Separovatelné a ambitové časoprostorové bodové procesy

Antonín Koubek

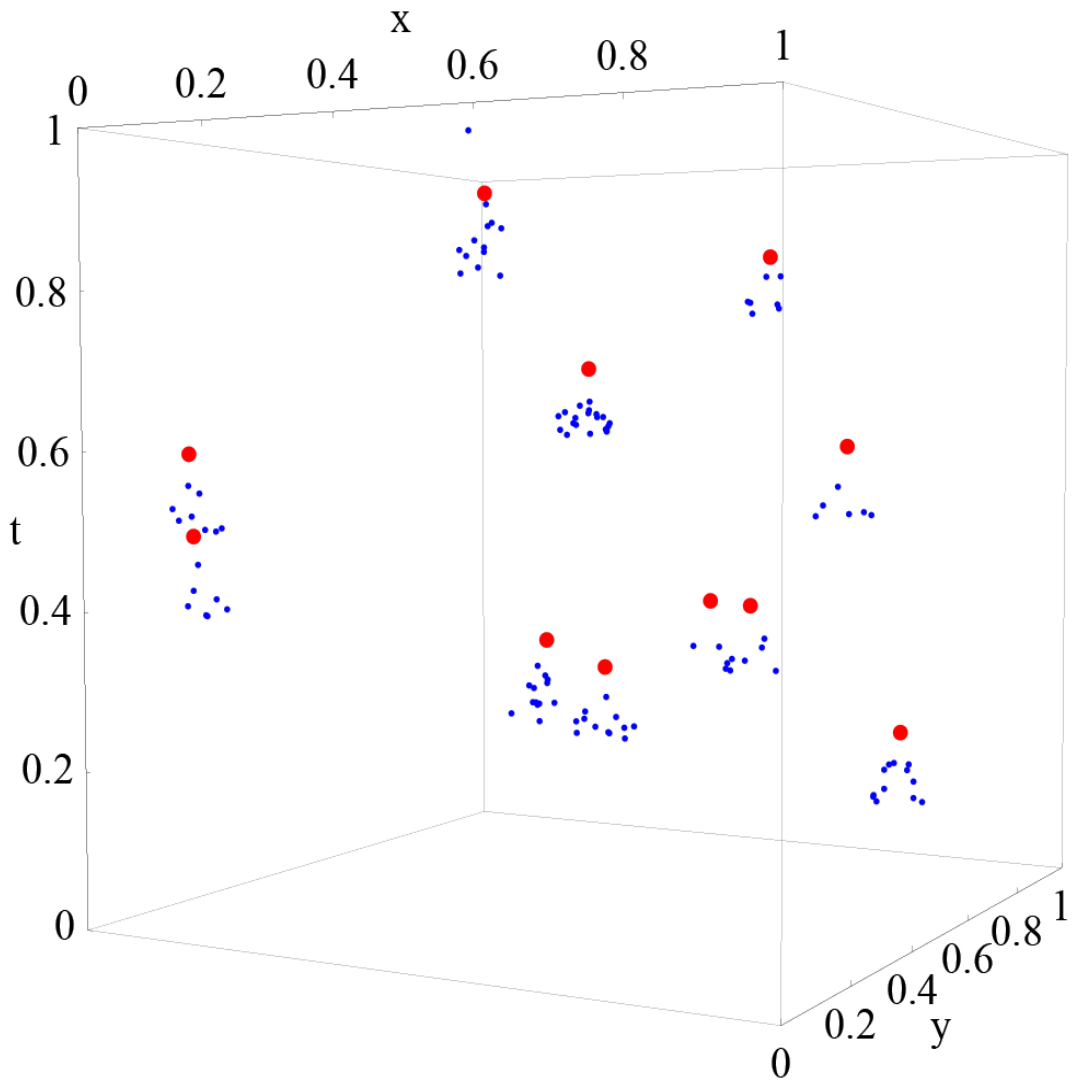
Katedra pravděpodobnosti a matematické statistiky, MFF UK Praha

Robust, Jetřichovice, 20.9.2014

Úvod a přehled

- Časoprostorový bodový proces. . . bodový proces v $\mathbb{R}^2 \times \mathbb{R}$
- Separovatelnost. . . možnost separovat prostor a čas (v dále definovaném smyslu)
- Testování separovatelnosti
- Ukázka na vybraných modelech SNCP
- Porovnání teorie a simulací

Ukázka – ambitový model



Obrázek: Realizace ambitového modelu. Svislá osa je časová, ostatní dvě osy jsou prostorové.

SNCP v $\mathbb{R}^2 \times \mathbb{R}$

Definice

(Časoprostorový shot noise Coxův proces (SNCP))

Nechť X je Coxův proces na $\mathbb{R}^2 \times \mathbb{R}$ s řídící intenzitou

$$Z(u,t) = \lambda(u,t) \sum_{(\xi,\eta,r)\in\Phi_p} rk((u,t),(\xi,\eta)) \tag{1}$$

kde $u \in \mathbb{R}^2$, $t \in \mathbb{R}$, $k(\cdot,\cdot)$ je pravděpodobnostní hustota, Φ_p je Poissonův bodový proces na $\mathbb{R}^2 \times \mathbb{R} \times (0,\infty)$ s lokálně integrovatelnou funkcí intenzity ζ a $\lambda(u,t)$ je nezáporná omezená funkce na $\mathbb{R}^2 \times \mathbb{R}$. Pak X se nazývá *časoprostorový shot noise Coxův proces (SNCP)*.

Podmínky na SNCP

- Pozorujeme proces X v kompaktní množině $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$
- $\zeta(u, t, r) = \nu p(r)$, $(u, t, r) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}_+$, $0 < \nu < \infty$, kóty jsou nezávislé se společnou hustotou $p(r)$
- $k((u, t), (v, s)) = k((u - v, t - s), 0)$ for all $(u, t), (v, s) \in \mathbb{R}^2 \times \mathbb{R}$

Separovatelnost

Definice

(Časoprostorová separovatelnost)

SNCP X v $\mathbb{R}^2 \times \mathbb{R}$ je časoprostorově separovatelný, pokud pro jádro k procesu X platí

$$k(u,t) = k_1(u)k_2(t), \quad u \in \mathbb{R}^2, t \in \mathbb{R},$$

kde k_1 je hustota v $\mathbb{R}^2$ a k_2 je hustota v $\mathbb{R}$.

Posouzení separovatelnosti

- Párová korelační funkce (za daných předpokladů) má tvar

$$g(u,t) = 1 + \frac{\int_{\mathbb{R}_+} r^2 p(r)dr \int_{\mathbb{R}} \int_{\mathbb{R}^2} k(x,y)k(x-u,y-t)dxdy}{\nu(\int_{\mathbb{R}_+} rp(r)dr)^2}$$

- Definujme funkce

$$g_1(u) = \frac{1}{|T|^2} \int_T \int_T g(u, t-s)dsdt, \quad u \in \mathbb{R}^2, \tag{2}$$

$$g_2(t) = \frac{1}{|W|^2} \int_W \int_W g(u-v, t)dudv, \quad t \in \mathbb{R}, \tag{3}$$

$$K_1(r) = \int_{\|u\| \le r} g_1(u)du, \quad r > 0, \tag{4}$$

$$K_2(t) = \int_{-t}^{t} g_2(s)ds, \quad t > 0. \tag{5}$$

Posouzení separovatelnosti – pokračování

- Definujme funkci

$$F(r,t) = \frac{K(r,t) - 2\pi r^2 t}{(K_1(r) - \pi r^2)(K_2(t) - 2t)}, \quad r,t > 0, \tag{6}$$

  kde položíme $\frac{a}{0} = 0$ pro $a \in \mathbb{R}$.
- F funkce je konstantní pro separovatelné procesy

Odhady

- $$\widehat{K}(r,t) = \frac{1}{|W|\,|T|} \sum_{i \neq j} \frac{I[\|u_i - u_j\| \leq r, |t_i - t_j| \leq t]}{\omega_1(u_i,u_j)\omega_2(t_i,t_j)\widehat{\rho}(u_i,t_i)\widehat{\rho}(u_j,t_j)},$$

  $(r,t) \in W \times T$, $\omega_1(u_i,u_j)$ je Ripleyho isotropická korekce okrajů, $\omega_2(t_i,t_j)$ je časová korekce okrajů.

- $$\widehat{K}_1(r) = \frac{1}{|W|\,|T|^2} \sum_{i \neq j} \frac{I[\|u_i - u_j\| \leq r]}{\omega_1(u_i,u_j)\widehat{\rho}(u_i,t_i)\widehat{\rho}(u_j,t_j)}, \quad r \geq 0,$$

- $$\widehat{K}_2(t) = \frac{1}{|W|^2\,|T|} \sum_{i \neq j} \frac{I[|t_i - t_j| \leq t]}{\omega_2(t_i,t_j)\widehat{\rho}(u_i,t_i)\widehat{\rho}(u_j,t_j)}, \quad t \geq 0,$$

- $$\widehat{F}(r,t) = \frac{\widehat{K}(r,t) - 2\pi r^2 t}{(\widehat{K}_1(r) - \pi r^2)(\widehat{K}_2(t) - 2t)}, \quad \text{kde } r,t > 0.$$

Simulace

Funguje pro výše definovaný prostorovo-časový SNCP.

- Simulace rodičovského Poissonova bodového procesu Φ v rozšířeném okně (okrajové efekty)
- Simulace kóty pro každý bod Φ
- Simulace bodů ve shluku pro každý bod Φ, počet bodů závisí na kótě
  - rovnoměrně na omezené množině aproximující shluk
  - nezávislé ztenčení pro získání požadované hustoty rozdělení

Model 1 - separovatelný model

- Rodičovský Poissonův bodový proces má intenzitu $\zeta(u,t,r) = 10\delta_1(r)$, $(u,t,r) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}_+$, takže bodový proces Φ je stacionární bodový proces v $\mathbb{R}^2 \times \mathbb{R}$ s intenzitou $\nu = 10$ a kótami rovnajícími se $r = 1$ s pravděpodobností jedna.
- Jádro určující rozdělení bodů ve shluku má součinový tvar $k(u,t) = k_1(u)k_2(t)$, $(u,t) \in \mathbb{R}^2 \times \mathbb{R}$, kde

$$k_1(u) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad u = (x,y) \in \mathbb{R}^2, \tag{7}$$

$$k_2(t) = \frac{\alpha}{1-\exp(-\alpha t^*)}\exp(-\alpha t), \quad 0 \le t \le t^*, \quad \text{jinak} \quad k_2(t) = 0, \tag{8}$$

kde $\sigma = 0.025$, $\alpha = 20$, $t^* = 0.1$.

- Funkce intenzity pro Model 1 je

$$\rho(u,t) = \frac{200}{(1-e^{-1})(e-1)(e^2-1)}\exp(-x+y+2t), \quad (u,t) \in \mathbb{R}^2 \times \mathbb{R}.$$

Model 2 - ambitový model

Definice

(Ambitová množina)

Nechť (u, t) je bod v $\mathbb{R}^2 \times \mathbb{R}$, nechť $t_1 > 0$ a $0 < \gamma < \frac{\pi}{2}$ jsou parametry. Potom *ambitová množina* $A_u(t)$ v $\mathbb{R}^2 \times \mathbb{R}$ je

$$A_u(t) = \left\{(v, s) \in \mathbb{R}^2 \times \mathbb{R} : t - t_1 \leq s \leq t, \|v - u\| \leq (t - s)\,\mathrm{tg}(\gamma)\right\}.$$

- Rodičovský bodový proces je stejný jako u Modelu 1
- Funkce intenzity je

$$\rho(u, t) = \nu\lambda(u, t) = 100, \quad (u, t) \in \mathbb{R}^2 \times \mathbb{R}.$$

- Body ve shluku mají rovnoměrné rozdělení na ambitu s vrcholy v bodech rodičovského procesu, tedy

$$k(u, t) = \frac{1}{|A_0(0)|}\,\mathrm{I}[(u, t) \in A_0(0)],$$

$(u, t) \in \mathbb{R}^2 \times \mathbb{R}$.

Výsledky – teoretické charakteristiky pro Model 1

- Funkce F pro Model 1.

$$F(r,t) = \frac{|W|^2 \, |T|^2 \, \nu}{\int_T \int_T k_2 * \widetilde{k}_2(s-t) ds dt \int_W \int_W k_1 * \widetilde{k}_1(u-v) du dv} = 10.9103,$$

- Párová korelační funkce pro Model 1.

$$g(u,t) = 1 + \frac{\alpha e^{-\frac{x_0^2+y_0^2}{4\sigma^2}} \left(e^{-\alpha|t|} - e^{\alpha(|t|-2t^*)} \right)}{8\nu\pi\sigma^2 \left(1 - e^{-\alpha t^*}\right)^2} \mathrm{I}[0 \le |t| \le t^*], \quad (9)$$

$(u,t) \in \mathbb{R}^2 \times \mathbb{R}$.

- K-funkce

$$K(r,t) = \int_{\|u\| \le r} \int_{-t}^{t} g(u,s) \, \mathrm{I}[0 \le |s| \le t^*] ds du.$$

Výsledky – teoretické charakteristiky pro Model 2

- Model 2

$$g(u,t) = 1 + \frac{|A_u(t) \cap A_0(0)|}{\nu \, |A_0(0)|^2} = 1 + c_1 \, |A_u(t) \cap A_0(0)| \,,$$

  kde $(u,t) \in \mathbb{R}^2 \times \mathbb{R}$.

- Konstantu c_1 spočítáme snadno, pro objem průniku dvou ambitů máme vzorec

$$\text{Intersection}(x,y,z) = \left|A_{(x,y)}(z) \cap A_0(0)\right| \tag{10}$$

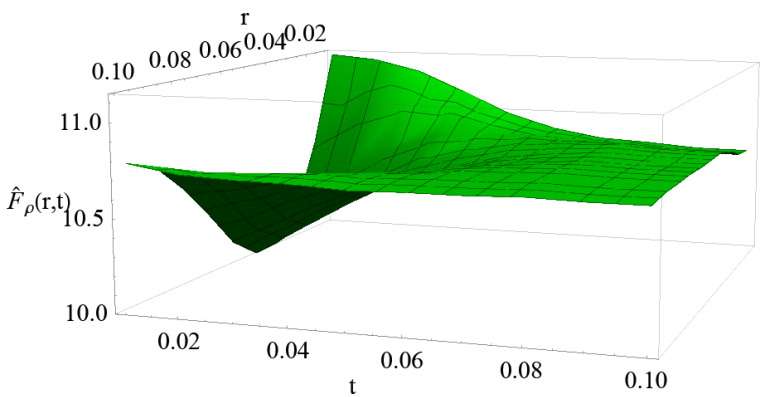
- K funkci vypočítáme integrováním g funkce, teoretickou F funkci podle vzorce

$$F(r,t) = \frac{K(r,t) - 2\pi r^2 t}{(K_1(r) - \pi r^2)(K_2(t) - 2t)}, \quad r,t > 0.$$

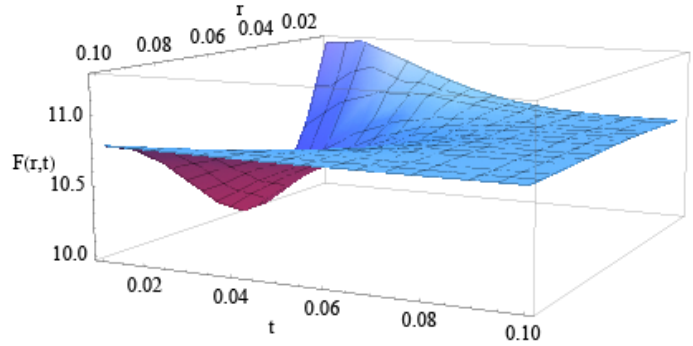
Výsledky – odhad F funkce

$$\widehat{F}(r,t) = \frac{\widehat{K}(r,t) - 2\pi r^2 t}{(\widehat{K}_1(r) - \pi r^2)(\widehat{K}_2(t) - 2t)}, \quad \text{kde } r,t > 0. \tag{11}$$

- Slouží pro posouzení prostorovo-časové separability. V separabilním případě je F funkce konstantní.



Obrázek: F-funkce s použitím ρ.



Obrázek: Teoretická F-funkce pro Model 2.

Testování separovatelnosti

- Kvůli rozptylu odhadu $\hat{F}$ nemusí být zřejmé, zda odpovídá konstantní funkci F
- Testujeme H_0: model je separovatelný, proti H_1: model není separovatelný
- Návrh: sestrojit obálky pro separovatelný model a porovnat s nimi odhad $\hat{F}$
- Návrh: získat v každém bodě (r, t) více odhadů $\hat{F}_i(r, t)$, $i = 1 \ldots N$ a provést analýzu rozptylu (testovat shodu středních hodnot, Welchův test pro různé rozptyly výběrů)

Diskuse

- Podařilo se získat kvalitní simulace pro separovatelné i ambitové modely.
- Odhady intenzity i K funkce mají malé vychýlení - velmi dobře se shodují s teoretickými hodnotami pro všechny modely.
- Problematické je používání odhadnuté intenzity pro další odhady, např. K funkci. Intenzita je odhadnutá jen v diskrétních bodech.
- Odhad F funkce je výpočetně náročný. Má velký rozptyl závisející na parametrech modelu. Je třeba počítat průměr nejméně z 1000 simulací.
- Dokázali jsme vypočítat teoretické hodnoty pro funkci intenzity, párovou korelační funkci, K funkci a F funkci pro všechny zkoumané modely.
- Dalo by se uvažovat o lépe fungujícím testování separovatelnosti. Obálkový test není dostatečně silný. Analýzu rozptylu lze využít, ale vyžaduje mnoho replikací.

Literatura

- Møller J., Waagepetersen R. P.:*Statistical Inference and Simulation for Spatial Point Processes*, Chapman & Hall/CRC, 2004.
- Møller J., Ghorbani M.: Aspects of second–order analysis of structured inhomogeneous spatio–temporal point processes, In: Statistica Neerlandica 66, 4, 472-491, 2012.
- Pawlas Z.: Prostorové modelování, prostorová statistika 1, Učební text MFF UK, Praha, 2010.
- Prokešová M., Hellmund G., Jensen E.B.V.: On spatio–temporal Lévy based Cox processes, In: Proceedings S^4G International Conference on Stereology, Spatial Statistics and Stochastic Geometry, Prague, Editors: Lechnerová R., Saxl I., Beneš V., Union of Czech math. and physicists, 111-116, 2006.