Variogramm-Anpassung:
Methode der kleinsten Quadrate

Universität Ulm

Daniel Wolfmüller
Dezember 2002

1 Bezeichnungen und Begriffe

Gegeben: Zufallsfeld $Z(x, \omega)$ mit

$\omega \in \Omega$ Elementarereignis

$x \in \mathbb{R}^d$ Vektor

$Z(x, \cdot)$ Zufallsvariable,

$Z(\cdot, \omega)$ Realisierung der zufälligen Funktion, auch kurz: $z(x)$

$x_1, ..., x_n$ seien die betrachteten Messstellen, $z(x_1), ..., z(x_n)$ die entsprechenden Messwerte.

Zur Vereinfachung: isotropes Zufallsfeld vorausgesetzt, d.h die Eigenschaften des Zufallsfeldes ändern sich nicht bei Drehung des Zufallsfeldes:

$$Z(x+h,\cdot) \sim Z(x+|h|e_1,\cdot) \quad \forall\ h \in \mathbb{R}^d \quad ,$$

wobei e_1 der erste euklidische Einheitsvektor sei.

2 Motivation

Von **Interesse** ist die Betrachtung der Abweichungen der Zufallswerte in Abhängigkeit von Richtungsvektoren $h \in \mathbb{R}^d$ - hier: in Abhängigkeit von den Abständen der beobachteten Orte zueinander $|h|$.
Als **Maßzahl** dafür betrachtet man das Variogramm.

Betrachten hier Zufallsfelder, die **intrinsisch stationär** sind, d.h. die Zufallsvariable $Y_h(x) = Z(x+h) - Z(x)$ ist stationär zweiter Ordnung. Dabei heißt ein Zufallsfeld intrinsisch stationär, wenn gilt:

(i) $E(Y_h(x)) = \; < a, h >$ mit $a \in \mathbb{R}^d$

(ii) $Var(Y_h(x)) = 2\gamma(h)$, wobei $\gamma(h)$ **theoretisches Variogramm** genannt wird

2 Motivation

Besitzt das Zufallsfeld keinen **Drift**, so ist $a = \mathbf{0}$ und damit $E(Y_h(x)) = 0$.
Dann gilt für das theoretische Variogramm: $\gamma(h) = \frac{1}{2} E\left((Y_h(x))^2\right)$.

Aus den (gemessenen) Daten lässt sich das **empirische Variogramm** erstellen:

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z(x_i) - z(x_j))^2$$

wobei $N(h) := \{(x_i,\ x_j) \ \mid \ \tilde{h}_k \leq |x_i - x_j| < \tilde{h}_{k+1},\ h \in [\tilde{h}_k, \tilde{h}_{k+1})\}$ mit $0 = \tilde{h}_0 < \tilde{h}_1 < \tilde{h}_2 < \ldots < \tilde{h}_r < \infty$ und $h \in [0, \tilde{h}_r)$

Ziel ist es nun eine Funktion $\gamma(h)$ zu finden, die das obige empirische Variogramm approximiert.

Die Wahl dieser Funktion sollte grundsätzlich folgenden Prinzipien folgen:

(i) Vernünftige Übereinstimmung des Variogramms mit den Daten
abhängig von den beobachteten Daten
Forderungen:

- geringe Abweichungen
- Monotonieerhaltung
- Asymptotik im Bereich der oberen Schwelle (des **sills**)

(ii) Einfluss anderer Informationen, die nicht aus den ermittelten Daten ableitbar sind

abhängig von Erfahrung und Zusatzinformationen des Beobachters

mögliche Forderungen:

- qualitativ bekannter Verlauf (exponentieller,linearer,... Anstieg)
- Verhalten im Unendlichen
- Stetigkeit im Ursprung

3 Methode der kleinsten Quadrate

Idee: Zunächst Festlegung einer geeigneten Messgröße für die Güte eines Variogramm-Modells $\gamma(h)$

Wähle dazu aufsummierte quadrierte Differenzen:

$$Q(\mathbf{h}) = \sum_{i=1}^{r} (\hat{\gamma}(h_i) - \gamma(h_i))^2 \quad , \qquad \text{wobei } \mathbf{h} = (h_1, h_2, ..., h_r)^T$$

$$\text{und} \qquad h_i = \frac{\tilde{h}_i - \tilde{h}_{i-1}}{2}, \quad i = 1, ...r$$

Ziel: Finden einer Funktion, die diesen quadratischen Fehler **minimiert**.

Wähle dazu als Variogramm-Modell $\gamma(h)$ eine geeignet erscheinende parametrische Familie von Funktionen $\{\gamma_\theta(h) \mid \theta \in \Theta\}$, wobei $\Theta \subset \mathbb{R}^m$ der dazugehörige Parameterraum sei.

$\Longrightarrow$ Reduktion des Problems auf Wahl des Parameters θ, so dass der quadratische Fehler minimiert wird.

Variogramm-Modell-Ansätze sind z. B.

- Nugget-Effekt-Modell:

$$\gamma_{nug}(h) = \begin{cases} 0, & \text{für } h = 0 \\ \theta, & \text{sonst} \end{cases} \quad , \text{ wobei } \theta > 0$$

- lineares Modell:

$$\gamma_{lin}(h) = \theta h \quad , \text{ wobei } \theta > 0$$

- exponentielles Modell:

$$\gamma_{exp}(h) = \theta_1(1 - exp(-\frac{h}{\theta_2})) \quad , \text{ wobei } \theta = (\theta_1, \theta_2)^T > \mathbf{0}$$

Dabei entspricht θ_1 dem sill, $3\theta_2$ dem Range.

- sphärisches Modell:

$$\gamma_{sph}(h) = \begin{cases} \theta_2 \left(\frac{3h}{2\theta_1} - \frac{h^3}{2\theta_1^3} \right), & \text{für } 0 \leq h \leq \theta_1 \\ \theta_2, & \text{sonst} \end{cases}$$

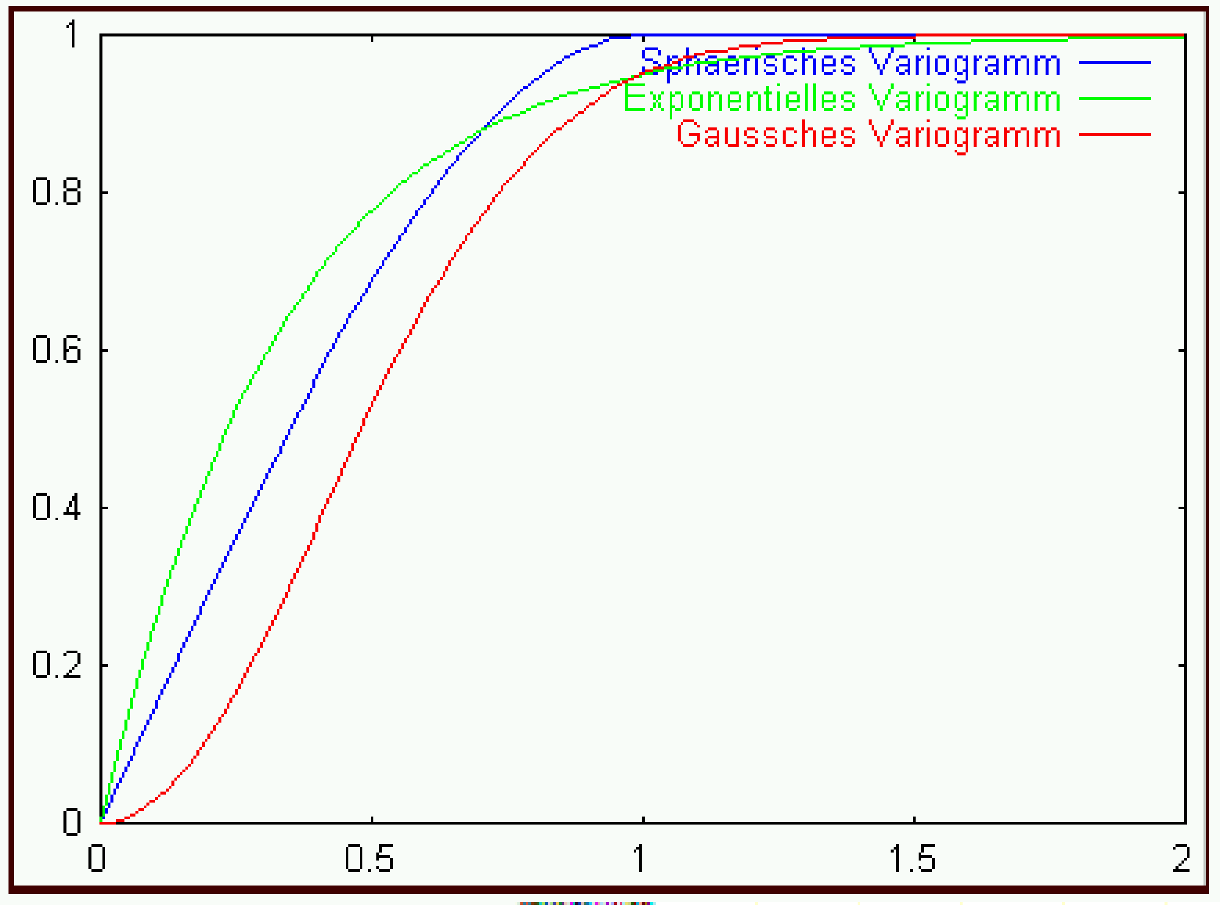
3 Methode der kleinsten Quadrate

- Gauß'sches Modell:

$$\gamma_{Gauß}(h) = \theta_1 \left(1 - exp(\frac{h^2}{\theta_2^2})\right)$$

Auch Linearkombinationen dieser Modelle sind denkbar, beispielsweise ein exponentielles Modell mit Nugget-Effekt.

Variogrammmodelle



3.1 Generalized Least Squares GLS

Allgemeiner Ansatz der Methode der kleinsten Quadrate

Einfluss der gemessenen Punkte untereinander ausgedrückt in Kovarianzmatrix $V = (Cov(\hat{\gamma}(h_i), \hat{\gamma}(h_j)))_{i,j}$

Minimiere nun $Q(\mathbf{h},\ \theta)$ mit

$$Q(\mathbf{h},\ \theta) = [\hat{\gamma} - \gamma_\theta]^T V^{-1} [\hat{\gamma} - \gamma_\theta] =$$

$$= \sum_{i,j=1}^{r} w_{ij} (\hat{\gamma}(h_i) - \gamma_\theta(h_i))(\hat{\gamma}(h_j) - \gamma_\theta(h_j)),$$

wobei $\hat{\gamma} = (\hat{\gamma}(h_1), ..., \hat{\gamma}(h_r))$, $\gamma_\theta = (\gamma_\theta(h_1), ..., \gamma_\theta(h_r))$ und $w_{ij} = (V^{-1})_{ij}$.

Vorteile:

- Unterschiedliche Anzahl von Werten innerhalb eines Intervalles $[h_k, h_{k+1})$ werden berücksichtigt
- große Aussagefähigkeit des Fehler $Q(\mathbf{h},\ \theta)$ durch Berücksichtigung von Korrelationen

Problem: im Allg. lässt sich V nicht angeben

3.2 Weighted Least Squares WLS

vereinfachender Ansatz ohne Berücksichtigung der Korrelationen; setze bei GLS-Methode $w_{ii} = \frac{1}{Var(\hat{\gamma}(h_i))} := \hat{w}_i,\ w_{ij} = 0,\ i \neq j,\ \ i,j = 1,...r$
Minimiere $Q(\mathbf{h}, \theta)$ mit

$$Q(\mathbf{h}, \theta) = \sum_{i=1}^{n} \hat{w}_i [\hat{\gamma}(h_i) - \gamma_\theta(h_i)]^2$$

Vorteile:

- geringerer Rechenaufwand als bei GLS
- Berücksichtigung der Varianzen

Problem: Berechnung der Varianzen, falls keine zusätzlichen Informationen über das Zufallsfeld bekannt sind

Spezialfall: Gauß'sches Zufallsfeld

Ein Zufallsfeld $Z(x,\omega)$ heißt Gauß'sches Zufallsfeld, falls jeder Zufallsvektor $(Z(x_1),...,Z(x_n))$ normalverteilt ist.

Dann gilt:

$$\frac{1}{2}(Z(x+h)-Z(x))^2 \;\sim\; \gamma(h)\chi_1^2$$

Annahme: Korrelationen der $|N(h_k)|$ Differenzen von Zufallsvariablen, die bei der Berechnung von $\hat{\gamma}(h_k)$ einfließen, sind vernachlässigbar

$$\Longrightarrow \quad Var(\hat{\gamma}(h_k)) = Var\left(\frac{1}{2}\frac{1}{|N(h_k)|}\sum_{N(h_k)}(Z(x_i)-Z(x_j))^2\right) \approx$$

$$\approx \frac{1}{|N(h_k)|^2}\sum_{N(h_k)} Var\left(\frac{1}{2}(Z(x_i)-Z(x_j))^2\right) \approx$$

3 Methode der kleinsten Quadrate

$$\approx \frac{1}{|N(h_k)|^2}|N(h_k)|Var\left(\frac{1}{2}(Z(x_i)-Z(x_j))^2\right) =$$

$$= 2\frac{\hat{\gamma}_\theta^2(h_k)}{|N(h_k)|},$$

wobei für eine ZV $X, X \sim \chi_1^2$ gilt:

$$Var\,(\gamma(h_k)\cdot X) = \gamma^2(x_k)\cdot VarX = 2\cdot\gamma^2(x_k)$$

Dann reduziert sich das Minimierungsproblem auf die Minimierung von

$$Q(\mathbf{h},\theta) = \sum_{i=1}^{n}\hat{w}_i^2[\hat{\gamma}(h_i)-\gamma_\theta(h_i)]^2 = \frac{1}{2}\sum_{i=1}^{r}|N(h_k)|\left(\frac{\hat{\gamma}(h_i)}{\gamma_\theta(h_i)}-1\right)^2$$

3.3 Ordinary Least Squares OLS

Einfachster Ansatz ohne Berücksichtigung der Varianzen und Korrelationen; setze bei GLS $w_{ii} = 1,\ w_{ij} = 0,\ i \neq j,\ \ i, j = 1, ...r$

Minimiere $Q(\mathbf{h},\ \theta)$ mit

$$Q(\mathbf{h},\ \theta) = \sum_{i=1}^{n} [\hat{\gamma}(h_i) - \gamma_\theta(h_i)]^2$$

Vorteile:

- keine Kenntnis über Verteilung von $\hat{\gamma}$ erforderlich
- Einfachheit der Rechnung

Problem: kein Einfluss von Informationen über Korrelation innerhalb des betrachteten Feldes

4 Beispiel: Algorithmus

Algorithmus zur Bestimmung des Parametervektors $\theta \in \Theta \subset \mathbb{R}^m$ mittels WLS:

Zu minimierender Ausdruck: $Q(\mathbf{h}, \theta) = \sum_{i=1}^{r} \hat{w}_i [\hat{\gamma}(h_i) - \gamma_\theta(h_i)]^2$

1. Sei k=0. Wähle initialen Parametervektor $\theta_k = \theta_0$ (z.B. mittels OLS) und berechne damit $\hat{w}_{i,k} = \hat{w}_{i,0}, \quad i = 1, ..., r$.
2. Finde Parametervektor θ_{k+1}, der $Q_k(\mathbf{h}, \theta_{k+1}) := \sum_{i=1}^{r} \hat{w}_{i,k} [\hat{\gamma}(h_i) - \gamma_{\theta k+1}(h_i)]^2$ minimiert.
3. Falls $|Q_k(\mathbf{h}, \theta_{k+1}) - Q_{k-1}(\mathbf{h}, \theta_k)| < \varepsilon$ für ein vorgegebenes $\varepsilon > 0$ und $k \geq 1$: STOPP.
   Sonst k := k+1 und wiederhole Schritt 2.

5 Beispiel: Boolsches Modell

Betrachten einen Poisson-Prozess $\Phi = \{x_1, x_2, ...\}$ in einem rechteckigen Simulationsfenster $A \in \mathbb{R}^2$.

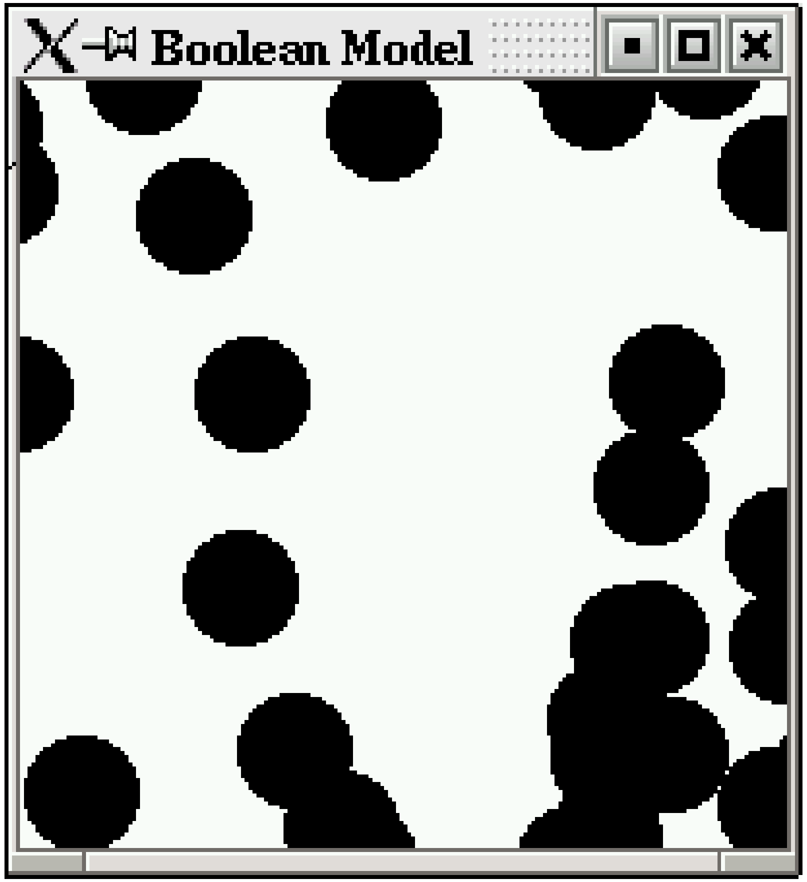
Simulation: Simuliere N gleichverteilte Punkte in A, wobei N eine Realisierung einer Poisson-verteilten Zufallsvariablen mit Parameter $\lambda|A|$ sei.

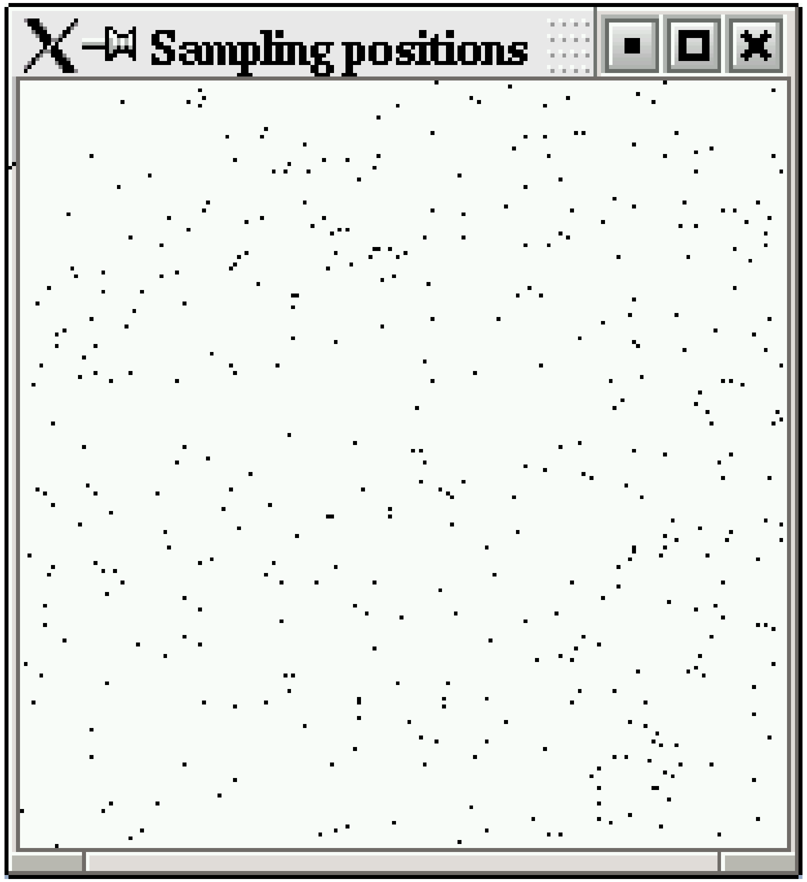
Hierbei sei das Boolsche Modell Ξ definiert durch

$$\Xi = \bigcup_{i=1}^{\infty} (\Xi_0 + x_i) \quad ,$$

wobei das Primärkorn $\Xi_0 = \{x \in \mathbb{R}^2 \;\mid\; |x| \le r\}$ ein Kreis mit Radius r ist.

5 Beispiel: Boolsches Modell





Betrachte nun das Zufallsfeld $Z = \{Z(x),\ x \in \mathbb{R}^2\}$ mit $Z(x) = \mathbf{1}_{\{x \in \Xi\}}$.

5 Beispiel: Boolsches Modell

Z besitzt folgende Eigenschaften:

1. Z nimmt nur die Werte 0 oder 1 an
2. Z ist stationär zweiter Ordnung, die Variogramm-Funktion lässt sich daher folgendermaßen berechnen: $\gamma(h) = C(0) - C(h)$

Dann gilt:

$$\gamma(h) = e^{-\lambda \pi r^2}(1 - e^{-\lambda(\pi r^2 - F(h))}),$$

wobei $F(h)$ der Flächeninhalt der Menge $\Xi_0 \cap (\Xi_0 - h)$ ist.

6 Zusammenfassung

Fazit: Allgemeine Methode der kleinsten Quadrate (GLS) und gewichtete Methode der kleinsten Quadrate (WLS) liefern **gute Approximationen**, da die Korrelationen der ZVen in die Berechnung einfließen. Einfache Methoder der kleinsten Quadrate (OLS) ist einfach anzuwenden und setzt **keine** Kenntnis über die Verteilung des betrachteten Zufallsfeldes voraus.
Problematisch bleibt im Allg. die Wahl des Variogrammmodells, des Modells für das betrachtete Zufallsfeld und die Wahl des Untersuchungsradius' h_r.

Je mehr also über das betrachtete Zufallsfeld bekannt ist, desto genauer sind die Ergebnisse.

7 Literatur

- J. Chiles, P. Delfiner; Geostatistics - Modelling spatial uncertainty; Wiley, 1999
- N. A. Cressie; Statistics for spatial data; Wiley, 1993
- P. K. Kitanidis; Introduction to geostatistics; Cambridge University Press, 1997
- H. Wackernagel; Multivariate geostatistics; Springer, 1998
- H. v. Waldow; Einführung in die Geostatistik; Institut für geographische Wissenschaften, Berlin, 1997