Partition Modelling

Sebastian Keim, Benedikt Jahn

Seminar „Stochastische Geometrie und ihre Anwendungen - Zufällige Netzwerke", 2008

1. Motivation
2. Das kollektive Modell
3. Partition Modelling
   - Problemstellung
   - Das Voronoi Mosaik
   - Partition Model
   - Bayes Schätzer
4. Simulationsalgorithmus
   - Markov-Ketten
   - Reversible-Jump-Sampler
   - Equivalent Kernel
   - Beispiel
5. Literatur
6. Anhang

1. Motivation
2. Das kollektive Modell
3. Partition Modelling
   - Problemstellung
   - Das Voronoi Mosaik
   - Partition Model
   - Bayes Schätzer
4. Simulationsalgorithmus
   - Markov-Ketten
   - Reversible-Jump-Sampler
   - Equivalent Kernel
   - Beispiel
5. Literatur
6. Anhang

Beispiel aus der Schadenversicherung - der Orkan Kyrill

Quelle: Münchner Rück

Bilanz des Orkans (18. Januar 2007)

Der Orkan Kyrill verursachte einen Schaden von ca. 6 Millarden Euro

$\Rightarrow$ große Risiken für Schadenversicherer

$\Rightarrow$ Modellierung von Risiken

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Das kollektive Modell der Schadenversicherungsmathematik:

Vorraussetzungen:

- Sei N die zufällige Anzahl der Schäden innerhalb der Periode t
- Seien die Schadenhöhen $\{U_i : i \in \mathbb{N}\}$ unabhängig und identisch verteilt
- $U_i > 0$ f.s.
- Seien $\{U_i : i \in \mathbb{N}\}$ und N voneinander unabhängig

Gesamtschaden im kollektiven Modell

$$X^{col} = \sum_{i=1}^{N} U_i \tag{1}$$

Kritik am kollektiven Modell:

Beide Unabhängigkeitsannahmen sind in der Realität problematisch:

- Schäden U_i sind korreliert, da sie gemeinsame Risikofaktoren haben (z.B. das Wetter in der Autoversicherung)
- Schäden U_i sind mit der Schadenanzahl N korreliert (z.B. Elementarschaden Versicherung)

Kritik am kollektiven Modell:

Beide Unabhängigkeitsannahmen sind in der Realität problematisch:

- Schäden U_i sind korreliert, da sie gemeinsame Risikofaktoren haben (z.B. das Wetter in der Autoversicherung)
- Schäden U_i sind mit der Schadenanzahl N korreliert (z.B. Elementarschaden Versicherung)

$\Rightarrow$ Verwende ein Modell, das räumliche Abhängigkeitsstrukturen beachtet

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

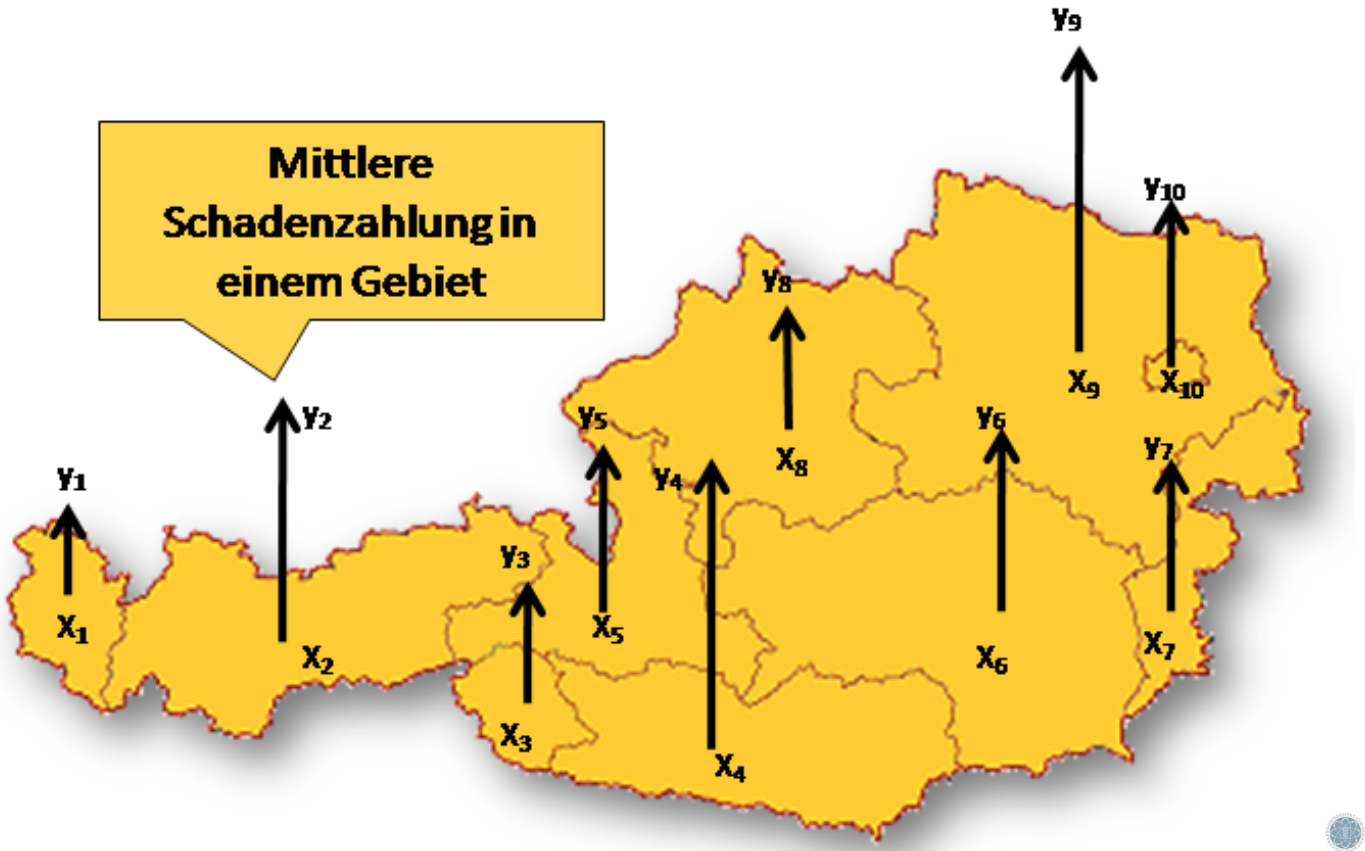
4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Problemstellung:

Seien die Zielwerte $y_1,\ldots,y_k$ (Höhe der mittleren Schadenzahlungen in einem PLZ Gebiet) an den Positionen $x_1,\ldots,x_k$ gegeben:



Das Modell:

$$y(x_i) = \overline{y} + z(x_i) + \epsilon_i \quad \text{für i} = 1,\ldots,n$$

wobei

- $\overline{y} = \frac{1}{n}\sum y_i$
- $\mathrm{E}(\epsilon_i) = 0$, $\mathrm{Var}(\epsilon_i) = \sigma^2$
- z(x) ist eine deterministische Funktion, die die Abweichung des lokalen Mittelwerts vom globalen empirischen Mittelwert $\overline{y}$ beschreibt

Das Modell:

$$y(x_i) = \overline{y} + z(x_i) + \epsilon_i \quad \text{für i} = 1,\ldots,n$$

wobei

- $\overline{y} = \frac{1}{n}\sum y_i$
- $\mathrm{E}(\epsilon_i) = 0$, $\mathrm{Var}(\epsilon_i) = \sigma^2$
- z(x) ist eine deterministische Funktion, die die Abweichung des lokalen Mittelwerts vom globalen empirischen Mittelwert $\overline{y}$ beschreibt

Schätze z(x) mit Hilfe der Daten $D = \{y_i, x_i\}_{i=1}^{n}$

Jedoch gilt für $\widetilde{z}(x)$

Entweder höhere Varianz oder höherer Bias von $\widetilde{z}(x)$:
Denn für den Mean squared error gilt:

$$MSE(\widetilde{z(x)}, z(x)) = E(\left\{(\widetilde{z(x)} - z(x))^2\right\}) = var(\widetilde{z(x)}) + bias(\widetilde{z(x)})^2 \quad (2)$$

$\Rightarrow$ Es muss ein Kompromiss gefunden werden

Je näher zwei Positionen x_1 und x_2 beieinander liegen, desto stärker sind sie (positiv) korreliert

Hat x_1 einen großen Zielwert y_1, so sollte der Zielwert y_2 bei x_2 ebenso relativ groß sein.

Je näher zwei Positionen x_1 und x_2 beieinander liegen, desto stärker sind sie (positiv) korreliert

Hat x_1 einen großen Zielwert y_1, so sollte der Zielwert y_2 bei x_2 ebenso relativ groß sein.

- Erster Ansatz: Feste Korrelationsstrukur (Korrelationsfunktion: $C(h) = Cov(y(x+h), y(x))$)
- Große Einschränkung: Einige Modelle können auf diese Weise nicht dargestellt werden

Je näher zwei Positionen x_1 und x_2 beieinander liegen, desto stärker sind sie (positiv) korreliert

Hat x_1 einen großen Zielwert y_1, so sollte der Zielwert y_2 bei x_2 ebenso relativ groß sein.

- Erster Ansatz: Feste Korrelationsstrukur (Korrelationsfunktion: $C(h) = Cov(y(x+h), y(x))$)
- Große Einschränkung: Einige Modelle können auf diese Weise nicht dargestellt werden
- Anderer Ansatz, ohne die Korrelationsstruktur explizit zu bestimmen

$$\Rightarrow \text{Partition Modeling}$$

1 Motivation

2 Das kollektive Modell

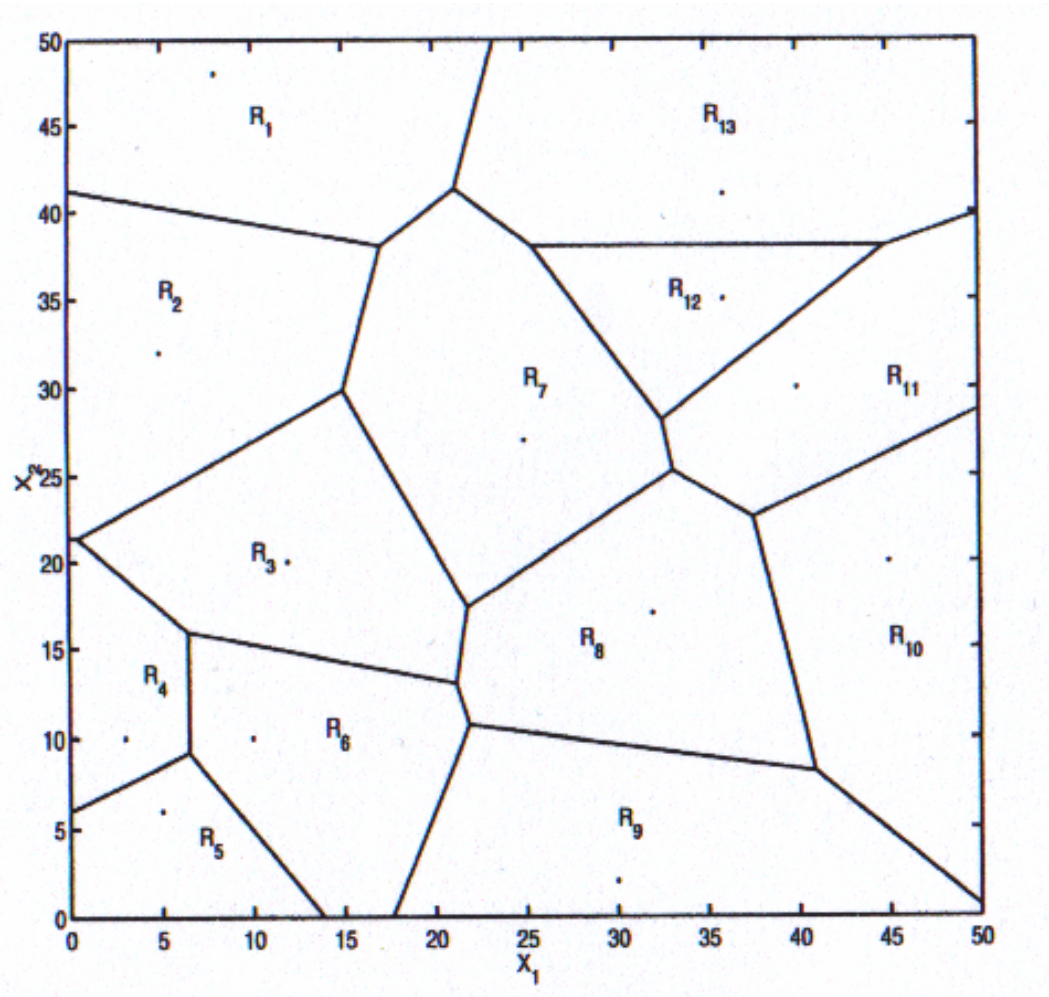
3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Das Voronoi Mosaik:



Das Voronoi Mosaik:

Sei folgende Zerlegung von χ gegeben:

Disjunkte Gebiete $R_1,..,R_k$, deren Vereinigung den gesamten Raum ergeben. (d.h. $R_i^o \cap R_j^o = \emptyset$,i $\neq$ j und $\bigcup_1^k R_i = \chi$)

$$R_i = \{x \in \chi : D(x, c_i) \leq D(x, c_j)\},\ \text{i} \neq \text{j}$$

- Die Zentren c $= (c_1, ..., c_k)$ der Gebiete legen das Mosaik von χ eindeutig fest.
- Verwende den euklidischen Abstand $D^2(x_1,x_2) = \sum_{i=1}^{p}(x_{1i} - x_{2i})^2$ (p=2 oder p=3)

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

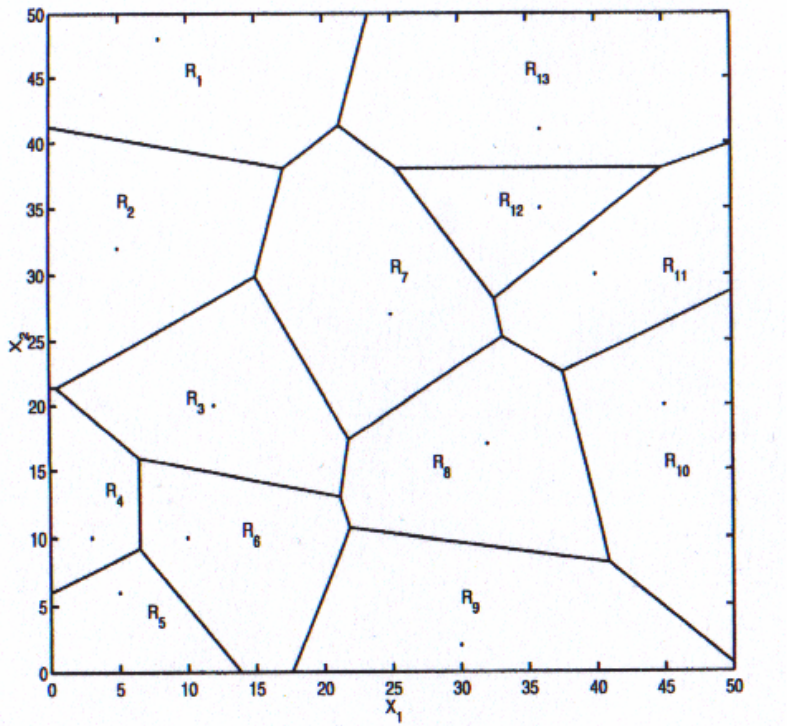
5 Literatur

6 Anhang

Partition Model:

In jeder Region R_i ist die Zielwert-Funktion $z(x) = \mu_i \; \forall \; x \in R_i$

$\Rightarrow$ Jedoch eventuell hoher Bias und niedrige Varianz



Verbesserungen:

- Berücksichtige die Daten: Lege mehr Zentren an Positionen mit hoher Variabilität und weniger Zentren an Positionen, an denen die Funktion relativ flach ist.

Verbesserungen:

- Berücksichtige die Daten: Lege mehr Zentren an Positionen mit hoher Variabilität und weniger Zentren an Positionen, an denen die Funktion relativ flach ist.
- Betrachte die gewichtete Summe von Partition Models:

$$\widetilde{z}(x) = \sum_{j=1}^{J} \omega_{(j)} \widetilde{z}^{(j)}(x|c^{(j)}) \tag{3}$$

wobei J die Anzahl der Partition Models und $\sum_{j=1}^{J} \omega^{(j)} = 1$

1 Motivation
2 Das kollektive Modell
3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur
6 Anhang

Bayes Schätzer

Wir nutzen die Daten $D = \{y_i, x_i\}_{i=1}^n$ als Update des Parameter(vektors) θ in einem parametrischen Modell

Nach dem Theorem von Bayes gilt:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} \tag{4}$$

wobei

- $p(\theta)$ a priori Verteilung
- $p(D|\theta)$ Likelihood Funktion
- $p(\theta|D)$ a posteriori Verteilung

Bayes Schätzer

Wir nutzen die Daten $D = \{y_i, x_i\}_{i=1}^n$ als Update des Parameter(vektors) θ in einem parametrischen Modell

Nach dem Theorem von Bayes gilt:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} \tag{4}$$

wobei

- $p(\theta)$ a priori Verteilung
- $p(D|\theta)$ Likelihood Funktion
- $p(\theta|D)$ a posteriori Verteilung

Die Berechenbarkeit des Integrals hängt von der a priori Verteilung ab
$\Rightarrow$ Verwende konjugierte a priori Verteilung (a priori Verteilung und a posteriori Verteilung gehören zur selben Verteilungsfamilie)

Hierarchischer Bayes Schätzer:

Zerlegt die a priori Verteilung in bedingte Verteilungen:

$$p_1(\theta|\theta_1), p_2(\theta_1|\theta_2), \ldots, p_n(\theta_{n-1}|\theta n) \tag{5}$$

mit

$$p(\theta) = \int_{\Theta_1 \times \ldots \times \Theta_n} p_1(\theta|\theta_1) p_2(\theta_1|\theta_2) \ldots p_{n+1}(\theta_n) d\theta_1 \ldots d\theta_{n+1} \tag{6}$$

$\Rightarrow$ Steigert die Robustheit des Schätzers

Ziel

Predicitve Inference

$$\text{Bestimme } p(y|x, D)$$

Vorgehensweise im Partition Model:

- Was ist gegeben, was gesucht?
- a priori Verteilung
- Likelyhood-Funktion
- Marginal Likelyhood

Im Partition Model:

Gegeben sind die Daten $D = \{y_i, x_i\}_{i=1}^{n}$

Im Partition Model:

Gegeben sind die Daten $D = \{y_i, x_i\}_{i=1}^n$

Im Modell sind folgende Parameter unbekannt:

- Die Lage der Zentren c $= (c_1, ..., c_k)$
- $\mu = (\mu_1, ..., \mu_k)$
- Die Länge k der Vektoren c und μ
- Regressionsvariable σ^2

Die a priori Verteilung:

Hierarchischer Bayes Schätzer: Zerlegt die a priori Verteilung in bedingte Verteilungen

$$p(\theta) = p(k, c, \mu, \sigma^2) = p(\mu|c, \sigma^2)p(\sigma^2)p(c|k)p(k) \tag{7}$$

Wir wählen:

$$p(k) = \frac{1}{K}, \quad k = 1, ..., K \tag{8}$$

$$p(c|k) = \frac{1}{\{Area(\tau)\}^k}, \quad c \in \tau^k \tag{9}$$

Die a priori Verteilung:

Hierarchischer Bayes Schätzer: Zerlegt die a priori Verteilung in bedingte Verteilungen

$$p(\theta) = p(k, c, \mu, \sigma^2) = p(\mu|c, \sigma^2)p(\sigma^2)p(c|k)p(k) \tag{7}$$

Wir wählen:

$$p(k) = \frac{1}{K}, \quad k = 1, ..., K \tag{8}$$

$$p(c|k) = \frac{1}{\{Area(\tau)\}^k}, \quad c \in \tau^k \tag{9}$$

$$p(\sigma^{-2}) = Gamma(\sigma^{-2}|a, b), \quad \sigma^2 > 0 \tag{10}$$

Die a priori Verteilung:

Hierarchischer Bayes Schätzer: Zerlegt die a priori Verteilung in bedingte Verteilungen

$$p(\theta) = p(k, c, \mu, \sigma^2) = p(\mu|c, \sigma^2)p(\sigma^2)p(c|k)p(k) \tag{7}$$

Wir wählen:

$$p(k) = \frac{1}{K}, \quad k = 1, ..., K \tag{8}$$

$$p(c|k) = \frac{1}{\{Area(\tau)\}^k}, \quad c \in \tau^k \tag{9}$$

$$p(\sigma^{-2}) = Gamma(\sigma^{-2}|a, b), \quad \sigma^2 > 0 \tag{10}$$

$$p(\mu|c, \sigma^2) = \prod_{i=1}^{k} N(\mu_i|0, \sigma^2 v), \quad \mu \in \Re^k \tag{11}$$

Die Likelyhood-Funktion

Außerdem können wir $p(D|\theta)$ bestimmen:

- $\theta = (c, \mu, \sigma^2)$
- $y_i = \overline{y} + z(x_i) + \epsilon_i$ für i = 1,...,n
- $\epsilon_i \sim N(0, \sigma^2)$

Die Likelyhood-Funktion

Außerdem können wir $p(D|\theta)$ bestimmen:

- $\theta = (c, \mu, \sigma^2)$
- $y_i = \overline{y} + z(x_i) + \epsilon_i$ für i = 1,...,n
- $\epsilon_i \sim N(0, \sigma^2)$

Damit gilt:

$$p(D|\theta) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \overline{y} - \mu_{r(i)})^2\right\} \qquad (12)$$

k = 1,...,K, und r(i) Index des i-ten Gebiets

Marginal Liklihood

Da es sich um konjugierte Prior handelt, können wir die Likelihood Funktion vereinfachen:

$$\begin{aligned} P(D|c) &= \int_0^\infty \int_{\Re^k} p(D|\mu,\sigma^2,c)p(\mu|\sigma^2,c)p(\sigma^2)d\mu d\sigma^2 \\ &= \frac{(2b)^a\Gamma(a^*)}{\pi^{n/2}\Gamma(a)} \frac{(2b^*)^{-a^*}}{\prod_1^k (n_i+1)^{1/2}}, \end{aligned}$$

für k = 1,...,K

$a^* = \mathrm{a} + \mathrm{n}/2,$

$b^* = b + \frac{1}{2}\sum_{i=1}^k \left\{ \sum_{j:r(j)=i}(y_j - \overline{y})^2 - \frac{vn_i^2}{1+vn_i}\overline{y_i})^2 \right\}$

$n_i = \sum_{j:r(j)=i} 1,$

$\overline{y_i} = n_i^{-1} \sum_{j:r(j)=i}(y_i - \overline{y})$

Vorhersage der Schadenhöhen an vorgegebenen Positionen

Sei $\epsilon \sim N(0, \sigma^2 I)$
Da $y = \overline{y} + z(x) + \epsilon$

$$\Rightarrow p(y|x, c, \mu, \sigma^2, D) = N(y|\mu_x, \sigma^2) \tag{13}$$

Vorhersage der Schadenhöhen an vorgegebenen Positionen

Sei $\epsilon \sim N(0, \sigma^2 I)$
Da $y = \overline{y} + z(x) + \epsilon$

$$\Rightarrow p(y|x, c, \mu, \sigma^2, D) = N(y|\mu_x, \sigma^2) \quad (13)$$

Durch Integration nach μ, σ^2 erhalten wir:

$$p(y|x, c, D) = Student\left\{\frac{n_x \overline{y_x}}{n_x + v^{-1}}, b^*(1 + \frac{1}{n_x + v^{-1}}), a^*\right\} \quad (14)$$

wobei n_x die Anzahl der Schäden y_i beschreibt, die im selben Gebiet wie x liegen

$$p(y|x, D) = \int p(y|x, c, D)p(c|D)dc \tag{15}$$

Wir wissen:

$$p(y|x, c, D) = Student\left\{\frac{n_x\overline{y_x}}{n_x + v^{-1}}, b^*(1 + \frac{1}{n_x + v^{-1}}), a^*\right\} \tag{16}$$

$$p(y|x, D) = \int p(y|x, c, D)p(c|D)dc \tag{15}$$

Wir wissen:

$$p(y|x, c, D) = Student\left\{\frac{n_x\overline{y_x}}{n_x + v^{-1}}, b^*(1 + \frac{1}{n_x + v^{-1}}), a^*\right\} \tag{16}$$

$$p(c|D) \propto p(D|c)p(c) \tag{17}$$

$$\begin{aligned} p(c) &= \frac{1}{K}\frac{1}{\{Area(\tau)\}}, \quad c \in \bigcup_{k=1}^{K} \tau^k \\ p(D|c) &= \frac{(2b)^a\Gamma(a^*)}{\pi^{n/2}\Gamma(a)}\frac{(2b^*)^{-a^*}}{\prod_1^k(n_i + 1)^{1/2}} \end{aligned}$$

Problem

Das Integral ist nicht analytisch zu bestimmen.

Problem

Das Integral ist nicht analytisch zu bestimmen.

Idee

Erzeuge Zentren $c^{(1)},\ldots,c^{(J)}$ als Realisierungen von $p(c|D)$

Problem

Das Integral ist nicht analytisch zu bestimmen.

Idee

Erzeuge Zentren $c^{(1)},\ldots,c^{(J)}$ als Realisierungen von $p(c|D)$

$$\tilde{y}(x) = \overline{y} + \tilde{z}(x) \tag{18}$$

mit $\tilde{z}(x) = \frac{1}{J}\sum_{j=1}^{J} \tilde{z}^{(j)}(x)$,
wobei $\tilde{z}^{(j)}(x) = E(y|x, c^{(j)}, D) = \frac{n_x \overline{y_x}}{n_x + v^{-1}}$

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Wiederholung(1)

- Für den Algorithmus wird eine Markov-Kette benötigt
- Geben zunächst eine Wiederholung der wichtigsten Begriffe
- Der Übersicht halber betrachten wir nur den Fall eines endlichen Zustandsraumes E = {1,2,...,I}

Wiederholung(2)

Definition: Markov-Kette

- Sei $X_0, X_1, ... : \Omega \rightarrow$ E eine Folge von Zufallsvariablen über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit Werten in E = {1,2,...,I}
- X_i heißt (homogene) Markov-Kette mit Anfangsverteilung $\alpha = (\alpha_1, ..., \alpha_I)$ und der Übergangsmatrix P = (p_{ij}),falls

$$P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = \alpha_{i_0} p_{i_0 i_1} ... p_{i_{n-1} i_n} \qquad (19)$$

für beliebige n=0,1,... und $i_0, i_1, ..., i_n \in$ E gilt

Wiederholung(3)

- α_i = Wahrscheinlichkeit, dass sich das betrachtete System zur Zeit n = 0 im Zustand i befindet
- $\alpha_i \in [0,1]$, $\sum_{i=1}^{I} \alpha_i = 1$
- I × I Matrix $P = (p_{ij})_{i,j=1,\ldots,I}$ der Übergangswahrscheinlichkeiten mit

$$p_{ij} > 0, \qquad \sum_{j=1}^{I} p_{ij} = 1 \quad \forall i \tag{20}$$

Wiederholung(4)

Definition: Stationäre Verteilung

Eine Verteilung μ heißt stationär, wenn sie der Gleichung

$$\mu^T = \mu^T P, \quad \mu_i \geq 0, \quad \sum_i \mu_i = 1 \tag{21}$$

genügt

Wiederholung(5)

Bemerkung

- Definition gilt nur für endliche Zustandsräume
  $\Rightarrow$ Nur bei Simulation auf einem Gitter gegeben
- Soll im Beobachtungsraum χ stetig simuliert werden ist eine Verallgemeinerung des Begriffs der Markow-Kette für überabzählbare Zustandsräume nötig
  $\Rightarrow$ Markov-Prozesse

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Problemstellung:

$\widetilde{z}(x)$ in einem Partition Model aus den Daten zu berechnen

Vorgehensweise:

- Verwenden Metropolis-Hastings Algorithmus
- Erhalten eine Markov-Kette von Tesselationen des Beobachtungsraumes χ
- stationäre Verteilung dieser Markov-Kette ist a posteriori Verteilung $p(c|D)$
- Berechnung von $\widetilde{z}(x)$ mit Hilfe von $p(c|D)$
- Es genügt die Lage der Zentren zu simulieren $\Rightarrow$ Tesselation eindeutig bestimmt
- Variable Anzahl von Zentren $\Rightarrow$ Reversible-Jump-Sampler

Bezeichnungen im Algorithmus:

- k: aktuelle Anzahl der Zentren
- K: maximale Anzahl der Zentren
- d_k: Wahrscheinlichkeit für Death-Step
- b_k: Wahrscheinlichkeit für Birth-Step
- $m_k = (1 - d_k - b_k)$: Wahrscheinlichkeit für Move-Step

Algorithmus(1)

Wähle Startwert $c^{(0)}$; Setze t $= 0$;
REPEAT
Erzeuge u_1 gemäß U(0,1) Verteilung;
IF $u_1 \leq b_k$
c' = Birth-proposal($c^{(t)}$);
ELSE IF $b_k < u_1 \leq b_k + d_k$
c' = Death-proposal($c^{(t)}$);
ELSE
c' = Move-proposal($c^{(t)}$);
ENDIF;

Algorithmus(2)

Erzeuge u_2 gemäß U(0,1) Verteilung;

IF $u_2 < min\{1, BF(c', c^{(t)})\}$=min$\{1, \frac{p(D|c')}{p(D|c^{(t)})}\}$

$c^{(t+1)} = c'$;

ELSE

$c^{(t+1)} = c^{(t)}$;

ENDIF;

t = t+1;

Nehme jeden m-ten Wert von $c^{(t)}$ nach burn-in;

END REAPEAT;

Birth-Proposal($c^{(t)}$)

- Erzeuge z gemäß U$\{\tau \backslash$c$\}$ Verteilung
- Setze c' = ($c^{(t)}$,z)

Death-Proposal($c^{(t)}$)

- Erzeuge z gemäß U$\{$1,...,k$\}$ Verteilung
- Setze c' = $(c^{(t)}_{-z})$

Notation: $x_{-z} = (x_1, ..., x_{z-1}, x_{z+1}, ..., x_k)$

Move-Proposal($c^{(t)}$)

- c^* = Death-Proposal($c^{(t)}$)
- c^{**} = Birth-Proposal(c^*)
- Setze c' = c^{**}

Der Bayes-Faktor

$\frac{p(D|c')}{p(D|c^{(t)})}$ nur definiert, wenn $p(D|c^{(t)}) > 0$

$\Rightarrow$ wähle Startwert $c^{(0)}$ so, dass $p(D|c^{(0)}) > 0$

$\Rightarrow$ $p(D|c^{(t)}) > 0 \ \forall t > 0$, da $p(D|c') = 0$ zur Ablehnung führt.

Der Bayes-Faktor: Berechnung

Für den Bayes-Faktor BF gilt:

$$\mathrm{BF}(c', c^{(t)}) = \frac{p(D|c')}{p(D|c^{(t)})} =$$

$$= \frac{(2b)^a \Gamma(a^*)}{\pi^{n/2}\Gamma(a)} \frac{(2(b')^*)^{-a^*}}{\prod_{i=1}^{k}(n_i'+1)^{1/2}} \Big\backslash \frac{(2b)^a \Gamma(a^*)}{\pi^{n/2}\Gamma(a)} \frac{(2(b^{(t)})^*)^{-a^*}}{\prod_{i=1}^{k}(n_i^{(t)}+1)^{1/2}} =$$

$$= \frac{(2(b')^*)^{-a^*}}{\prod_{i=1}^{k}(n_i'+1)^{1/2}} \frac{\prod_{i=1}^{k}(n_i^{(t)}+1)^{1/2}}{(2(b^{(t)})^*)^{-a^*}} =$$

$$= \left(\frac{(b^{(t)})^*}{(b')^*}\right)^{a^*} * \frac{\prod_{i=1}^{k}(n_i^{(t)}+1)^{1/2}}{\prod_{i=1}^{k}(n_i'+1)^{1/2}}$$

Anmerkung: Parameter a,b und $a^* = a + n/2$ hängen nicht von der Wahl der Zentren c ab!!!

Wahl der Parameter:

- $b_K = d_0 = m_0 = 0$
- $b_0 = 1$
- $d_K = m_K = \frac{1}{2}$
- $b_k = d_k = m_k = \frac{1}{3}$ für alle k $\neq$ 0,K

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Definition Equivalent Kernel

Für ein x und eine Folge von Zentren $c^{(1)}, .., c^{(J)}$ ist der equivalent Kernel gegeben durch:

$$K(x,x') = \frac{1}{J} \sum_{j=1}^{J} \frac{\mathbb{1}(r_{x'}^{(j)} = r_x^{(j)})}{Area^{(j)}(r_x^{(j)})}, \qquad x' \in \chi$$

wobei:

- $r_x^{(j)}$: Index der Region in welcher x bei Tesselation j liegt, analog für $r_{x'}^{(j)}$
- $Area^{(j)}(i)$: Flächeninhalt des i-ten Gebietes bei der j-ten Tesselation

$\rightarrow$ Equivalent Kernel Kenngröße für die Korrelation der beiden Punkte x und x'

1 Motivation

2 Das kollektive Modell

3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur

6 Anhang

Datensatz

- Lage und Höhe der jährlichen Sturmschäden in Österreich
- betrachtete Jahre: 1996 - 2006

Aufgabenstellung

- Modellierung der Schadenshöhe
- Untersuchung von Abhängigkeiten zwischen den Schadensfällen
- Verwenden Reversible-Jump-Sampler zur Simulation der a posteriori Verteilung $p(c|D)$
- Berechnen $\widetilde{z}(x)$ als Schätzer für die Schadenshöhe am Punkt x
- Wähle χ als Rechteck, welches Österreich einschließt

Beispiel: eine vom Reversible-Jump-Sampler erzeugte Tesselation

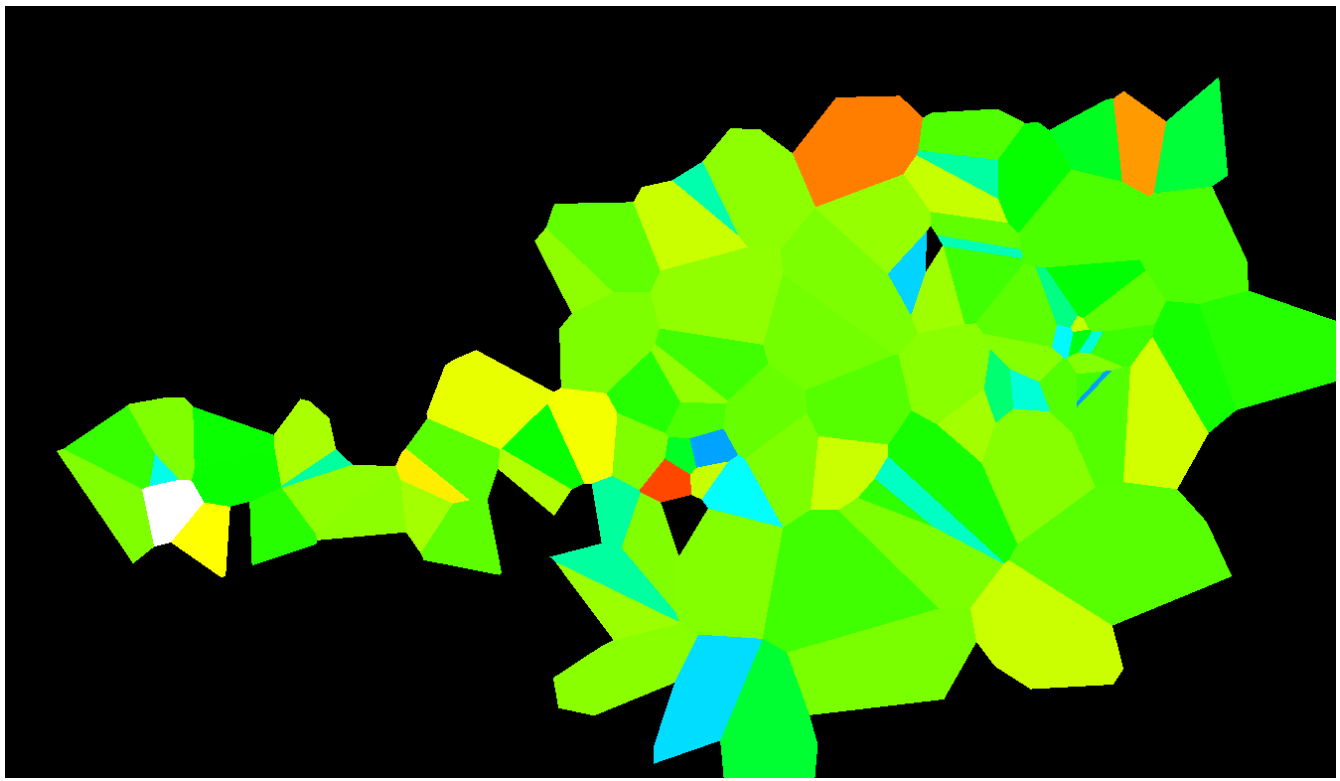


Abbildung: logarithmische Durchschnitts-Schäden in den einzelnen Gebieten einer Tesselation im Jahr 1999

Beispiel: eine weitere vom Reversible-Jump-Sampler erzeugte Tesselation

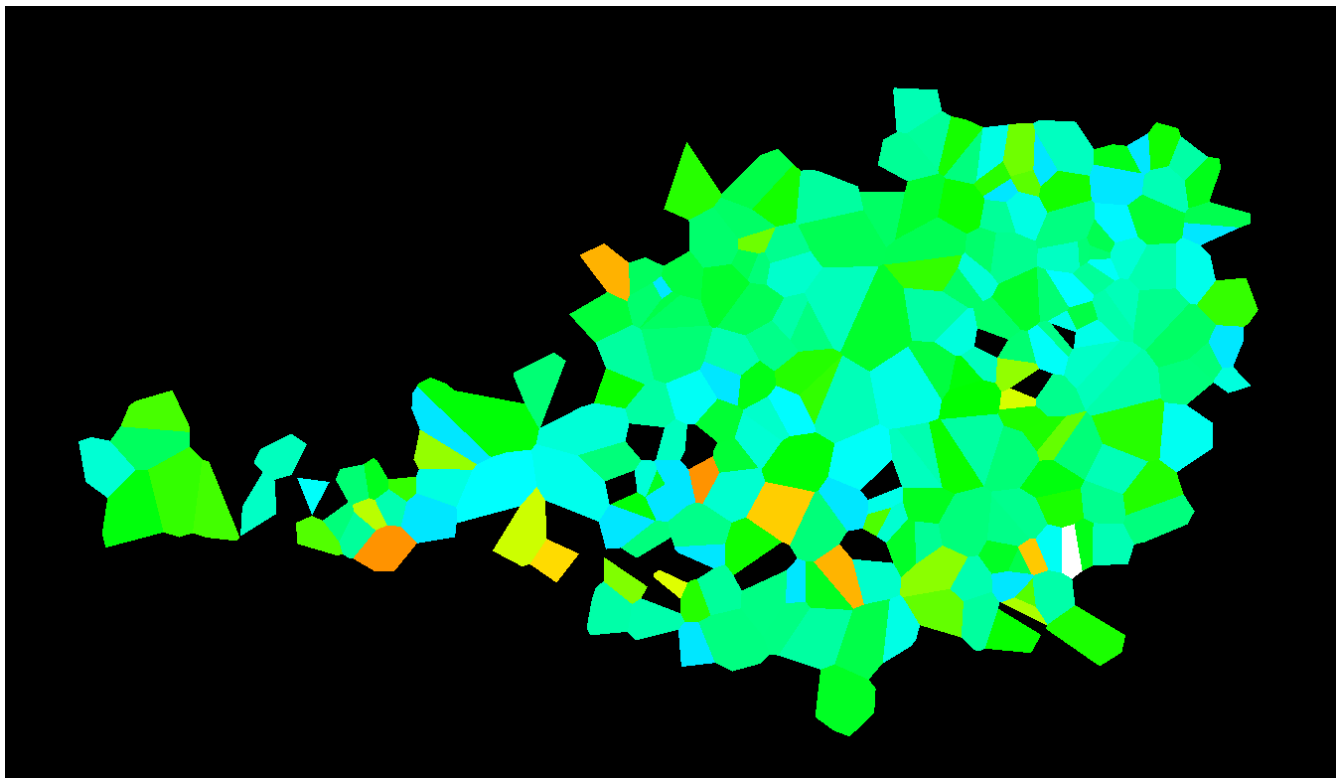


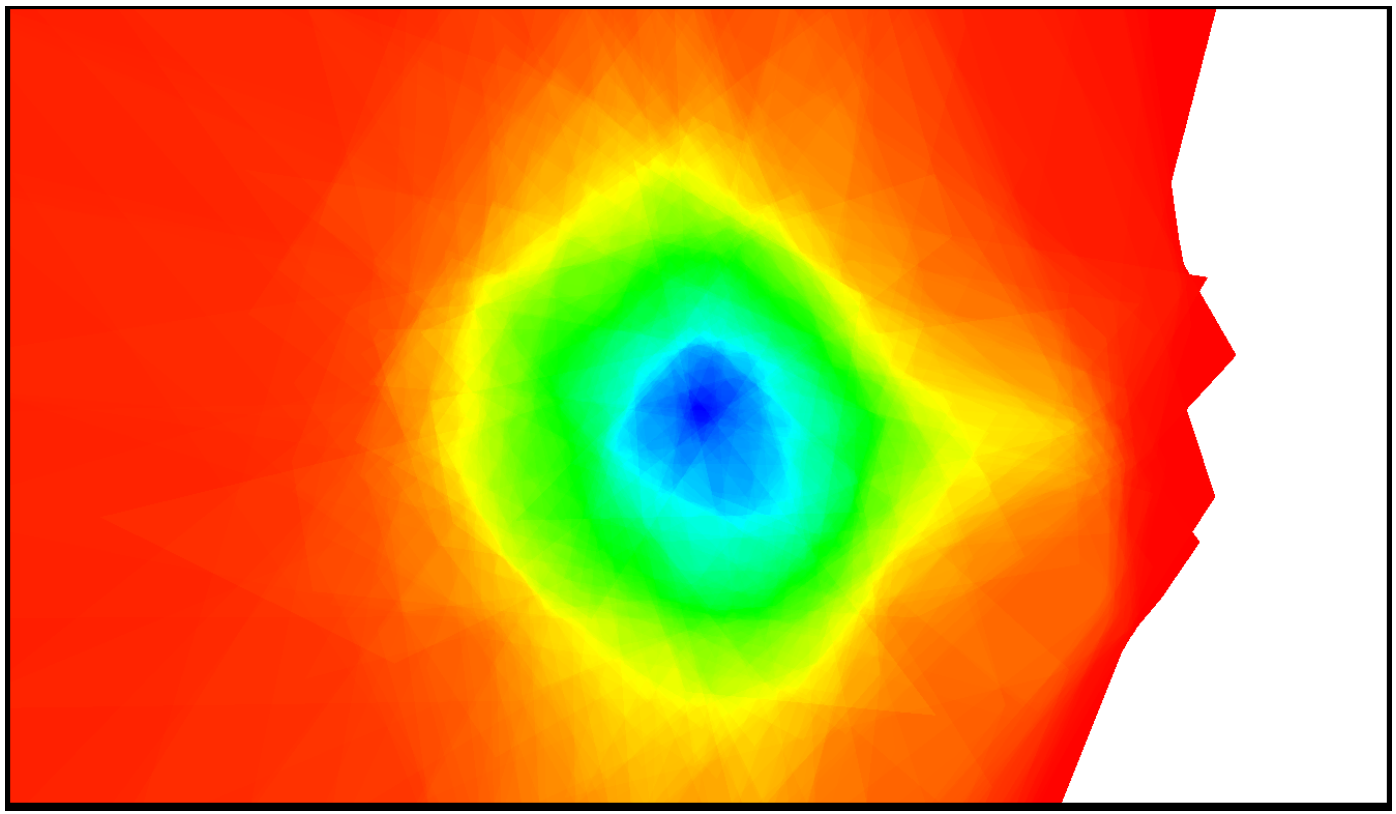
Abbildung: logarithmische Durchschnitts-Schäden in den einzelnen Gebieten einer Tesselation im Jahr 2002

Bestimmung der Abhängigkeiten

- Equivalent Kernel K(x,x') Kenngröße für die Abhängigkeit der Punkte x und x'
- Um die Abhängigkeit eines Punktes x von der Umbebung darzustellen legt man über χ ein Gitter von Punkten
- Berechnen K(x,x') für alle x' aus dem Gitter
- Der Flächeninhalt einer Region wird mit der Summe der Gitterpunkte in dieser Region approximiert
- Interpretation von K(x,x') als Pixelwert $\rightarrow$ farbliche Darstellung der Abhängigkeit
- Im Folgenden steht ein Pixel für ein Fläche von 500 * 500 Metern

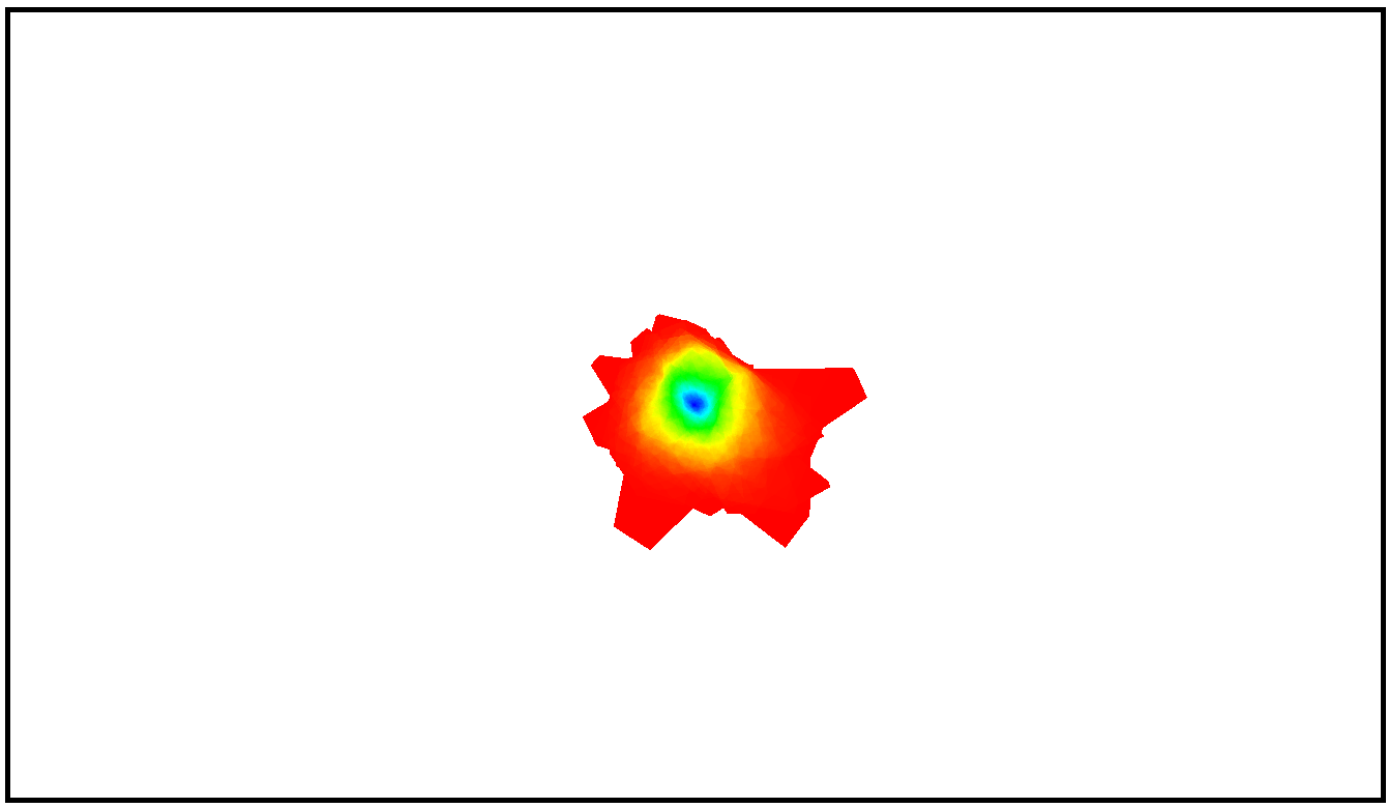
Kernel 1 bei im Jahr 1996

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



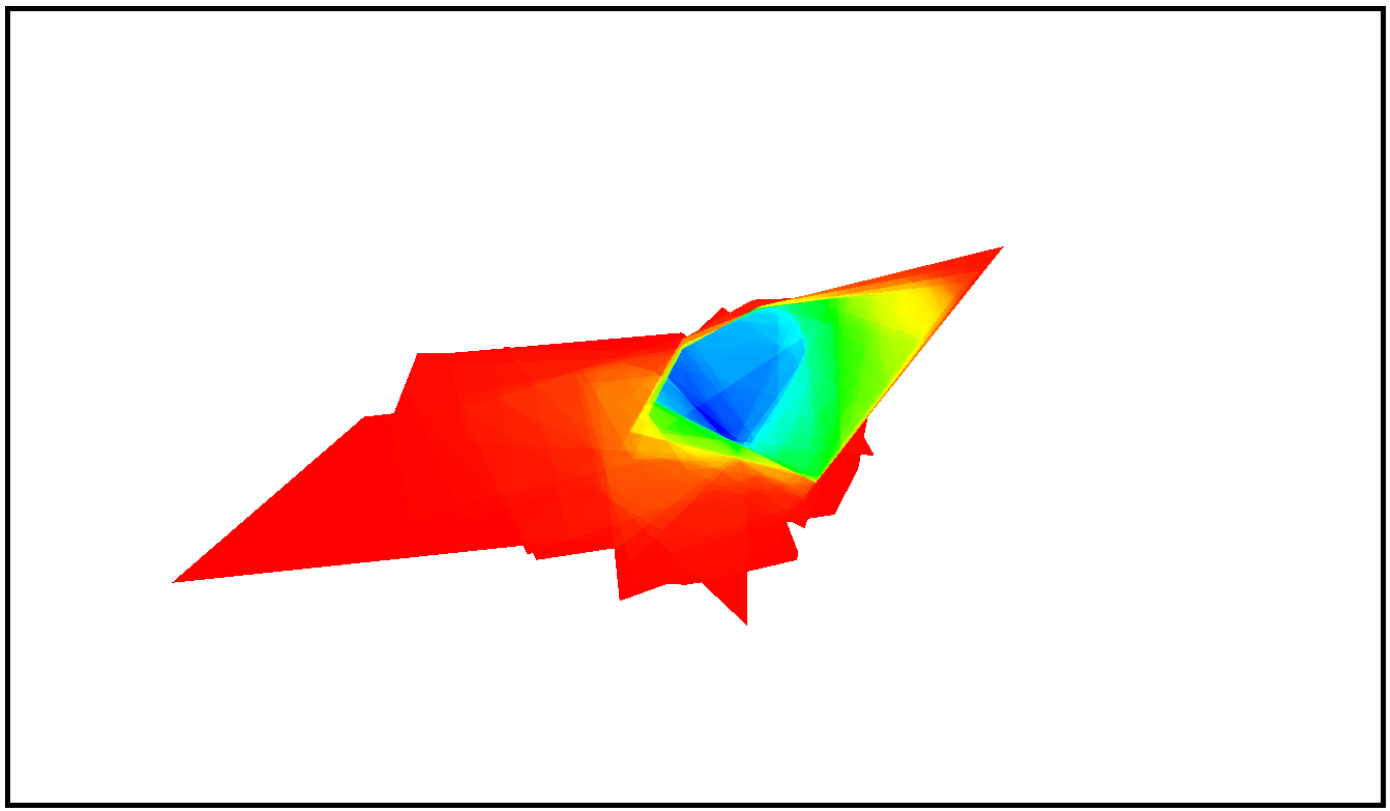
Zum Vergleich: Kernel 1 bei im Jahr 1999

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



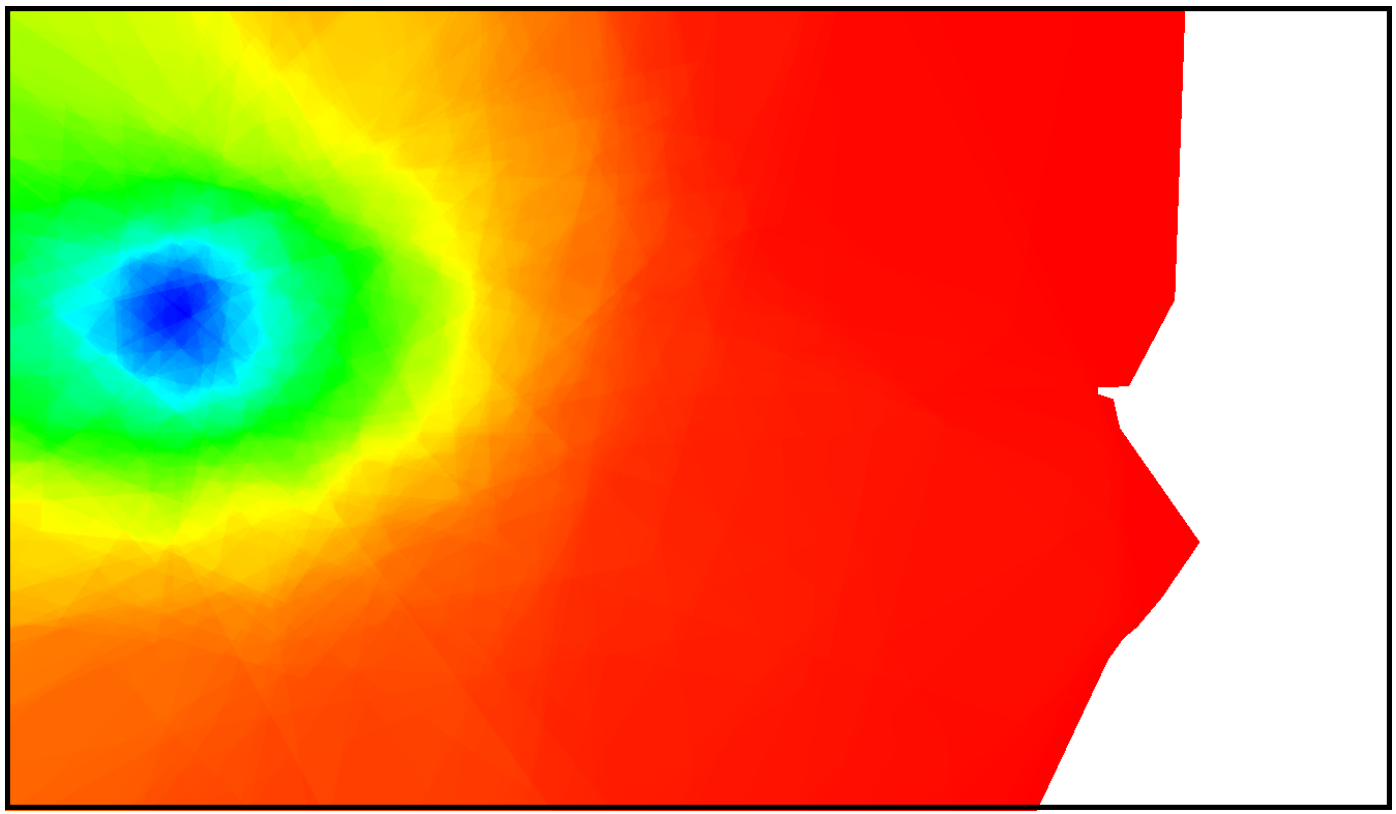
Kernel 1 bei im Jahr 2002 - Achtung: Weniger Iterationen

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



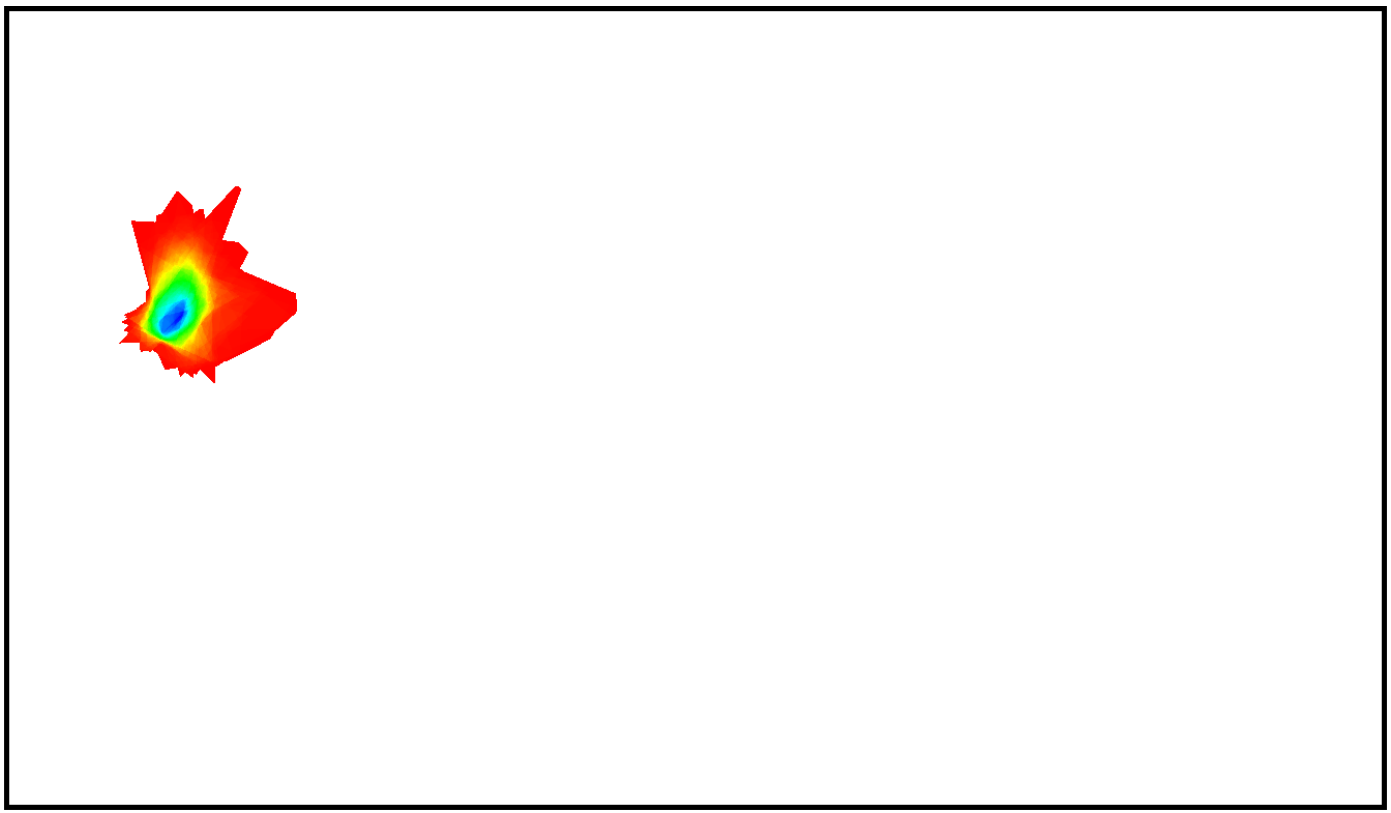
Kernel 2 im Jahr 1996

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



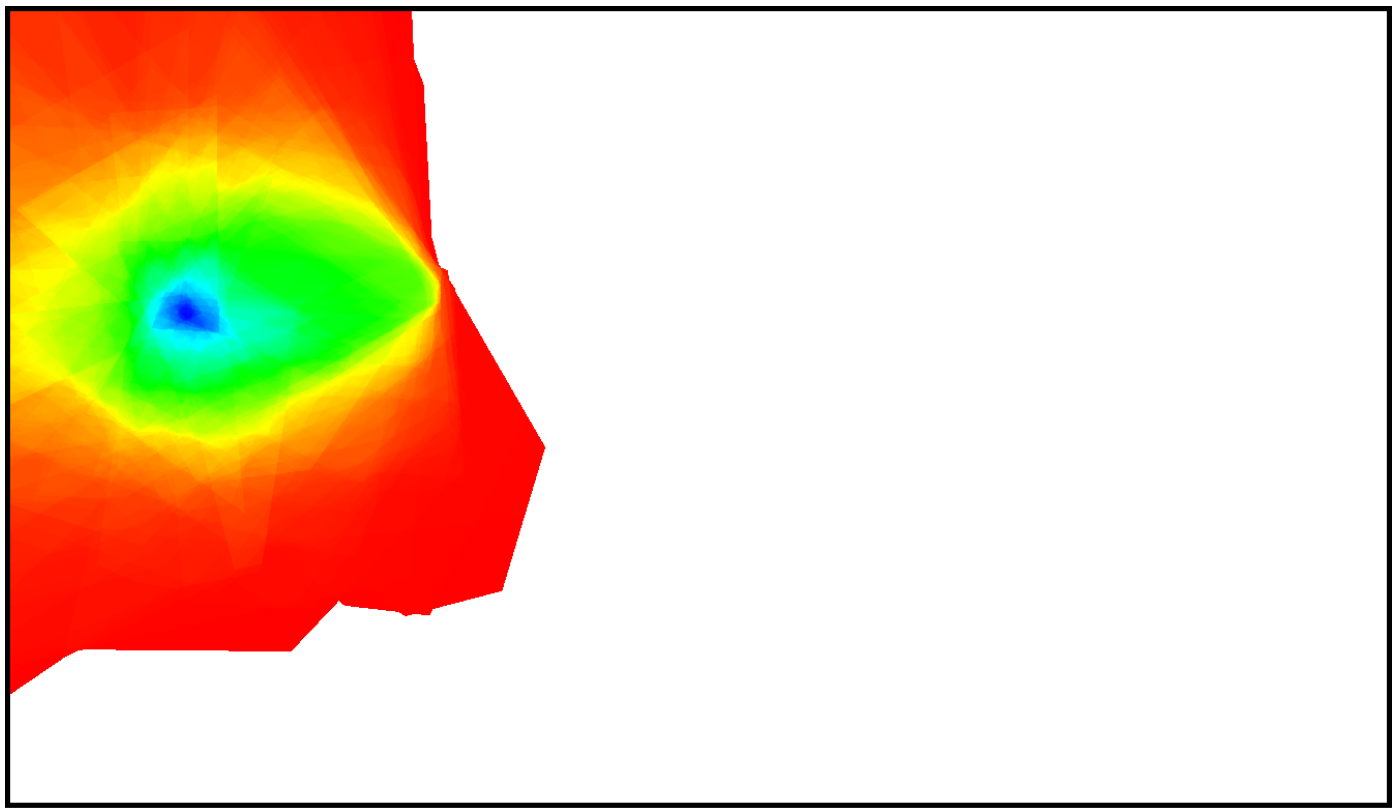
Zum Vergleich: Kernel 2 im Jahr 1999

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



Kernel 2 im Jahr 2002 - Achtung weniger Iterationen

blau: hohe Abhängigkeit bis rot: niedrige Abhängigkeit
weiß: keine Abhängigkeit



1 Motivation
2 Das kollektive Modell
3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur
6 Anhang

Literatur

- David G.T. Dension, Andrew B. Lawson
  Spatial Cluster Modelling, Chapter 7
  Chapmann & Hall/CRC
- David G.T. Dension and Christopher C.Holmes
  Bayesian Methods for Nonlinear Classification and Regression
  John Wiley and sons, LTD
- Christian P. Robert
  The Bayesian Choise A Decision-Thearetic Motivation
  Springer Texts in Statistics, 1997
- Volker Schmidt
  Vorlesungsskript Wahrscheinlichkeitstheorie
  Ulm, im Juli 2006

1 Motivation
2 Das kollektive Modell
3 Partition Modelling
- Problemstellung
- Das Voronoi Mosaik
- Partition Model
- Bayes Schätzer

4 Simulationsalgorithmus
- Markov-Ketten
- Reversible-Jump-Sampler
- Equivalent Kernel
- Beispiel

5 Literatur
6 Anhang

Die Dichte der multivariaten p-dimensionalen Student Verteilung mit Parametern Σ, μ, n:

$$\frac{\Gamma(\frac{n+p}{2})}{\Gamma(\frac{n}{2})n^{\frac{p}{2}}\pi^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}(1+\frac{1}{n}(x-\mu)^T\Sigma^{-1}(x-\mu))^{\frac{n+p}{2}}} \tag{22}$$