Vaja 3 - 5: **Velika mreža Strunjan - iskanje grobih pogreškov**

V Veliki mreži Strunjan boste dobili načrtno pokvarjeno:

- eno dolžino in
- en kot

za približno 10-kratno vrednost standardnega odklona opazovanja. Predpostavite, da je kovariančna matrika opazovanj znana. S postopki iskanja grobih pogreškov poiščite morebitne prisotne grobe pogreške in ocenite njihovo velikost. Izračunajte:

- popravke opazovanj,
- referenčno varianco a-posteriori $\hat{\sigma}_0^2$,
- izvedite globalni test modela, ob stopnji značilnosti $\alpha = 0.05$,
- ob stopnji značilnosti $\alpha = 0.01$ izvedi postopek pregleda opazovanj ("data snooping") posameznih popravkov opazovanj,
- ob stopnji značilnosti $\alpha = 0.01$ izvedi τ-test posameznih popravkov opazovanj (predpostavite, da natančnosti opazovanj ne poznate),
- določite število nadštevilnosti r_i za vsako opazovanje,
- ocenite velikost odkritih grobih pogreškov opazovanj.

Iskanje grobih pogreškov

Iskanje grobih pogreškov je sestavljeno iz treh korakov, in sicer:

- izvedbe globalnega testa,
- izvedba iskanja grobih pogreškov (pregled opazovanj ali τ-test) in
- ocena velikosti grobih pogreškov.

Globalni test

Izvedbo globalnega testa smo detaljno prikazali pri predmetih *Izravnalni račun 2* in *Analiza opazovanj v geodeziji 2*, zato bo tu postopek testa prikazan v manjšem obsegu. Pri globalnem testu ugotavljamo skladnost referenčne variance a-posteriori $\hat{\sigma}_0^2$ z referenčno varianco a-priori σ_0^2. Nastavimo ničelno (H_0) in alternativno (H_a) hipotezo statističnega testa:

$$H_0 : \hat{\sigma}_0^2 = \sigma_0^2 \qquad \text{oz.} \qquad H_a : \hat{\sigma}_0^2 \neq \sigma_0^2 \tag{1}$$

Testna statistika, ki jo bomo sestavili ima obliko:

$$Z = r\frac{\hat{\sigma}_0^2}{\sigma_0^2} \sim \chi_r^2 \tag{2}$$

kar lahko preoblikujemo v:

$$\frac{Z}{r} = \frac{\hat{\sigma}_0^2}{\sigma_0^2} \sim \frac{\chi_r^2}{r} \tag{3}$$

Ničelne hipoteze ne moremo zavrniti, ko velja:

$$\frac{\chi^2_{\alpha/2,r}}{r} < \frac{\hat{\sigma}_0^2}{\sigma_0^2} < \frac{\chi^2_{1-\alpha/2,r}}{r} \tag{4}$$

Kjer je α stopnja tveganja (značilnosti) testa. V tem primeru tako lahko rečemo, da sta si referenčni varianci statistično enaki. V splošnem, predvsem ko je število opazovanj večje in je veliko nadštevilnih opazovanj, lahko meji grobo nastavimo kot:

$$0.6 < \frac{\hat{\sigma}_0^2}{\sigma_0^2} < 1.6 \tag{5}$$

Pregled opazovanj ("angl. data snooping")

Postopku iskanja grobih pogreškov pregledu opazovanj rečemo tudi "data snooping" ali tudi Baardova metoda, po geodetu, ki je postopek prvi definiral. V tem primeru izhajamo iz predpostavke, da so natančnosti opazovanj znane, torej da poznamo variančno-kovariančno matriko opazovanj $\mathbf{\Sigma}$ in referenčno varianco a-priori σ_0^2. Za vsak izračunan popravek opazovanj v_i ($i = \{1, 2, \ldots, n\}$) sestavimo testno statistiko ω_i, ki ima obliko:

$$\omega_i = \frac{v_i}{\sigma_{v_i}} = \frac{v_i}{\sigma_0\sqrt{q_{v_iv_i}}} \sim N(0,1) \tag{6}$$

Ključno pri enačbi 6 je, da natančnost popravka σ_{v_i} izračunamo na osnovi referenčne variance a-priori σ_0^2. Za testno statistiko ω_i, ki ji pravimo tudi "standardizirani popravek", velja, da se porazdeljuje po standardni normalni porazdelitvi. Sestavimo ničelno in alternativno hipotezo:

$$H_0 : \omega_i = 0 \qquad \text{oz.} \qquad H_a : \omega_i \neq 0 \tag{7}$$

Ničelne hipoteze iz enačbe 7 ne moremo zavrniti, ko velja:

$$|\omega_i| < N_{\alpha/2} \tag{8}$$

V enačbi 8 $N_{\alpha/2}$ predstavlja kritično mejo, ki jo dobimo iz porazdelitve, pri stopnji značilnosti $\alpha/2$. Iz testa tudi vidimo, da je test dvostranski, saj so popravki lahko tako negativni, kot tudi pozitivni. V praksi se je pojavila dokaj dobra meja za iskanje grobih pogreškov, in sicer, popravek vsebuje grobi pogrešek, če velja:

$$|\omega_i| > 3 \tag{9}$$

Za grobo pogrešena opazovanja obravnavamo vsa tista opazovanja, katerih popravek opazovanj je po absolutni vrednosti vsaj 3-krat večji od pripadajoče natančnosti.

τ-test

V primeru, ko natančnosti opazovanj ne poznamo, ne poznamo ne kovariančne matrike $\mathbf{\Sigma}$, ne referenčne variance a-priori σ_0^2. V tem primeru seveda globalnega testa ne moremo izvesti, pa tudi pregleda opazovanj ne. Tu uporabimo τ-test. Za vsak izračunan popravek opazovanj v_i ($i = \{1, 2, \ldots, n\}$) sestavimo testno statistiko T_i, ki ima obliko:

$$T_i = \frac{v_i}{\sigma_{v_i}} = \frac{v_i}{\hat{\sigma}_0 \sqrt{q_{v_i v_i}}} \sim \tau_r \tag{10}$$

Standardnizirani popravek T_i v enačbi 10 se izračuna zelo podobno, kot v enačbi 6 z eno **zelo pomembno razliko**, natančnost popravka se izračuna na osnovi referenčne variance a-posteriori $\hat{\sigma}_0^2$. Ker je izračun natančnosti popravka odvisen od popravkov samih (izračun $\hat{\sigma}_0^2$), se testna statistika T_i porazdeljuje po τ porazdelitvi z r prostostnimi stopnjami. Sestavimo ničelno in alternativno hipotezo:

$$H_0 : T_i = 0 \qquad \text{oz.} \qquad H_a : T_i \neq 0 \tag{11}$$

Ničelne hipoteze iz enačbe 11 ne moremo zavrniti, ko velja:

$$|T_i| < \tau_{r,\alpha/2} \tag{12}$$

Ocena velikosti grobega pogreška

V primeru, da grobi pogrešek s postopkom pregleda opazovanja ali τ-testa odkrijemo, lahko izračunamo njegovo velikost ∇v_i kot:

$$\nabla v_i = -\frac{v_i}{r_i} \tag{13}$$

Velikost grobega pogreška iz enačbe 13 je odvisna od velikosti popravka, a tudi od nadštevilnosti r_i opazovanja, ki mu ta popravek pripada. Na iskanje grobih potreškov tako vpliva natančnost opazovanj in geometrija geodetske mreže.