

Názov prednášky:

Vybrané interpolačné metódy - Kriging

Osnova prednášky:

- Úvod do krigingu
- Semivariancia a kovariancia
- Semivariogram

Odporúčaná literatúra

KAŇUK, J., 2015: Priestorové analýzy a modelovanie. Vysokoškolské učebné texty. Prírodovedecká fakulta UPJŠ v Košiciach, 106 s.

HLÁSNY, T., 2005: Geoštatistický koncept priestorovej závislosti pre geografické aplikácie. Geografický časopis. Vol. 57, no. 2 (2005), p. 97-116. ([pdf](#))

ATKINSON, P. M., LLOYD, C. D., 2014: Geostatistical models and spatial interpolation (kapitola). Geostatistical models and spatial interpolation. In Handbook of Regional Science (pp. 1461-1476).

Úvod

- Kriging patrí medzi stochastické metódy
- je súčasťou geoštatistických metód.
- Metóda dostala názov krigeage (kriging) na počesť D. G. Krigea, ktorý sa snažil odhadnúť najpravdepodobnejšie rozloženie nálezísk zlata na základe dát z niekoľkých vrtov.
- Kriging sa používa na opis a hodnotenie prírodných javov opísaných premennými hodnotami v priestore.
- táto metóda môže byť použitá pri modelovaní javov, pri ktorých sa predpokladá spojitý priebeh hodnôt alebo aspoň určitý vzťah medzi hodnotami javu lokalizovaného v priestore
 - vstupné dáta majú byť priestorovo autokorelované.
- základnou myšlienkou metódy je nájsť priestorovú závislosť, ktorá charakterizuje vlastnosti javu v priestore, ako to vyplýva z meraných dát a tieto odhalené vlastnosti použiť vo forme matematickej funkcie aj na zvyšný priestor.

Kriging – všeobecný zápis

- kriging ako interpolačná metóda je tvorená z troch častí, pričom všetky časti sú analyzované samostatne.
- Prvú časť predstavuje tzv. všeobecný trend, ktorý sa odvodzuje na základe trendovej funkcie.
- Druhú časť predstavuje kolísanie hodnôt, tzv. regionalizovaná premenná, ktorú je možné vyjadriť matematickou funkciou, avšak vyjadruje určitú priestorovú závislosť. Táto časť je určená na základe **variogramov**.
- Tretiu zložku tvorí náhodný šum.
- Z uvedeného je možné zapísať nasledujúci vzťah:

$$\mathbf{z}(\mathbf{x}) = \boldsymbol{\mu}(\mathbf{x}) + \boldsymbol{\gamma}(\mathbf{h}) + \boldsymbol{\varepsilon}'$$

$\mathbf{z}(\mathbf{x})$ – je interpolovaná hodnota v bode $\mathbf{x}$ so súradnicami x a y .

- ide o určenú funkciu a nie o hodnotu určenú na základe merania.

$\boldsymbol{\mu}(\mathbf{x})$ – (čítaj mý) predstavuje deterministickú zložku, ktorá je určená na základe trendovej funkcie v bode $\mathbf{x}$ so súradnicami x a y .

$\boldsymbol{\gamma}(\mathbf{h})$ – (čítaj gama) reprezentuje stochastickú zložku, ktorú sa označuje ako regionalizovanú premenná avšak priestorovo nezávislú od trendovej funkcie $\boldsymbol{\mu}(\mathbf{x})$, $\mathbf{h}$ je hodnota vektora vzdialená od bodu $\mathbf{x}$ so súradnicami x a y .

$\boldsymbol{\varepsilon}'$ - (čítaj epsílon) náhodná, priestorovo nezávislá zložka, tzv. gaussovský šum s nulovým priemerom a rozptylom σ^2 (čítaj sigma).

Princíp

- základom pre určenie interpolovanej hodnoty $z(x)$ je štruktúrna analýza.
- **Geoštatistická štruktúrna analýza** (tzv. variografia) je proces zahŕňajúci výpočet štrukturálnych funkcií, výber a konštrukciu zodpovedajúcich teoretických modelov, ich aplikáciu a interpretáciu priebehu štrukturálnych funkcií.
- Cieľom štruktúrnej analýzy je popísanie vlastností vstupného bodového poľa, ako sú **kontinuita, homogenita, stacionarita a anizotropia**.
- Tieto vlastnosti sú určované na základe mier priestorovej autokorelácie a priestorovej variability.
- Štruktúrna analýza je východiskom pre geoštatistické modelovanie, avšak sama o sebe poskytuje širokú paletu dôležitých informácií o štruktúre náhodného poľa ako modelu konkrétneho objektu v krajinnej sfére.
- Ku kvantifikácii priestorovej autokorelácie slúži štruktúrna analýza, ktorá určuje silu korelačného vzťahu pre súbor pravidelne rozmiestnených bodov na línii.
- Charakteristiky, ktoré ju popisujú môžu byť napr. **priemer**, **rozptyl**, **kovariancia** a **semivariancia**.
- Jednou z najpoužívanějších charakteristík v geoštatistike je semivariancia a kovariancia.

Semivariancia

- Pre zjednodušenie uvedieme vysvetlenie semivariancii na 1D priestore, kde semivarianciu môžeme definovať ako:

$$\gamma(x_i x_j) = \frac{1}{2} \text{var} \left(z(x_i) - z(x_j) \right)$$

kde $\text{var} \left(z(x_i) - z(x_j) \right)$ je rozptyl, teda rozdiel hodnôt sledovaného javu v bodoch x_i a x_j .

- ak sú dva body x_i a x_j blízko seba, bude rozdiel hodnôt $z(x_i)$ a $z(x_j)$ malý
- s rastom vzdialenosti sa podobnosť hodnôt bude znižovať
- grafické znázornenie závislosti semivariácie na vzdialenosti je štruktúrna funkcia, ktorá sa nazýva **semivariogram**
- Semivariogram je jeden z najpoužívanejších štruktúrnych funkcií v geoštatistike, ktorý vyjadruje mieru nepodobnosti

Kovariancia

- Iný typ štrukturálnej funkcie:

$$C_{(x_i x_j)} = cov \left(z(x_i), z(x_j) \right)$$

kde $cov \left(z(x_i), z(x_j) \right)$ vyjadruje mieru podobnosti

- ak sú dva body x_i a x_j blízko seba, bude ich kovariancia veľká
 - hodnota kovariancie je pri nulovej vzdialenosti rovná rozptylu množiny spracovávaných dát
-
- **Vzťah** medzi **semivariogramom** a **kovariačnou funkciou** je možné vyjadriť pomocou vzťahu:

$$\gamma_{(x_i x_j)} = sill - C_{(x_i x_j)}$$

kde $sill$ označuje tzv. **prahovú hodnotu**, na ktorej má semivariogram vodorovný priebeh

Semivariogram

- Aby bolo možné popísať závislosť hodnôt v 2D priestore, určíme hodnoty semivariacie pre všetky dvojice bodov na základe vzťahu:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [z(x_i) - z(x_i + h)]^2$$

kde

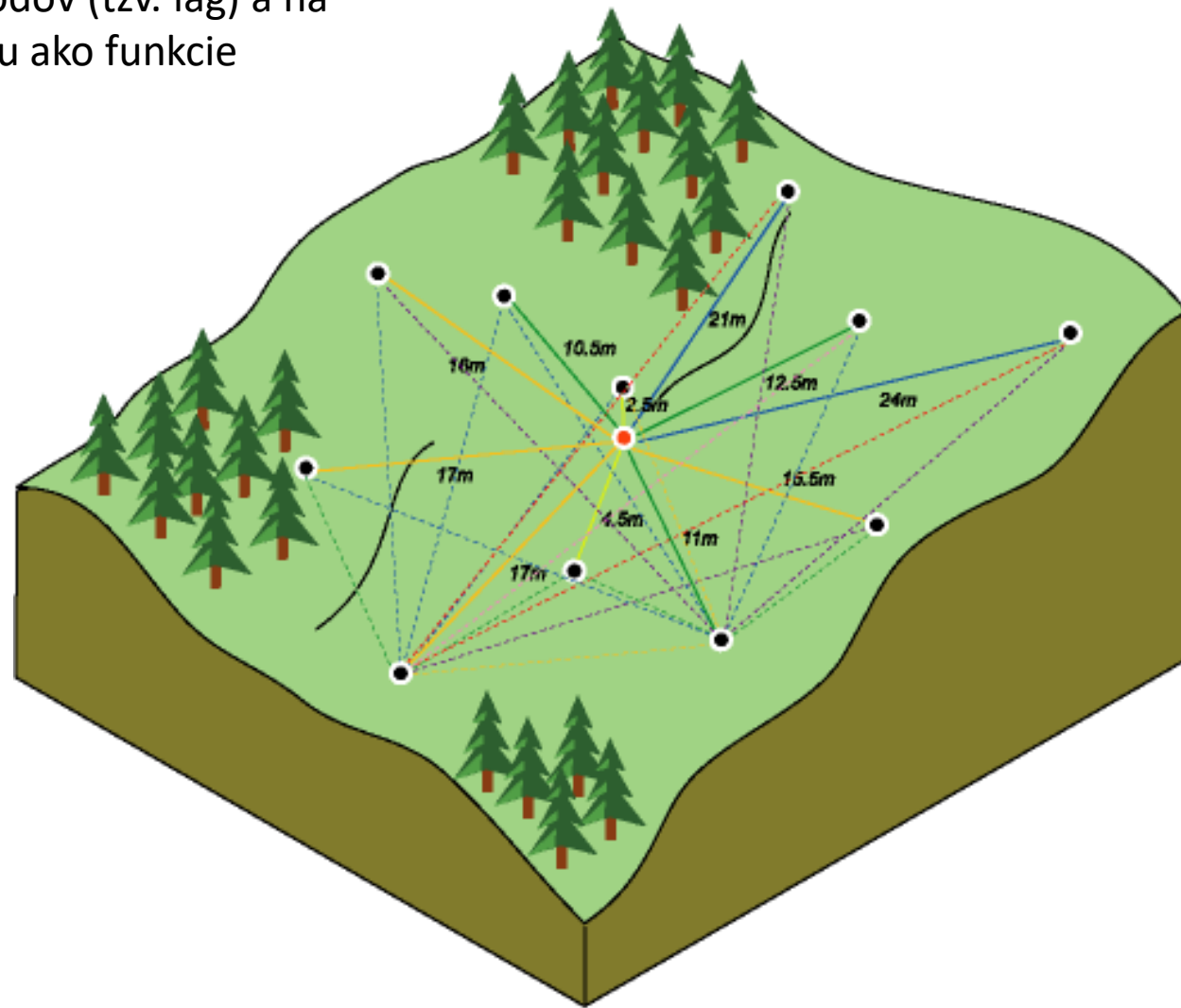
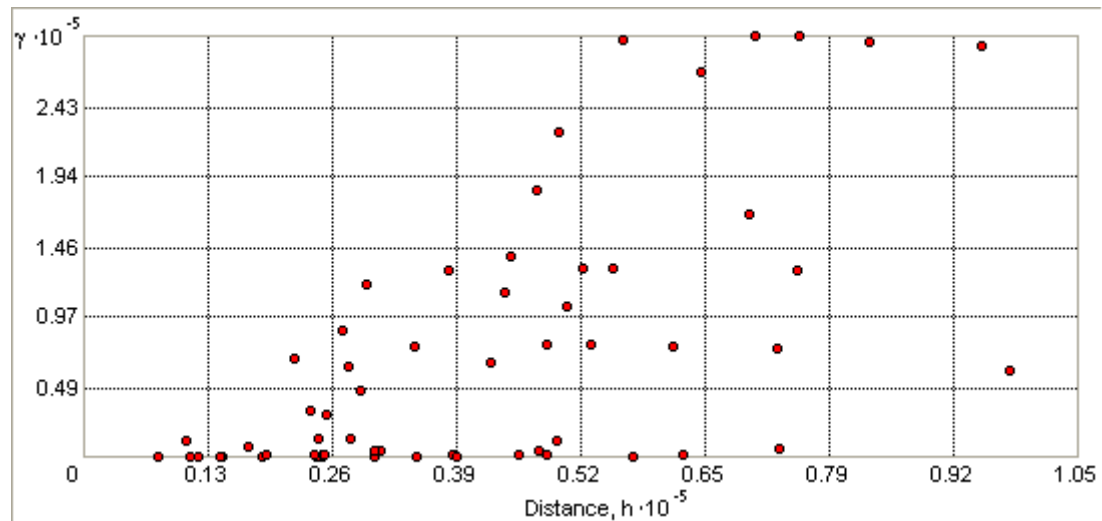
n – je počet dvojíc z bodov vstupného bodového poľa s hodnotou z vzdialených o hodnotu h

h - (*lag*) vzdialenosť danej dvojice bodov,

a následne určíme semivariogramy, cez ktoré sa snažíme preložiť funkciu vyjadrujúcu závislosti

Experimentálny (empirický) semivariogram

- Graf hodnôt $\gamma(h)$ sa označuje ako experimentálny semivariogram
- Na horizontálnu os sú vynášané vzdialenosti danej dvojice bodov (tzv. lag) a na vertikálnu os sú nanášané hodnoty rozptylu sledovaného javu ako funkcie vzájomných vzdialeností jednotlivých meraných bodov



Semivariogram

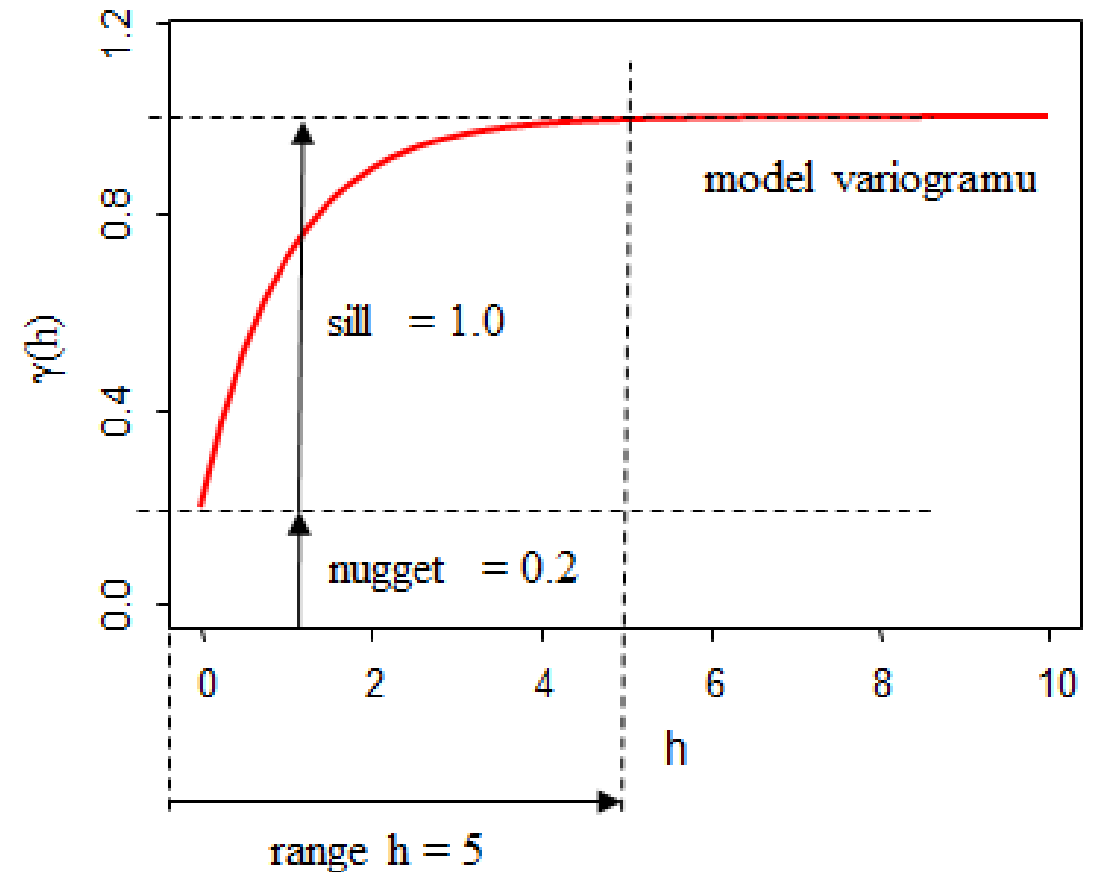
nugget – zvyškový rozptyl alebo veľkosť posunutia semivariogramu v jeho začiatku;

- ide o hodnotu vyjadrujúcu tzv. nespojitosť na začiatku, resp. hodnotu variancie pri veľmi malých hodnotách veľkosti vektora h väčšiu ako nula.

- tento pojem pochádza z banského inžinierstva

(kde pri porovnaní výsledkov analýzy koncentrácie určitej rudy dvoch vrtov oddelených minimálnou vzdialenosťou by mala byť hodnota variancie nulová, alebo zanedbateľná – ale len v prípade, že jeden z vrtov neobsahoval zlatý nugget. V tomto prípade bol rozdiel v zrudnení týchto dvoch lokalít významný, čo sa prejavilo v zmenení priebehu štruktúrnej funkcie, ktorá pri veľkosti $h = 0$ nadobudne hodnotu variancie väčšiu ako nula)

- Ďalšie z príčin posunutia variogramu sú napr. chyby pri zbere údajov, ktoré môžu mať viaceré príčiny (limitujúca schopnosť zariadenia merať presnejšie, nepresnosť v spracovaní dát, subjektivnosť pri zbere dát a pod.).
- Zvyškový rozptyl je tak vyjadrením náhodného šumu ε' .



sill – prah alebo lepšie prahová hodnota variogramu vo vzdialenosti, pri ktorej je dosiahnutý *range*.

- Ideálnym prípadom je, keď sa pri väčších vzdialenostiach jednotlivé regionalizované premenné stanú priestorovo nezávislé a dosiahnutá hodnota variogramu zostáva viac-menej konštantná.
- Hodnota prahu je súčasne odhadom výberovej variancie (rozptylu) sledovaného javu.

Kriging

Ak teda vyselektujeme ľubovoľný súbor regionalizovaných premenných s rozsahom n , tak platí vzťah:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = C$$

pričom

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Prahová hodnota variogramu

Z hľadiska spôsobu, akým variogramy dosahujú hodnotu prahu, rozlišujeme teoretické modely variogramov na:

1. bez prahové – lineárny, logaritmický
 2. s prahom – sférický, exponenciálny, Gaussov
 3. s nuggetovým efektom
- v prípade exponenciálneho a Gaussovho variogramu nie je dosiahnutý skutočný prah
 - ako prah sa v týchto prípadoch označuje vzdialenosť, v ktorej variogram dosiahol 95 % hodnoty dosahu, čo sa označuje ako účinný alebo praktický dosah

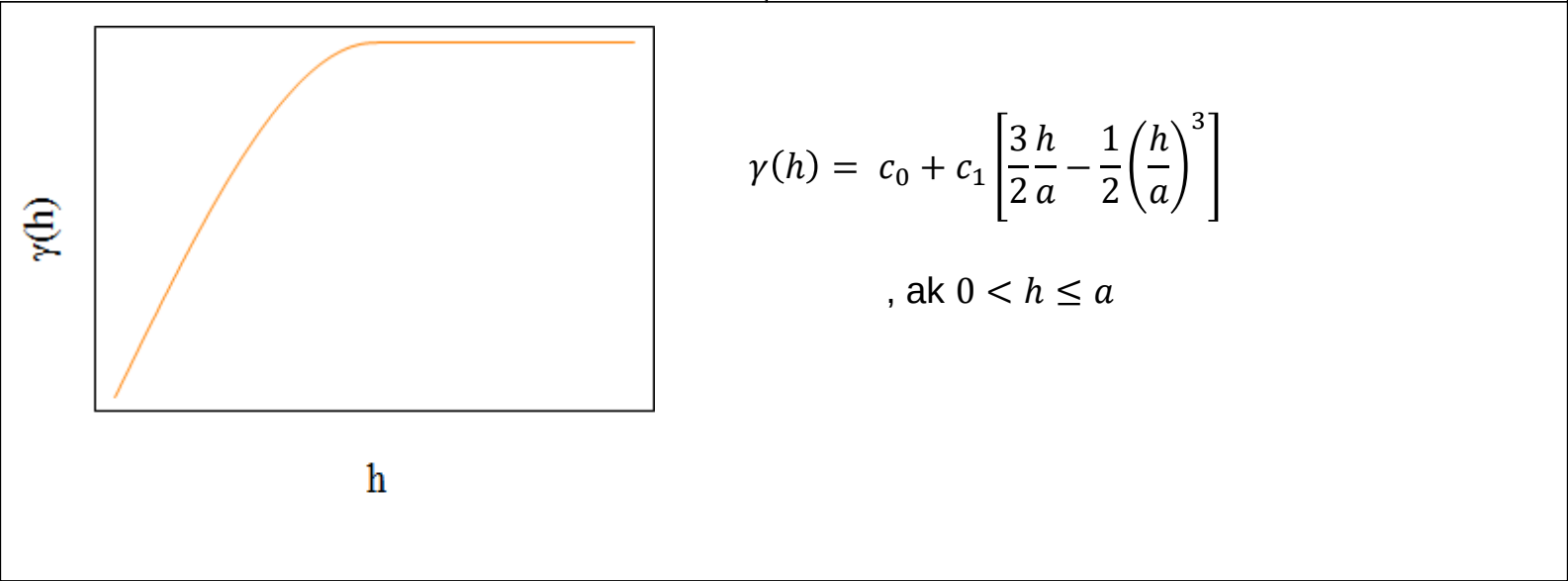
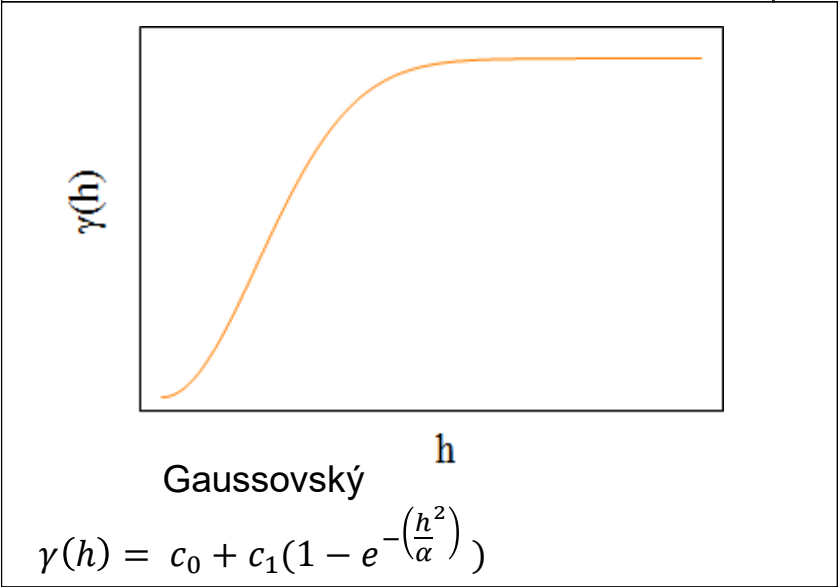
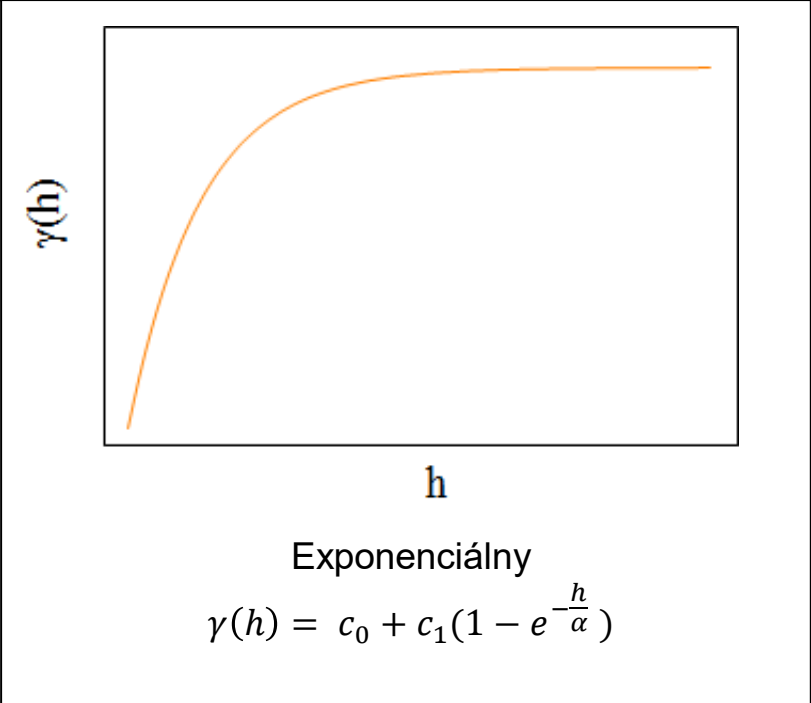
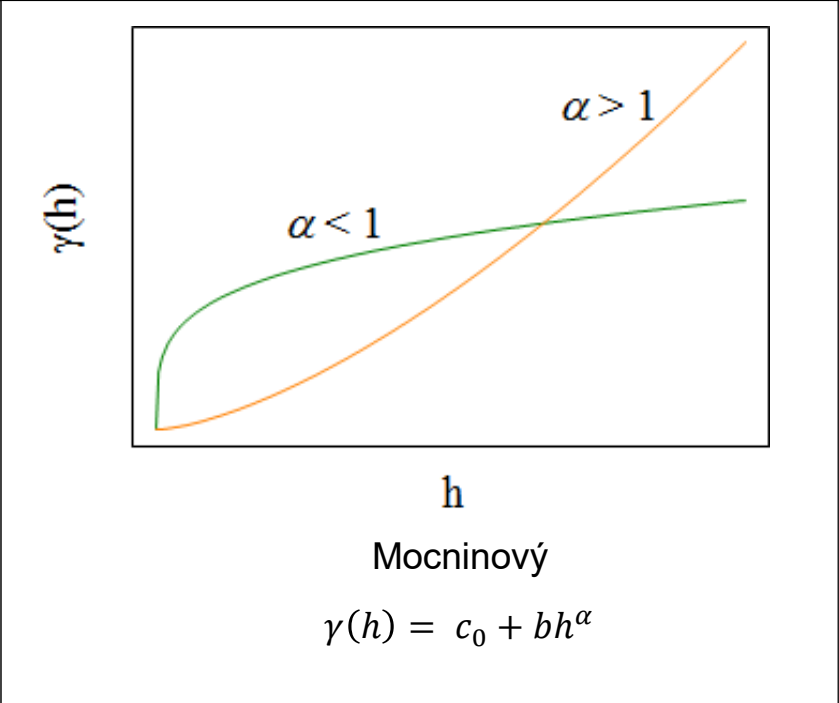
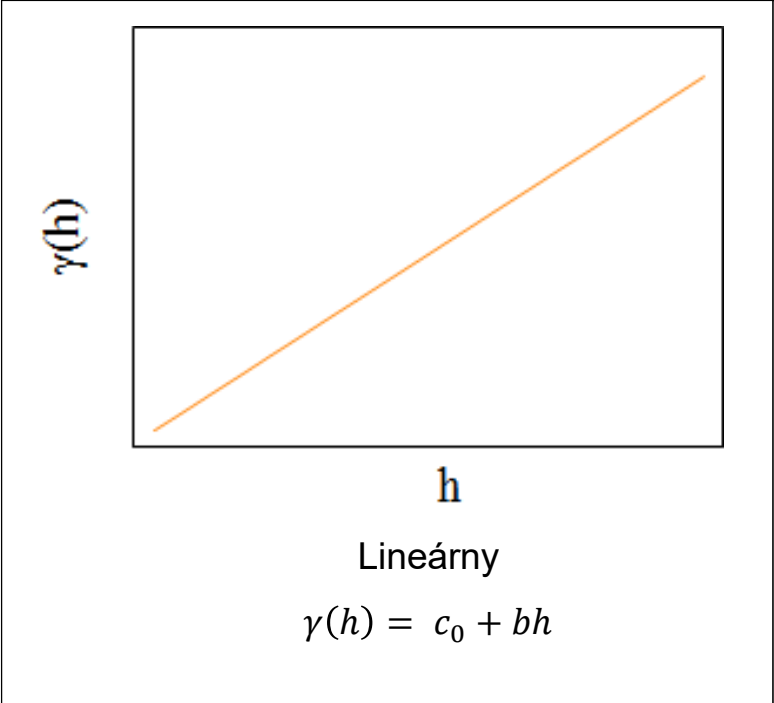
range – dosah, je hodnota vzdialenosti, v ktorom je rozdiel variogramu už zanedbateľný, resp. v ktorej sa štruktúrna funkcia stabilizuje

- za touto hodnotou sa nie sú polohy sledovaného javu korelované
- hodnota dosahu môže byť interpretovaná ako zóna vplyvu regionalizovaných premenných
- za špecifický prípad teoretického modelu variogramu je možné považovať tzv. účinok prepadu (hole effect), ktorý sa prejavuje obnovením priestorovej závislosti po tom, ako hodnota dosahu už bola dosiahnutá

Regionalizované premenné

- na základe variogramov je možné kvantitatívne popísať regionalizované premenné
- experimentálnym semivariogramom sa v nasledujúcom kroku prekladá teoretický model semivariogramu
- existuje mnoho teoretických modelov semivariogramov (napr. sférický, tetrasférický, pentasférický, exponenciálny, gaussovský, všeobecný lineárny, K-Besselov, J-Besselov, logaritmický a ďalšie).
- jeho výber určuje používateľ, a preto je značne subjektívny, čo sa často uvádza ako nedostatok tejto metódy

Teoretické modely semivariogramov



Záver

- ak parametre semivariogramu sú rôzne v rôznych smeroch, tak študované pole je anizotrópne. Môže ísť o geometrickú alebo zonálnu anizotrópiu.
- ide o výpočtovo náročnú úlohu, do ktorej vstupujú aj ďalšie parametre
- **Kokriging** je geoštatistická interpolačná metóda, ktorá umožňuje pri interpolácii skúmanej premennej využívať informácie o priestorovom rozložení doplnkovej, sekundárnej premennej
- využíva sa najmä v situáciách, keď sa predpokladá vzťah medzi skúmanou (interpolovanou) a doplnkovou premennou v situácii, keď informácie o priestorovom rozložení doplnkovej premennej sú podrobnejšie ako pri interpolovaných dátach.
- Pri kokrigingu sa vychádza podobne ako pri krigingu z výsledkov geoštatistickej analýzy.
- Na rozdiel od krigingu je však potrebná konštrukcia nielen semivariogramu skúmanej premennej, ale aj semivariogramu doplnkovej premennej a krížového semivariogramu medzi skúmanou a doplnkovou premennou.

Aplikácie:

- Najčastejšie aplikácie krigingových metód v praxi sú najmä v geológii, chémii, geomorfológii a hydrológii.
- V súčasnosti väčšina softvérov GIS umožňujúca spracovanie rastrových dát obsahuje moduly na interpoláciu, resp. geoštatistiku.
- Geoštatistické metódy poskytujú veľké možnosti analýzy skúmaného javu, avšak aby boli vhodne použité a interpretované, vyžadujú si znalosti z geoštatistiky.