Von Simpel Kriging zu Universal Kriging

Universität Ulm
Abteilung Stochastik

Kathrin Kötting
November 2002

1 Notation

Ein **Zufallsfeld** ist eine zufällige Funktion $\{Z(x,w) : x \in \mathbb{R}^d, w \in \Omega\}$, dabei bezeichnet

- $Z(x,\cdot)$ **Zufallsvariable**, kurz $Z(x)$
- $Z(\cdot,w)$ **regionalisierte Variable** (Realisierung der zufälligen Funktion), kurz $z(x)$

Sei $W \subset \mathbb{R}^d$ das Beobachtungsfenster,

dann bezeichnen $x_1,\ldots,x_n$ die Messstellen mit $x_i \in W$, $i = 1,\ldots,n$, und $z(x_1),\ldots,z(x_n)$ die Messwerte an den Messstellen.

Ein Zufallsfeld heißt **streng stationär**, wenn jede k-dimensionale Verteilung von $Z(x)$ invariant gegenüber beliebigen Translationen im $\mathbb{R}^d$ ist, d.h. wenn gilt

$$\mathbb{P}(Z(x_1) < z_1, Z(x_2) < z_2, \ldots, Z(x_k) < z_k) = \mathbb{P}(Z(x_1 + h) < z_1, Z(x_2 + h) < z_2, \ldots, Z(x_k + h) < z_k),$$

$\forall k \in \mathbb{R}$, $\forall x_1, \ldots, x_k \in \mathbb{R}^d$ und $\forall h \in \mathbb{R}^d$.

Ein Zufallsfeld heißt **stationär 2. Ordnung (schwach stationär)**, wenn gilt

$$\mathbb{E}\,[Z(x+h)] \,=\, \mathbb{E}\,[Z(x)] \,=\, m\,,$$

$$\mathrm{cov}[Z(x),\, Z(x+h)] \,=\, \mathbb{E}\,[\{Z(x)-m\}\{Z(x+h)-m\}] \,=\, C(h).$$

Ein Zufallsfeld heißt **intrinsisch stationär 2. Ordnung**, wenn gilt

$$\mathbb{E}\left[Z(x+h) - Z(x)\right] = m(h) = 0$$

$$\operatorname{Var}\left(Z(x+h) - Z(x)\right) = \mathbb{E}\left[(Z(x+h) - Z(x))^2\right] = 2\gamma(h).$$

Theoretisches Variogramm:

$$\gamma(h) = \frac{1}{2}\mathbb{E}\left[(Z(x+h) - Z(x))^2\right]$$

Existiert die Kovarianzfunktion, so gilt

$$\gamma(h) = C(0) - C(h).$$

NUR wenn das Variogramm durch einen endlichen Wert beschränkt ist, existiert eine Kovarianzfunktion.

2 Simple Kriging

Annahmen:

Das Zufallsfeld $Z(x)$ ist stationär 2. Ordnung, d.h.

$$\mathbb{E}\left[Z(x+h)\right] = \mathbb{E}\left[Z(x)\right] = m\,,$$

$$\mathrm{cov}[Z(x),\, Z(x+h)] = C(h)$$

und der Erwartungswert, sowie die Kovarianzfunktion sind bekannt.

Kriging-Schätzer:

$$Z^*(x_0) = m + \sum_{i=1}^{n} w_i\left(Z(x_i) - m\right)$$

Der Schätzer soll erwartungstreu sein, d.h. es soll gelten

$$\begin{aligned}
\mathbb{E}\left[Z^*(x_0) - Z(x_0)\right] &= 0 \\
&= m + \sum_{i=1}^{n} w_i \left(\mathbb{E}\left[Z(x_i)\right] - m\right) - \mathbb{E}\left[Z(x_0)\right] \\
&= m + \sum_{i=1}^{n} w_i \left(m - m\right) - m \;=\; 0.
\end{aligned}$$

Des weiteren soll die Varianz des Schätzfehlers minimal sein.

$$\begin{aligned}
\operatorname{Var}\left(Z^*(x_0) - Z(x_0)\right) &= \mathbb{E}\left[(Z^*(x_0) - Z(x_0))^2\right] \\
&= \mathbb{E}\left[(Z^*(x_0) - m)^2 + (Z(x_0) - m)^2 - 2\,(Z^*(x_0) - m)(Z(x_0) - m)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C(x_i - x_j) + C(x_0 - x_0) - 2\sum_{i=1}^{n} w_i C(x_i - x_0)
\end{aligned}$$

Das Optimierungsproblem lautet:

$$\{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C(x_i - x_j) + C(0) - 2\sum_{i=1}^{n} w_i C(x_i - x_0)\} \to min$$

Die gesuchten Gewichte erhält man durch das Lösen des Gleichungssystems

$$\sum_{j=1}^{n} w_j C(x_i - x_j) \;=\; C(x_i - x_0), \quad i = 1, \ldots, n.$$

Die Varianz des Schätzfehlers:

$$\sigma^2_{SK} \;=\; \mathbb{E}\left[(Z^*(x_0) - Z(x_0))^2\right] \;=\; C(0) - \sum_{i=1}^{n} w_i C(x_i - x_0)$$

Charakteristische Merkmale von $Z^*(x)$ und σ_{SK}^2 :

- Genaues Interpolationsverfahren: $Z^*(x_i) = z(x_i), \ \forall i = 1, \dots, n$
- Die Varianz des Schätzfehlers nimmt an den Messstellenorten $x_i, \ i = 1, \dots, n$, den Wert 0 an.

3 Kriging the Mean

Annahme:

$Z(x)$ ist stationär 2. Ordnung mit **unbekanntem** Erwartungswert, aber bekannter Kovarianz.

Erster Ansatz:

$$M_A^* = \frac{1}{n} \sum_{i=1}^{n} Z(x_i)$$

Zweiter Ansatz:

$$M^* = \sum_{i=1}^{n} w_i Z(x_i)$$

Der Schätzer M^* soll erwartungstreu sein, d.h. es soll gelten

$$\begin{aligned}\mathbb{E}\left[M^* - m\right] &= 0\\ &= \mathbb{E}\left[\sum_{i=1}^{n} w_i Z(x_i) - m\right] = \sum_{i=1}^{n} w_i\, m - m\\ &= \left(\sum_{i=1}^{n} w_i - 1\right) m.\end{aligned}$$

Anforderung an die Gewichte:

$$\sum_{i=1}^{n} w_i = 1$$

Das Optimierungsproblem mit Hilfe der Lagrange Funktion dargestellt lautet:

$$L(w_i, \mu) \;=\; \mathrm{Var}\,(M^* - m) + \mu(\sum_{i=1}^{n} w_i - 1).$$

Kriging System:

$$\begin{cases} \sum_{j=1}^{n} w_j C(x_i - x_j) + \mu \;=\; 0, \quad \text{für } i = 1, \ldots, n \\ \\ \sum_{j=1}^{n} w_j \;=\; 1 \end{cases}$$

Beispiel (Nugget-Effekt):

Der Schätzer des Kriging of the Mean ist äquivalent zum arithmetischen Mittel:

$$M^* = \sum_{i=1}^{n} w_i Z(x_i) = \frac{1}{n} \sum_{i=1}^{n} Z(x_i) = M_A^*.$$

Folglich gilt:

- Wenn es keine Korrelation zwischen den Messstellen gibt, ist der arithmetische Mittelwert der beste lineare erwartungstreue Schätzer des Mittelwertes.
- Kriging of the Mean ist eine Verallgemeinerung des arithmetischen Mittel Schätzers im Falle räumlich korrelierter Messstellen.

Frage:

Ist es zulässig zuerst mit Kriging the Mean den Erwartungswert zu schätzen und dann diesen Schätzwert in Simple Kriging zu verwenden?

Antwort:

JA. $\Longrightarrow$ Ordinary Kriging

Es gilt ferner:

$$Z^*_{SKM}(x_0) = Z^*_{OK}(x_0) \text{ und}$$

$$\sigma^2_{OK} = \sigma^2_{SK} + \overline{w}^2 \sigma^2_{KM},$$

wobei $\overline{w} = 1 - \sum_{i=1}^{n} w_i^{SK}$.

4 Nicht-stationäre Methoden

In der Praxis erfüllen manche Parameter nicht die stationäre Hypothese.

Die bisher vorgestellten Methoden berücksichtigen keinen systematischen Trend und würden so unakzeptable Resultate liefern.

Methoden, die einen systematischen Trend des Parameters berücksichtigen:

- Universal Kriging
- External Drift
- Intrinsic random functions of order k
- Residual Kriging

5 Universal Kriging

Annahmen:

Die Zufallsvariable $Z(x)$ läßt sich wie folgt zerlegen:

$$\begin{aligned} \mathbb{E}\left[Z(x)\right] &= m(x) \\ Y(x) &= Z(x) - m(x) \\ \mathbb{E}\left[Y(x)\right] &= 0 \end{aligned}$$

$$Z(x) = m(x) + Y(x)\,,$$

wobei $m(x)$ die deterministische Komponente ist und als *Drift* bezeichnet wird und $Y(x)$ die stochastische Komponente ist und als *Fluktuation* bezeichnet wird.

Annahmen an $m(x)$:

$$m(x) = \sum_{l=0}^{L} a_l f_l(x) \,,$$

wobei $f_l(x)$, $l = 0, \ldots, L$, bekannte Funktionen sind
und a_l unbekannte Koeffizienten, mit $a_l \neq 0$ für $l = 0, \ldots, L$.
Es gilt $f_0(x) = 1$.

Annahmen an $Y(x)$:

$Y(x)$ ist stationär 2. Ordnung mit $\mathbb{E}\left[Y(x)\right] = 0$ und Kovarianzfunktion $C(h)$.

Kiging-Schätzer:

$$Z^*(x_0) = \sum_{i=1}^{n} w_i Z(x_i)$$

Der Schätzer soll erwartungstreu sein, d.h. es soll gelten

$$\begin{aligned}
\mathbb{E}\left[Z^*(x_0) - Z(x_0)\right] &= 0 \\
&= \sum_{i=1}^{n} w_i m(x_i) - m(x_0) \\
&= \sum_{l=0}^{L} a_l \left(\sum_{i=1}^{n} w_i f_l(x_i) - f_l(x_0)\right).
\end{aligned}$$

Universelle Bedingungen:

$$\sum_{i=1}^{n} w_i f_l(x_i) \;=\; f_l(x_0),\; \text{für } l = 0, \dots, L$$

Sind die Universellen Bedingungen erfüllt, so gilt

$$\begin{aligned}\mathrm{Var}\,(Z^*(x_0) - Z(x_0)) \;&=\; \mathbb{E}\,[(Z^*(x_0) - Z(x_0))^2] \\ &=\; E[(\sum_{i=1}^{n} w_i Y(x_i) - Y(x_0))^2]\end{aligned}$$

$$= \textstyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C(x_i - x_j) - 2\sum_{i=1}^{n} w_i C(x_i - x_0) + C(x_0 - x_0).$$

Kriging System:

$$\begin{cases} \sum_{j=1}^{n} w_j C(x_i - x_j) + \sum_{l=o}^{L} \mu_l f_l(x_i) \; = \; C(x_i - x_0), & \text{für } i = 1, \ldots, n \\ \\ \sum_{j=1}^{n} w_j f_l(x_j) \; = \; f_l(x_0) & \text{für } l = 0, \ldots, L \end{cases}$$

In Matrix Notation

$$\begin{pmatrix} \mathbf{C} & \mathbf{F} \\ \mathbf{F}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{w} \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{f} \end{pmatrix}$$

Das Kriging System hat eine eindeutige Lösung, wenn die Kovarianzmatrix streng positive definit ist und die Matrix $\mathbf{F} = (\mathbf{f_0}, \ldots, \mathbf{f}_L)$ den Rang L+1 hat.

Bemerkungen:

- Der Einfluss des Drift-Modells auf das Kriging hängt von der Relation zwischen L+1 und der Anzahl n der Messstellen ab.
- Es gilt: $Z^*(x_i) = z(x_i)$, für $i = 1, \ldots, n$.
- Ferner gilt: $\mathbb{E}\left[(Z^*(x_i) - Z(x_i))^2\right] = 0$, für $i = 1, \ldots, n$.

Schätzung des Drifts:
Unter der Annahme, dass die Koeffizienten a_l zufällig sind, kann auch der Drift als zufällig angesehen werden:

$$M(x) = \sum_{l=0}^{L} A_l f_l(x) \text{ und } \mathbb{E}[A_l] = a_l, \text{ für } l = 0, \ldots, L.$$

Schätzer der Koeffizienten:

$$A_l^* = \sum_{i=1}^{n} \lambda_l^i Z(x_i), \text{ für } l = 0, \ldots, L$$

Der Schätzer $A_{l_0}^*$, $l_0 = 0, \ldots, L$, soll erwartungstreu sein, deshalb

$$\sum_{i=1}^{n} \lambda_{l_0}^i f_l(x_i) = \delta_{ll_0}.$$

Die Gewichte erhält man durch das Lösen des Gleichungssystems:

$$\begin{cases} \sum_{j=1}^{n} \lambda_{l_0}^{j} C(x_i - x_j) - \sum_{l=0}^{L} \mu_{ll_0} f_l(x_i) = 0, & \text{für } i = 1, \ldots, n \\ \\ \sum_{j=1}^{n} \lambda_{l_0}^{j} f_l(x_j) = \delta_{ll_0}, & \text{für } l = 0, \ldots, L \end{cases}$$

Der Lagrange Parameter $\mu_{l_0 l_1}$ drückt die Kovarianz zwischen zwei Driftkoeffizienten mit Indices l_0 und l_1 aus.

Geschätzter Drift am Ort x_0 :

$$M^*(x_0) = \sum_{l=0}^{L} A_l^* f_l(x_0)$$

Annahme:

Kovarianzfunktion von $Y(x) = Z(x) - m(x)$ ist unbekannt.

Theoretische Variogramm: $\gamma(h) = \frac{1}{2}\mathbb{E}\left[(Y(x+h) - Y(x))^2\right]$

Berechnung des empirischen Variogramms:

- Plotten der Variogrammwolke mittels $\gamma_{ij}^* = \frac{(y(x_i)-y(x_j))^2}{2}$.
- Variogrammwolke in Klassen von $h = x_i - x_j$ unterteilen und für jede Klasse den Mittelwert berechnen.

Erster Ansatz:

$$R^*(x_i) = Z(x_i) - M^*(x_i) = Z(x_i) - \sum_{l=0}^{L} A_l^* f_l(x_i)$$

Zweiter Ansatz: Iterative Prozedur

1. Bestimmung der Art des Drifts.
2. Mit Hilfe eines theoretischen Variogramms die Driftkoeffizienten bestimmen.
3. Berechnung des empirischen Variogramms.
4. Vergleich von theoretischem und empirischem Variogramm.
   Wenn Kurven gut übereinstimmen $\Longrightarrow$ STOP.
   Andernfalls die Prozedur vom 2. Schritt an mit einem neuen theoretischen Variogramm wiederholen.