Punktprozesse

Andreas Frommknecht

Seminar Zufällige Felder
Universität Ulm

20.01.2009

Inhalt

Punktmuster

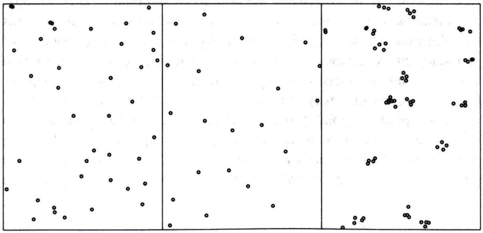


Figure 1.1 Three simulated point patterns: (left) random, (centre) regular, (right) clustered.

Kapitel 1 [Illian,Penttinen,Stoyan,Stoyan (2008)]

- Ziel: Analyse und Modellierung geometrischer Strukturen von Mustern, die durch zufällig verteilte Objekte im Raum entstehen.

Anwendung

Beispiele für Anwendungen in verschiedenen Fachgebieten

- Biologie: Verteilung von Bäumen
- Medizin: Zellkulturen
- Astronomie: Verteilung von Sternen

Vorbemerkungen

- Raum:
  - Definition allgemein in abstrakten Räumen möglich.
  - Wir betrachten hier den $\mathbb{R}^d$.
- Prozess normalerweise zeitliche Entwicklung, hier meist zeitunabhängig.

Zufälliges Zählmaß

- Vorbemerkung:
  - $Q^d := \{B \in \mathbb{R}^d : B = (a_1, b_1] \times ... \times (a_d, b_d], a_i, b_i \in \mathbb{R}, a_i \leq b_i, \forall i = 1, ..., d\}$
  - Sei Z die Familie aller lokal endlichen Zählmaße $\varphi : \mathfrak{B}(\mathbb{R}^d) \to \mathbb{N}_0 \cup \infty$, d.h. $\varphi(B) < \infty, \forall B \in Q^d$ und $\varphi(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \varphi(B_i)$ für paarweise disjunkte $B_i \in \mathfrak{B}(\mathbb{R}^d)$.
  - Sei $\mathfrak{N}$ die kleinste σ-Algebra von Teilmengen von Z, so dass $\varphi \mapsto \varphi(B), \forall$ beschränkte $B \in \mathfrak{B}(\mathbb{R}^d)$ eine $(\mathfrak{N}, \mathfrak{B}(\mathbb{R}))$ messbare Abbildung ist.

Zufälliges Zählmaß

Definition

Ein zufälliges Zählmaß $\mu : \Omega \to Z$ *ist eine Zufallsvariable über einem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, P)$ *mit Werten in* $(Z, \mathfrak{N})$.

Definitionen eines Punktprozesses

Bezeichnung: N bzw. N'

1. Definition

$N' = \{x_1, x_2, ..., x_n\}$ *ist eine beliebige Menge von Zufallsvektoren über einem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, P)$*, wobei* $x_i \in \mathbb{R}^d$, $\#\{x_i, x_i \in B\} < \infty \ \forall B \in \mathfrak{B}_0(\mathbb{R}^d)$ *und* $n \in \mathbb{N}_0 \cup \infty$ *zufällig .*

Alternativ:

2. Definition

$N(B) := \sum_{x \in N'} \mathbb{1}_B(x) \ \forall B \in \mathfrak{B}(\mathbb{R}^d)$

Bemerkungen

- N ist ein lokal endliches zufälliges Zählmaß.
- $N(B)$ gibt die zufällige Anzahl der Punkte in B an.
- Unter einer Punktprozess-Summe versteht man $S_f = \sum_{x \in N'} f(x)$, wobei f eine reellwertige messbare Funktion ist.

Beispiel: Dichte von Pflanzensamen

- Annahmen:
  - Jeder Punkt $x_i \in N'$ entspricht einer Pflanze.
  - Jede Pflanze produziert unabhängig Samen.
  - Die Samen sind zufällig um die Pflanze verteilt.
  - m ist der Mittelwert der Planzensamen pro Pflanze.
  - d ist eine monoton fallende Funktion, z.B. $d(r) = \exp(-r)$.
  - Die Samendichte an einem Punkt y von einer Pflanze x_i lässt sich durch $f_y(x_i) = md(||y - x_i||)$ beschreiben.
- Frage: Wie hoch ist die Samendichte an einem Punkt y?

Beispiel: Dichte von Pflanzensamen

- Annahmen:
  - Jeder Punkt $x_i \in N'$ entspricht einer Pflanze.
  - Jede Pflanze produziert unabhängig Samen.
  - Die Samen sind zufällig um die Pflanze verteilt.
  - m ist der Mittelwert der Planzensamen pro Pflanze.
  - d ist eine monoton fallende Funktion, z.B. $d(r) = \exp(-r)$.
  - Die Samendichte an einem Punkt y von einer Pflanze x_i lässt sich durch $f_y(x_i) = md(||y - x_i||)$ beschreiben.
- Frage: Wie hoch ist die Samendichte an einem Punkt y?
- Antwort: $S_f(y) = \sum_{x_i \in N'} f_y(x_i)$

Verteilungen

- Ein Punktprozess ist eindeutig bestimmt durch die Familie der endlich-dimensionalen Wahrscheinlichkeiten:

$$\{\mathbf{P}(N(B_1) = n_1, ..., N(B_k) = n_k),\ n_i, k \in \mathbb{N}_0\ \forall B_i \in Q^d\}$$

- Bemerkung: Da unendlich viele Mengen B_i existieren, existieren auch unendlich viele Verteilungen.

Verteilungen

- Punktprozessverteilung (allgemeiner)

$$\mathbf{P}(N' \in \mathcal{A})$$

  wobei $\mathcal{A} \subset \mathfrak{B}_0(\mathbb{R}^d)$ und N' einfach ($x_i \neq x_j \; \forall i \neq j$).
- Falls z.B. $\mathcal{A}$ die Menge aller Punktmuster ist die keinen Punkt in B haben, gilt $\mathbf{P}(N' \in \mathcal{A}) = \mathbf{P}(N(B) = 0)$.

Erwartungswert

Definition

$\boldsymbol{E}(N(B)) :=$ *Erwartete Anzahl der Punkte von* N' *in* B

- Bezeichnung: $\Lambda(B) := \mathbf{E}(N(B))$
- Bemerkung: Λ heißt Intensitätsmaß.
- Falls Λ absolutstetig bzgl. des Lebesgue-Maßes ist, existiert eine Funktion $\lambda(x)$ mit folgender Eigentschaft: $\Lambda(B) = \int_B \lambda(x)dx$
- Bezeichnung: λ heißt Intensitätsfunktion.

Stationarität

- Ein Punktprozess N' heißt stationär, falls $N' \stackrel{d}{=} N'_x$, $\forall x \in \mathbb{R}^d$ gilt, wobei $N' = \{x_1, x_2, ...\}$ und $N'_x = \{x_1 + x, x_2 + x, ...\}$.

Äquivalent:

- Ein Punktprozess N heißt stationär, falls $(N(B_1), ..., N(B_n)) \stackrel{d}{=} (N(B_{1+x}), ..., N(B_{n+x})), \forall x \in \mathbb{R}^d$, $n \in \mathbb{N}$ und $B_1, ..., B_n \in \mathfrak{B}(\mathbb{R}^d)$ gilt, wobei $B_{i+x} = \{y + x, y \in B_i\}$.

Isotropie

- Ein Punktprozess N' heißt isotrop, falls $N' \stackrel{d}{=} RN'$, $\forall R$ gilt, wobei $N' = \{x_1, x_2, ...\}$, $RN' = \{Rx_1, Rx_2, ...\}$ und R eine Drehung um den Ursprung ist.

Äquivalent:

- Ein Punktprozess N heißt isotrop, falls $(N(B_1), ..., N(B_n)) \stackrel{d}{=} (N(RB_1), ..., N(RB_n))$, $\forall R$, $n \in \mathbb{N}$ und $B_1, ..., B_n \in \mathfrak{B}(\mathbb{R}^d)$ gilt, wobei $RB_i = \{Ry, y \in B_i\}$.

Folgerung aus Stationarität

- Falls N stationär, gilt $\Lambda(B) = \lambda\nu(B)$, wobei λ eine Konstante und $\nu(B)$ das Volumen von B ist.
- **Beweis**
  $N(B) \stackrel{d}{=} N(B_x) \Rightarrow \mathbf{E}(N(B)) = \mathbf{E}(N(B_x))$
  $\Rightarrow \Lambda(B) = \Lambda(B_x)\ \forall x, B \underset{\textbf{Maßtheorie}}{\Rightarrow} \Lambda$ ist ein vielfaches des Lebesguemaßes $\Rightarrow \Lambda(B) = \lambda\nu(B)$
- Bemerkung:
  - λ wird als Intensität oder Punktintensität bezeichnet.
  - λ kann als erwartete Anzahl der Punkte pro Volumeneinheit betrachtet werden.

Markierte Punktprozesse

- Vorbemerkung:
  - Ein Punkt $x_i \in \mathbb{R}^d$ heißt markiert, wenn ihm ein Wert $m(x_i) \in \mathbb{R}^l$ zugewiesen ist.
  - Die Menge der $m(x_i)$ ist eine Menge von Zufallsvektoren über $(\Omega, \mathcal{A}, P)$.
  - Ein markierter Punktprozess wird mit M bzw. M' bezeichnet.

Definition Markierter Punktprozesse

1. Definition

$M' = \{[x_1; m(x_1)], ..., [x_n; m(x_n)]\}$ ist eine Menge von Tupeln von Zufallsvektoren, wobei $n \in \mathbb{N}_0 \cup \infty$ zufällig.

2. Definition

Lokal endliches zufälliges Zählmaß
$M(B \times C) := \sum_{x \in M'} \mathbb{1}_{B \times C}(x)$, wobei $B \in \mathfrak{B}(\mathbb{R}^d)$ und $C \in \mathfrak{B}(\mathbb{R}^l)$ ist.

Bemerkung & Beispiel: Dichte von Pflanzensamen

- Bemerkung:
  - Falls $C = \mathbb{R}^l$ gilt folgt $M(B \times C) = M(B \times \mathbb{R}^l) = N(B)$.
  - Summe von markierten Punktprozessen: $S_f = \sum_{[x;m(x)] \in M'} f(x, m(x))$, wobei f eine reellwertige Funktion ist.
- Beispiel Dichte von Pflanzensamen:
  - Annahmen: Analog zu Beispiel auf Folie 10.
  - Zusätzliche Annahme: $m(x_i)$ ist die Anzahl der Planzensamen der Pflanze x_i.
  - Frage: Wie hoch ist die Samendichte an einem Punkt y?

Bemerkung & Beispiel: Dichte von Pflanzensamen

- Bemerkung:
  - Falls $C = \mathbb{R}^l$ gilt folgt $M(B \times C) = M(B \times \mathbb{R}^l) = N(B)$.
  - Summe von markierten Punktprozessen: $S_f = \sum_{[x;m(x)] \in M'} f(x, m(x))$, wobei f eine reellwertige Funktion ist.
- Beispiel Dichte von Pflanzensamen:
  - Annahmen: Analog zu Beispiel auf Folie 10.
  - Zusätzliche Annahme: $m(x_i)$ ist die Anzahl der Planzensamen der Pflanze x_i.
  - Frage: Wie hoch ist die Samendichte an einem Punkt y?
  - Antwort: $S_f(y) = \sum_{[x_i;m(x_i)] \in M'} f_y(x_i) = \sum_{(i)} m(x_i) d(||y - x_i||)$

Inhalt

Definition homogener Poisson-Prozess

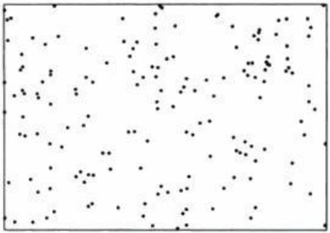
Definition

Ein Punktprozess N mit der Verteilung

$$\boldsymbol{P}(N(B) = n) = \frac{\lambda^n (\nu(B))^n}{n!} \exp(-\lambda \nu(B)), \; n \in \mathbb{N}_0, \; B \in \mathfrak{B}_0(\mathbb{R}^d)$$

heißt homogener Poisson-Punktprozess mit Intensität λ, falls $N(B_1), N(B_2), \ldots$ unabhängig für paarweise disjunkte $B_1, B_2, \ldots \in \mathfrak{B}_0(\mathbb{R}^d)$

Realisierung



Realisierung eines homogenen Poissonprozesses
Kapitel 3 [Möller,Waagepeterson (2003)]

Bemerkung

- Bemerkung
  - In diesem Fall gilt:

$$\Lambda(B) = \lambda\nu(B)$$

und

$$\mathbf{P}(N(B_1) = n_1, ..., N(B_k) = n_k)$$

$$= \frac{\lambda^{n_1+\ldots+n_k}(\nu(B_1))^{n_1}\ldots(\nu(B_k))^{n_k}}{n_1!\ldots n_k!} \exp(-\sum_{i=1}^{k}\lambda\nu(B_i))$$

$$n_i \in \mathbb{N}_0, B_i \in \mathfrak{B}_0(\mathbb{R}^d) \textit{ disjunkt}$$

  - Stationarität und Isotropie sind offensichtlich.

Definition inhomogoner Poisson-Prozess

Definition

Ein Punktprozess N mit der Verteilung

$$\boldsymbol{P}(N(B) = n) = \frac{\Lambda(B)^n}{n!} \exp(-\Lambda(B)), \ n \in \mathbb{N}_0, \ B \in \mathfrak{B}_0(\mathbb{R}^d)$$

und

$$\Lambda(B) = \int_B \lambda(x) dx$$

heißt inhomogener Poisson-Punktprozess mit Intensitäts-funktion $\lambda(x)$*, falls* $N(B_1), N(B_2), \ldots$ *unabhängig für paarweise disjunkte* $B_1, B_2, \ldots \in \mathfrak{B}_0(\mathbb{R}^d)$

ulm university universität uulm

Bemerkung

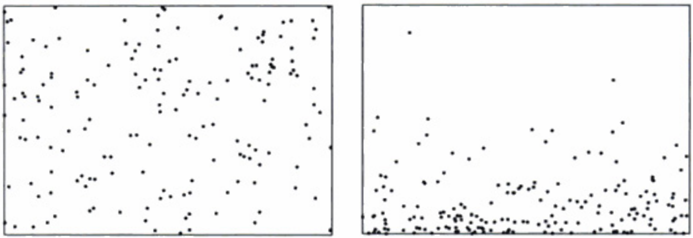
- Bemerkung
  - Im Gegensatz zum homogenen Poissonprozess ist hier die Intensitätsfunktion $\lambda(x)$ nicht konstant.
  - Es gilt:

$$\mathbf{P}(N(B_1) = n_1, ..., N(B_k) = n_k)$$

$$= \frac{\Lambda(B_1)^{n_1} ... \Lambda(B_k)^{n_k}}{n_1! ... n_k!} \exp(-\sum_{i=1}^{k} \Lambda(B_i))$$

$$n_i \in \mathbb{N}_0, B_i \in \mathfrak{B}_0(\mathbb{R}^d) \textit{ disjunkt}$$

Vergleich Poissonprozesse



Realisierung eines homogenen(links) und eines inhomogenen (rechts) Poissonprozesses

Kapitel 3 [Möller,Waagepeterson (2003)]

Shot-Noise Felder

- Voraussetzungen:
  - $M' = \{[x_1, m(x_1)], ..., [x_n, m_n]\}$ sei ein markierter Punktprozess .
  - $s(y - x_i, m(x_i)) := m(x_i) f(||y - x_i||), \; y \in \mathbb{R}^d$
  - $S(y) := \sum_{[x_i; m(x_i)] \in M'} s(y - x_i, m(x_i))$
  - $\mathbb{E}(|S(y)|) < \infty$, $\forall y \in \mathbb{R}^d$

Definition

Definition

Falls vorherige Voraussetzungen erfüllt sind heißt

$$S(y) = \sum_{[x_i; m(x_i)] \in M'} s(y - x_i, m(x_i))$$

Shot-Noise Feld.

Bemerkung

- Bemerkung:
  - s wird als Impulsfunktion oder Dämpfungsfunktion bezeichnet.
  - Die Dichte von Pflanzensamen kann als Shot-Noise Feld aufgefasst werden:

$$S_f = S_f(y) = \sum_{[x_i;m(x_i)]\in M'} s(y - x_i, m(x_i))$$

$$= \sum_{[x_i;m(x_i)]\in M'} m(x_i) f(||y - x_i||) = \sum_{[x_i;m(x_i)]\in M'} m(x_i) d(||y - x_i||)$$

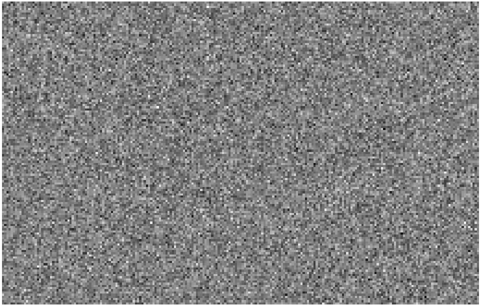
Inhalt

Modell

- Vorbemerkung:
  - Papier besteht aus Fasern.
  - Anhäufung von Fasern werden als Flocken bezeichnet.
- $Y(x) = Y_1(x) + Y_2(x) : x \in \mathbb{R}^2$ reellwertiger stochastischer Prozess, der Deckfähigkeit (Lichtdurchlässigkeit) des Papiers simulieren soll.
- $Y_1, \ Y_2$ unabhängig.

Weißes Rauschen

- $Y_1(x)$ Weißes Rauschen (unabhängige $N(\alpha, \tau^2)$ verteilte Zufallsvariablen).
- α durchschnittliche Deckkraft ohne Flockenbildung.

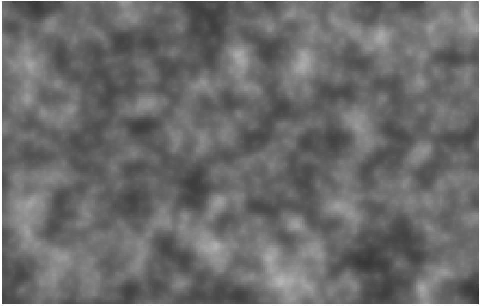


Weißes Rauschen
[Sebastian Lochbrunner]

Shot-Noise

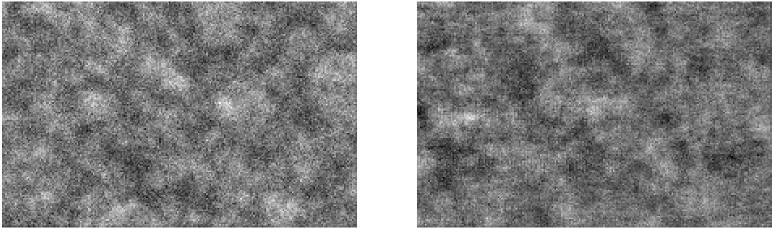
- $Y_2(x) = \beta \sum_j f(x - X_j)$ Shot-Noise Feld.
- X_j Punkte eines homogenen Poissonprozesses.
- $f(x)$ beschreibt Oberflächenprofil einer Flocke.
- $\beta f(0)$ durchschnittliche Deckkraft am Zentrum einer Flocke (konstante Markierungsfunktion $m(x_i)$).

Shot-Noise Bild



Shot-Noise
[Sebastian Lochbrunner]

Vergleich zwischen Papier und Realisierung



Realisierung (links), Papier (rechts)
[Sebastian Lochbrunner]

Mögliche Wahl von f

- $f(x) = \begin{cases} (\pi\sigma^2)^{-1} & \text{für } ||x|| \leq \sigma \\ 0 & \text{für } ||x|| > \sigma \end{cases}, \ \sigma > 0$
- $f(x) = t\, exp(-s||x||), \ t, s > 0$

Mögliche Wahl von f

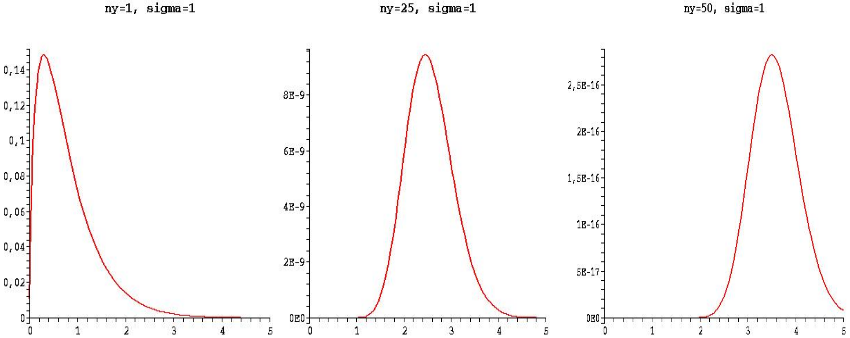
- Matérn-Funktion:

$$f(x) = \frac{\nu 2^{1-\nu}}{\pi\sigma^2\Gamma(\nu+1)}\left(\frac{||x||}{\sigma/2\sqrt{\nu}}\right)^{\nu} K_\nu\left(\frac{||x||}{\sigma/2\sqrt{\nu}}\right), \ \sigma > 0$$

- Bemerkung:
  - K_ν ist eine modifizierte Bessel-Funktion zweiter Ordnung.
  - ν und σ sind Skalierungsparameter.
  - ν bestimmt wie oft f differenzierbar ist und damit die Rauhigkeit des Papiers.

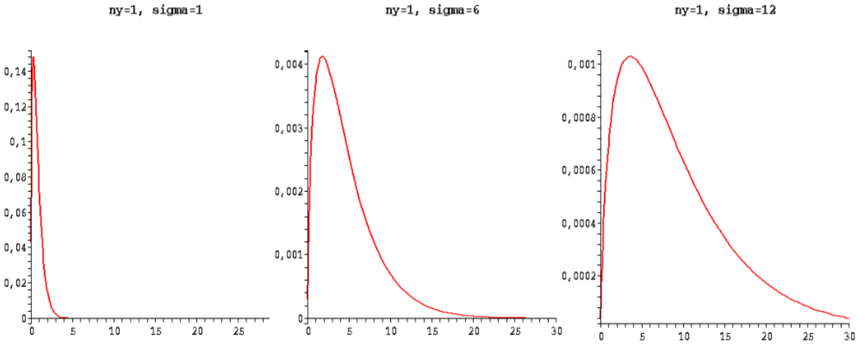
ulm university universität uulm

Einfluss von ν auf Matérn-Funktion



ν verschiebt Funktion und verkleinert Amplitude

Einfluss von σ auf Matérn-Funktion



σ staucht Funktion

Literatur

V. Schmidt
Räumliche Statistik
Vorlesungsskript, Wintersemester 2007/08

J. Illian, A. Penttinen, H. Stoyan und D. Stoyan
Statistical Analysis and Modelling of Spatial Point Patterns
Wiley, 2008

J. Möller, R. Waagepeterson
Statistical Inference and Simulation for Spatial Point Processes
Chapman & Hall, 2003

D.J. Daley, D. Vere-Jones
An Introduction to the Theory of Point Processes: Elementary Theory and Methods
Springer, 2003

P.E. Brown, P.J. Diggle und R. Henderson
A Non-Gaussian Spatial Process Model For Opacity Of Flocculated Paper
Scandinavian Journal Of Statistics, 30: 355–368, 2003.