Geostatistik

Variogramme und Kriging

Prof. van den Boogaart

boogaart@math.tu-freiberg.de

Gliederung

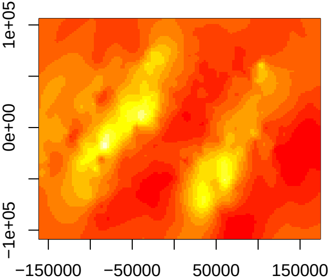
Gliederung

Regionalisierte Variablen

$$
\begin{aligned}
Z(s) &= \text{Messwert an Stelle s} \\
s &\in R^2, \quad \text{Ort}
\end{aligned}
$$

Eine regionalisierte Variable (Zufallsfeld) hat überall einen Wert.

Toblers erstes Gesetz der Geographie: Alles ist abhängig. Nahes ist stärker abhängig.



Geostatistische Daten

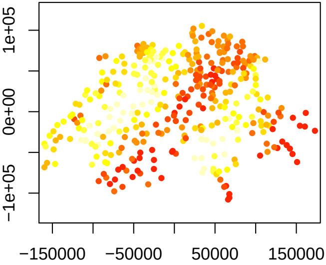
x	**y**	$Z\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$
x_1	y_1	z_1
x_2	y_2	z_2
$\vdots$	$\vdots$	$\vdots$

Messorte:

$$s_i, i = 1, \ldots, n$$

Beobachtungen / Messwerte

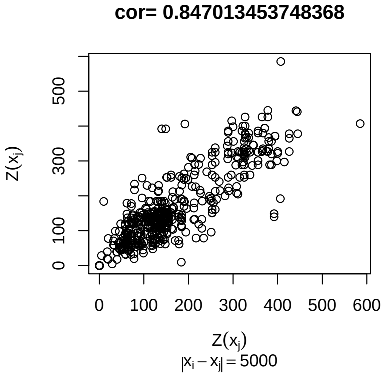
$$z_i = Z(s_i)$$



h-Scatterplot

X	Y
$Z(s_1)$	$Z(s_1)$
$Z(s_2)$	$Z(s_1)$
$\vdots$	$\vdots$
$Z(s_1)$	$Z(s_2)$
$\vdots$	$\vdots$

mit $\|s_1 - s_2\| \approx h = 5000$

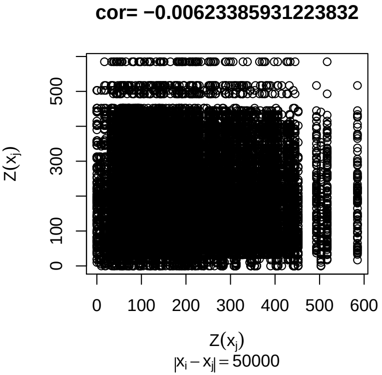


h-Scatterplot

X	**Y**
$Z(s_1)$	$Z(s_1)$
$Z(s_2)$	$Z(s_1)$
$\vdots$	$\vdots$
$Z(s_1)$	$Z(s_2)$
$\vdots$	$\vdots$

mit $\|s_1 - s_2\| \approx h = 55000$
Kovarianzfunktion:

$$c(s_1, s_2) = cov(Z(s_1), Z(s_2))$$



Stationarität

- Starke Stationarität
  Verteilung bleibt unter Verschiebung gleich.
- Schwache Stationarität
  Erwartungswert, Varianz und Kovarianz bleibt unter Verschiebung gleich.
- Intrisische Stationarität
  Differenzen haben Erwartungswert 0 und behalten unter Verschiebung ihre Varianz.

Variogram

Variogram

$$2\gamma(s_1, s_2) := E\left[(Z(s_1) - Z(s_2))^2\right]$$

Semivariogram:

$$\gamma(s_1, s_2) := \frac{1}{2} E\left[(Z(s_1) - Z(s_2))^2\right]$$

Stationäres Variogram

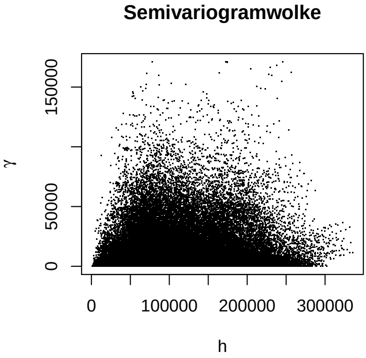
$$\gamma_s(s_1 - s_2) = \gamma(s_1, s_2)$$

Isotropes Variogram

$$\gamma_I(\|s_1 - s_2\|) = \gamma(s_1, s_2)$$

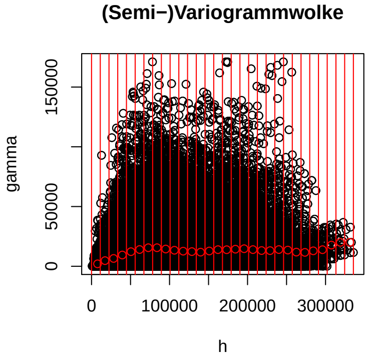
Gliederung

Variogram Wolke

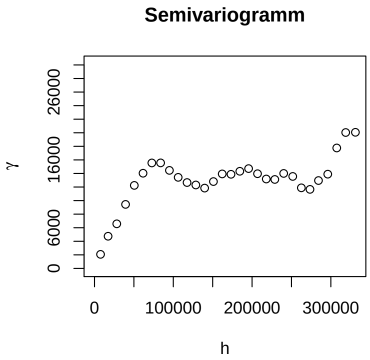


Empirisches Variogramm erklärt

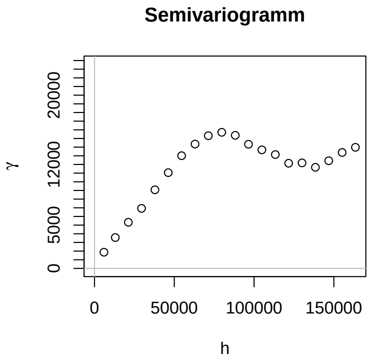
-



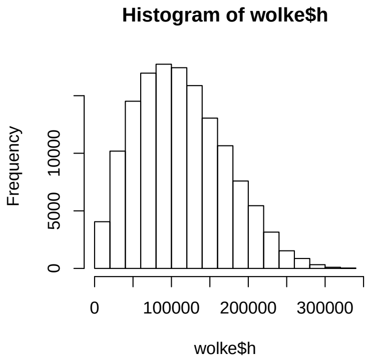
Rohes Empirisches Variogramm



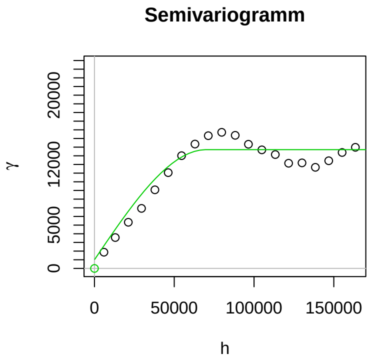
Empirisches Variogramm



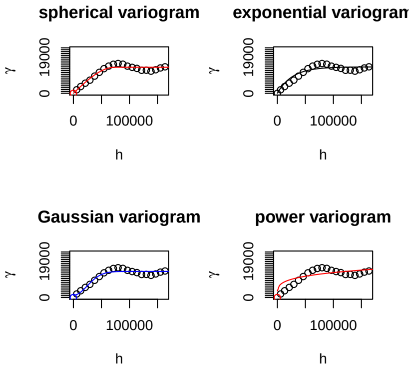
Verlässlichkeit des Empirische Variogram



Nugget, Sill und Range



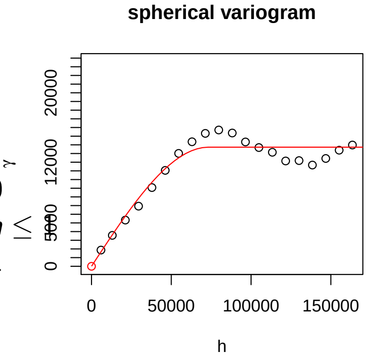
Variogrammodelle



Sphärisches Variogram

$$\gamma(h) = n + (s - n) * \gamma_{sph}(h/r)$$

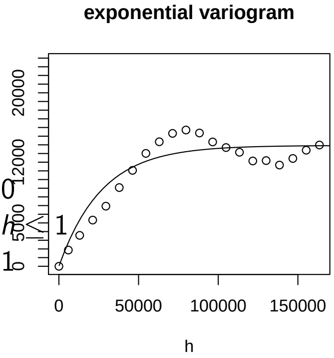
$$\gamma_{sph}(h) = \begin{cases} 0, & h = 0 \\ \frac{3}{2}h - \frac{1}{2}h^3, & 0 \le h \le \\ 1, & h \ge 1 \end{cases}$$



Exponentiells Variogram

$$\gamma(h) = n + (s - n) * \gamma_{exp}(h/r)$$

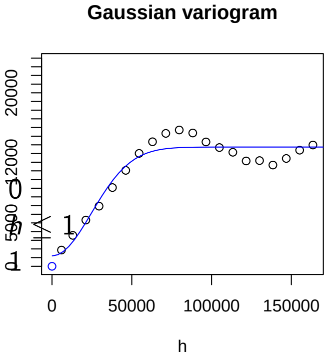
$$\gamma_{exp}(h) = \begin{cases} 0, & h = 0 \\ 1 - \exp\{-h\}, & 0 \leq h \leq 1 \\ 1, & h \geq 1 \end{cases}$$



Gausssches Variogram

$$\gamma(h) = n + (s-n) * \gamma_{gau}(h/r)$$

$$\gamma_{gau}(h) = \begin{cases} 0, & h = 0 \\ 1 - \exp\{-h^2\}, & 0 \leq h \leq 1 \\ 1, & h \geq 1 \end{cases}$$

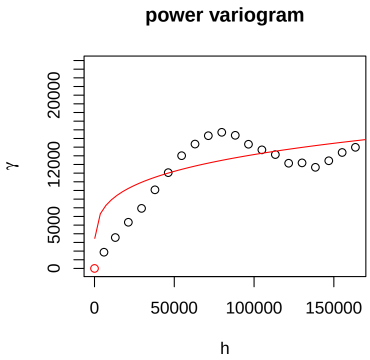


Power Variogram

$$\gamma(h) = n + (s-n)\gamma_{pow}(h)$$

$$\gamma_{pow}(h) = \begin{cases} 0, & h = 0 \\ ||h||^{\lambda}, & h > 0 \end{cases}$$

$$0 \leq \lambda \leq 2$$



Variogrammodelle anpassen

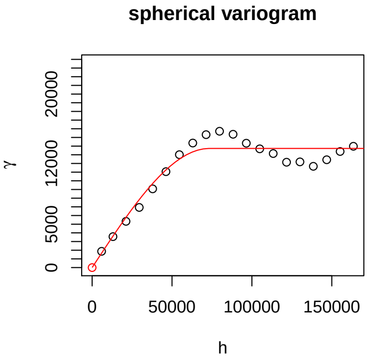
Methode der kl. Quadrate

$$\operatorname*{argmin}_{\theta} \sum_{j=1}^{N} (\gamma_\theta(h) - \hat{\gamma}(h_j))^2$$

ML/REML (Restricted Maximum Likelihood)

$$\operatorname*{argmax}_{\theta} L(z_1, \ldots, z_n; \theta)$$

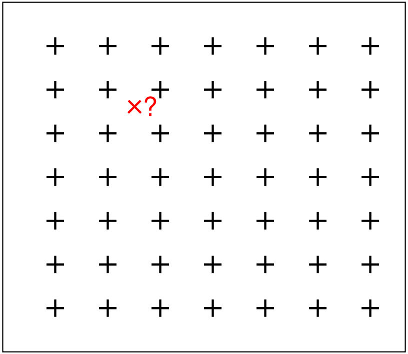
$L(\ldots)$ = Wahrscheinlichkeitsdichte der Abweichungen vom Modell



Gliederung

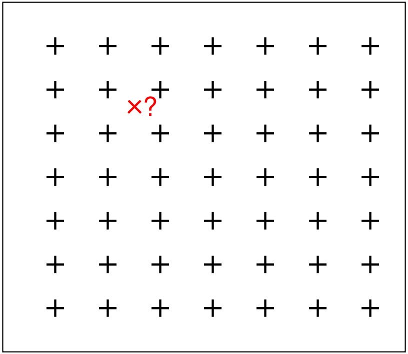
Kriging Aufgabe

Gemessen wurden die Werte $z_1, \ldots, z_n$ eines geostatistischen Feldes an n Punkten $x_1, \ldots, x_n$. Frage wie kann man den Wert an einem weiteren Punkt x_{n+1} möglichst gut vorhersagen?



Kriging Aufgabe

Gemessen wurden die Werte $z_1, \ldots, z_n$ eines geostatistischen Feldes an n Punkten $x_1, \ldots, x_n$. Frage wie kann man den Wert an einem weiteren Punkt x_{n+1} möglichst gut vorhersagen?



Kriging Gleichungssystem

Ideen:
linear(, weil einfach):

$$\hat{Z}(s) = \sum_{i=1}^{n} w_i(s) Z(s_i)$$

Unverzerrtheitsbedingung (im Mittel treffen):

$$E[\hat{Z}(s) - Z(s)] = 0$$

Fehler so klein wie möglich:

$$\mathrm{var}(\hat{Z}(s) - Z(s)) \to \min$$

Kriging Gleichungssystem

Bezeichungen:

$$\mathbf{z} = \begin{pmatrix} Z(s_1) \\ \vdots \\ Z(s_n) \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma(s_1, s) \\ \vdots \\ \gamma(s_n, s) \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \text{`} = \begin{pmatrix} \gamma(s_1, s_1) & \cdots & \gamma(s_1, s_n) \\ \vdots & \ddots & \vdots \\ \gamma(s_n, s_1) & \cdots & \gamma(s_n, s_n) \end{pmatrix}$$

Gleichungssystem Ausgangspunkt

Unverzerrtheitsbedingung (im Mittel treffen):

$$E[\hat{Z}(s) - Z(s)] = \mathbf{1}^t\mathbf{w} = 0$$

Fehler so klein wie m"glich:

$$\mathrm{var}(\hat{Z}(s) - Z(s)) = \mathbf{w}^t\Gamma\mathbf{w} \to \min$$

Kriging Gleichungssystem

$$\begin{pmatrix} \gamma \\ 1 \end{pmatrix} = \begin{pmatrix} \Gamma & 1 \\ 1^t & 0 \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ l \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{w} \\ l \end{pmatrix} = \begin{pmatrix} \Gamma & 1 \\ 1^t & 0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma \\ 1 \end{pmatrix}$$

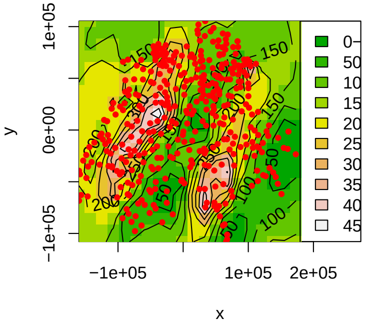
Kriging Vorhersage

$$\hat{Z}(s) = \begin{pmatrix} \Gamma & 1 \\ 1^t & 0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ 0 \end{pmatrix}$$

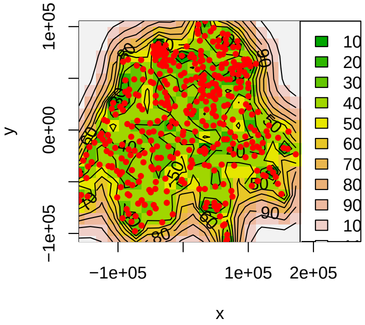
$$\mathrm{var}(\hat{Z}(s) - Z()) = \sigma_K^2(s)$$

$$\mathrm{var}(\hat{Z}(s) - Z()) \approx \mathrm{var}(Z(s)) - \mathrm{var}()$$

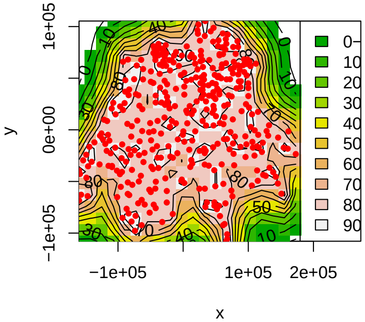
Kriging Karte



Kriging Fehler

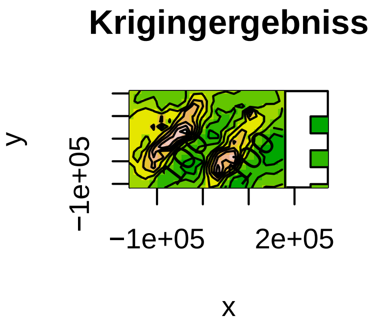


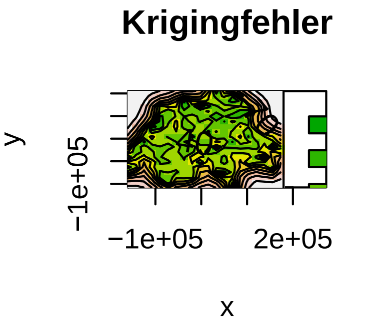
Kriging Gain



Gewinn durch Kriging in % der Varianz

Vorhersageintervall



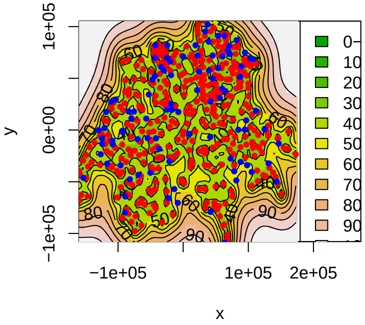


Kriging Eigenschaften

- Kriging interpoliert (aber nicht unbedingt stetig am Messpunkt)
- Kriging ist das beste lineare erwartungstreue Verfahren,
  - wenn das Variogram und die Voraussetzungen (Stationarität) stimmen.
  - Aber das Beste ist nicht unbedingt gut.
  - Linear und Erwartungstreu sind Bug und Feature zugleich
- Kennt seine eigene Genauigkeit (Besser wissen das man etwas nicht weiss als irren)

Validation

- Einfache Validation: Bisher unbeobachtete Punkte Vorhersagen und prüfen
- Kreuvalidation: Jeden Punkt einmal weglassen und vorhersagen.



Kriging Schritt für Schritt

<-+>Probenplan erstellen (Versuchsplanung später)Räumliche Daten erheben (Koordinaten und Werte)Datenexploration / Voraussetzungen untersuchenVariographieKriging durchführenKriging FehlerkarteKreuzvalidationSchlüsse ziehen

Gliederung

Universelle Kriging

Herausforderung: Was tun wenn ein Trend vorliegt und das Zufallsfeld daher nicht stationär ist.

Lokales Kriging

Herausforderung: Was tun, wenn man man zu viele Daten hat oder die Mittelwerte sich lokal verändern.

Indikator Kriging

Herausforderung: Was tun, wenn man die Verteilung braucht und nicht nur den Mittelwert.

Duales Kriging

Herausforderung: Was tun, wenn man sehr viele Punkte interpolieren muss und der Computer nicht hinterherkommt.

Lognormal Kriging

Herausforderung: Was tun, wenn man ein positives Zufallsfeld schiefe Verteilungen hat.

Transgaussches Kriging

Herausforderung: Was tun, wenn weder normal noch log-normal stimmen.

Disjunktives Kriging

Herausforderung: Was tun, wenn linear so gar nicht optimal ist, weildie Verteilung so gar nicht normal ist.

Einfaches Kriging

Herausforderung: Was tun, wenn der Trend bekannt ist.

Block Kriging

Herausforderung: Was tun, wenn man den Mittelwert/Gesamtmenge in einem Gebiet braucht.

Kriging linearer Funktionale

Herausforderung: Was tun, wenn man statt der Höhe das Gefälle braucht.

Gliederung