} &= \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T(\mathbf{I}_p - \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T) = \\ &= \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T - \underbrace{\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T}_{=\mathbf{I}_p} = 0. \end{aligned}$$

\mathbb{X} je dle předpokladu Gaussovské neboli $\mathbb{X}(s) \sim N(\bar{0}, \mathbf{I}_p)$ z čehož plyne, že $\mathbf{Q}_\mathbf{B}\mathbb{X}(s)$ a $\mathbf{P}_\mathbf{B}\mathbb{X}(s)$ jsou nezávislé a tedy

$$\mathbb{E}[\mathbf{Q}_\mathbf{B}\mathbb{X}(s)|\mathbf{P}_\mathbf{B}\mathbb{X}(s)] = 0.$$

Dále

$$\begin{aligned} \mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))\mathbb{X}(s)] &= \mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))(\mathbf{P}_\mathbf{B} + \mathbf{Q}_\mathbf{B})\mathbb{X}(s)] = \\ &= \mathbf{P}_\mathbf{B}\mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))\mathbb{X}(s)] + \mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))\mathbb{X}(s)]\mathbb{E}[\mathbf{Q}_\mathbf{B}\mathbb{X}(s)|\mathbf{P}_\mathbf{B}\mathbb{X}(s)] = \\ &= \mathbf{P}_\mathbf{B}\mathbb{E}[f_1(\mathbf{B}^T\mathbb{X}(s))\mathbb{X}(s)]. \end{aligned}$$

Tedy $\mathcal{S}(\mathbf{B}_{\text{SIR}}) \subset \mathcal{S}(\mathbf{B}) = \mathcal{S}_1$ a důkaz je hotov.

1.5 SIR - Plátková inverzní regrese

Mějme realizaci bodového procesu závislého (nějakou funkcí intenzity) na realizaci náhodného pole. Jednou z metod hledání centrálního podprostoru redukce dimenze je tzv. "Plátková inverzní regrese" (nebo *SIR* z ang. Sliced inverse regression). Metodu popíšeme a na základě poznatků z předchozí sekce odůvodníme její funkčnost pro Coxův bodový proces řízený náhodným polem.

Popis metody SIR.

V popisu metody volně převzaté z [1] předpokládejme kontext předešlé sekce. Nechť tedy:

- (i) $W \subset \mathbb{R}^2$ je omezené okno pozorování, $|W| > 0$.
- (ii) $\tilde{\mathbb{X}}$ je realizace stacionárního Gaussovského p -rozměrného náhodného pole \mathbb{X} .
- (iii) \mathbf{N} je realizace bodového procesu pozorovaného v okně W , jehož funkce intenzity λ splňuje $\lambda(s) = f(\beta_i^T \mathbb{X}(s))$, kde $i = 1, \dots, p$, $k \leq p$, $s \in W$.

Metoda SIR hledá odhady $\hat{\beta}_i$ vektorů β_i následujícím algoritmem:

1. Označme $\Sigma_X = Cov(\tilde{\mathbb{X}})$. Standardizujme $\tilde{\mathbb{X}}$ jako

$$\mathbf{Z} = \Sigma_X^{-1/2} [\tilde{\mathbb{X}} - \mathbb{E}[\tilde{\mathbb{X}}]].$$

2. Přiřadíme každému bodu realizace \mathbf{N} nějakou kótu. Rozdělme realizaci \mathbf{N} do S disjunktních množin podle jejich kóty H_j ; $j = 1, \dots, S$, tak aby v každé H_j bylo zhruba stejně bodů realizace \mathbf{N} . Tyto množiny budeme dále nazývat *plátky*. Počet bodů v j -tém plátku označíme jako n_j .

3. Spočteme $\hat{m}_j = \frac{1}{n_j} \sum_{s \in H_j} \mathbf{Z}(s)$ výběrové průměry j -tého plátku.

4. Položíme

$$V = \sum_{j=1}^S n_j \hat{m}_j \hat{m}_j^T.$$

5. Provedeme analýzu hlavních komponent (viz např. [2]) matice V , tj. nalezneme vlastní čísla λ_l a vlastní vektory η_l , $l = 1, \dots, p$ matice V . Pak

$$\hat{\beta}_i = \Sigma_X^{-1/2} \eta_i,$$

kde η_i , $i = 1, \dots, k$ je k vlastních vektorů příslušných k největším vlastním číslům matice V .

1.6 Plátkování pomocí kót

Metoda SIR předpokládá rozdělení realizace bodového procesu na plátky pomocí kót jednotlivých bodů. Jaká veličina však bude onou kótou pro jednotlivé body realizace není bezprostředně zřejmé.

Při zachování předpokladů minulé sekce aplikujme nejprve metodu SIR omezeným způsobem: Standardizované náhodné pole \mathbb{X} rozdělme (jako v [5]) na dva plátky $\{\mathbb{X}(s); s \in \mathbf{N}\}$ a $\{\mathbb{X}(s); s \notin \mathbf{N}\}$. Protože $s \notin \mathbf{N}$ pro skoro všechna $s \in W$ a $\mathbb{E}[\mathbb{X}(s)] = \bar{0}$, zanedbáme $\mathbb{E}[\mathbb{X}(s); s \notin \mathbf{N}]$. Metoda SIR potom přímo dává odhad $\mathbb{E}[\mathbb{X}(s); s \in \mathbf{N}]$ jako

$$\hat{\mathbf{B}}_{\text{SIR}} = \frac{1}{\mathbf{N}(W)} \sum_{s \in \mathbf{N} \cap W} \mathbb{X}(s).$$

což přesně vede k postupu v sekci 1.4.

Kótou jednotlivých bodů zde je tedy přímo jejich indikátor. Mohli bychom však, podobně jako v [5], rozdělit realizaci \mathbf{N} na více plátek, dle hodnoty $\hat{\beta}_1^T \mathbb{X}$ v bodech realizace \mathbf{N} . Označíme tento způsob jako **plátkování \mathbf{A}** . Jiný, námi navrhovaný přístup, nazveme **plátkování \mathbf{B}** : Zde kótou rozumíme vzdálenost od jiného nejbližšího bodu procesu.

1.7 Způsoby měření kvality nalezených směrů redukce dimenze

V praktické části otestujeme všechny tři navrhované postupy. Cílem je nalezení takových $\hat{\beta}_i$, aby se jimi generovaný podprostor co nejvíce přiblížil skutečnému podprostoru redukce dimenze. Jejich kvalitu měříme dvěma způsoby a to:

(i) spočtením korelace R^2 vektorů získaného $\hat{\beta}$ a předpokládaného β_0 :

$$R^2(\hat{\beta}, \beta_0) = \frac{(\hat{\beta}^T \beta_0)^2}{\hat{\beta}^T \hat{\beta} \beta_0^T \beta_0}$$

(ii) spočtením jejich vzdálenosti:

$$\Delta(\hat{\beta}, \beta_0) = \left\| \beta_0 (\beta_0^T \beta_0)^{-1} \beta_0^T - \hat{\beta} (\hat{\beta}^T \hat{\beta})^{-1} \hat{\beta}^T \right\|_{\max}$$

kde $\|\mathbf{A}\|_{\max}$ značí největší absolutní hodnotu ze složek matice \mathbf{A} .

Kapitola 2

Praktická část

Zvolíme $W = [0, 1]^2$. V programu R budeme příkazem *GaussRF* z knihovny *RandomFields* simulovat čtyřrozměrná Gaussovská stacionární náhodná pole $\mathbb{X}(s) \in \mathbb{R}^4$; $s \in W$ s kovarianční funkcí $C(x) = \exp(-x^\beta)$, $\beta \in (0, 2]$ (standardní parametry s modelem¹ *stable* viz. příloha).

Pro dva modely (řídící funkce intenzity λ) vygenerujeme 200 opakování realizace náhodného pole a závislého bodového procesu. Odhadneme popsányi metodami vektory redukce dimenze a dvěma způsoby budeme měřit, zda získané vektory efektivně redukuje dimenzi náhodného pole pro zvolený model.

2.1 Volba modelu

Generujeme Coxův proces pro tyto řídící funkce intenzity prvního řádu:

Model I. $\lambda(s) = \alpha \exp\{\mathbf{X}_1(s)\}$

Model II. $\lambda(s) = \alpha \exp\{\mathbf{X}_1(s) + \mathbf{X}_2(s)\}$

Konstanta $\alpha > 0$ je pokaždé zvolena tak, aby počet bodů simulovaného procesu v okně W byl vhodný ke zkoumání. V obou případech funkce λ splňují předpoklady definice 15. a podmínku (1.9) z Tvzení 2. Platí tedy, že hledaná $\mathcal{S}_{\mathbb{N}^{\mathbb{X}}} = \mathcal{S}_1$.

¹Zde a jen zde je slovo model použito ve smyslu parametru příkazu *GaussRF* v programu R.

2.2 Popis konstrukce simulace

Teoretická část práce popisuje konstrukci náhodného pole a metody SIR velmi obecným způsobem. Pro praktickou část jsme zvolili jisté omezující podmínky, které se drží v teoretickém rámci práce. Je tedy vhodné detailně popsat vlastní konstrukci simulace. Následující část shrnuje postup, který se dá následovat v kódu pro program R v příloze práce. Práce si neklade za cíl časově nebo jinak optimalizovat výpočty. Kód je tedy psán a komentován tak, aby jasně následoval postup výpočtu, někdy za cenu vysoké náročnosti výpočtů.

2.2.1 Náhodné pole

Náhodné pole \mathbb{X} je definované pro všechny hodnoty uvnitř okna W . Pro účely simulace musíme přistoupit k „diskretizaci“ hodnot náhodného pole. Knihovna *RandomFields* v programu R generuje hodnoty náhodného pole na čtvercové mříži. Pro rozměr jednotkového čtverce (naše W) jsme pro simulaci zvolili mříž obsahující 100×100 bodů. Symbol $\mathbb{X}(s)$; $s \in W$ tedy v praktické části znamená „hodnota náhodného pole v nejbližším bodě jeho generované mříže“. Stejně tak střední hodnoty náhodného pole jsou počítány jako střední hodnoty přes všechny body jeho mříže.

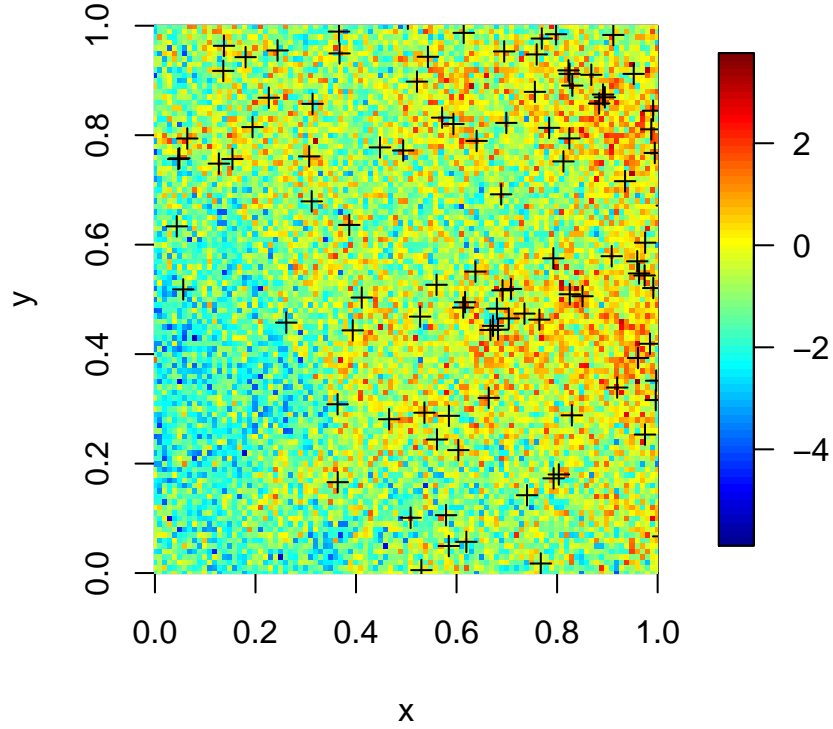
2.2.2 Bodový proces

Bodový proces N je generován funkcí *rpoispp* z knihovny *fields* s použitím funkce λ odpovídající zvolenému modelu se stejným přístupem diskretizace v okně W jako tomu je u náhodného pole. Příklad první složky náhodného pole a jím řízeného bodového procesu (**Model I.**) vidíme na obrázku 2.1.

2.2.3 SIR

Metodu SIR používáme ve shodě se sekci 1.5. Plátky vytváříme nejdříve způsobem popsaným v sekci 1.4, dále dvěma způsoby popsanými v sekci 1.6, tedy **plátkováním A** a **plátkováním B**. V druhém případě se však musíme vyrovnat s tzv. okrajovým efektem, ke kterému při výpočtu kóty bodu dochází, a který by výsledky nepřirozeně zkresloval: totiž u bodů blízko hrany okna pozorování bychom vzdálenost od nejbližšího souseda uvažovali pouze pro ty v okně, i když těsně za hranou okna pozorování by mohl být soused bližší. Tomuto nedostatku se vyhneme simulací většího okna než je okno pozorování. Příslušný bodový proces a kótu počítáme i z bodů za hranicí okna pozorování. Vzhledem k tomu, že hledáme vždy nejbližšího souseda stačí, při dostatečném počtu bodů (zajištěného naší volbou α), volba velikosti okraje jako poloviny okna pozorování.

Dále již postupujeme přesně dle algoritmu popsaném v sekci 1.5 s volbou 2 a 4



Obrázek 2.1: Příklad bodového procesu řízeného náhodným polem

vrstev, které obsahují vždy stejný počet prvků (až na jeden). Zajímá nás směr odpovídající největšímu vlastnímu číslu, neboli $\hat{\beta}_1$. Ten porovnááme s předpokládaným směrem redukce dimenze v závislosti na zvoleném modelu (řídící funkci intenzity). V případě **Modelu I.** máme $\lambda(s)$ závislou na první souřadnici náhodného pole \mathbb{X} , nezávislou na ostatních, tedy předpokládaná $\beta_0 = (1, 0, 0, 0)^T$. U **Modelu II.** předpokládáme $\beta_0 = (1, 1, 0, 0)^T$.

2.3 Výsledky

2.3.1 Model I.

Z ukazatelů $R^2(\hat{\beta}_1, \beta_0)$ a $\Delta(\hat{\beta}_1, \beta_0)$ uvádíme v tabulce výsledků jejich průměr (směrodatnou odchylku) po 200 opakování simulace náhodného pole a příslušného bodového procesu. V příloze práce uvádíme histogramy ukazatelů.

	$R^2(\hat{\beta}_1, \beta_0)$	$\Delta(\hat{\beta}_1, \beta_0)$
1 plátek:	0.9840 (0.02210689)	0.0614368 (0.05003994)
plátkování A:		
2 plátky:	0.9797 (0.02467742)	0.071175 (0.05732581)
4 plátky:	0.9785 (0.02686904)	0.0739883 (0.05951091)
plátkování B:		
2 plátky:	0.9839 (0.0216527)	0.0612799 (0.05031933)
4 plátky:	0.9733 (0.0351632)	0.0785618 (0.06027275)

Tabulka 2.1: Výsledky simulace - Model I.

2.3.2 Model II.

	$R^2(\hat{\beta}_1, \beta_0)$	$\Delta(\hat{\beta}_1, \beta_0)$
1 plátek:	0.9864 (0.02666433)	0.055306 (0.04406592)
plátkování A:		
2 plátky:	0.9813 (0.03413159)	0.0655523 (0.05039011)
4 plátky:	0.9793 (0.03645645)	0.07015 (0.05303043)
plátkování B:		
2 plátky:	0.9850 (0.03817084)	0.057139 (0.04806754)
4 plátky:	0.9822 (0.02567831)	0.064170 (0.04701532)

Tabulka 2.2: Výsledky simulace - Model II.

Závěr

V práci byly zavedeny tři způsoby aplikace metody plátkové inverzní regrese pro určení centrálního podprostoru redukce dimenze pro Coxův bodový proces řízený vícerozměrným Gaussovským náhodným polem v okně pozorování. Pro dva modely bylo v prostředí programu R vygenerováno 200 opakování náhodného pole a závislého bodového procesu. Pro dva různé modely byly dvěma způsoby porovnány předpokládané a nalezené směry redukce dimenze.

Z výsledků můžeme učinit závěry v několika směrech:

Posouzením průměrů výsledků můžeme s jistotou říci, že navržená metoda hledání centrálního podprostoru redukce dimenze funguje spolehlivě. Ve všech případech se nalezené směry dobře shodují se směry předpokládanými, nezávisle na způsobu měření. Praktická část je tedy ve shodě s teoretickou částí práce.

Srovnáme-li spolu způsoby měření vidíme, že oba způsoby se od sebe liší jen málo. Průměry jsou ve všech případech vzájemně v nepřímé úměře: čím větší je jejich korelace, tím menší je jejich vzdálenost. Porovnáním směrodatných odchylek způsobů měření můžeme usoudit, že způsob měření korelací předpokládaného a získaného směru je metoda vhodnější.

Srovnáme-li výsledky obou testovaných modelů, není patrná žádná konkrétní závislost. Nezdá se tedy, že by metoda byla citlivá k tomu, zda je bodový proces závislý pouze na jedné složce, nebo na dvou složkách náhodného pole. Jinými slovy, pro zvolené modely fungují všechny naše způsoby aplikace stejně dobře.

Konečně srovnáním jednotlivých způsobů plátkování dostáváme v případech obou modelů zhruba stejnou situaci: Základní způsob bez použití více plátek dosahuje (až na jednu výjimku) vždy nejlepšího výsledku. S větším množstvím plátek se nezávisle na způsobu plátkování výsledek téměř vždy (opět až na jednu výjimku) nezlepší. Dále vidíme, že způsob plátkování B je ve většině případů lepší než způsob plátkování A.

Závěrem práce je tedy tvrzení, že metoda plátkové inverzní regrese je pro Coxův bodový proces závislý na Gaussovském stacionárním náhodném poli efektivním způsobem hledání centrálního podprostoru redukce dimenze. Nově navržený způsob plátkování B se zdá oproti způsobu A použitého v [3] mírně lepší. Zlepšení však není značné a použití pouze jednoho plátku dopadá ve všech případech (modelech, způsobech měření i větším počtu plátek) na námi testovaných datech

lépe.

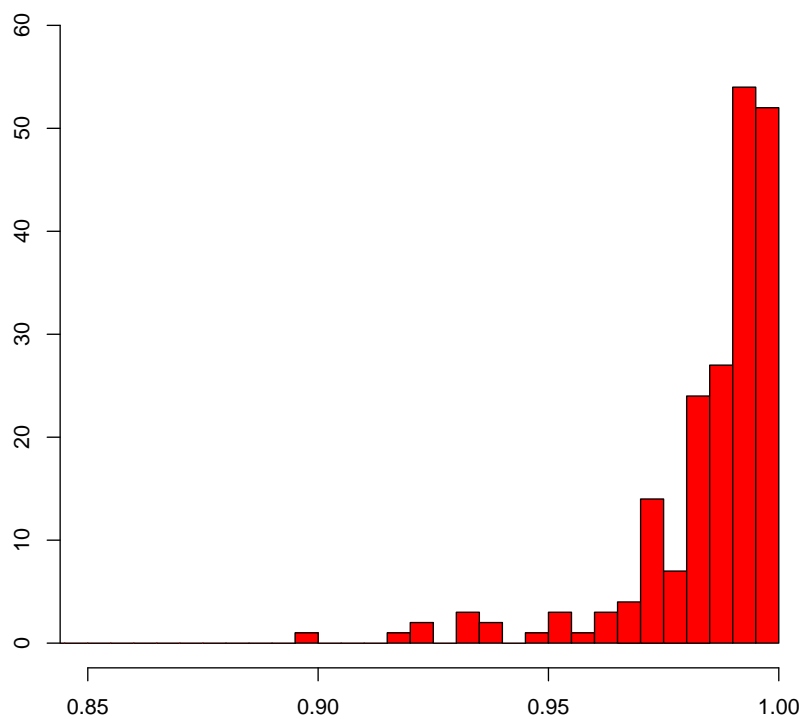
Nelze nicméně vyloučit, že druhá část závěru by se s jinak zvolenými modely, zvětšením rozsahu testování, nebo případnou aplikací na reálná data, nezměnila. Poznamenejme ještě, že změna modelu, zvětšení rozsahu, nebo i jiný způsob vkládání testovaných dat, by se dala jednoduše implementovat do přiloženého zdrojového kódu vytvořeného programu.

Seznam použité literatury

- [1] HÄRDLE, SIMAR. *Applied Multivariate Statistical Analysis*. 3. vydání. Springer, 2012. ISBN 978-3-642-17228-1.
- [2] ANDĚL, Jiří. *Matematická statistika*. 2. vydání. SNTL, 1985.
- [3] GUAN, Yongtao; WANG, Hansheng. *Sufficient dimension reduction for spatial point processes directed by Gaussian random field*. J. Royal Statist. Soc. B (2010). 72, Part 3, 367-387.
- [4] STOYAN, Dietrich; KENDALL, Wilfrid S.; MECKE, Joseph. *Stochastic Geometry and its Applications*. Second Edition. John Wiley & Sons Ltd, 1995. ISBN 0-471-95099-8.
- [5] GUAN, Yongtao. *On Consistent Nonparametric Intensity Estimation for Inhomogenous Spatial Point Processes*. Journal of the American Statistical Association, September 2008, Vol. 103, No. 483, 1239-1247.

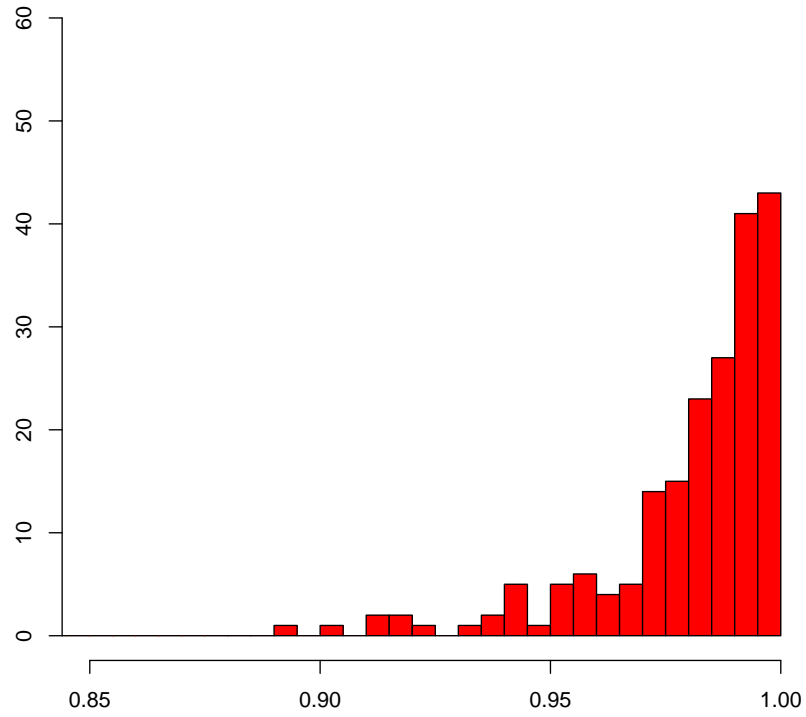
Přílohy

Příloha A - Histogramy modelu I.

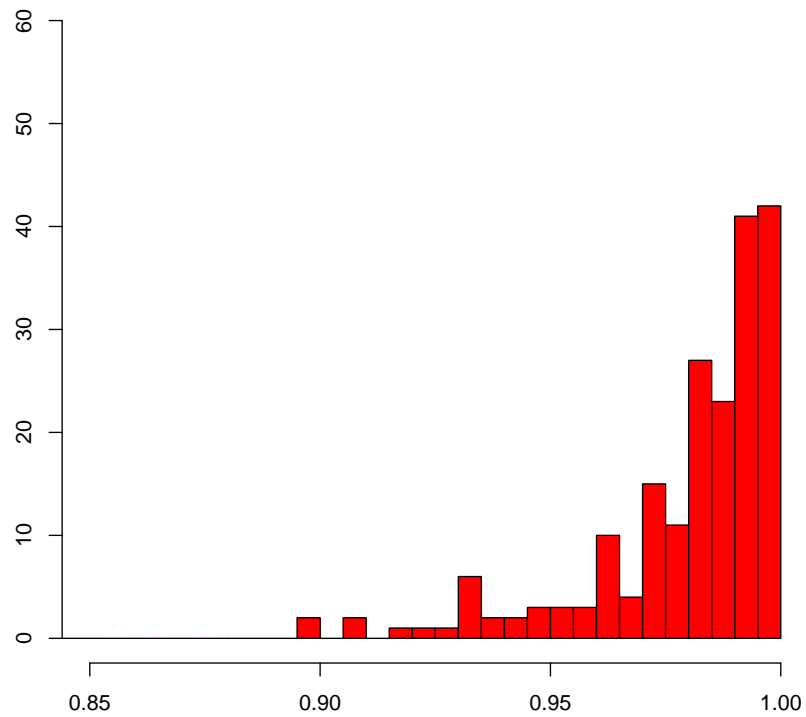


Obrázek 2.2: Histogram $R^2(\hat{\beta}_1, \beta_0)$, 1 plátek

Příloha A - Histogramy modelu I.

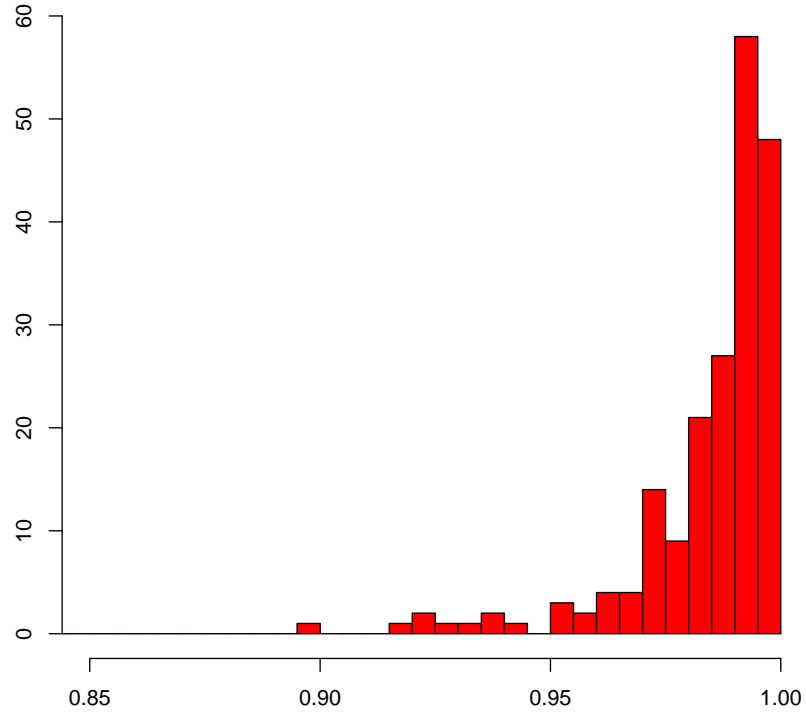


Obrázek 2.3: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování A, 2 plátky

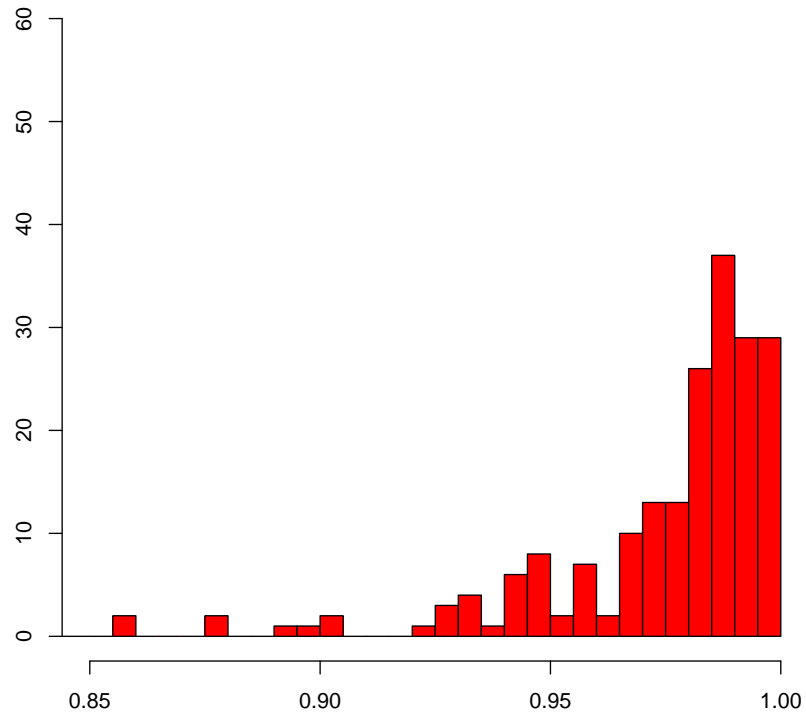


Obrázek 2.4: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování A, 4 plátky

Příloha A - Histogramy modelu I.

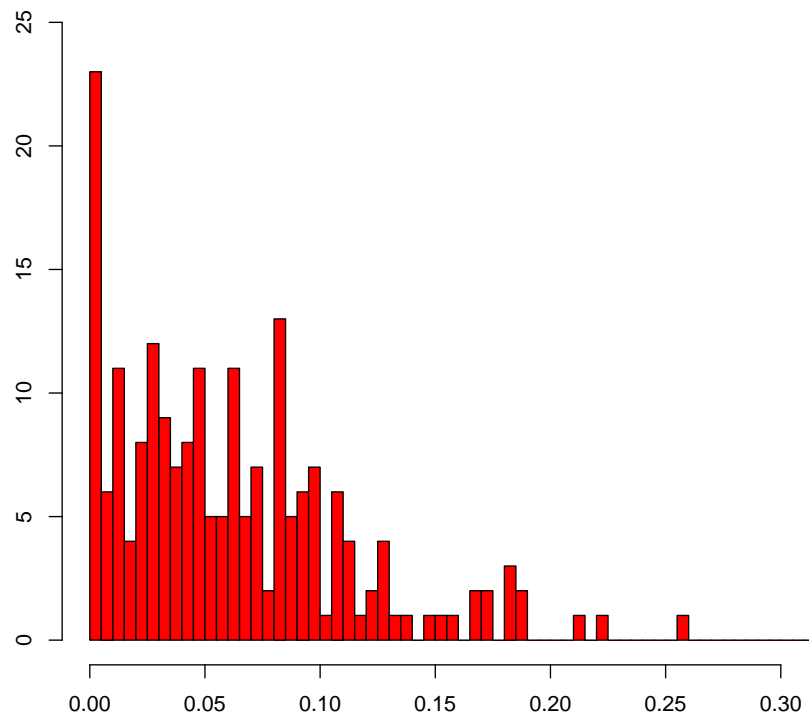


Obrázek 2.5: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování B, 2 plátky



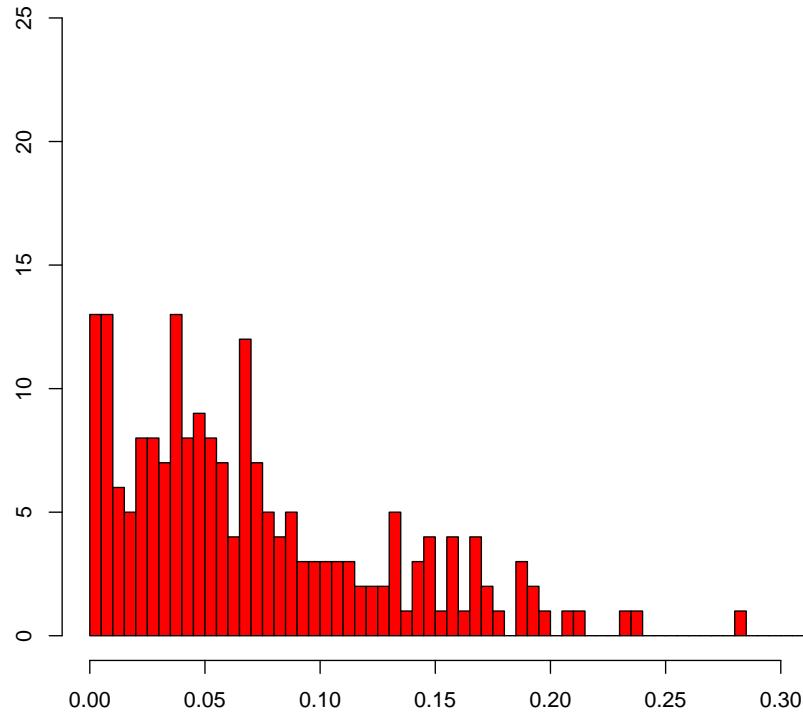
Obrázek 2.6: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování B, 4 plátky

Příloha A - Histogramy modelu I.

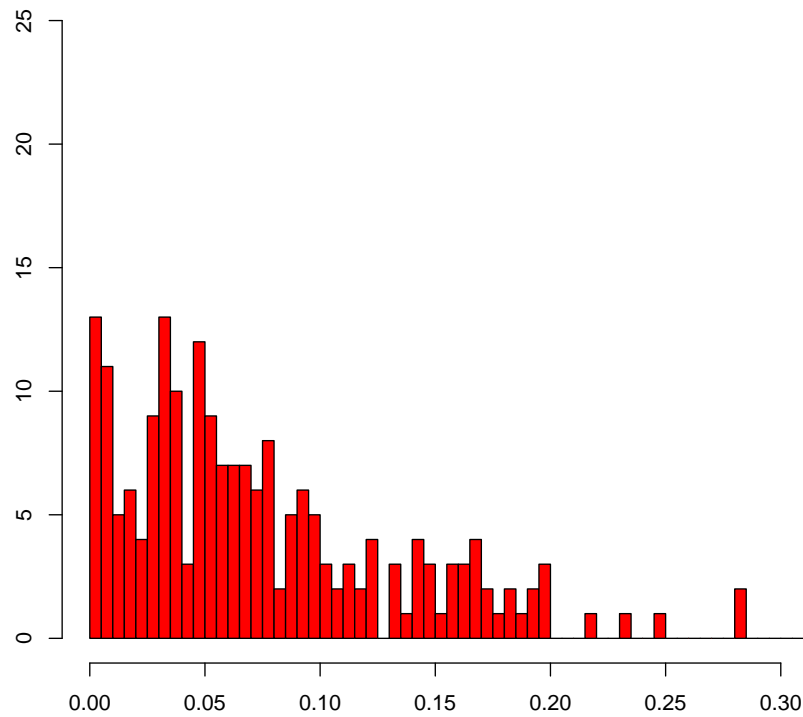


Obrázek 2.7: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, 1 plátek

Příloha A - Histogramy modelu I.

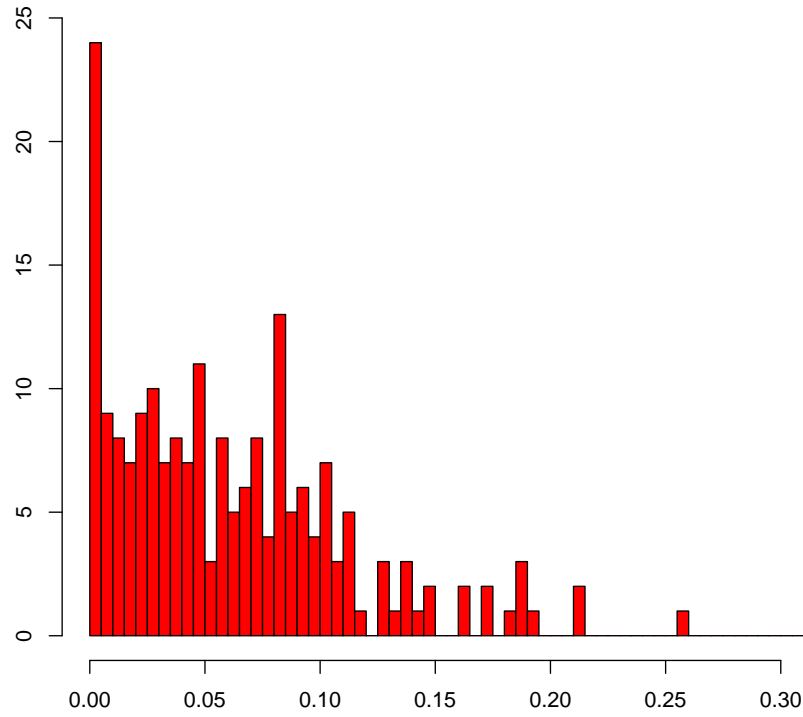


Obrázek 2.8: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování A, 2 plátky

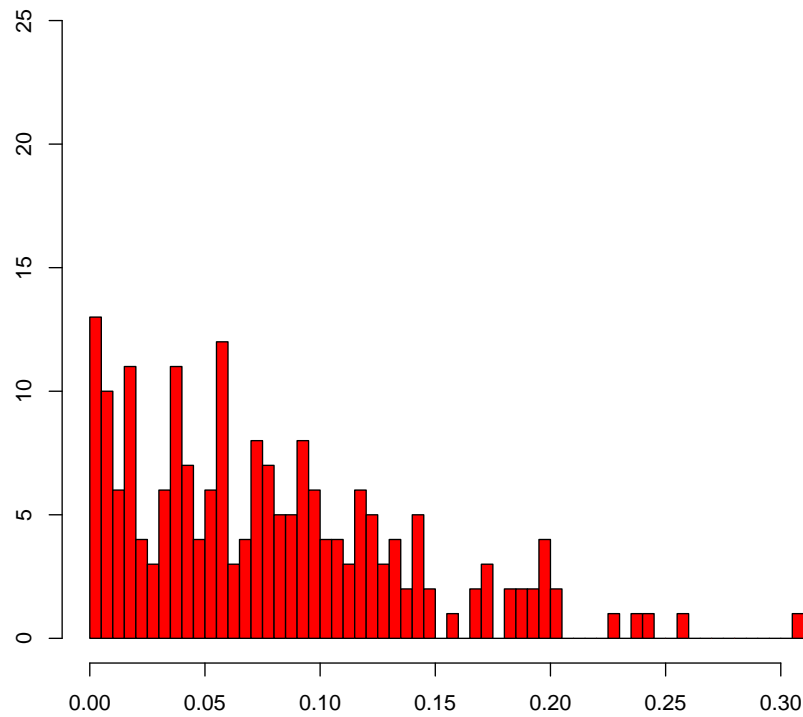


Obrázek 2.9: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování A, 4 plátky

Příloha A - Histogramy modelu I.

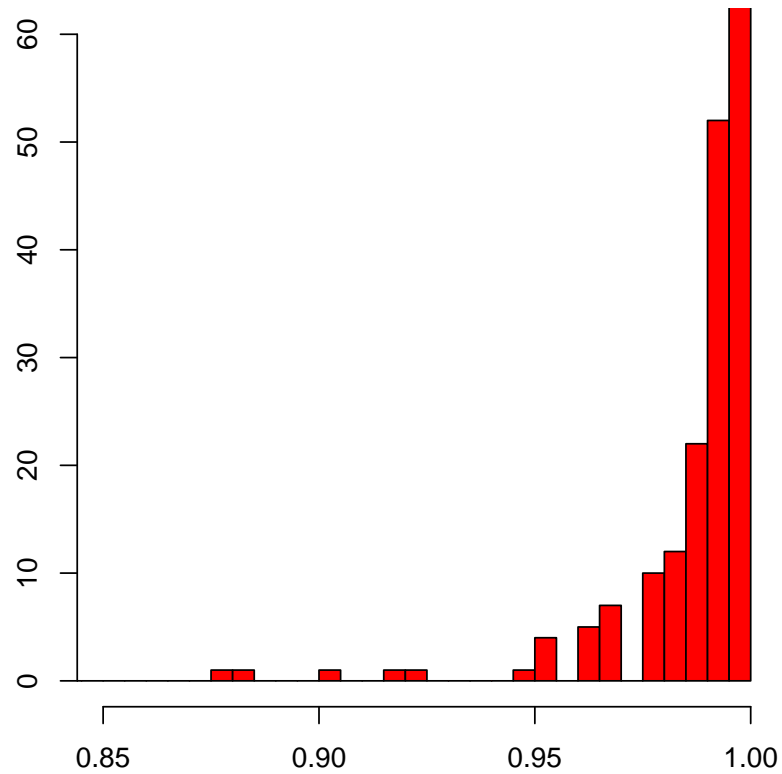


Obrázek 2.10: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování B, 2 plátky

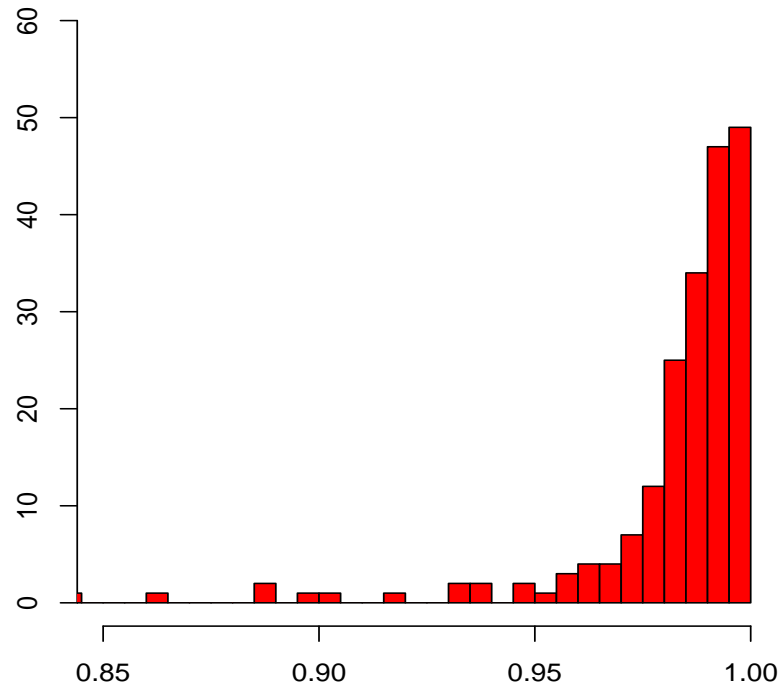


Obrázek 2.11: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování B, 4 plátky

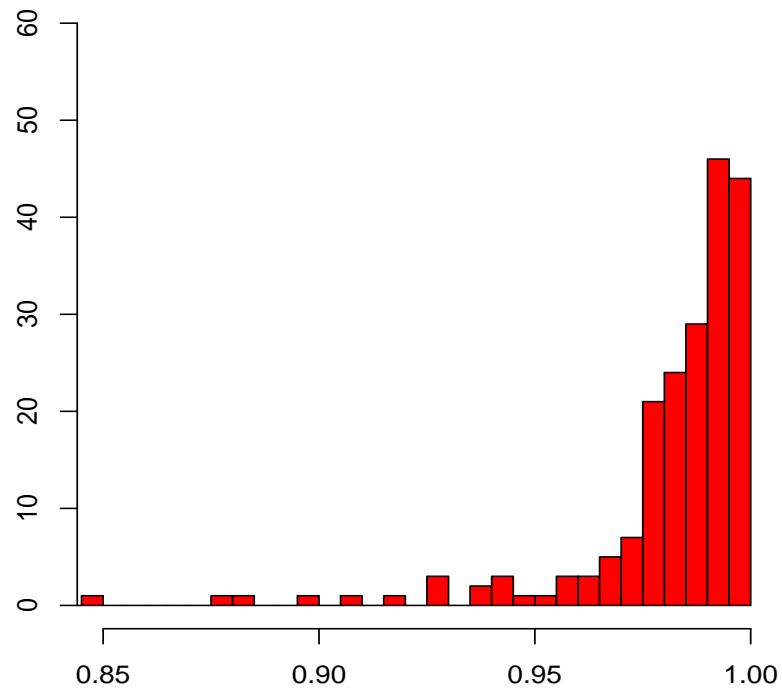
Příloha B - Histogramy modelu II.



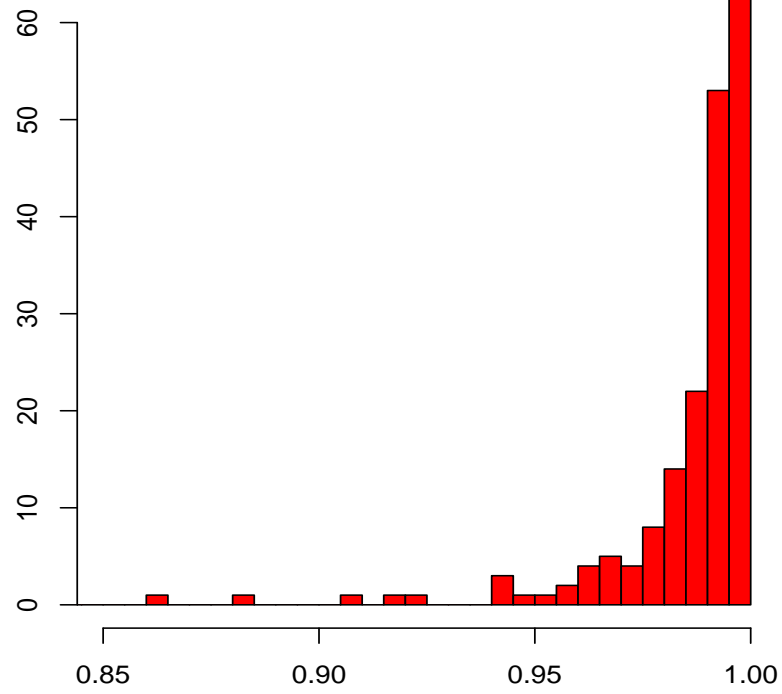
Obrázek 2.12: Histogram $R^2(\hat{\beta}_1, \beta_0)$, 1 plátek



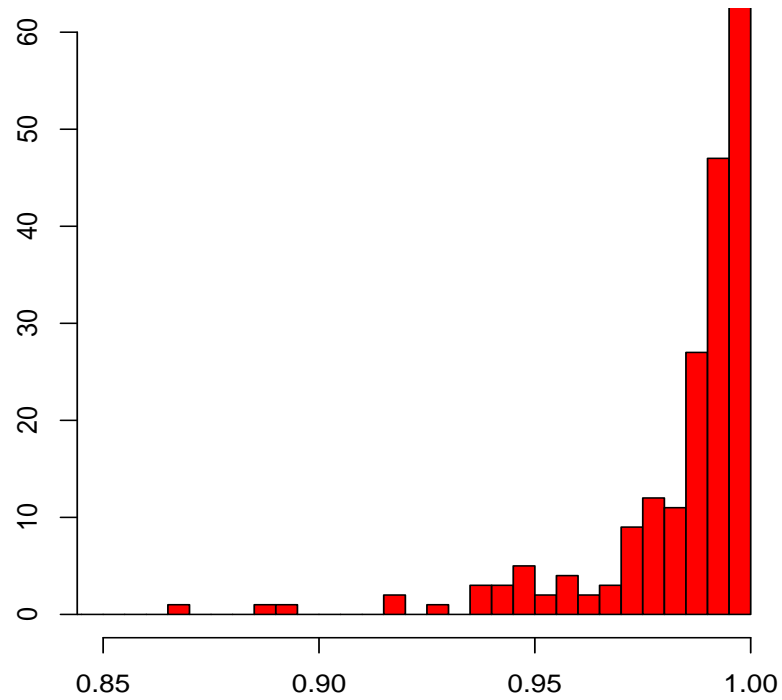
Obrázek 2.13: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování A, 2 plátky



Obrázek 2.14: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování A, 4 plátky

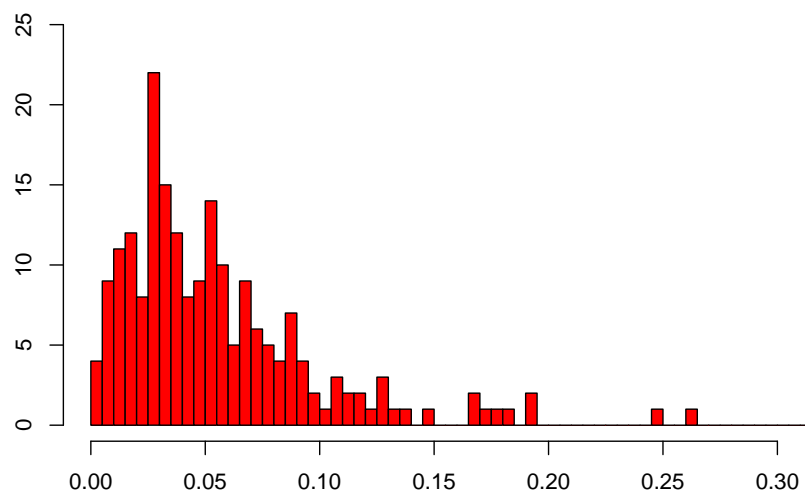


Obrázek 2.15: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování B, 2 plátky



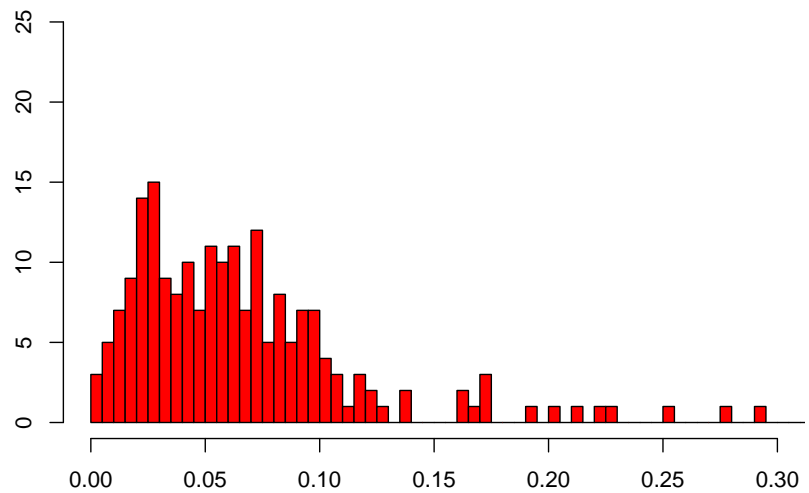
Obrázek 2.16: Histogram $R^2(\hat{\beta}_1, \beta_0)$, plátkování B, 4 plátky

Příloha B - Histogramy modelu II.

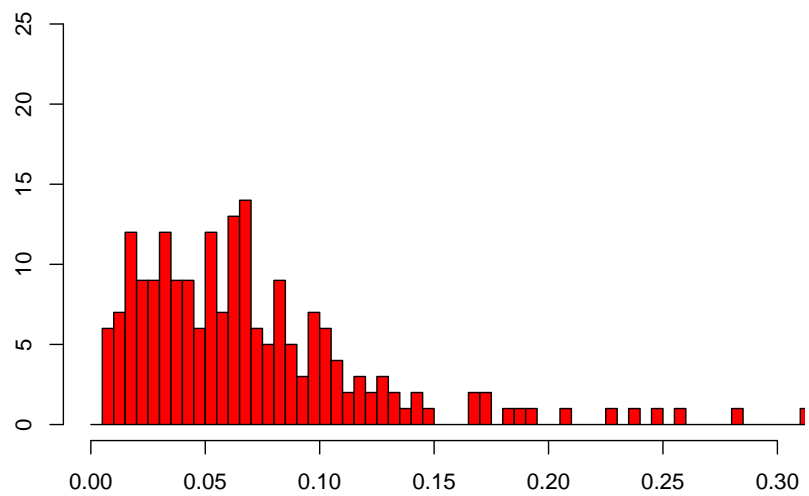


Obrázek 2.17: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, 1 plátek

Příloha B - Histogramy modelu II.

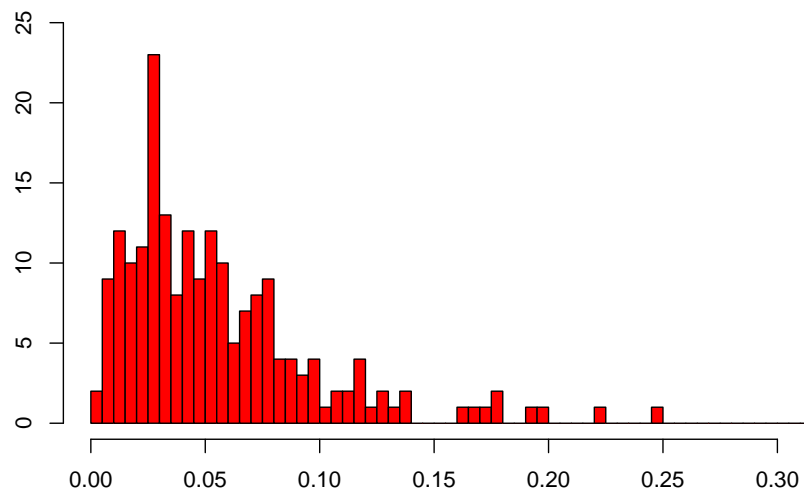


Obrázek 2.18: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování A, 2 plátky

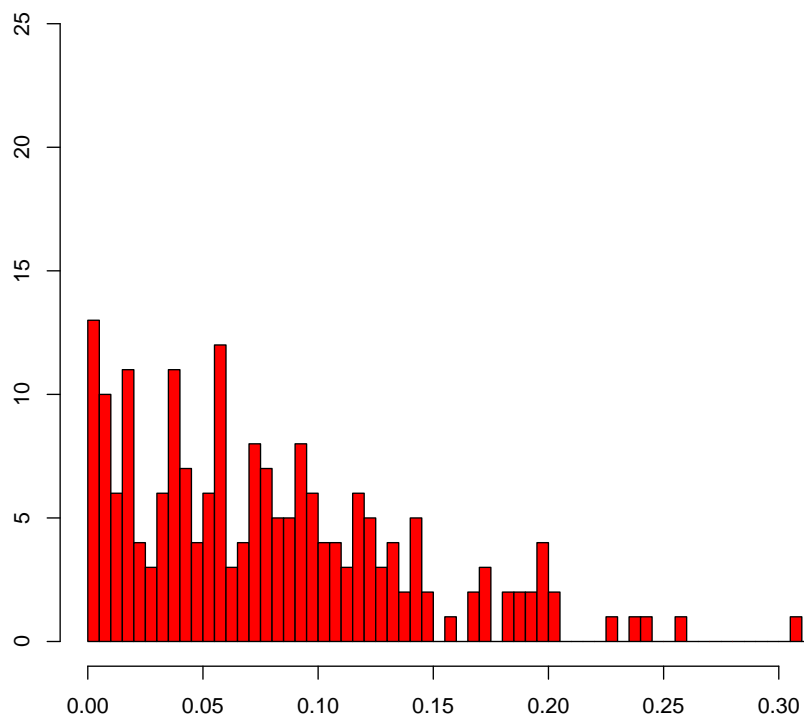


Obrázek 2.19: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování A, 4 plátky

Příloha B - Histogramy modelu II.



Obrázek 2.20: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování B, 2 plátky



Obrázek 2.21: Histogram $\Delta(\hat{\beta}_1, \beta_0)$, plátkování B, 4 plátky

Příloha C zdrojový kód pro aplikaci v programu R

```
library(RandomFields)
library(MASS)
library(spatstat)
library(fields)

model <- "stable"
mean <- 0
variance <- 1
nugget <- 1
scale <- 1
alpha <- 1
step <- 0.01
realizace<-4
x <- seq(0, 2, step)
y <- seq(0, 2, step)
N<-length(x)
SIGMA<-matrix(0,realizace,realizace)
beta1<-matrix(0,realizace,1)
beta1[1]<-1
vysledek<-array(0,c(200,34))
vysledekbeta<-matrix(0,4,200)
window<-owin(c(0,2),c(0,2))

# Funkce korelace (R^2):
korelace<-function(u,v){
kor<-((t(u)%*%v)^2)/(t(u)%*%u)%*%t(v)%*%v)
return(kor)}

# Funkce vzdalenost vektorů:
vzdalenost<-function(u,v){
dist<-max((u %*% ((t(u) %*% u)^(-1)) %*% t(u))
- (v %*% ((t(v) %*% v)^(-1)) %*% t(v)))
return(dist)}

# L - funkce vracející (transformovanou) hodnotu pole v nejbližším bodě mříže
L<-function(x,y){
vec<-numeric(length(x))
for (i in 1:length(x)){
vec[i]<-Lambda[round(x[i]/step)+1,round(y[i]/step)+1]
}
return(vec)
}

# funkce kótování bodového procesu
kota<-function(i,B){
vzdalenostodnejblizsiho<-4
```

```

for (j in 1:PP$n) {
  if ((PP$x[j]!=B[1,i]) & (PP$y[j]!=B[2,i])) {
    if (vzdalenostodnejblizsiho > sqrt((PP$x[j]-B[1,i])^2+(PP$y[j]-B[2,i])^2)) {
      vzdalenostodnejblizsiho <- sqrt((PP$x[j]-B[1,i])^2+(PP$y[j]-B[2,i])^2)
    }
  }
}
return(vzdalenostodnejblizsiho)
}

```

Vlastní smyčka generující 200 krát náhodné pole a závislý bodový
 #### proces, provede celý průběh programu (zde pro Model I.) a výsledky
 #### ukládá do promenne s nazvem "vysledek"

```

for (l in 1:200){

  # deklarace náh. pole X, zde X[, ,i] i-tá realizace pole:
  X <- GaussRF(n=realizace, x=x, y=y, model=model, grid=TRUE,
  param=c(mean, variance, nugget, scale, alpha))

  # Lambda - transformovaná funkce intenzity závislá na první složce pole:
  a<-100
  Lambda <- a*exp(X[, ,1])

  # Bodový proces závislý na fci intenzity L
  PP<-rpoispp(L, win=window)

  # Transformace do "okna"
  Z <- array(0, c(101,101,4))
  for (i in 1:101) {
    for (j in 1:101) {
      Z[i,j,]<-X[i+50,j+50,]
    }
  }

  #Y - normované pole:
  Y <- array(0, c(101,101,4))
  for(k in 1:realizace) {
    Y[, ,k]<-(Z[, ,k]-mean(Z[, ,k]))}

  # Nasčítání: SIGMA matice p*p přes jednotlivé body normovaného
  pole a zprůměrování
  for(i in 1:101)
    for(j in 1:101) {
      D<-Y[i,j,] %*% t(Y[i,j,])
      SIGMA<-SIGMA+D}
  SIGMA<-SIGMA * 101^(-2)
}

```

```

# SIGMAodm: Odmocnina z matice SIGMA
E<-eigen(SIGMA)
D<-matrix(0,realizace,realizace)
D<-diag(E$values)
C<-matrix(0,realizace,realizace)
C[,]<-E$vectors
for (i in 1:realizace) D[i,i]<-sqrt(D[i,i])
SIGMAodm<- C %*% D %*% ginv(C)

# SIGMAodminv: SIGMA(-1/2)
SIGMAodminv <- ginv(SIGMAodm)

# Zvlnka: Z s vlnkou (normované pole jako v Guan 2008)
Zvlnka <- array(0, c(101,101,4))
for(i in 1:101)
  for(j in 1:101){
Zvlnka[i,j,] <- SIGMAodminv %*% Y[i,j,]}

# převedení bodového procesu pro pohodlnější zacházení:
# struktura Body (pouze ty v okně):
# souřadnice 1: poloha na ose x
# souřadnice 2: poloha na ose y
# souřadnice 3: kóta bodu
# souřadnice 4: vrstva bodu podle hodnoty pole v bodě
# souřadnice 5: vrstva bodu podle kóty
# souřadnice 6: hodnota V

Body <- array(0,c(6,TRUE))
for (i in 1:PP$n) {
if (0.5 <= PP$x[i] & PP$x[i] <= 1.5 & 0.5 <= PP$y[i] & PP$y[i] <= 1.5)
Body <- cbind(Body, c(PP$x[i],PP$y[i],NA,NA,NA,NA,NA))
}
Body<-Body[,-1]
nBody<-dim(Body)[2]
Body[1,]<-Body[1,]-0.5
Body[2,]<-Body[2,]-0.5

# m vektor, m[i] pro i-tou realizaci pole průměr hodnot pole
v bodech procesu (v nejbližším bodě mříže)
m<-matrix(0,1,realizace)
for (i in 1:realizace) {
sumapole<-0
for (j in 1:nBody) {
sumapole<-sumapole+Zvlnka[round(Body[1,j]/step)+1,
round(Body[2,j]/step+1,i)]
}
m[i]<-sumapole/nBody}

```

```

# beta: inverze odmocněné matice sigma krát transponované m
beta<-SIGMAodminv %*% t(m)

vysledek[1,1]<-korelace(beta,beta1)
vysledek[1,34]<-vzdalenost(beta,beta1)

# V
for(i in 1:nBody) {
Body[6,i] <- t(beta) %*% Zvlnka[round(Body[1,i]/step)+1,round(Body[2,i]/step)+1,]
}

# kótování bodového procesu pro všechny body (nejen ty v okně)
for (m in 1:nBody) Body[3,m]<-kota(m,Body)

#2 plátky:

# rozdělení Bodů do dvou plátek podle hodnoty pole V v bodě
medV <- median(Body[6,])
for(i in 1:nBody) {
if (Body[6,i]>medV)
Body[4,i]<-1
else
Body[4,i]<-2
}

# rozdělení Bodů do dvou plátek podle kóty
medVk<-median(Body[3,])
for(i in 1:nBody) {
if (Body[3,i]>medVk)
Body[5,i]<-1
else
Body[5,i]<-2
}

# m1,m2 vektory, mj[i] pro i-tou realizaci pole průměr hodnot
pole v bodech procesu v j-tém plátku (v nejbližším bodě mříže)
m1<-array(0,c(realizace,1))
m2<-array(0,c(realizace,1))
for (i in 1:realizace) {
sumaplatku1<-0
sumaplatku2<-0
n1<-0
n2<-0
for (j in 1:nBody) {
if (Body[4,j]==1)
{sumaplatku1 <- sumaplatku1 +

```

```

    Zvlnka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n1<-n1+1 }
  if (Body[4,j]==2)
    {sumaplatku2 <- sumaplatku2 +
      Zvlnka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n2<-n2+1 }
  }
  m1[i]<-sumaplatku1/n1
  m2[i]<-sumaplatku2/n2
}

# A jako v (iv) pro dva plátky
A1 <- n1* m1 %*% t(m1)
A2 <- n2* m2 %*% t(m2)
A <- A1 + A2

Avlastnihodnoty <- eigen(A)

beta2p1<-SIGMAodminv %*% Avlastnihodnoty$vector[,1]
beta2p2<-SIGMAodminv %*% Avlastnihodnoty$vector[,2]
beta2p3<-SIGMAodminv %*% Avlastnihodnoty$vector[,3]
beta2p4<-SIGMAodminv %*% Avlastnihodnoty$vector[,4]

# zápis výsledků verze (A)
vysledek[1,2]<-korelace(beta2p1,beta1)
vysledek[1,6]<-korelace(beta2p2,beta1)
vysledek[1,10]<-korelace(beta2p3,beta1)
vysledek[1,14]<-korelace(beta2p4,beta1)
vysledek[1,3]<-vzdalenost(beta2p1,beta1)
vysledek[1,7]<-vzdalenost(beta2p2,beta1)
vysledek[1,11]<-vzdalenost(beta2p3,beta1)
vysledek[1,15]<-vzdalenost(beta2p4,beta1)

# m1,m2 vektory, mj[i] pro i-tou realizaci pole průměr hodnot
  pole v bodech procesu v j-tém plátku (dle kóty)
m1<-array(0,c(realizace,1))
m2<-array(0,c(realizace,1))
for (i in 1:realizace) {
  sumaplatku1<-0
  sumaplatku2<-0
  n1<-0
  n2<-0
  for (j in 1:nBody) {
    if (Body[5,j]==1)
      {sumaplatku1 <- sumaplatku1 +
        Zvlnka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
        n1<-n1+1 }
    if (Body[5,j]==2)

```

```

    {sumaplatku2 <- sumaplatku2 +
      Zvlnka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n2<-n2+1 }
  }
  m1[i]<-sumaplatku1/n1
  m2[i]<-sumaplatku2/n2
}

# A jako v (iv) pro dva plátky
A1 <- n1* m1 %*% t(m1)
A2 <- n2* m2 %*% t(m2)
A <- A1 + A2

Avlastnihodnoty <- eigen(A)

beta2p1<-SIGMAodminv %*% Avlastnihodnoty$vector[,1]
beta2p2<-SIGMAodminv %*% Avlastnihodnoty$vector[,2]
beta2p3<-SIGMAodminv %*% Avlastnihodnoty$vector[,3]
beta2p4<-SIGMAodminv %*% Avlastnihodnoty$vector[,4]

# zápis výsledků verze (B)
vysledek[1,4]<-korelace(beta2p1,beta1)
vysledek[1,8]<-korelace(beta2p2,beta1)
vysledek[1,12]<-korelace(beta2p3,beta1)
vysledek[1,16]<-korelace(beta2p4,beta1)
vysledek[1,5]<-vzdalenost(beta2p1,beta1)
vysledek[1,9]<-vzdalenost(beta2p2,beta1)
vysledek[1,13]<-vzdalenost(beta2p3,beta1)
vysledek[1,17]<-vzdalenost(beta2p4,beta1)

#4 plátky:

# Přeznačení plátek 1 -> 11, 2 -> 22
for (i in 1:nBody) {
  if (Body[4,i]==1) Body[4,i]<-11
  if (Body[4,i]==2) Body[4,i]<-22
}

# rozdělení dvou plátek do čtyř podle hodnoty pole V v bodě
pmedianV1<-c()
pmedianV2<-c()
for (i in 1:nBody)
  if (Body[4,i]==11)

```

```

{Body[6,i]->pmedianV1[length(pmedianV1)+1]} else
{Body[6,i]->pmedianV2[length(pmedianV2)+1]}
medV1 <- median(pmedianV1)
medV2 <- median(pmedianV2)
for(i in 1:nBody) {
if (Body[4,i]==11)
  {if (Body[6,i] <= medV1) Body[4,i]<-12}
if (Body[4,i]==22)
  {if (Body[6,i] > medV2) Body[4,i]<-21}
}

```

```

# Přeznačení plátek 1 -> 11, 2 -> 22
for (i in 1:nBody) {
if (Body[5,i]==1) Body[5,i]<-11
if (Body[5,i]==2) Body[5,i]<-22
}

```

```

# rozdělení dvou plátek do čtyř plátek podle kóty
pmedianV1k<-c()
pmedianV2k<-c()
for (i in 1:nBody)
if (Body[5,i]==11)
{Body[3,i]->pmedianV1k[length(pmedianV1k)+1]} else
  {Body[3,i]->pmedianV2k[length(pmedianV2k)+1]}
medV1k <- median(pmedianV1k)
medV2k <- median(pmedianV2k)
for(i in 1:nBody) {
if (Body[5,i]==11)
  {if (Body[3,i] <= medV1k) Body[5,i]<-12}
if (Body[5,i]==22)
  {if (Body[3,i] > medV2k) Body[5,i]<-21}
}

```

```

# m11,m12,m21,m22 vektory, mj[i] pro i-tou realizaci pole průměr hodnot
pole v bodech procesu v j-tém plátku (v nejbližším bodě mříže)
# V11>V12>V21>V22 (ve smyslu průměru bodového procesu)
m11<-array(0,c(realizace,1))
m12<-array(0,c(realizace,1))
m21<-array(0,c(realizace,1))
m22<-array(0,c(realizace,1))
for (i in 1:realizace) {
  sumaplatku11<-0
  sumaplatku12<-0

```

```

sumaplatku21<-0
sumaplatku22<-0
n11<-0
n12<-0
n21<-0
n22<-0
for (j in 1:nBody) {
  if (Body[4,j]==11)
    {sumaplatku11 <- sumaplatku11 +
      Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n11<-n11+1 }
  if (Body[4,j]==12)
    {sumaplatku12 <- sumaplatku12 +
      Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n12<-n12+1 }
  if (Body[4,j]==21)
    {sumaplatku21 <- sumaplatku21 +
      Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n21<-n21+1 }
  if (Body[4,j]==12)
    {sumaplatku22 <- sumaplatku22 +
      Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
      n22<-n22+1 }

}
m11[i]<-sumaplatku11/n11
m12[i]<-sumaplatku12/n12
m21[i]<-sumaplatku21/n21
m22[i]<-sumaplatku22/n22

}

# A4 jako v (iv) pro čtyři plátky
A41 <- n11* m11 %*% t(m11)
A42 <- n12* m12 %*% t(m12)
A43 <- n21* m21 %*% t(m21)
A44 <- n22* m22 %*% t(m22)
A4 <- A41 + A42 + A43 + A44

A4vlastnihodnoty <- eigen(A4)

beta4p1<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,1]
beta4p2<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,2]
beta4p3<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,3]
beta4p4<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,4]

# zápis výsledků verze (A)
vysledek[1,18]<-korelace(beta4p1,beta1)

```



```

vysledek[1,22]<-korelace(beta4p2,beta1)
vysledek[1,26]<-korelace(beta4p3,beta1)
vysledek[1,30]<-korelace(beta4p4,beta1)
vysledek[1,19]<-vzdalenost(beta4p1,beta1)
vysledek[1,23]<-vzdalenost(beta4p2,beta1)
vysledek[1,27]<-vzdalenost(beta4p3,beta1)
vysledek[1,31]<-vzdalenost(beta4p4,beta1)

# m11,m12,m21,m22 vektory, mj[i] pro i-tou realizaci pole průměr hodnot
# pole v bodech procesu v j-tém plátku (v nejbližším bodě mříže)
# V11>V12>V21>V22 (ve smyslu průměru bodového procesu)
m11<-array(0,c(realizace,1))
m12<-array(0,c(realizace,1))
m21<-array(0,c(realizace,1))
m22<-array(0,c(realizace,1))
for (i in 1:realizace) {
  sumaplatku11<-0
  sumaplatku12<-0
  sumaplatku21<-0
  sumaplatku22<-0
  n11<-0
  n12<-0
  n21<-0
  n22<-0
  for (j in 1:nBody) {
    if (Body[5,j]==11)
      {sumaplatku11 <- sumaplatku11 +
        Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
        n11<-n11+1 }
    if (Body[5,j]==12)
      {sumaplatku12 <- sumaplatku12 +
        Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
        n12<-n12+1 }
    if (Body[5,j]==21)
      {sumaplatku21 <- sumaplatku21 +
        Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
        n21<-n21+1 }
    if (Body[5,j]==22)
      {sumaplatku22 <- sumaplatku22 +
        Zvluka[round(Body[1,j]/step)+1,round(Body[2,j]/step)+1,i]
        n22<-n22+1 }
  }
  m11[i]<-sumaplatku11/n11
  m12[i]<-sumaplatku12/n12
  m21[i]<-sumaplatku21/n21
  m22[i]<-sumaplatku22/n22
}

```

```

}
# A4 jako v (iv) pro čtyři plátky
A41 <- n11* m11 %*% t(m11)
A42 <- n12* m12 %*% t(m12)
A43 <- n21* m21 %*% t(m21)
A44 <- n22* m22 %*% t(m22)
A4 <- A41 + A42 + A43 + A44

A4vlastnihodnoty <- eigen(A4)

beta4p1<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,1]
beta4p2<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,2]
beta4p3<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,3]
beta4p4<-SIGMAodminv %*% A4vlastnihodnoty$vectors[,4]

# zápis výsledků verze (B)
vysledek[1,20]<-korelace(beta4p1,beta1)
vysledek[1,24]<-korelace(beta4p2,beta1)
vysledek[1,28]<-korelace(beta4p3,beta1)
vysledek[1,32]<-korelace(beta4p4,beta1)
vysledek[1,21]<-vzdalenost(beta4p1,beta1)
vysledek[1,25]<-vzdalenost(beta4p2,beta1)
vysledek[1,29]<-vzdalenost(beta4p3,beta1)
vysledek[1,33]<-vzdalenost(beta4p4,beta1)

# konec velké smyčky
}

```