Teorie chyb a vyrovnávací počet 2

Téma č. 6+7: **Aproximace vztahů. Regresní a korelační analýza. Vyrovnávací přímka a rovina. Aproximace empirickým polynomem.**

1. Regrese.
2. Lineární regrese.
3. Korelace.
4. Koeficient korelace.
   1. Výpočet
   2. Výpočet z kovarianční matice
   3. Testování empirického koeficientu korelace
5. Vyrovnávací rovina.
6. Aproximace empirickým polynomem.
7. Aproximace v parametrickém tvaru
   1. Kružnice
   2. Koule
   3. Přímka

1. Regrese

V přírodě často probíhají jevy jako funkce jedné nebo více proměnných $y = f(x)$, $z = f(x, y)$, atd., u nichž předem neznáme přesně typ a konstanty (parametry) funkce a teprve je zjišťujeme empiricky.

K empirickému určení analytického typu funkce a číselných hodnot konstant sledujeme průběh jevu měřením hodnot závisle proměnné y při měnících se hodnotách argumentu x. Grafické znázornění průběhu jevu dá vlivem měřických chyb nebo jiných rušivých vlivů nepravidelnou řadu bodů (empirický polygon). Úkolem je najít takovou funkční závislost mezi proměnnými x a y, aby průběh funkce co nejlépe vyjadřoval měřený průběh jevu, tj. aby se vyrovnávací křivka při jednoduchém tvaru funkce dostatečně přimkla empirickému polygonu. Zpravidla k tomu použijeme metodu nejmenších čtverců.

Měřické chyby nebo rušivé vlivy a neznalost přesného analytického typu funkce způsobují, že typ funkce a její konstanty neurčíme s absolutní přesností. Mluvíme proto o aproximaci empirických funkcí a výsledkem bude tzv. regresní křivka.

Druhou, neméně důležitou otázkou, kterou řešíme při hledání aproximačních funkcí, je síla (věrohodnost) jejich platnosti. Regresní analýza sice určí tvar funkce, ale korelační analýza určí věrohodnost platnosti tohoto vztahu v rozmezí absolutní nezávislosti až po funkční vztah dvou veličin.

1. Regrese

- Aproximace měřených dat matematickou křivkou $y = f(x)$ (plochou $z = f(x, y)$, vícerozměrnou funkcí $y = f(x_1, x_2, \dots, x_n)$) .
- Výpočet MNČ
- Opravy lze přiřazovat x; y; x a y, atd., každé toto řešení je jiné.
- Obvykle neznáme přesně model (= funkci vysvětlující vztah proměnných a konstant), lze odhadovat např. na základě grafického znázornění. Pro mnohá měření ani přímý matematický vztah neexistuje , čili se jedná o aproximaci (odhad, nahrazení).

2. Lineární regrese

- Nejjednodušší případ, aproximace měřených dat matematickou křivkou - přímkou
- $y = f(x) = ax + b$
- Výpočet MNČ
- Opravy lze přiřazovat x; y; x a y, každé toto řešení je jiné, použije se případ, který vyhovuje původu chyb.
- Často vyžívané, a to i v nelineárních případech pro popis malých úseků, kde s dostatečnou přesností platí lineární zjednodušení.

- Výpočet: (viz. odvození z 1. semestru)
    - Vyrovnání zprostředkujících

3. Korelace 4. Korelační koeficient

Funkční vztah

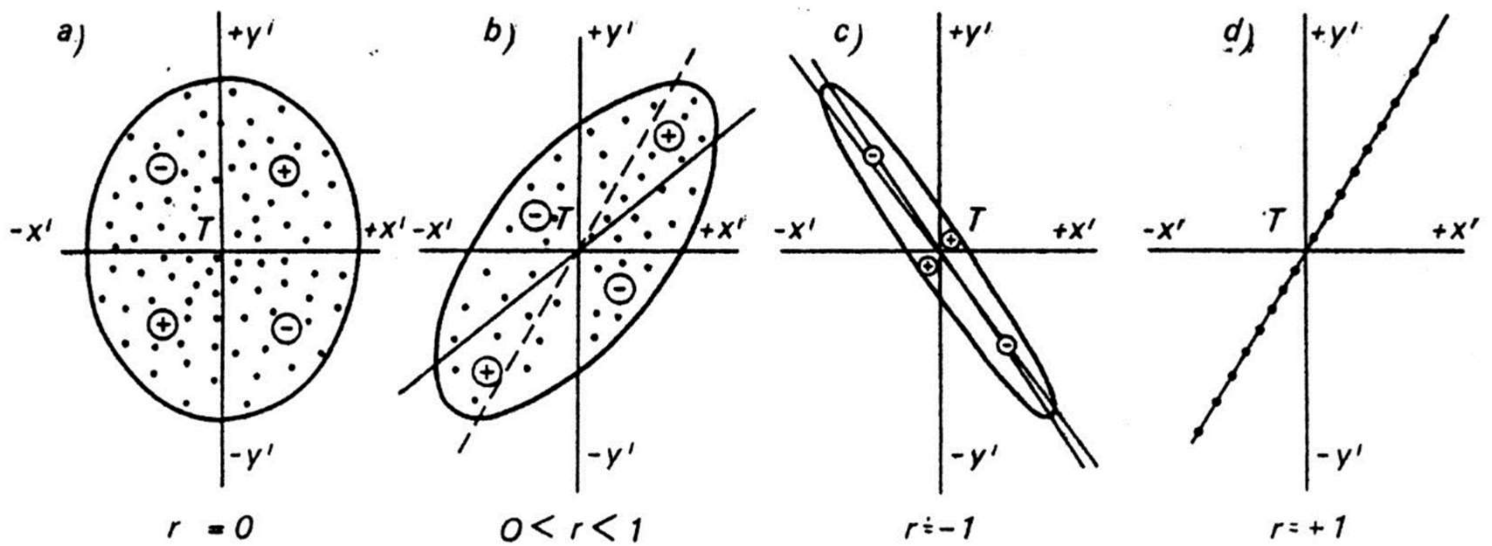
- vztah dvou (či více) veličin x a y lze popsat matematickou funkcí, od které nevykazují žádné odchylky. Pro jednu hodnotu x existuje dle předpisu pouze jedna hodnota y.

Korelační závislost, stochastický (statistický) vztah

- mezi veličinami (měřeními) je vztah, nelze jej však beze zbytku vyjádřit.

Korelační koeficient

- Popisuje míru statistického vztahu, rozsah $\langle -1; 0; 1 \rangle$



4. Korelační koeficient

Korelační koeficient

- Popisuje míru statistického vztahu, rozsah $\langle -1; 0; 1\rangle$
- Pro lineární vztah:

$$r = \frac{[px\prime y\prime]}{\sqrt{[px\prime \quad][py\prime y\prime]}}$$

Z kovarianční matice:

$$\boldsymbol{M} = \begin{pmatrix} \sigma_{x1}^2 & Cov_{12} & Cov_{13} & \dots & Cov_{1n} \\ Cov_{21} & \sigma_{x2}^2 & Cov_{23} & \dots & Cov_{2n} \\ Cov_{31} & Cov_{32} & \sigma_{x3}^2 & \dots & Cov_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Cov_{1n} & Cov_{2n} & Cov_{3n} & \dots & \sigma_{xn}^2 \end{pmatrix}$$

$$r_{ij} = \frac{Cov_{ij}}{\sqrt{\sigma_{xi}^2 \cdot \sigma_{xj}^2}} = \frac{\sigma_{xy}}{\sigma_x \sigma_x}$$

$$Cov = E\left(X - E(X) \cdot \left(Y - E(Y)\right)\right) = \frac{[v_x v_y]}{n}$$

4. Korelační koeficient

Testování empirického korelačního koeficientu

- Různé možnosti, zde testování nenulovosti korelačního koeficientu r
  - r … výběrový
  - ρ … základní
  - Nulová hypotéza: $\rho = 0$, za tohoto předpokladu má veličina t_r Studentovo rozdělení s $(n-2)$ stupni volnosti
  - $t_r = r \cdot \frac{\sqrt{n-2}}{\sqrt{1-r^2}}$
  - Při překročení kritické hranice zamítáme nulovou hypotézu.

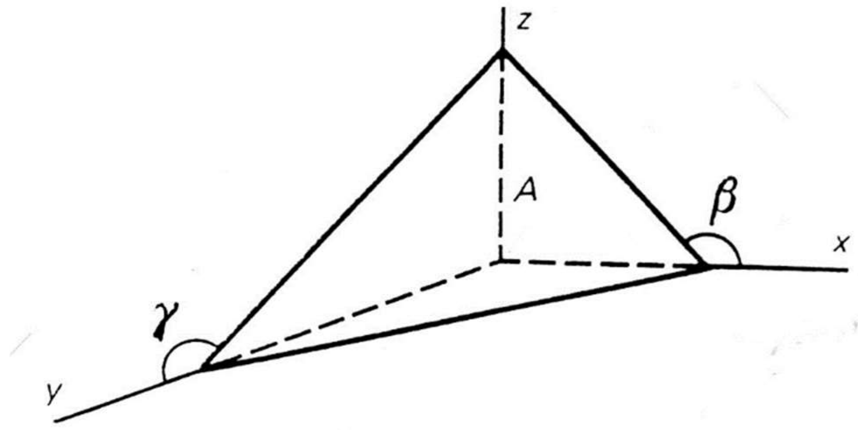
5. Vyrovnávací rovina

Pro technická řešení úloh v terénu bude nejvhodnější lineární funkcí v prostoru taková vyrovnávací rovina, která dá nejmenší odchylky ve směru vertikální osy z (v nadmořských výškách), tj. např. umožní minimální přemisťování zemin. Podmínka bude mít proto tvar (předpokládáme váhy rovné jedné, tj. $p_i = 1$). $[v_z v_z] = min$.

$$\boldsymbol{ax + by + cz - d = 0}$$

Vodorovná rovina

Obecná rovina



6. Aproximace empirickým polynomem

Volba stupně (tvaru)

Výpočet

$$\boldsymbol{A}, \boldsymbol{L}, \boldsymbol{X}, \boldsymbol{s_0}$$

Kvalita aproximace

Určení významnosti koeficientů

Bivariantní polynom

Ortogonální polynom

7. Aproximace v parametrickém tvaru

Přímka

Kružnice

Koule

☺ Konec ☺