KRIGING
geostatistické metody interpolace

Vlastnosti krigování

Metoda krigování se často označuje akronymem **BLUE** (Best Linear Unbiased Estimator – tedy nejlepší lineární nezkreslený odhad).

Výchozí podmínky krigování:

Odhadovaná hodnota je vypočtena jako **lineární** kombinace vstupních hodnot

$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i \cdot z(x_i)$ kde pro váhy platí $\sum_{i=1}^{n} \lambda_i = 1$

Nezkreslený (**nestranný**) odhad značí, že průměrná chyba tohoto odhadu je rovna nule

$$\sum (\hat{z}_i - z_i) = 0$$

Odhad je **nejlepší**, protože je minimalizován rozptyl odhadu

$$\sum (\hat{z}_i - z_i)^2 = \min.$$

Vlastnosti krigování

Pokud prostorově závislá náhodná kolísání nejsou překryta nekorelovaným šumem, potom může být semivariogram využit k určení vah λ_i potřebných pro interpolaci.

Procedura je podobná jako v případě metody vážených klouzavých průměrů s tím rozdílem, že právě váhy jsou odhadnuty geostatistickými metodami.

Váhy λ_i jsou zvoleny tak, aby odhad byl nestranný a odhad rozptylu byl menší, než jakákoliv jiná lineární kombinace pozorovaných hodnot (minimální).

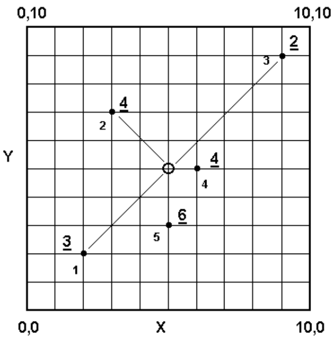
Přitom pro minimální rozptyl hodnot $\hat{z}(x_0)$ platí výraz :

$$\hat{\sigma}_e^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \phi \qquad \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_j) + \phi = \gamma(x_j, x_0)$$

$\gamma(x_i, x_j)$ semivariance proměnné *z* mezi body x_i a x_j.

ϕ Lagrangeův multiplikátor, který zajišťuje požadavek minimalizace odchylek a zároveň podmínku, že suma vah je rovna jedné.

PŘÍKLAD: Výpočet neznámé hodnoty v bodě metodou základního krigingu.



Postup řešení:

1. Na základě předem provedené strukturní analýzy použijeme vhodnou strukturní funkci (např. sférický semivariogram) s příslušnými hodnotami parametrů
2. Řešíme soustavu rovnic, kde jednotlivé členy mají následující význam:

$$A \cdot \lambda = b$$

A – matice semivariancí mezi všemi dvojicemi bodů
b – vektor semivariancí mezi všemi body a bodem predikovaným
λ – vektor vah jednotlivých bodů
Φ – tzv. Lagrangeův člen

3. Soustavu rovnic řešíme pro váhy λ, tak, aby byla splněna podmínka

$\sum \lambda = 1$ (proto je v soustavě použit člen *Φ)*

$$A^{-1} \cdot b = \begin{bmatrix} \lambda \\ \phi \end{bmatrix}$$

Postup řešení:

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & . & . & . & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & . & . & . & \gamma_{2n} & 1 \\ . & . & & & & & . \\ . & . & & & & & . \\ . & . & & & & & . \\ \gamma_{n1} & \gamma_{n2} & . & . & . & \gamma_{nn} & 1 \\ 1 & 1 & . & . & . & 1 & 0 \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ . \\ . \\ . \\ \lambda_n \\ \phi \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ . \\ . \\ . \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

Postup řešení:

4. K určení hodnot semivariancí je zapotřebí vytvořit matici vzdáleností mezi datovými body a vektor vzdáleností mezi měřenými body a bodem predikovaným

5. Řešením soustavy rovnic získáme hodnoty vah λ a hodnotu *Φ*

6. Vypočteme predikovanou hodnotu v bodě (i=0):

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i \cdot z(x_i)$$

$Z(x_{i=0})$ **=4,560**

7. Vypočteme rozptyl odhadu:

$$\hat{\sigma}_e^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \phi$$

σ_e^2 **= 4,008**

Typy krigování

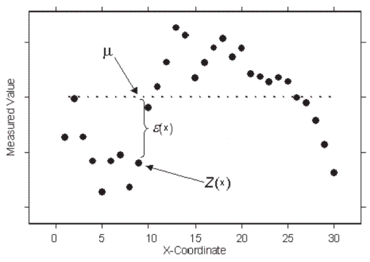
- krigování s bodovým odhadem
- krigování s blokovým odhadem

- základní (ordinary) krigování
- jednoduché (simple) krigování
- univerzální krigování
- indikátorové krigování
- pravděpodobnostní krigování
- co-kriging
- lognormální krigování.

Základní krigování (ordinary kriging)

$$Z(x_i) = \mu(x_i) + \varepsilon(x_i)$$

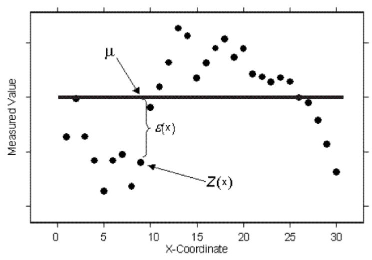
kde μ je neznámá hodnota trendu.



Jednoduché krigování (Simple kriging)

$$Z(x_i) = \mu(x_i) + \varepsilon(x_i)$$

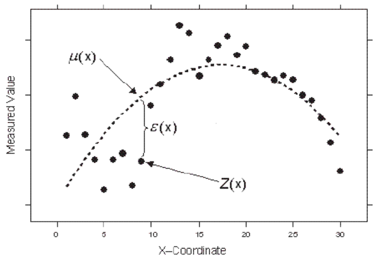
kde μ je známá konstanta.



Univerzální krigování (Universal kriging)

$$Z(x_i) = \mu(x_i) + \varepsilon(x_i)$$

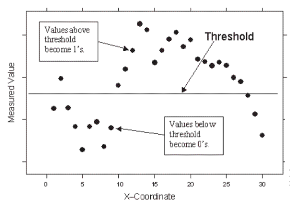
kde $\mu(x)$ je deterministická funkce



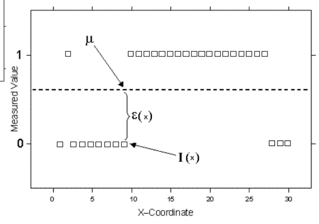
Indikátorové krigování

$$I(x_i) = \mu + \varepsilon(x_i)$$

kde μ neznámá konstanta, *ε(x)* autokorelovaná náhodná proměnná a *I(x)* je binární proměnná, která nabývá hodnot 0 nebo 1 - indikátor



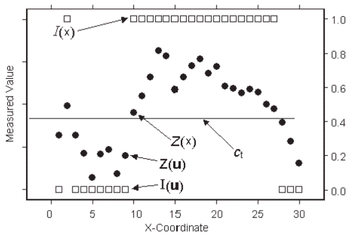
Výsledkem odhadu je pravděpodobnost překročení určité prahové hodnoty



Pravděpodobnostní krigování (probalility kriging)

$$I(x) = I(Z(x) > c_1) = \mu_1 + \varepsilon_1(x) \qquad Z(x) = \mu_2 + \varepsilon_2(x)$$

kde μ_1 a μ_2 jsou neznámé konstanty. *I(x)* je binární proměnná vytvořená indikátorovým prahováním (*I(Z(x) > c_1)*) a *$\varepsilon_1(x)$* a *$\varepsilon_2(x)$* jsou dvě náhodné chyby.

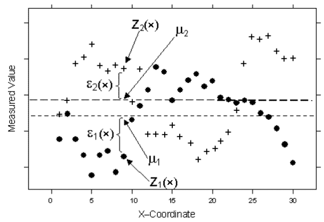


Co-kriging

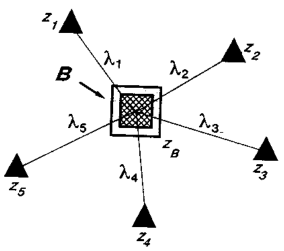
Proměnné z_1 a z_2 vykazují prostorovou korelaci.
Proměnná z_2 je snáze získatelná
Hodnot proměnné z_2 je využito k interpolaci hodnot proměnné z_1.
Základní co-kriging využívá následujících modelů:

$$Z_1(x) = \mu_1 + \varepsilon_1(x)$$
$$Z_2(x) = \mu_2 + \varepsilon_2(x)$$

kde μ_1 a μ_2 jsou neznámé konstanty a *$\varepsilon_1(x)$* a *$\varepsilon_2(x)$* dvě náhodné chyby.



Blokový odhad při základním krigování (Block kriging)



Hodnocení a verifikace modelů

Krigování jako interpolační metoda umožňuje pro každý interpolovaný bod odhadnout potenciální velikost chyby odhadu.

Mapy druhé mocniny σ_e^2

tzv. **směrodatné chyby** (odchylky) **krigingu (Standard error map)**,

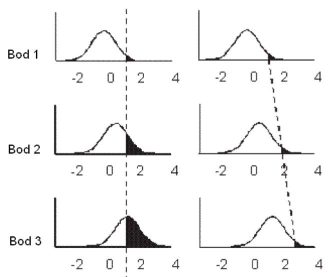
Mají-li chyby predikce normální rozdělení, potom 95% **interval spolehlivosti predikovaných hodnot** lze určit z následujícího vztahu:

$$Z(x_0) \pm 1{,}96 \cdot \sqrt{\sigma_e^2}$$

kde *$Z(x_0)$* je odhad hodnoty proměnné *z* v bodě x_0 a σ_e^2 je rozptyl odhadu.

Při opakovaném použití stejného modelu padne 95 % odhadovaných hodnot do uvedeného intervalu

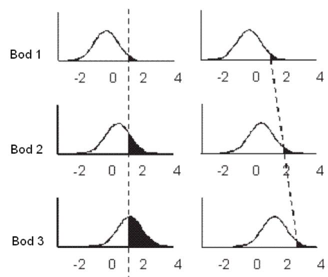
Probability map



Pravděpodobnost, že predikovaná hodnota bude větší než prahová hodnota - např. 1.

Při konstantní prahové hodnotě se její pravděpodobnost výskytu pro jednotlivé body mění – tedy lze z ní vytvořit mapu pravděpodobností (**probability map**).

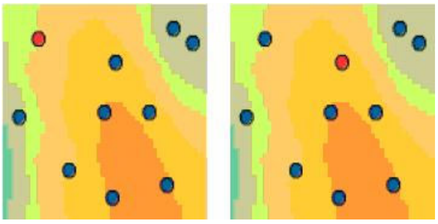
Quantile map



Určení kvantilu s např. 5 procentní pravděpodobností výskytu.

Při konstantní pravděpodobnosti se budou měnit hodnoty kvantilů a lze je opět prezentovat ve formě kvantilové mapy (**quantile map**).

Křížová validace modelu



Jednotlivé body měření (červené) jsou po jednom postupně vynechány ze vstupní množiny dat

Ze zbývajících (modrých) je vypočtena hodnota v místě vynechaného bodu.

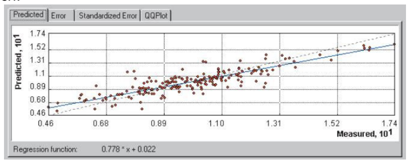
Následně jsou porovnány pozorované a vypočtené hodnoty

Validace modelu

Vstupní soubor měřených hodnot rozdělí na dvě části – data trénovací a testovací.

Trénovací množina dat se použije pro odhad trendu a autokorelačního modelu.

Pokud sestavený model vyhovuje trénovacím datům, je ověřen na datech testovacích.



Korelační pole měřených a predikovaných hodnot

Obecnou vlastností krigingu jako interpolační metody je podhodnocení vysokých hodnot a naopak nadhodnocení hodnot nízkých.

Sumární statistika rozdílů pozorovaných a vypočtených hodnot

Rozdíly pozorovaných a vypočtených hodnot lze hodnotit dále uvedenými měrami:

MPE – mean prediction error - průměr rozdílů měřených a předikovaných hodnot - hodnoty chyb odhadů by měly být nestranné – tedy jejich průměr by se měl rovnat nule.

$$MPE = \frac{\sum_{i=1}^{n}(\hat{Z}(x_i) - z(x_i))}{n}$$

RMSPE (root mean square prediction error) – druhá odmocnina průměrného čtverce vzdálenosti vypočtených hodnot od teoretických. Slouží k porovnání několika různých modelů. Čím menší je RMSPE, tím vhodnější je model (tím bližší jsou vypočtené hodnoty hodnotám měřeným).

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Z}(x_i) - z(x_i))^2}{n}}$$