

Bayessche Ansätze in der Geostatistik

Seminar

Bayessche Ansätze in der Bildanalyse

Bernhard Clasen

Fakultät für Mathematik und Wirtschaftswissenschaften

3. Juli 2006



Übersicht

1. Einführung und Motivation für modell-basierte Geostatistik
2. Anwendungen des Bayesschen Paradigmas bei räumlichen Daten
 - i. Modelle stationärer, räumlicher Prozesse
 - ii. Modelle mit Flächendaten
3. Modell-basierte Geostatistik
 - I. Das Gauß-Modell
 - II. Parametrische Schätzung der Kovarianz-Struktur
 - III. Prognosen durch Plug-In
 - IV. Bayessche Folgerungen für das Gauß-Modell
4. Eine Fallstudie: Schweizer Regenfall-Daten



Einführung

- Fokus liegt auf der Untersuchung räumlicher Phänomene $s(x)$
- $s(x)$ Realisierungen eines stochastischen Prozesses $S(\bullet) = \{S(x) : x \in A\}$
- im Allgemeinen ist $S(\bullet)$ nicht beobachtbar



Einführung

Erhebung und Interpretation der Daten

- Daten bestehen aus Messungen Y_1, \dots, Y_n an Orten s_1, \dots, s_n
- Y_i ist verrauschte Version von $S(s_i)$
- Das sog. „sampling Design“ s_1, \dots, s_n ist deterministisch oder stochastisch, aber unabhängig vom Prozess $S(-)$



Motivation

- ursprünglich kommt die Geostatistik aus dem Bereich der Bergbauindustrie
- erstmals im Zusammenhang mit Schätzungen von Erzreserven geostatistische Methoden angewandt
- vorerst unabhängige Entwicklung von der eigentlichen räumlichen Statistik
- Cressie (1993) bezeichnet Geostatistik als ein der drei Unterbereiche der räuml. Statistik
- der Begriff „modell-basierte Geostatistik“ wurde 1998 durch Diggle, Tawn und Moyeed geprägt



Motivation

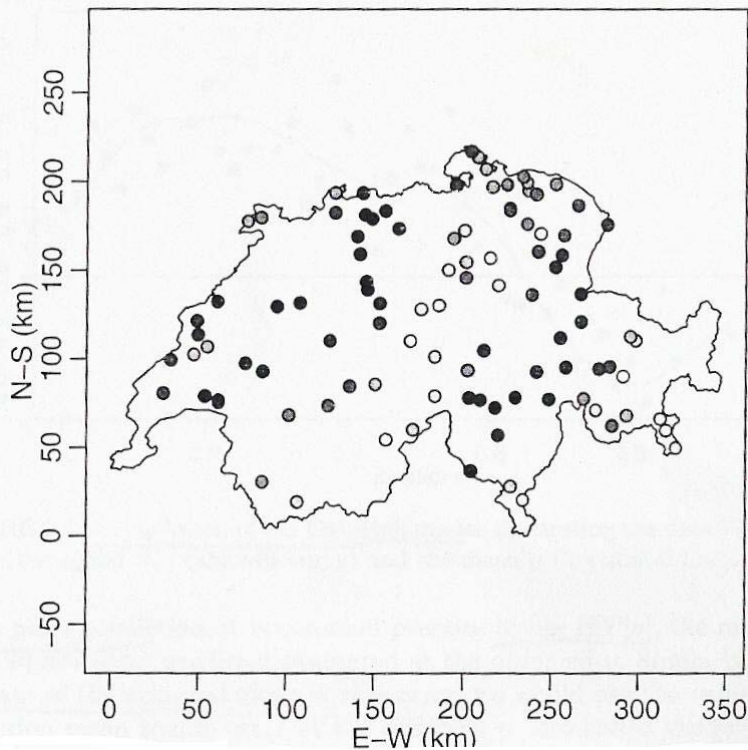
Zwei einführende Beispiele für geostatistische Probleme

1. Schweizer Regenfall-Daten:

- Daten bestehen aus Messungen des Regens am 8. Mai 1986 an 467 Orten in der Schweiz
- Ziel: eine kontinuierliche räumliche Karte der Niederschlagswerte zu schaffen, wobei man eine diskrete Menge von Messorten zugrunde legt
- 100 Messorte werden dazu genutzt, um ein Modell zu erstellen und Aussagen machen zu können; die restlichen 367 dazu, diese Aussagen/Prognosen empirisch zu überprüfen



Motivation



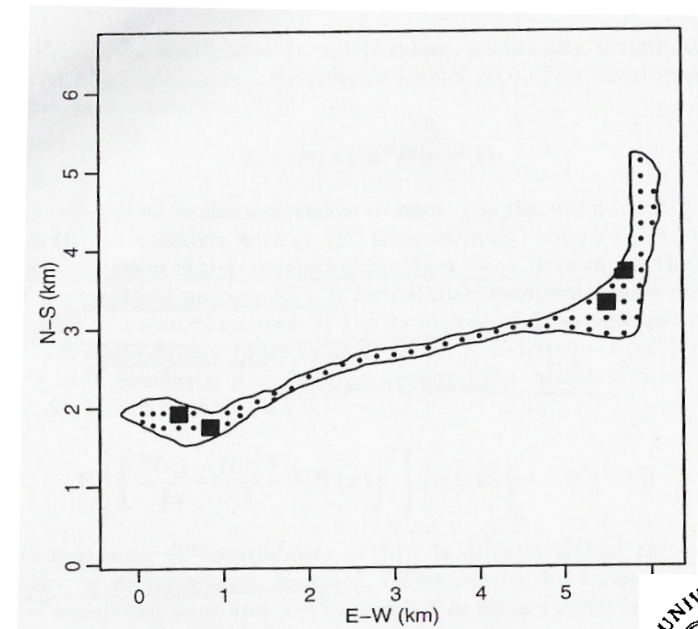
Schweizer Niederschlags-Daten
an den Orten s_1, \dots, s_{100} .

Die Grauskalierung von weiß
(wenig) bis schwarz (viel) gibt
die Werte des beobachteten
Regens wieder

Motivation

2. Residuale Kontamination der Rongelap Insel:

- Messungen auf Rongelap (Pazifik) im Zuge der Atomtests der USA in der 50er Jahren
- Ziel: die räumliche Variation der residualen Kontamination auf der Insel zu schätzen
- Verrauschte Messung Y_i der radioaktiven Cäsium-Konzentration wurde zunächst in 200-Meter Abständen s_i , dann in Quadraten mit 40m Abständen erhoben



Anwendungen des Bayesschen Paradigmas bei räumlichen Daten

1. Modelle stationärer, räumlicher Prozesse

- Isotropes Modell
- Anisotropes Modell

2. Modelle mit Flächendaten



Modelle stationärer, räumlicher Prozesse

Das grundlegende Modell:

$$Y(s) = \mu(s) + \omega(s) + \varepsilon(s)$$

wobei $\mu(s) = \mathbf{x}^T(s)\beta$ (Mittelwert)

- $\omega(s)$: Realisierung eines null-zentrierten, stationären Gauß-Prozesses
- $\varepsilon(s)$: unkorrelierter Fehlerterm, Varianz τ^2



Modelle stationärer, räumlicher Prozesse

Verschiedene Interpretationen von $\varepsilon(s)$:

- reiner Fehlerterm (im Gegensatz zu $\omega(s)$):

τ^2 wird als Komponente der Varianz von $Y(s)$ gesehen

keine räumlich stetigen Residuen, um einen
Messfehler auszudrücken, da zusätzliche Variabilität
bei $Y(s)$ befürchtet wird („Rauschen“).

- räumlicher Prozess :

τ^2 repräsentiert die sog. „microscale“ Variabilität

Cressie: $\varepsilon(s)$ und τ^2 in zwei Teile geteilt; ein Teil reflektiert den reinen
Fehler, der andere die „microscale“ Variabilität

Modelle stationärer, räumlicher Prozesse

Isotropes Modell

- Daten: $Y = (Y(s_1), \dots, Y(s_n))^T$
- Problem: passende Definition der Matrix $\Sigma = \sigma^2 H(\phi) + \tau^2 I$,
wobei $H_{ij} = \rho(s_i - s_j; \phi)$, ρ eine Korrelationsfunktion sei
- Modellparameter: $\theta = (\beta, \sigma^2, \tau^2, \phi)$
- Schätzung für Parameter durch a-posteriori-Verteilung

$$p(\theta | y) \propto f(y | \theta) p(\theta), \text{ wobei}$$

$$Y | \theta \sim N(X\beta, \sigma^2 H(\phi) + \tau^2 I)$$

Modelle stationärer, räumlicher Prozesse

Isotropes Modell

- unabhängige a-priori-Verteilungen, dh.

$$p(\theta) = p(\beta)p(\sigma^2)p(\tau^2)p(\phi)$$

- passende a-priori-Verteilungen :

- β : multivariate Normalverteilung
- σ^2, τ^2 : inverse Gamma-Verteilung
- ϕ : (oft) Gamma-Verteilung (hängt von
Spezifikation von ρ ab)



Modelle stationärer, räumlicher Prozesse

Isotropes Modell

- oftmals separate Betrachtung der Parameter von Interesse
- dazu marginale a-posteriori Verteilung nötig
- ein Beispiel: Punktschätzer bzw. denkbares Intervall für β
erhält man aus

$$p(\beta | y) = \iiint p(\beta, \sigma^2, \tau^2, \phi | y) d\sigma^2 d\tau^2 d\phi$$
$$\propto p(\beta) \iiint f(y | \theta) p(\sigma^2) p(\tau^2) p(\phi) d\sigma^2 d\tau^2 d\phi$$

Modelle stationärer, räumlicher Prozesse

Isotropes Modell

- Prognose für Y an einem neuen Wert s_0 mit zugehörigem Kovarianzvektor $x(s_0)$
- sog. Bayessches „Kriging“
- $Y_0 \equiv Y(s_0), x_0 \equiv x(s_0)$
- Bayessche Lösung besteht aus dem Finden der Prognoseverteilung

$$p(y_0 | y, X, x_0) = \int p(y_0 | y, \theta, x_0) p(\theta | y, X) d\theta$$



Modelle stationärer, räumlicher Prozesse

Anisotropes Modell

- räumliche Korrelation basiert auf dem Separationsvektor zwischen zwei Punkten
- $\rho(h; \phi) = \text{Cov}(Y(s+h), Y(s))$
- prominente Anwendung: *geometrische Anisotropie*
 - Koordinatenraum kann linear transformiert werden in einen isotropen Raum
 - $\rho(h; \phi) = \rho_0(\|Lh\|; \phi)$, wobei L eine dxd Matrix ist (lineare Transformation)
- stationäres Normalmodell:
 - $Y(s) = \mu + \omega(s) + \varepsilon(s); Y \sim N(\mu \mathbf{1}, \sum(\alpha))$, wobei $\alpha = (\tau^2, \sigma^2, B)^T$
 $B = L^T L$ und $\sum(\alpha) = \tau^2 \mathbf{1} + \sigma^2 H((h' B h)^{1/2})$



Modelle stationärer, räumlicher Prozesse

Modelle räumlicher Daten

- häufige Methode: „Disease mapping“

Y_i = beobachtete Anzahl von Fällen einer Krankheit im Staat i (ZV)

E_i = erwartete Anzahl von Fällen einer Krankheit im Staat i (feste Funktionen)

- Annahme:

$$E_i = n_i r \equiv n_i \left(\frac{\sum_i y_i}{\sum_i n_i} \right) \quad n_i = \text{Anzahl von Risikopersonen im Staat } i$$

- Alternativ:

$$E_i = \sum_j n_{ij} r_j \quad n_{ij} = \text{Anzahl der Personen in Altersgruppe } j \text{ die in Region } i \text{ leben mal die Anzahl der Studienjahre}$$



Modelle stationärer, räumlicher Prozesse

Modelle räumlicher Daten

- falls E_i nicht zu groß => Poisson Modell für Y_i

$$Y_i | \eta_i \sim Poi(E_i \eta_i)$$

wobei η_i = relatives Risiko der Krankheit in der Region i

- ML-Schätzer (MLE) für η_i

$$\hat{\eta} \equiv SMR_i = \frac{Y_i}{E_i}$$

Modelle stationärer, räumlicher Prozesse

-Modelle räumlicher Daten

- *Hierarchische Bayes-Methoden:* Versuch, die zugrundeliegenden, relative Risiken $\{n_i, i = 1, \dots, I\}$ abzubilden

- *Poisson-Gamma-Modell*

$$Y_i | \eta_i \sim \text{Poi}(E_i \eta_i), i = 1, \dots, I$$

$$\eta_i \sim G(a, b) \text{ (Gamma Verteilung mit MW } \mu = a/b \text{ und Varianz } \sigma^2 = a/b^2)$$

- passender (Bayesscher) Punktschätzer für η_i

$$E(\eta_i | y) = \omega_i \text{SMR}_i + (1 - \omega_i) \mu$$

$$\text{wobei } \omega_i = E_i / [E_i + (\mu / \sigma^2)]$$



Modell-basierte Geostatistik

1. Allgemeines geostatistisches Modell und Gauß-Modell
2. Parametrische Schätzung der Kovarianz-Struktur
3. Prognosen durch Plug-In
4. Bayessche Folgerungen für das Gauß-Modell



Das allgemeine geostatistische Modell

- Daten (s_i, y_i) ; s_1, \dots, s_n Messorte, y_1, \dots, y_n Messdaten
- $\{s_i : i = 1, \dots, n\}$ „*sampling design*“
- y_i Realisierungen von $Y_i = Y(s_i)$, wobei $Y() = \{Y(s) : s \in A\}$ der sog. *Messungsprozess* ist
- $S() = \{S(s) : s \in A\}$ „*Signal Prozess*“
(Y_i als verrauschte Version von $S(s_i)$)



Das allgemeine geostatistische Modell

Ein geostatistisches Modell ist eine Spezifikation der gemeinsamen Verteilung des Messungsprozesses und des Signalprozesses

$$[S(), Y()] = [Y()|S()][S()]$$

- Prognose integriert in geostatistische Analysen
- Ziel einer Prognose $T=T(S)$
- *Prädiktor* für T ist eine Funktion $\hat{T} = \hat{T}(Y)$
- Prädiktor des minimalen, quadratischen Fehlers im Mittelwert minimiert

$$MSE(\hat{T}) = E[(T - \hat{T})^2]$$



Das allgemeine geostatistische Modell

Proposition

Sei $Var[T] < \infty$, dann ist der minimale MSE-Prädiktor von T

$$\hat{T} = E_T[T | Y]$$

mit dem zugehörigen prognostiziertem MSE

$$E[(T - \hat{T})^2] = E_T Var_T[T | Y]$$



Das allgemeine geostatistische Modell

- für punktuelle Prognosen wird oft $E[T|y]$ verwendet
- als Schätzer für den MSE:

$\text{Var}[T|y]$

Prognosevarianz

- die komplette Lösung eines Prognoseproblems wird in einer Wahrscheinlichkeitsverteilung $[T|y]$ (*Prognoseverteilung*) ausgedrückt. Diese fällt, unter Annahme des Bayesschen Paradigmas, mit der *a-posteriori Verteilung* von T zusammen.



Gauß-Modell

- S stationärer Gauß Prozess mit

$$E[S(s)] = \mu, \text{ Var}[S(s)] = \sigma^2 \text{ und Korrelationsfunktion } \rho(u) = \text{Corr}[S(s), S(s')] \\ , \text{ wobei } u = \|s - s'\|$$

- Verteilung von Y_i bei gegebenem S ist eine Gaußverteilung mit MW $S(s_i)$ und Varianz τ^2
- Äquivalente Formulierung:

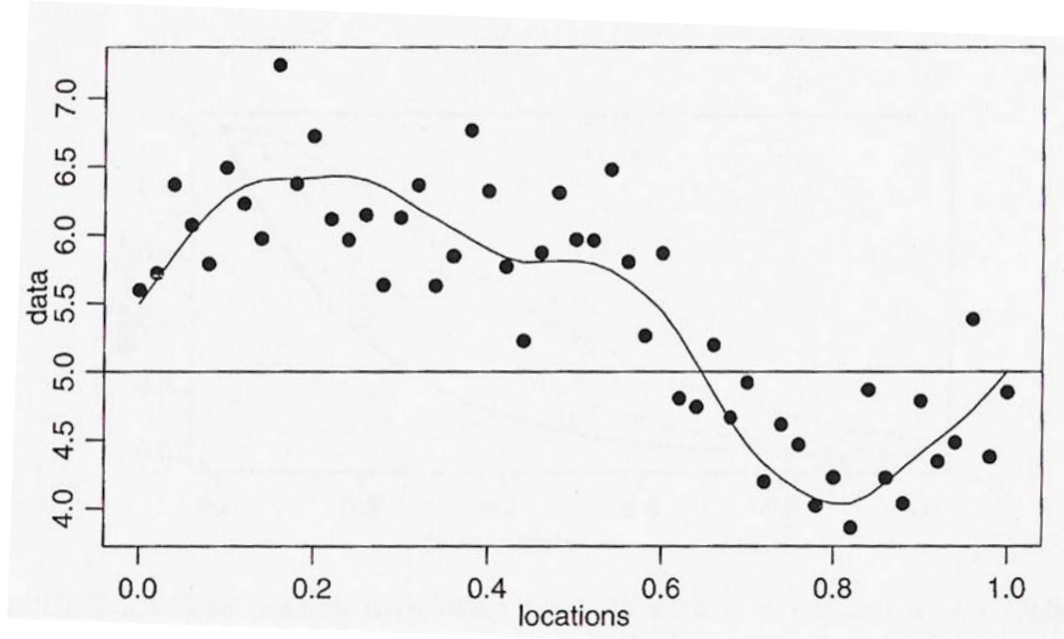
$$Y_i = S(s_i) + Z_i : i = 1, \dots, n$$

wobei Z_1, \dots, Z_n iid sind mit $Z \sim N(0, \tau^2)$ und Y multivariat gauß-verteilt ist mit

$$Y \sim N(\mu \mathbf{1}, \sigma^2 R + \tau^2 \mathbf{I}) \quad , \text{ wobei } R_{ij} = \rho(\|s_i - s_j\|)$$



Gauß-Modell



Simulation eines Gaußschen Modells mit Daten Y_1, \dots, Y_n (Punkte),
Signal $S(\cdot)$ (Kurve) und dem Mittelwert μ (horizontale Linie)

Gauß-Modell

Verschiedene Familien von Korrelationsfunktionen

1. Matérn Familie:

$$\rho(u) = \{2^{k-1} \Gamma(k)\}^{-1} (u/\phi)^k K_k(u/\phi) \quad K(\cdot) \text{ Bessel Funktion}$$

2. Exponential Familie:

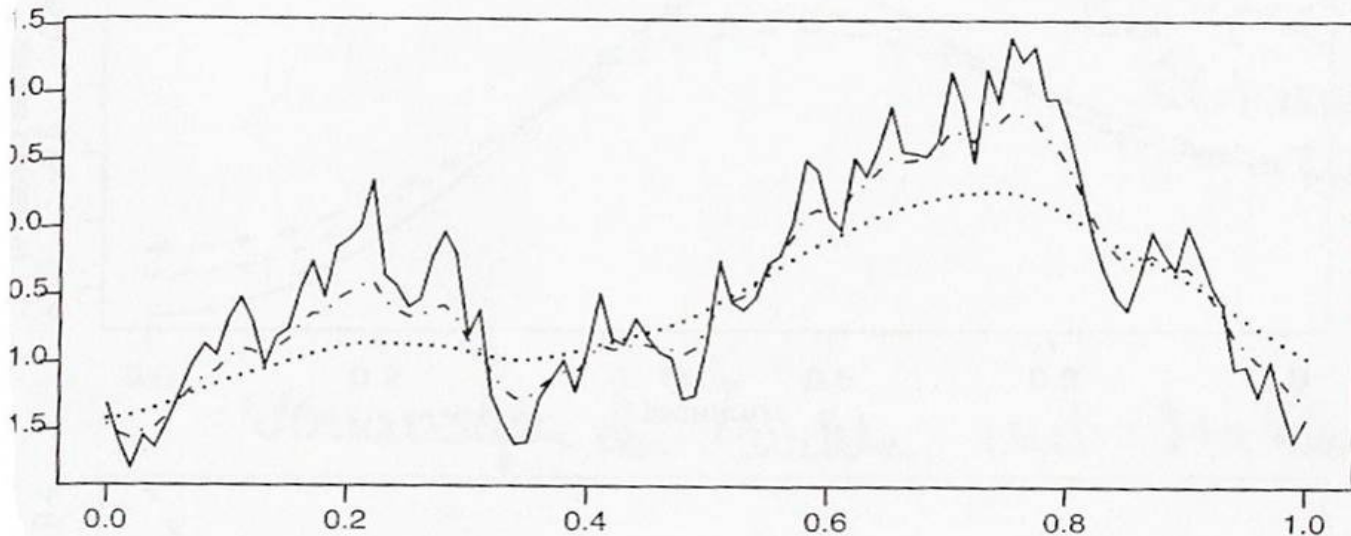
$$\rho(u) = \exp\{-(u/\phi)^k\}$$

3. Sphärische Familie:

$$\rho(u; \phi) = \begin{cases} 1 - \frac{3}{2}(u/\phi) + \frac{1}{2}(u/\phi)^3 & : 0 \leq u \leq \phi \\ 0 & : u > \phi \end{cases}$$



Gauß-Modell



Simulation eines Gaußprozesses mit Matérn Korrelationsfunktionen, wobei $\phi=0,2$ und $K=0,5$ (durchgehende Linie), $k=1$ (gestrichelte Linie) und $k=2$ (gepunktete Linie)

Gauß-Modell

Prognosen mit Hilfe des Gauß-Modells

- Ziel der Prognose ist $T = S(s_0)$, Wert des Signalprozesses an einem Punkt s_0 , wobei s_0 nicht notwendigerweise im Sampling Design enthalten sein muss.

- Prognoseverteilung $[T, Y]$ ist multivariat gauß-verteilt mit MW μ^1 und Varianzmatrix

$$\begin{pmatrix} \sigma^2 & \sigma^2 r^T \\ \sigma^2 r & \tau^2 I + \sigma^2 R \end{pmatrix}$$

wobei $r_i = \rho(\|s_0 - s_i\|) : i = 1, \dots, n$

Gauß-Modell

Prognosen mit Hilfe des Gauß-Modells

- MSE-Prädiktor für $T = S(s_0)$ ist $\hat{T} = \mu + \sigma^2 r^T (\tau^2 \mathbf{I} + \sigma^2 R)^{-1} (y - \mu \mathbf{1})$
mit Prognosevarianz $\text{Var}[T | y] = \sigma^2 - \sigma^2 r^T (\tau^2 \mathbf{I} + \sigma^2 R)^{-1} \sigma^2 r$
(falls \hat{T} für jeden Punkt s_0 so gegeben ist, spricht man vom sog. „*simple kriging*“)
- der MSE-Prädiktor kann für $S(s_0)$ explizit als lineare Funktion von y dargestellt werden:

$$\hat{T} = \left\{1 - \sum_{i=1}^n \omega_i(s_0)\right\} \mu + \sum_{i=1}^n \omega_i(s_0) \mu_i$$
wobei $\omega_1(s_0), \dots, \omega_n(s_0)$ die sog. Prognosegewichtungen sind



Gauß-Modell

Prognosen mit Hilfe des Gauß-Modells

Anmerkungen:

1. \hat{T} interpoliert die Daten gdw. $\tau^2 = 0$ (da hier $Y(s_i) = S(s_i)$)
2. $\hat{T}(s_0)$ wird stärker von y_i beeinflusst, falls s_0 nahe bei einem oder mehreren Punkten s_i liegt.



Parametrische Schätzung der Kovarianz Struktur

Variogramm Analyse

- Überblick über die Struktur eines räumlichen stochastischen Prozesses bzgl. des zweiten Momentes gibt sein *Variogramm*
- Variogramm eines stoch. Prozesses $Y(\cdot)$:

$$V(s, s') = \frac{1}{2} \text{Var}\{Y(s) - Y(s')\}$$

- für das Gauß Modell mit $u = \|s - s'\|$:

$$V(u) = \tau^2 + \sigma^2 \{1 - \rho(u)\}$$



Parametrische Schätzung der Kovarianz Struktur

Variogramm Analyse

- **Variogrammwolke**

Scatterplot von Punkten (u_{ij}, v_{ij}) , welche man aus den Größen

$$u_{ij} = \|s_i - s_j\| \quad \text{und} \quad v_{ij} = \frac{(y_i - y_j)^2}{2} \quad \text{erhält.}$$

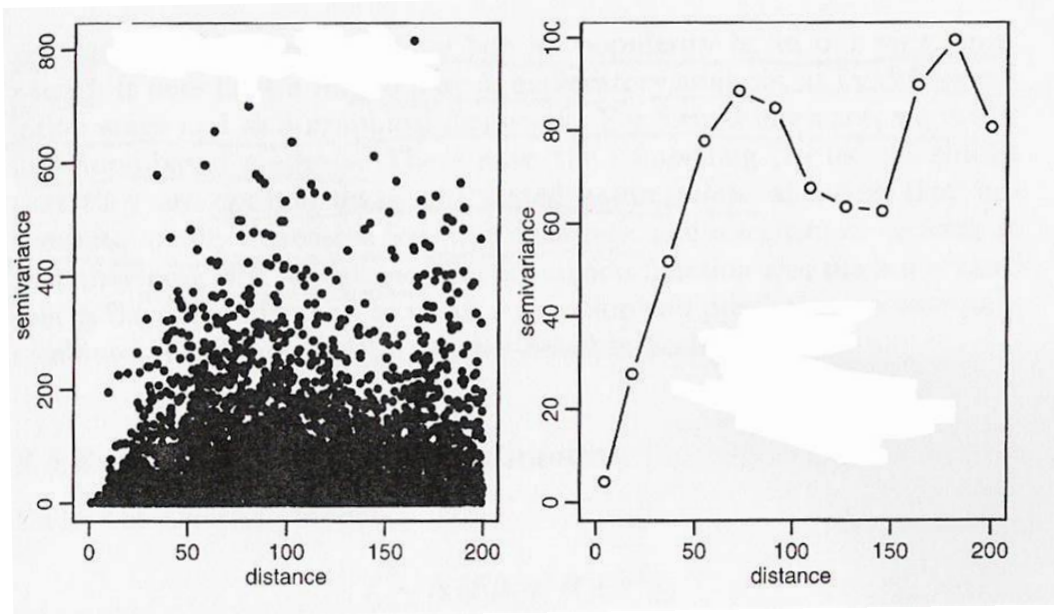
- **Empirisches Variogramm**

stabilere Variante der Variogrammwolke

bilden des Durchschnittes über die v_{ij} , für die $|u - u_{ij}| < \frac{h}{2}$, wobei u ein Separationsabstand sei und h bzgl. der Darstellung gewählte Weite sei.



Parametrische Schätzung der Kovarianz Struktur



Links: Variogrammwolke für Schweizer Niederschlagsdaten

Rechts: Empirisches Variogramm davon

Parametrische Schätzung der Kovarianz Struktur

Variogramm Analyse

- Schätzung der Kovarianz Parameter θ durch hinzufügen einer parametrischen Kovarianzfunktion zum empirischen Variogramm.
- Alternativ: Schätzung durch gewichtete, kleinste MW Methoden, dh. Minimiere

$$S(\theta) = \sum_k \omega_k [U(u_k) - V(u_k; \theta)]^2$$

, wobei ω_k die Anzahl der Paare von Messungen sind, die in das empirische Variogramm $U(u_k)$ eingehen.



Parametrische Schätzung der Kovarianz Struktur

Maximum-Likelihood Schätzung

- $Y \sim N(F\beta, \sigma^2 R + \tau^2 I)$
- β Vektor der Parameter
- F $n \times p$ Matrix („Designmatrix“)
- R hängt von (ϕ, k) ab
- log-LH Funktion

$$l(\beta, \tau^2, \sigma^2, \phi, k) = -0.5 \{ \log |(\sigma^2 R + \tau^2 I)| + (y - F\beta)^T (\sigma^2 R + \tau^2 I)^{-1} (y - F\beta) \}$$



Parametrische Schätzung der Kovarianz Struktur

Maximum Likelihood Schätzung

Berechnung der ML-Schätzer:

- Reparemetrisierung: $v^2 = \frac{\tau^2}{\sigma^2}; V = (R + v^2 I)$
- $\beta(V) = (F^T V^{-1} F)^{-1} F^T V^{-1} y$ und $\sigma^2(V) = n^{-1} (y - F\beta)^T V^{-1} (y - F\beta)$
maximieren die log-LH Funktion
- Substituiere $(\beta(V), \sigma^2(V))$ in die log-LH Funktion
 $\Rightarrow l(v^2, \phi, k) = -0.5 \{n \log |\sigma^2(V)| + \log |V|\}$ (reduzierte log-LH Funktion)
- Optimierung numerisch gefolgt von Rücksubstituieren



Parametrische Schätzung der Kovarianz Struktur

Maximum Likelihood Schätzung

Populäre Variante : Eingeschränkte MLH-Schätzung (REML)

- Annahme: $E[Y] = F\beta$
- lineare Transformation zu $W=AY$, sd. die Verteilung von W nicht von β abhängt
- Schätzung von $\theta = (v^2, \sigma^2, \phi, k)$ durch MLH bzgl. W
- $A = I - F(F^T F)^{-1} F^T$
- Berechnung des REML Schätzers durch maximieren von

$$l(\theta) = -0.5\{\log |\sigma^2 V| - \log |F^T \{\sigma^2 V\}^{-1} F| + (y - F\beta)^T \{\sigma^2 V\}^{-1} (y - F\beta)\}$$

wobei $\beta = \beta(V)$ Maximierer der urspr. log-LH Funktion ist



Parametrische Schätzung der Kovarianz Struktur

Prognose durch Einsetzen

- Einsetzen von Schätzern für unbekannte Modellparameter in die Prognosegleichungen
- optimistischer Ansatz, da die Variabilität zwischen Schätzern und wahren Werten ignoriert wird
- Anwendung im Gauß Modell:

Ersetzen der wahren Parameter in den Prognosegleichungen für

$$\hat{T} = \mu + \sigma^2 r^T (\tau^2 \mathbf{I} + \sigma^2 \mathbf{R})^{-1} (y - \mu \mathbf{1}) \quad (\text{minimaler MSE Prädiktor})$$

$$\text{Var}[T | y] = \sigma^2 - \sigma^2 r^T (\tau^2 \mathbf{I} + \sigma^2 \mathbf{R})^{-1} \sigma^2 r \quad (\text{Prognosevarianz})$$

durch deren Schätzer. („simple kriging“ ist Einsetzen von Schätzern für MW- und Kovarianzparameter)



Parametrische Schätzung der Kovarianz Struktur

Prognose durch Einsetzen

- **Bemerkung:**

Bei konstanten Werten Y_1, \dots, Y_n , aber offensichtlich unpassendem Gauß-Modell, erhält man zusätzliche Flexibilität durch Einführung einer Box-Cox Transformation in Form des Parameters λ .

=> Transformation der Daten y_1, \dots, y_n

$$\tilde{y}_i = h_\lambda(y_i) = \begin{cases} (y_i^\lambda - 1) / \lambda; \lambda \neq 0 \\ \log y_i; \lambda = 0 \end{cases}$$

Parametrische Schätzung der Kovarianz Struktur

Prognose durch Einsetzen

- betrachten das transformierte Gauß-Modell
- $Y(x_0)$ Ziel der Prognose, $T_\lambda = h_\lambda(Y(x_0))$ Prädiktor
- MSE-Prädiktor \hat{T}_λ und Prognosevarianz $Var[T_\lambda | y]$ durch „simple kriging“ bestimmbar

- $\lambda = 0$:

$$\hat{T} = \exp(\hat{T}_0 + 0.5Var[T_0 | y])$$

$$Var[T | y] = \exp(2\hat{T}_0 + Var[T_0 | y])(\exp(Var[T_0 | y]) - 1)$$

- $\lambda > 0$ (Approximieren durch eine Summe von Momenten für die Normalverteilung, hier $\lambda = 0.5$)

$$\hat{T} \approx (0.5\hat{T}_{0.5} + 1)^2 + 0.25Var[T_{0.5} | y]$$

$$Var[T | y] \approx (0.5\hat{T}_{0.5} + 1)^4 + 1.5(0.5\hat{T}_{0.5} + 1)^2Var[T_{0.5} | y] + 3(Var[T_{0.5} | y])^2 / 16$$



Parametrische Schätzung der Kovarianz Struktur

Prognose durch Einsetzen (bei nicht-linearem Ziel $T(S^*)$)

- $T = T(S^*)$ nicht-lineares Ziel der Prognose
- S^* Werte von $S(\cdot)$ an bestimmten Punkten
- Berechnung der Prädiktors \hat{T} durch Simulation

Berechnen des Prädiktors

1. Berechnen $E[S^*|y]$ und $\text{Var}[S^*|y]$ durch simple kriging
2. Simuliere $s^*(1), \dots, s^*(m)$ von $[S^*|y]$ (multivariate Gaußverteilung)
3. Approximiere den minimalen MSE Prädiktor

$$E[T(S^*) | y] \approx \frac{1}{m} \sum_{i=1}^m T(s^*(i))$$



Bayessche Folgerungen für das lineare Gauß Modell

Der Bayessche Ansatz behandelt die Modellparameter wie Zufallsvariablen, weshalb keine formale Unterscheidung zwischen Parameter-Schätzung und Prognoseproblemen gemacht wird.



Bayessche Folgerungen für das lineare Gauß Modell

Feste Korrelationsparameter

- $\tau^2 = 0$, alle anderen Parameter der Korrelationsfunktion sind bekannt
- Prognoseverteilung kann analytisch bestimmt werden
- ϕ fest \Rightarrow die a-priori Verteilungsfamilie für (β, σ^2) ist die gauß-skalierte-Inverse X^2
- a-priori Verteilungen für β und σ^2 :

$$[\beta | \sigma^2, \phi] \sim N(m_b, \sigma^2 V_b) \quad \text{und} \quad [\sigma^2 | \phi] \sim X_{ScI}^2(n_\sigma, S_\sigma^2)$$

wobei $X_{ScI}^2(n_\sigma, S_\sigma^2)$ die Dichte der Form $\pi(z) \propto z^{-(n_\sigma/2+1)} \exp(-n_\sigma S_\sigma^2 / (2z))$, $z > 0$ hat.

(Kurzschreibweise: $[\beta, \sigma^2 | \phi] \sim NX_{ScI}^2(m_b, V_b, n_\sigma, S_\sigma^2)$)



Bayessche Folgerungen für das lineare Gauß Modell

Feste Korrelationsparameter

Unter Verwendung des Baysschen Theorems ist die a-priori verbunden mit der LH-Funktion und die sich ergebende a-posteriori Verteilung der Parameter ist:

$$[\beta, \sigma^2 | y, \phi] \sim NX_{Sci}^2(m_b, V_\beta, n_\sigma + n, S^2)$$

wobei $V_\beta = (V_\beta^{-1} + F^T R^{-1} F)^{-1}$, $\beta = V_\beta (V_\beta^{-1} m_b + F^T R^{-1} y)$ und

$$S^2 = \frac{n_\sigma S_\sigma^2 + m_b^T V_b^{-1} m_b + y^T R^{-1} y - \beta^T V_\beta^{-1} \beta}{n_\sigma + n}$$

Bayessche Folgerungen für das lineare Gauß Modell

Unsicherheit bei den Korrelationsparametern

- Unsicherheit in allen Modellparametern
- a-priori $\pi(\beta, \sigma^2, \phi) = \pi(\beta, \sigma^2 | \phi)\pi(\phi)$
- Verteilung von β und σ^2 wie zuvor
- a-posteriori ist dann $p(\beta, \sigma^2, \phi | y) = p(\beta, \sigma^2 | y, \phi)p(\phi | y)$ und

$$p(\phi | y) \propto \pi(\phi) |V_\beta|^{-\frac{1}{2}} |R|^{-\frac{1}{2}} (S^2)^{-\frac{n+n_\sigma}{2}}$$

wobei V_β, S^2 wie zuvor sind.

Bayessche Folgerungen für das lineare Gauß Modell

Unsicherheit in den Modellparametern

Simulation von Beispielen

- Berechne a-posteriori Wahrscheinlichkeiten $p(\phi | y)$
- Simuliere einen Wert ϕ mit $[\phi | y]$
- Füge diesen Wert zu $[\beta, \sigma^2 | y, \phi]$ und man erhält eine Simulation dieser Verteilung
- Wiederholen dieser Simulation genügend oft ergibt ein Muster von Triplets (β, σ^2, ϕ) der gemeinsamen a-posteriori Verteilung der Modellparameter

Bayessche Folgerungen für das lineare Gauß Modell

Unsicherheit in den Modellparametern

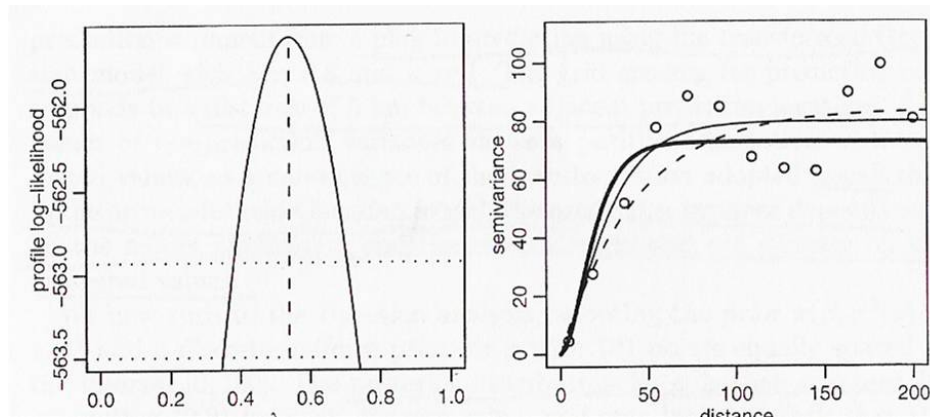
- Prognoseverteilung für den Wert des Signalprozesses $S_0 = S(s_0)$

$$p(s_0 | y) = \int p(s_0 | y, \phi) p(\phi | y) d\phi$$

- diskrete a-priori von ϕ erlaubt eine analytische Kalkulation der Momente
- Simulation mit dieser Prognoseverteilung:
 - Berechnung der a-posteriori Wahrscheinlichkeiten von $p(\phi | y)$
 - Simulieren der Werte von ϕ durch $[\phi | y]$
 - Einfügen der Werte von ϕ in $[S_0 | y, \phi]$ und simulieren mit dieser Verteilung => Realisierung der Prognoseverteilung

Eine Fallstudie

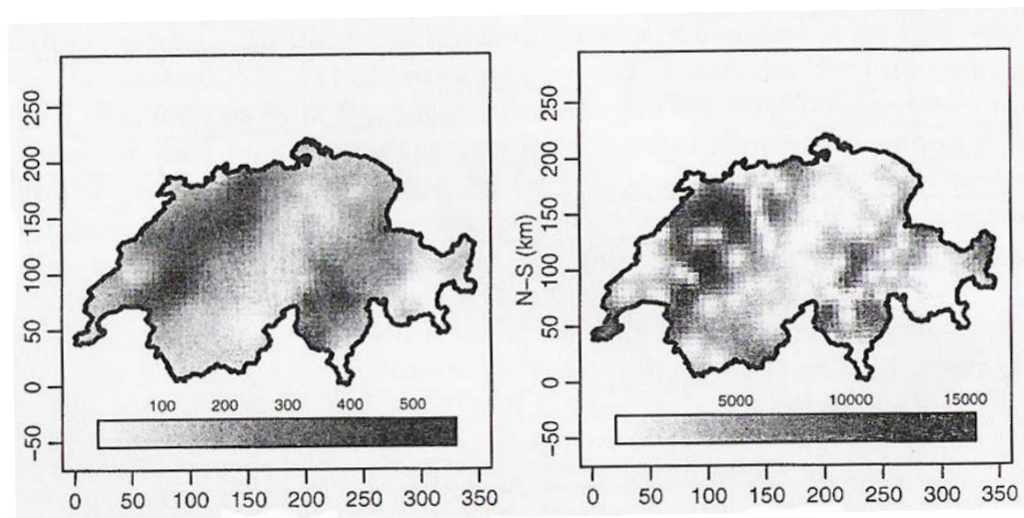
Verwendet werden 100 Daten der Schweizer Regenfalldaten zur Modellformulierung. Das Modell ist ein transformiertes Gaußmodell mit Matérn Korrelationsfunktion.



Links: Profil der log-LH zeigt hier, dass weder $\lambda = 0$ noch $\lambda = 1$ für dieses Modell geeignet sind

Rechts: Empirisches und eingebautes Variogramm für $k=0.5, 1, 2$

Eine Fallstudie



Punktprognosen der Regenfallwerte und zugehörigen Prognosevarianz einer Prognose durch Einsetzen unter Verwendung des transformierten Gaußmodells mit $\lambda = 0.5$ und $k=1$.

Aufteilung des Prognosenetzes entspricht einer Distanz von 5 km zwischen den angrenzenden Prognosepunkten.

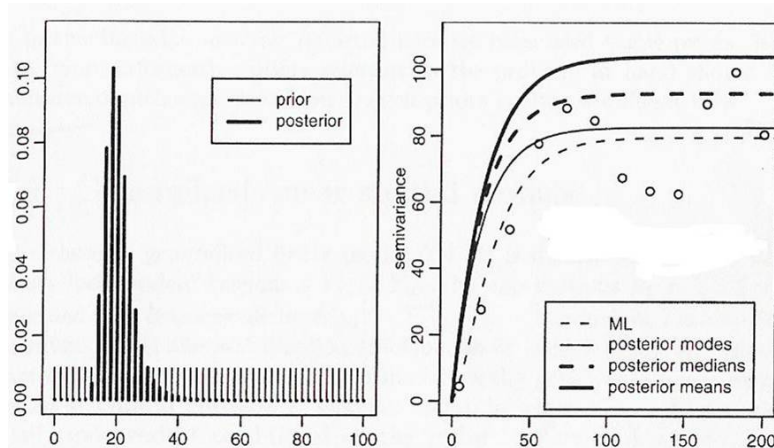
Die Werte der Prognosevarianz zeigen eine positive Korrelation mit den prognostizierten Werten.

Eine Fallstudie

Bayessche Analyse

A-priori Verteilung $\pi(\beta, \sigma^2 | \phi) \propto 1/\sigma^2$

- diskrete a-priori für ϕ mit 101 Punkten, gleichmäßig über ein Intervall [1;100] verteilt sind
- a-posteriori durch Berechnung von $p(\phi | y) \propto \pi(\phi) |V_\beta|^{-\frac{1}{2}} |R|^{-\frac{1}{2}} (S^2)^{-\frac{n+n_\sigma}{2}}$ für jeden diskreten Wert

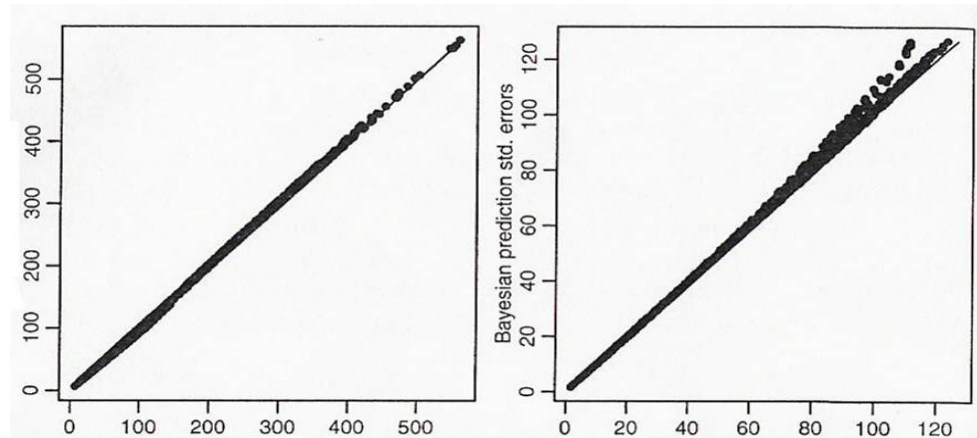


Links: angenommene a-priori und berechnete a-posteriori

Rechts: Variogramme, die auf verschiedenen Auswertungen der a-posteriori $[\sigma^2, \phi | y]$ beruhen

Eine Fallstudie

Prognose der Regenfallwerte



Prognose mit Hilfe des Modells mit festen Parametern und des Modells mit Unsicherheiten in den Parametern

- links: Vergleich zwischen Plug-In Methode und Bayesscher Punktprognose
- rechts: Standardfehler beider Methoden

Anhang

Literatur

- S. Banerjee, B. P. Carlin. *Hierarchical Modelling and Analysis for Spatial Data*. Chapman & Hall, 2004.
- J. Moeller, Ed. *Spatial Statistics and Computational Methods*. Springer, LNS 173. 2003.

