

9. REGRESNÁ A KORELAČNÁ ANALÝZA

Vo vedeckých a inžinierskych analýzách sa často stretávame s kvantitatívnym hodnotením dvoch a viac veličín, ktoré vyjadrujeme funkčným vzťahom

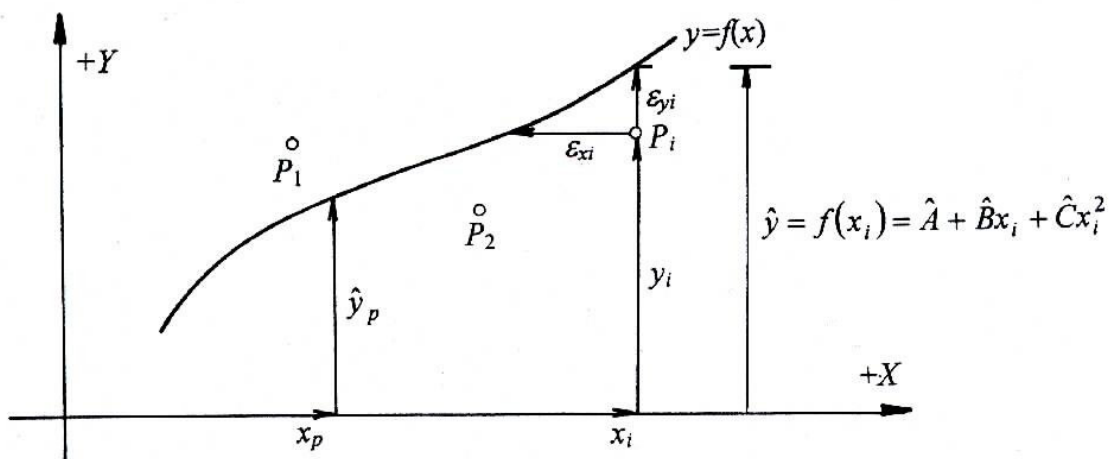
$$y = f(x), \quad z = \varphi(y, x). \quad (9.1)$$

Veličiny sú vzájomne štatisticky korelované (závislé). Pritom nepoznáme typ a konštanty funkcie, ktoré dodatočne určujeme na podklade empiricky zistených (odmeraných) údajov. Tento druh riešenia a problému nazývame **regresná analýza**. Tesnosť empirickej závislosti korelovaných veličín na štatisticky vyhodnotenom funkčnom vzťahu nazývame **korelačná analýza**.

V rovniciach (9.1) napr. namiesto presných hodnôt X, Y máme k dispozícii odmerané hodnoty x_i, y_i . Vyrovnávacia krivka $y = f(x)$ je spojitá a prechádza medzi bodmi empirického polygónu, ktorý je vytvorený odmeranými údajmi (obr. 9.1). Odstupy bodov P_i od krivky ε_i sú reziduá alebo regresné chyby.

Empirickým určením typu analytickej funkcie a jej číselných konštánt vyjadrujeme priebeh javu odmeraných hodnôt závislej premennej y pri meniacich sa hodnotách argumentu x . Grafické znázornenie priebehu javu, vplyvom meračských chýb alebo iných rušivých vplyvov, vyjadruje nepravidelný rad bodov (empirický polygón). Úlohou je nájsť takú funkčnú závislosť medzi premennými x, y , aby priebeh funkcie javu charakterizovaný **vyrovnávacou krivkou**, sa pri jednoduchom tvare funkcie optimálne primkol k empirickému polygónu. Zvyčajne máme k dispozícii nadbytočný počet meraní, vtedy koeficienty funkcie (9.1) určíme s vyrovnaním MNŠ.

Výsledkom bude tzv. **regresná krivka**. Aproximácia skutočného priebehu javu je nevyhnutná k interpolácii priebehu javu pre ľubovoľnú hodnotu argumentu. Používa sa často k číselnému vyjadreniu fyzikálnych vzťahov v geodézii a v iných vedných odboroch.



Obr. 9.1. Regresná krivka

Metódy regresnej a korelačnej analýzy, ako všetky metódy matematickej štatistiky, pri obmedzenom splnení podmienok budú mať aj obmedzenú platnosť záverov, v zásade iba na definovanom intervale $\in x_i, (i = 1, 2, \dots, n)$.

Metódy regresnej a korelačnej analýzy sú založené na výsledkoch meraní radu dvoch súčasne nezávislých premenných. Výsledkom bude rad hodnôt dvojíc, ktoré považujeme za meranie

dvojrozmernej veličiny. Pri korelácii budeme predpokladať, že obe premenné sú spojité náhodné veličiny. V regresnej analýze stačí iba predpoklad, že jedna z obidvoch premenných je spojitá náhodná veličina, u druhej nemusí byť tento predpoklad splnený.

Pri regresnej analýze považujeme zvyčajne hodnoty jednej premennej, napr. y , za meranie spojitej náhodnej veličiny pre dané hodnoty druhej premennej x . Pre každú z daných hodnôt x_i ($i = 1, 2, \dots, n$) bude mať náhodná veličina y určité rozdelenie so strednou hodnotou a varianciou, zodpovedajúcou príslušnej hodnote x_i .

Ak vyhodnocované kvantitatívne vzťahy riešime lineárnou funkciou o niekoľkých neznámych parametroch (regresných koeficientoch), riešený problém nazývame **lineárna regresia**. Nelineárna regresia vyžaduje špeciálne riešenie. Keď je počet analyzovaných prvkov (x, y) dva, riešený problém označujeme **jednopremenná regresná analýza**. Väčší počet analyzovaných prvkov ako dva označujeme **multipremenná regresná analýza**.

Existujú tri varianty riešenia:

1. Uvažujeme len regresné chyby ε_{yi} funkcie y a vyrovnanie vykonáme z podmienky $\varepsilon^T \mathbf{P} \varepsilon = \min$. Meračskými chybami sú zaťažené iba hodnoty y_i a hodnoty x_i považujeme za bezchybné

$$y_i + \varepsilon_{yi} = f_1(x_i), \quad \varepsilon_{xi} = 0, \quad \varepsilon_y^T \mathbf{P}_y \varepsilon_y = \min. \quad (9.2)$$

Je to najčastejší prípad v geodetických aplikáciách.

2. Uvažujeme len regresné chyby ε_{xi} argumentu x_i . Vyrovnanie vykonáme za podmienky

$$y_i = f_2(x_i + \varepsilon_{xi}), \quad \varepsilon_{yi} = 0, \quad \varepsilon_x^T \mathbf{P}_x \varepsilon_x = \min. \quad (9.3)$$

3. Uvažujeme regresné chyby ε_{yi} funkcie y ako aj regresné chyby argumentu x_i . Riešime podmienku

$$\varepsilon_y^T \mathbf{P}_y \varepsilon_y + \varepsilon_x^T \mathbf{P}_x \varepsilon_x = \min. \quad (9.4)$$

9.1 Lineárna regresia

Predpokladajme, že dve veličiny x a y sme odmerali n krát s údajmi: x_1, x_2, \dots, x_n a y_1, y_2, \dots, y_n . Úlohou je zistiť, či platí vzťah medzi veličinami vyjadrený rovnicou

$$y = A + Bx, \quad (9.5)$$

kde A, B sú neznáme (teoretické) regresné koeficienty. Ak rovnica (9.5) geometricky predstavuje priamku, nazývame ju **teoretická regresná priamka**. Vzťah medzi x a y môže byť ovplyvnený mnohými komplikovanými faktormi, okrem toho údaje ($x_i, y_i, i = 1, 2, \dots, n$) sú zaťažené meračskými chybami.

Rovnicu (8.5) upravíme pre odmerané údaje

$$y_i + \varepsilon_i = A + Bx_i, \quad (9.6)$$

pričom o zvyškových regresných chybách (reziduách) predpokladáme, že všetky chyby $\varepsilon_i (1 \leq i \leq n)$ sú navzájom nezávislé. Aby sme našli vyrovnané hodnoty regresných koeficientov \hat{A} , \hat{B} , ktoré by najlepšie vyhovovali vzťahu odmeraných údajov (x_i, y_i) stanovujeme podmienku

$$\sum_{i=1}^n \varepsilon_i = \sum_{i=1}^n (A + Bx_i - y_i)^2 = \min. \quad (9.7)$$

Regresná analýza formulovaná vzťahom (9.7) \hat{A} , \hat{B} neznámych regresných koeficientov, predstavuje aproximáciu s vyrovnaním MNS. Vyrovnané optimálne hodnoty regresných koeficientov A, B dostaneme deriváciou rovnice (9.7) a porovnáme s nulou

$$\frac{\partial}{\partial A} \left\{ \sum_{i=1}^n (A + Bx_i - y_i)^2 \right\}_{A=\hat{A}, B=\hat{B}} = \sum_{i=1}^n 2(\hat{A} + \hat{B}x_i - y_i)(+1) = 0,$$

$$\frac{\partial}{\partial B} \left\{ \sum_{i=1}^n (A + Bx_i - y_i)^2 \right\}_{A=\hat{A}, B=\hat{B}} = \sum_{i=1}^n 2(\hat{A} + \hat{B}x_i - y_i)(+x_i) = 0,$$

čo vedie k lineárnemu systému rovníc pre neznáme regresné koeficienty \hat{A} , \hat{B}

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} \hat{A} \\ \hat{B} \end{vmatrix} = \begin{vmatrix} b_1 \\ b_2 \end{vmatrix}, \text{ kde} \quad (9.8)$$

$$a_{11} = n,$$

$$a_{12} = a_{21} = \sum_{i=1}^n x_i,$$

$$a_{22} = \sum_{i=1}^n x_i^2, \quad (9.9)$$

$$b_1 = \sum_{i=1}^n y_i,$$

$$b_2 = \sum_{i=1}^n x_i y_i.$$

Rovnicu (9.8) môžeme maticovo zapísať aj v tvare

$$\mathbf{C}\mathbf{x} = \mathbf{D}. \quad (9.10)$$

Riešením rovnice (9.8), resp. (9.10) s inverziou matice \mathbf{C} pomocou determinatu, vyriešime regresné koeficienty \hat{A} , \hat{B}

$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{D} = \begin{vmatrix} \hat{A} \\ \hat{B} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}^{-1} \begin{vmatrix} b_1 \\ b_2 \end{vmatrix} = \frac{1}{a_{11}a_{22} - a_{12}^2} \begin{vmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{vmatrix} \begin{vmatrix} b_1 \\ b_2 \end{vmatrix} =$$

$$= \frac{1}{a_{11}a_{22} - a_{12}^2} \begin{vmatrix} a_{22}b_1 & -a_{12}b_2 \\ -a_{12}b_1 & a_{11}b_2 \end{vmatrix}. \quad (9.11)$$

Určením regresných koeficientov \hat{A} , \hat{B} vyrovnaním MNŠ, vypočítame vyrovnané reziduá $\hat{\varepsilon}_i$, ktoré jednotlivo vyjadrujú tesnosť empirického polygónu od regresnej priamky $y = \hat{A} + \hat{B}x$

$$\hat{\varepsilon}_i = \hat{A} + \hat{B}x_i - y_i. \quad (9.12)$$

Teoretické reziduá ε_i majú apriori rozdelenie $\varepsilon_i \sim N(0, \sigma^2)$. Výberovú (náhodnú) hodnotu σ^2 vypočítame z reziduí vyrovnaných MNŠ z rovnice

$$\hat{\sigma}_0^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{\varepsilon}_i^2 = \frac{1}{n-2} \sum_{i=1}^n (\hat{A} + \hat{B}x_i - y_i)^2. \quad (9.13)$$

Variancia $\hat{\sigma}_0^2$ je tiež mierou tesnosti všetkých bodov k vypočítanej regresnej priamke $y = \hat{A} + \hat{B}x$. Empirické stredné chyby vyrovnaných regresných koeficientov \hat{A} , \hat{B} vypočítame z rovníc:

$$\sigma_{\hat{A}} = \hat{\sigma}_0 \sqrt{\frac{a_{22}}{a_{11}a_{22} - a_{12}^2}}, \quad \sigma_{\hat{B}} = \hat{\sigma}_0 \sqrt{\frac{a_{11}}{a_{11}a_{22} - a_{12}^2}}. \quad (9.14)$$

Pre teoretické reziduá $\varepsilon_i \sim N(0, \sigma^2)$ aplikujeme t a χ^2 rozdelenie

$$\frac{\hat{A}}{\sigma_{\hat{A}}} \sim t(n-2), \quad \frac{\hat{B}}{\sigma_{\hat{B}}} \sim t(n-2), \quad (9.15)$$

$$(n-2) \frac{\hat{\sigma}_0^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{\hat{A} + \hat{B}x_i - y_i}{\sigma} \right)^2 \sim \chi^2(n-2). \quad (9.16)$$

Ak je známa variancia σ^2 , celková platnosť lineárnej regresie sa štatisticky testuje rovnicou (9.16).

Ak lineárnou regresiou určené koeficienty \hat{A} , \hat{B} boli štatisticky spoľahlivo určené (testom koeficienta korelácie r , kap. 9.4) môžu byť použité na predikciu hodnoty y_p pre hodnotu x_p veličiny x z intervalu $x_p \in (x_{imin}, x_{imax})$

$$\hat{y}_p = \hat{A} + \hat{B}x_p. \quad (9.17)$$

9.2 Nelineárna regresia

Vo všeobecnosti nelineárna regresia nie je taká jednoduchá ako lineárna regresia. Pri hodnotení empirického polygónu vzťahov medzi veličinami x_i a y_i často nachádzame nelineárnu závislosť. Niekedy nelineárny regresný model môžeme transformovať na lineárny model zavedením funkčných vzťahov medzi regresnými koeficientami tak, ako si to uvedieme v príkladoch.

Príklad 9.1: Predpokladajme, že dve série meraní x_i, y_i ($i = 1, 2, \dots, n$) vyhovujú nasledovnému nelineárnemu regresnému modelu

$$y_i = \frac{x_i}{Ax_i + B}, \quad (9.18)$$

kde A, B sú regresné koeficienty, ktoré je potrebné určiť. Rovnicu (9.18) v čitateli a menovateli vydelíme $\frac{1}{x_i}$, po úprave dostaneme

$$\frac{1}{y_i} = A + \frac{B}{x_i}.$$

Zavedieme nové veličiny

$$Y_i = \frac{1}{y_i}, \quad X_i = \frac{1}{x_i}. \quad (9.19)$$

Rovnica (9.18) sa zredukuje do lineárneho tvaru

$$Y_i = A + BX_i, \quad (9.20)$$

z ktorého vypočítame regresné koeficienty \hat{A} a \hat{B} .

Príklad 9.2: Iný nelineárny regresný model dvoch sérií meraní x_i, y_i ($i = 1, 2, \dots, n$) je

$$y_i = Ae^{-\frac{x_i^2}{B^2}}, \quad (9.21)$$

kde A a B sú regresné koeficienty, ktoré je potrebné určiť. Rovnicu (9.21) na oboch stranách linearizujeme prirodzenými logaritmami a dostaneme

$$\ln y_i = \ln A - \frac{1}{B^2} x_i^2. \quad (9.22)$$

Po substitúcii

$$X_i = x_i^2, \quad Y_i = \ln y_i, \quad a = \ln A, \quad b = \frac{1}{B^2}, \quad (9.23)$$

dostaneme lineárny regresný model dvoch nových regresných koeficientov a a b :

$$Y_i = a - bX_i. \quad (9.24)$$

Keď vyrovnaním MNŠ vypočítame regresné koeficienty a a b , originálne regresné koeficienty A a B určíme z rovníc (9.23)

$$\hat{A} = e^a, \quad -\hat{B} = \sqrt{\frac{1}{b}}. \quad (9.25)$$

Príklad 3: Z ďalších typov funkcií, ktoré sa často používajú v geodetických aplikáciách sú:

- exponenciálna funkcia

$$y = ae^{bx}, \quad (9.26)$$

lineárny tvar funkcie je

$$\ln y = \ln a + bx \Rightarrow Y = A + Bx, \text{ kde } \hat{A} = \ln a; \quad a = e^{\hat{A}}, \quad \hat{B} = b, \quad (9.27)$$

- logaritmická funkcia

$$y = a + b \ln x, \quad (9.28)$$

lineárny tvar funkcie je

$$Y = A + BX, \text{ kde } X = \ln x \quad \text{a} \quad x = e^X, \quad (9.29)$$

-mocninová funkcia

$$y = ax^b, \quad (9.30)$$

lineárny tvar funkcie je

$$\ln y = \ln a + b \ln x \Rightarrow Y = A + BX, \quad a = e^A, \quad b = B. \quad (9.31)$$

Príklad 4: Takmer univerzálnym typom regresnej funkcie na vyrovnanie v danom intervale je mocninový rad

$$y = A_0 + A_1x + A_2x^2 + A_3x^3 + \dots + A_nx^n. \quad (9.32)$$

Ak určujeme dva parametre (A_0, A_1) je to priamka, tri parametre (A_0, A_1, A_2) je kvadratická parabola, štyri parametre kubická parabola.

Optimálny odhad regresných koeficientov \hat{A}_0 a \hat{B}_i . ($i = 1, 2, \dots, n$) docielime pomocou vyrovnanie MNŠ, čo vyžaduje nadbytočný počet regresných vzťahov medzi veličinami x a y ako je počet regresných koeficientov.

Jednotkovú strednú chybu vypočítame zo vzťahu

$$m_0 = \sqrt{\frac{\mathbf{\epsilon}^T \mathbf{\epsilon}}{n - k}}, \quad (9.33)$$

kde k je počet regresných koeficientov.

Výpočtom regresných koeficientov vytvoríme interpolačnú funkciu $y = f(x)$.

Spoľahlivosť interpolácie odhadneme využitím jednotkovej strednej chyby m_0 . Vytvoríme pásmo spoľahlivosti interpolácie pri hladine významnosti α s kritickou hranicou $t_\alpha m_0$ od regresnej funkcie, keď t_α určíme zo Studentovho rozdelenia pre n' nadbytočných regresných vzťahov.

Pásmo interpolácie je v intervale $\langle x_{min}, x_{max} \rangle$. Extrapolácia nie je spoľahlivá.

Či použijeme lineárny alebo nelineárny regresný model závisí iba od charakteru vzťahu medzi veličinami. Predbežné rozhodnutie je možné urobiť pri vykreslení bodov $P_i(x_i, y_i)$ pre $i = 1, 2, \dots, n$ v rovinnom súradnicovom systéme súradníc a vizuálnom porovnaní regresného polygónu niekoľkými známymi funkciami (priamka, parabola, mocnonový rad, trigonometrická funkcia a pod.). Najvhodnejší tvar regresnej funkcie vyplynie po vyhodnotení korelačného koeficienta r .

9.3 Priestorová regresná analýza

Pri rôznych technických úlohách sa využíva priestorová regresná analýza. Príkladmi sú: minimalizácia presunu zeminy, určenie regresnej roviny na odvodnenie náklonu vysokých stavebných objektov, určenie priestorovej polohy diskontinuit, výpočet výšok na digitálnom modeli reliéfu a iné.

Tvary priestorových regresných funkcií sú napr.:

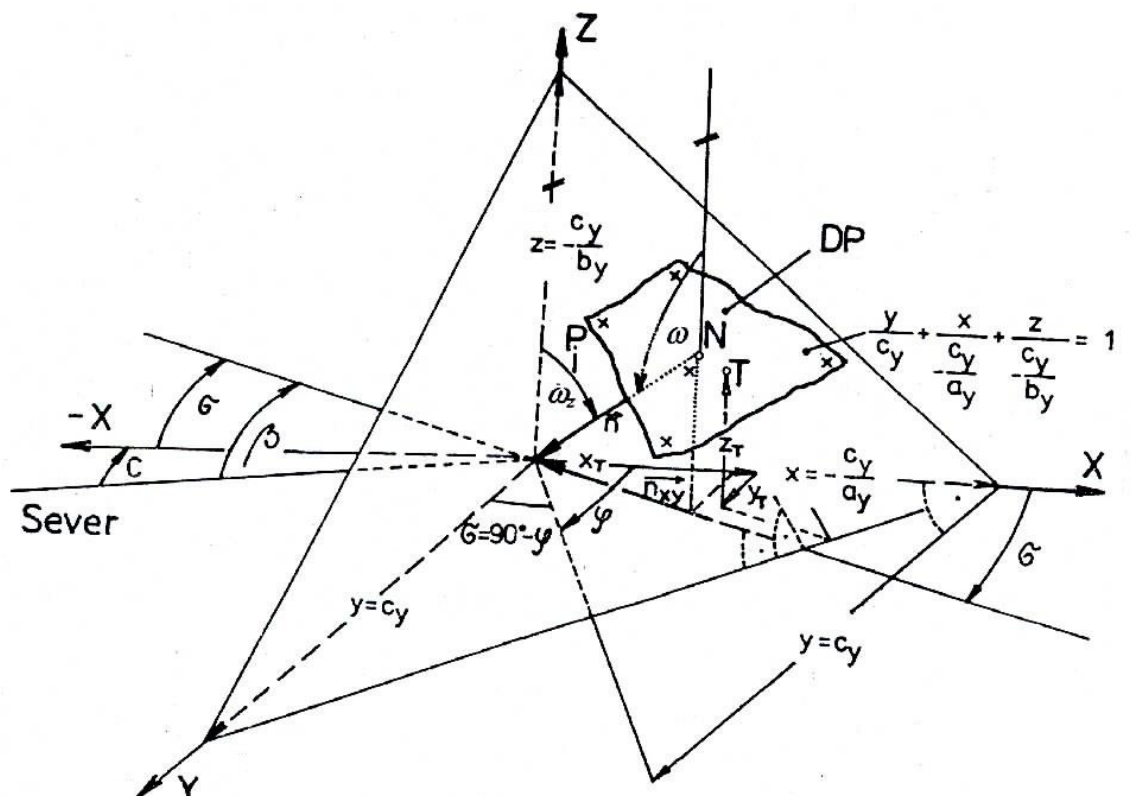
- rovina: $z = a_0 + a_1x + a_2y$, (9.34)

- plocha druhého stupňa priamková (hyperbolický paraboloid):

$$z = a_0 + a_1yx + a_2x + a_3y, \quad (9.35)$$

- plocha tretieho stupňa v tvare polynómu:

$$z = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2 + a_6x^2y + a_7xy^2 + a_8x^3 + a_9y^3. \quad (9.36)$$



Obr. 9.2 . Priestorová poloha roviny (diskontinuity)

Príkladom využitia priestorovej regresie v geotechnike je určenie priestorovej polohy diskontinuit (puklinových plôch) z výsledkov fotogrametrického vyhodnotenia diskontinuity s počtom charakteristických bodov $n > 3$. Cieľom riešenia je určiť vyrovnávaciu rovinu v smere niektorej z priestorových osí XYZ a z regresných koeficientov roviny vypočítať sklon normály ω a smerník priemetu σ rádius vektora $\overrightarrow{n_{xy}}$ (obr. 9.2).

Všeobecný tvar roviny, ktorý vyjadrujú reziduá v smere osi Y je

$$\varepsilon_{yi} = a_y x_i + b_y z_i + c_i - y_i \quad (9.37)$$

v maticovom zápise

$$\varepsilon_{y(n,1)} = \mathbf{A}_{(n,3)} \mathbf{y}_{(3,1)} + \ell_{(n,1)}, \quad (9.38)$$

kde je

\mathbf{A} – matica súradníc charakteristických bodov diskontinuity s členmi v stĺpcoch $x_i, z_i, 1$,

\mathbf{y} – stĺpcový vektor koeficientov a_y, b_y, c_y ,

ℓ - stĺpcový vektor s členmi $-y_i$, ($i = 1, 2, \dots, n$).

Podmienka $\varepsilon_y^T \varepsilon_y$, bude splnená, ak $\frac{\partial \varepsilon^T \varepsilon}{\partial \mathbf{y}} = 0$.

Stĺpcový vektor \mathbf{y} určíme z rovnice

$$\mathbf{y}_{(3,1)} = -(\mathbf{A}_{(3,n)}^T \mathbf{A}_{(n,3)})^{-1} \mathbf{A}_{(3,n)}^T \ell_{(n,1)}. \quad (9.39)$$

Súčet štvorcov reziduí $\varepsilon_y^T \varepsilon_y$ vypočítame z rovnice:

$$\varepsilon_y^T \varepsilon_{y(1,1)} = \ell_{(1,n)}^T \mathbf{A}_{(n,3)} \mathbf{y}_{(3,1)} + \ell_{(1,n)}^T \ell_{(n,1)}. \quad (9.40)$$

Presnosť aproximácie diskontinuity vyrovnávacou regresnou rovinou charakterizuje jednotková stredná chyba

$$m_{0y} = \sqrt{\frac{\varepsilon^T \varepsilon}{n-3}}. \quad (9.41)$$

Podobne sa určia regresné koeficienty regresných rovín, ktoré aproximujú diskontinuitu v smere osi X a Z . Sklon ω a smerník priemetu rádius vektora σ vypočítame z tých koeficientov regresnej roviny, u ktorej jednotková stredná chyba m_{0j} ($j = y, x, z$) mala minimálnu hodnotu.

Napr. podľa obr. 9.2 smerník σ priemetu rádius vektora $\overrightarrow{n_{xy}}$ vypočítame z koeficientov regresnej roviny upravenej do úsekového tvaru

$$\frac{y}{c_y} + \frac{x}{-\frac{c_y}{a_y}} + \frac{z}{\frac{c_y}{b_y}} = 1, \quad (9.42)$$

$$\text{z rovnice } \sigma = 90^\circ - \varphi = \operatorname{arctg} \frac{\frac{c_y}{b_y}}{-\frac{c_y}{a_y}}. \quad (9.43)$$

Sklon ω rádius vektora \vec{n}_{xy} diskontinuity vypočítame z rovnice

$$\omega = 180^\circ - \omega_z = 180^\circ - \arccos \frac{b_y}{\sqrt{a_y^2 + 1 + b_y^2}}. \quad (9.44)$$

Znamienko odmocniny v menovateli je vždy opačné ako u regresného koeficienta c_y vyrovnávacej roviny.

Pri aproximácii diskontinuity (roviny) v smere osi X a Z vypočítame uhol φ a ω_z rovníc:

- v smere osi X

$$\varphi = \operatorname{arctg} \frac{-\frac{c_x}{a_x}}{c_x}, \quad \omega_z = \arccos \frac{b_x}{\sqrt{1 + a_x^2 + b_x^2}}, \quad (9.45)$$

- v smere osi Z ,

$$\varphi = \operatorname{arctg} \frac{-\frac{c_z}{b_z}}{c_z}, \quad \omega_z = \arccos \frac{-1}{\sqrt{a_z^2 + b_z^2 + 1}}. \quad (9.46)$$

Uvedený postup výpočtu sklonu regresnej roviny a smerníka rádius vektora môžeme aplikovať aj na určenie náklonov vysokých stavebných objektov z výsledkov nivelačných meraní. Vtedy použijeme namiesto súradnice v smere osi Z rozdiely výšok pozorovaných bodov medzi dvoma etapovými meraniami.

9.4 Aproximácia bodového radu funkciou trigonometrickej rady (harmonická analýza)

V prírode a technických zariadeniach prebiehajú niektoré javy tak, že s určitým argumentom vplyvu ako je napr. čas, teplota, uhol, vlnová dĺžka, atď. plynulo narastá veľkosť meraného argumentu. Po dosiahnutí maxima bodový rad klesá na minimum a opäť sa vracia k pôvodnej hodnote. Po prvej perióde P sa priebeh javu opakuje v nasledujúcich periódach

$$y = f(x) = f(x + P) = f(x + 2P) = \dots \quad (9.47)$$

Preto stačí vyšetriť priebeh javu v jednej perióde. Jej rozsah $v_n - v_0 = P$ upravíme substitúciou

$$t = \frac{2\pi}{P} x. \quad (9.48)$$

Napr. na rozsah 2π u ročného priebehu strednej hodnoty teploty bude doba jedného mesiaca rovná 30° periódy.

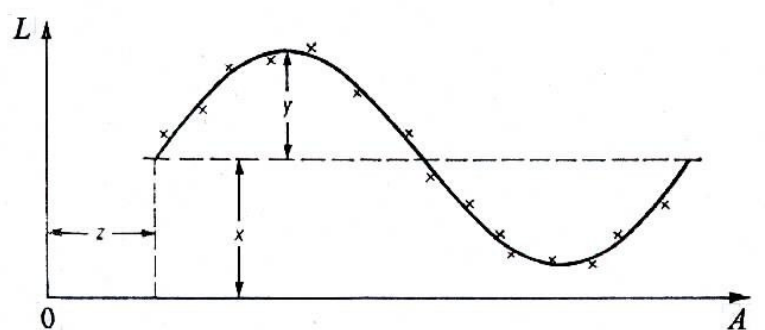
Jednoduchý jav, ktorý plynie z jednej príčiny (napr. chybu z excentricity alidády), je možné vyjadriť krivkou sinusoidy (obr. 9.3)

$$L_i = x + y \sin(A_i - z) \quad (9.49)$$

Máme odmeraných n hodnôt ℓ_i pri n hodnotách spojitkej premennej argumentu A_i . x , y a z sú hľadané tri konštanty vyrovňavacej funkcie, kde x je poradnica sinusoidy, y je amplitúda, z je posun počiatku sinusoidy oproti počiatku argumentu A_i .

Ak počet odmeraných hodnôt ℓ_i $i > 3$ aplikujeme vyrovnanie bodového radu s vyrovnaním MNŠ. Zostavíme rovnicu opráv

$$L_i = \ell_i + v_i = x + y(\sin A_i - z) \quad (9.50)$$



Obr. 9.3 Vyrovnanie bodového radu sinusoidou

Funkcia (9.50) je príliš zložitá na to, aby sme pomocou troch vhodne rozložených bodov na bodovom rade, určili približné hodnoty konštánt. Vhodné je postupovať tak, že si graficky znázorníme priebeh bodového radu. Hodnoty x_0 , y_0 a z_0 odčítame z grafu. Metódou vyrovňovania sprostredkujúcich meraní určíme opravy dx , dy a dz ,

$$x = x_0 + dx, \quad y = y_0 + dy, \quad z = z_0 + dz. \quad (9.51)$$

Funkciu (9.50) rozvineme do Taylorovho radu s členmi rozvoja

$$\ell_0 = x_0 + y_0 \sin(A_i - z_0), \quad \frac{\partial L_i}{\partial x} = 1, \quad \frac{\partial L_i}{\partial y} = a_i = \sin(A_i - z_0),$$

$$\frac{\partial L_i}{\partial z} = b_i = -y_0 \cos(A_i - z_0), \quad -\ell'_i = \ell_0 - \ell_i. \quad (9.52)$$

Pretvorené rovnice opráv budú mať tvar

$$v_i = dx + \sin(A_i - z_0)dy - y_0 \cos(A_i - z_0)dz + \ell_0 - \ell_i, \quad (9.53)$$

$$v_i = dx + a_i dy + b_i dz - \ell'_i. \quad (9.54)$$

Opravy dx , dy a dz k približným hodnotám x_0 , y_0 a z_0 určíme známym postupom vyrovnaní MNŠ.

Vypočítanými hodnotami (9.51) spresníme odčítané hodnoty z grafu a zopakujeme vyrovnanie. Vyžadovanú presnosť určenia konštánt môžeme limitovať porovnaním dvoch po sebe vykonaných výpočtov. Spravidla nám stačí jedno opakované vyrovnanie.

Jednotkovú strednú chybu m_0 a neznámych konštánt x , y a z vypočítame z rovníc

$$m_0 = \sqrt{\frac{v^T v}{n-3}}, \quad m_x = m_0 \sqrt{Q_{xx}}, \quad m_y = m_0 \sqrt{Q_{yy}}, \quad m_z = m_0 \sqrt{Q_{zz}}. \quad (9.55)$$

9.5 Analýza korelácie

Majme rad meraní dvojíc premenných x_i , y_i . Výsledky meraní môžu ukázať, že jednej hodnote veličiny x_i bude zodpovedať viac hodnôt y_{ij} ($j = 1, 2, \dots, n$) a naopak jednej hodnote y_i bude odpovedať viac hodnôt x_{ik} ($k = 1, 2, \dots, m$). S meniacou hodnotou sa mení stredná hodnota druhej premennej. Takúto závislosť medzi dvoma premennými označujeme pojmom **korelačná závislosť** a taký nefunkčný vzťah má názov stochastický (náhodný) alebo štatistický vzťah dvoch veličín. Ak vynesieme graficky odmerané hodnoty x_i , y_i nedostaneme bodový rad, ale plošný útvar – **korelačné pole**.

Príčinou vzniku korelačného poľa je existencia pôsobenia náhodných faktorov na premennú y a na argument x . Úlohou korelačného počtu je určiť vzájomný vzťah medzi premennou y a argumentom x , ktorý vyjadrujeme **koeficientom korelácie**.

V korelačnom poli dostaneme dve vyrovnávacie priamky podľa definície vyrovnaní MNŠ

$$(\mathbf{e}^T \mathbf{P} \mathbf{e})_y = \min. \quad \text{a} \quad (\mathbf{e}^T \mathbf{P} \mathbf{e})_x = \min. \quad (9.56)$$

Funkčné rovnice a rovnice reziduí budú pre prvú funkciu (9.56).

$$y_i + \varepsilon_{yi} = A_y + B_y x_i$$

$$\varepsilon_{yi} = A_y + B_y x_i - y_i \quad (9.57)$$

$$\varepsilon_{xi} = 0.$$

Normálne rovnice s použitím váhových koeficientov p majú tvar (rovnice 9.8 a 9.9)

$$\sum p \hat{A}_y + \sum p x \hat{B}_y - \sum p y = 0, \quad (9.58)$$

$$\sum p x \hat{A}_y + \sum p x x \hat{B}_y - \sum p x y = 0.$$

Z rovníc vypočítame regresné koeficienty \hat{A}_y , \hat{B}_y (napr. eliminačnou metódou, pomocou determinantu (9.11), alebo maticovým riešením) a stredné chyby vyrovnaných regresných