UDK 528.14
Pregledni rad

METODA NAJMANJIH KVADRATA U NAŠE VRIJEME

Miljenko LAPAINE — Zagreb*

1. UVOD

Najprikladnijim izborom vrijednosti nepoznatih veličina na temelju rezultata opažanja bavili su se krajem 18. i početkom 19. stoljeća mnogi naučnici. Među najznačajnija imena ubrajaju se Simpson, Lagrange, Bernoulli, Euler, Gauß, Legendre i Laplace. To je doba početaka i procvata metode najmanjih kvadrata, jedne od metoda koja iz rezultata opažanja daje »najbolje« procjene za nepoznate parametre.

»Man hat nach dieser Zeit die M. d. kl. Qu. rasch in die Wissenschaften eingefürt, wo es darauf ankommt, aus Beobachtungsresultaten möglichst zuverlässige Endwerthe zu ermitteln. Andrerseits ist es das Bestreben zahlreicher Mathematiker gewesen, der Grundlage der M. d. kl. Qu. zu beleuchten und es giebt in der That eine grössere Anzahl Ableitungen derselben aus den verschiedensten Hypothesen.« (Nakon ovog vremena, metoda najmanjih kvadrata je brzo uvedena u one znanosti, kod kojih se radi o tome da se iz rezultata opažanja odrede najpouzdanije moguće definitivne vrijednosti. S druge strane, težnja mnogobrojnih matematičara postala je rasvjetljavanje osnova metode najmanjih kvadrata i u stvari postoji veći broj njenih izvoda iz najrazličitijih hipoteza.) (Helmert 1872)

Zanimljiv je podatak da je Mansfield Merriman u radu »List of Writings relating to the Method of Least Squares«, Connecticut Transact., IV, još 1877. god. naveo 408 naslova (Czuber 1891).

Iako je metoda najmanjih kvadrata stvorena na temelju pretpostavke o normalnoj raspodjeli pogrešaka mjerenja, još je Gauß ustanovio i pokazao da to nije jedina alternativa i da mnogo općenitije pretpostavke vode k istim procjenama.

U našoj zemlji se proučavanju metode najmanjih kvadrata prilazilo redovito s aspekta teorije pogrešaka i normalnosti njihove raspodjele (Horvat 1937), (Čubranić 1948), (Čubranić 1958), (Muminagić, Jovanović 1965), Vranić 1971), (Macarol 1978), (Čubranić 1980), (Mihailović 1980), (Klak 1982), (Pašalić 1984). Taj bi pristup u najbližoj budućnosti trebalo modificirati (imajući u vidu knjige (Hadživuković i dr. 1981) i (Perović 1986) proizašle iz predavanja na postdiplomskim studijama), kako bi s jedne strane naši studenti i diplomirani inženjeri geodezije mogli lakše koristiti suvremenu

* Miljenko Lapaine, dipl. inž., Geodetski fakultet Sveučilišta u Zagrebu, Kačićeva 26.

literaturu, i kako s druge strane, ne bi baš sve rezultate dobivene primjenom metode najmanjih kvadrata nazivali najvjerojatnijima. Tome bi trebao dati svoj doprinos ovaj rad.

Pretpostavka za dalje čitanje ovog rada je poznavanje osnovnih pojmova linearne algebre. Najvažnije činjenice o metodi najmanjih kvadrata istaknute su u teoremima 1—7. Korištena je prednost matričnog zapisivanja koje omogućuje bolju preglednost, a posebna pažnja posvećena je što jednostavnijem i kraćem dokazivanju teorema. Tako se npr. važno svojstvo 9° iz teorema 2 dokazuje u ovom radu u nekoliko redaka, dok se izvođenje analognog svojstva nekada protezalo na nekoliko stranica. Spomenimo još da se potpuni dokaz posljednjeg teorema 7 u literaturi redovito izbjegava. Smatram da bi geodeti koji svakodnevno koriste metodu najmanjih kvadrata morali biti upoznati s potpunim dokazom (jedan mogući način dokazivanja izveden je u ovom radu) kako bi do kraja doživjeli odnos metode najmanjih kvadrata i pretpostavke o normalnoj distribuciji pogrešaka opažanja.

2. GEOMETRIJA METODE NAJMANJIH KVADRATA

Neka je R_n realni vektorski prostor jednostupčastih matrica koje ćemo zvati vektori. Za bilo koja dva vektora $x, y \in R_n$ može se definirati njihov *skalarni produkt* kao realan broj

$$x^T y \in R. \tag{1}$$

Za prostor R_n u kojem je ovako uveden skalarni produkt kažemo da je *euklidski vektorski prostor*. Realni broj

$$\|x\| = \sqrt{x^T x} \tag{2}$$

zove se norma ili duljina vektora x.

Neka su $a_1, a_2, \ldots, a_m$ $(m \leq n)$ linearno nezavisni vektori iz R_n. Sastavimo od ovih vektora matricu A tipa (n, m)

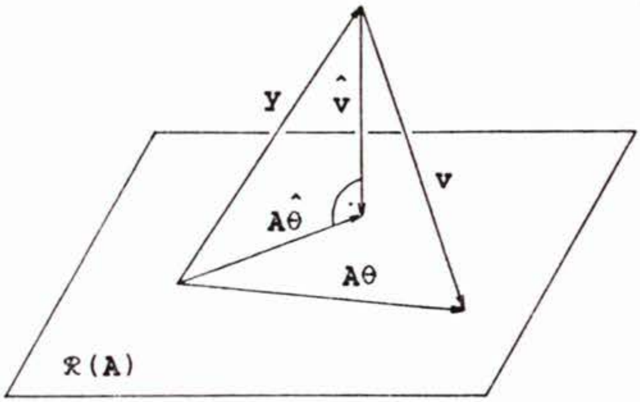
$$A = [a_1\ a_2 \ldots a_m]. \tag{3}$$

Skup svih linearnih kombinacija vektora $a_1, a_2, \ldots a_m$ je potprostor R (A) prostora R_n koji se zove *prostor stupaca matrice* A. Kako je rang(A) = m, to je dimenzija potprostora R (A) jednaka m. Ovaj potprostor

$$R(A) = \{z \in R_n \mid z = A\Theta,\ \Theta \in R_m\} \tag{4}$$

prikazan je na slici 1 kao ravnina. Neka je $y \in R_n$ bilo koji vektor. Ako je $y \in R(A)$, onda postoji $\Theta \in R_m$, tako da je $y = A\Theta$. Ako $y \notin R(A)$, tada ne postoji $\Theta \in R_m$, tako da je $y = A\Theta$. Međutim očito da za svaki $y \in R_n$ i za svaki $\Theta \in R_m$, postoji $v \in R_n$, tako da bude

$$y + v = A\Theta. \tag{5}$$



Slika 1. Geometrijski pristup metodi najmanjih kvadrata

Teorem 1. Neka je A matrica tipa (n,m), rang(A) = m. Za bilo koji vektor $y \in R_n$ jednoznačno je određen vektor $\hat{\Theta} \in R_m$

$$\hat{\Theta} = (A^T A)^{-1} A^T y, \tag{6}$$

koji je rješenje sistema linearnih jednadžbi

$$A^T A\hat{\Theta} - A^T y = 0, \tag{7}$$

tako da vrijedi

$$\forall \Theta \in R_m, \qquad \|A\Theta - y\|^2 \geq \|A\hat{\Theta} - y\|^2, \tag{8}$$

s time da umjesto znaka $\geq$ u (8) vrijedi = samo za

$$\hat{\Theta} = \Theta. \tag{9}$$

Dokaz: Koristit ćemo poznato svojstvo

$$\begin{aligned} \operatorname{rang}(A) = m \Leftrightarrow A^T A \quad & \text{je pozitivno definitna matrica} \\ \Leftrightarrow (A^T A)^{-1} \quad & \text{je pozitivno definitna matrica.} \end{aligned} \tag{10}$$

Lako se može vidjeti da je

$$\begin{aligned} \|A\Theta - y\|^2 &= (A\Theta - y)^T (A\Theta - y) = \\ &= (A^T A\Theta - A^T y)^T (A^T A)^{-1} (A^T A\Theta - A^T y) + \\ &+ y^T (I - A(A^T A)^{-1} A^T) y \geq y^T (I - A(A^T A)^{-1} A^T) y = \\ &= \|A\hat{\Theta} - y\|^2, \end{aligned} \tag{11}$$

gdje smo s I označili jediničnu matricu tipa (n,n) te koristili svojstvo (10). Odmah se vidi da u (11) umjesto znaka $\geq$ može pisati $=$ ako i samo ako je Θ rješenje sistema linearnih jednadžbi

$$A^T A\Theta - A^T y = 0, \tag{12}$$

tj., ako je

$$\Theta = \hat{\Theta} = (A^T A)^{-1} A^T y. \tag{9}$$

Time je dokaz teorema 1 završen. Označimo li

$$\hat{v} = A\hat{\Theta} - y, \tag{13}$$

možemo napisati (7) u ekvivalentnom obliku

$$A^T \hat{v} = 0, \tag{14}$$

kojim se izražava svojstvo okomitosti vektora $\hat{v}$ na sve vektore stupce matrice A, a time i na cijeli potprostor R(A).

Nadalje, koristeći (6) možemo napisati

$$\hat{v} = (A(A^T A)^{-1} A^T - I)\, y, \tag{15}$$

te prikazati kvadrat norme vektora $\hat{v}$ na više načina

$$\begin{aligned}\|\hat{v}\|^2 &= \hat{v}^T \hat{v} = (A\hat{\Theta} - y)^T (A\hat{\Theta} - y) = y^T y - \hat{\Theta}^T A^T A\hat{\Theta} = \\ &= y^T y - y^T A (A^T A)^{-1} A^T y = y^T (I - A(A^T A)^{-1} A^T)\, y.\end{aligned} \tag{16}$$

Prema (16) vidimo da vrijedi »Pitagorin poučak« za pravokutni trokut sa slike 1.:

$$\|y\|^2 = \|\hat{v}\|^2 + \|A\hat{\Theta}\|^2. \tag{17}$$

Zadana matrica A obično se zove *konfiguracijska matrica, dizajn matrica* ili *matrica regresije*. Elementi vektora y su (npr. u geodeziji) najčešće vrijednosti dobivene mjerenjem pa se vektor

$$\hat{y} = y + \hat{v} = A\hat{\Theta} = A(A^T A)^{-1} A^T y \tag{18}$$

zove *vektor popravljenih* ili *prilagođenih mjerenja*, a vektor $\hat{v}$ se zove *vektor procijenjenih pogrešaka* mjerenja, dok je geometrijski vektor $\hat{y}$ *ortogonalna projekcija* vektora y na potprostor R(A). Vektor $\hat{\Theta}$ zove se *vektor procije-*

njenih parametara. Izraz (13) predstavlja *jednadžbe pogrešaka*, a sistem linearnih jednadžbi (7) zove se *sistem normalnih jednadžbi.*

Opisani postupak ortogonalnog projiciranja zadanog vektora $y \in R_n$ na potprostor $R(A)$ zove se *metoda najmanjih kvadrata.* U radu (Kolmogorov 1946) po prvi puta je korištena geometrija n-dimenzionalnih vektorskih prostora radi zornijeg i jednostavnijeg načina izlaganja metode najmanjih kvadrata. U našoj literaturi tome je posvećen rad (Junašević 1970).

3. STOHASTIKA METODE NAJMANJIH KVADRATA

Neka je A matrica tipa (n,m), rang(A) = m. Za svaki vektor $y \in R_n$ i za svaki $\Theta \in R_m$, postoji $v \in R_n$, tako da bude

$$y + v = A\Theta. \tag{5}$$

Uzmimo jedan od parova Θ, v za koje vrijedi (5) i nazovimo ih *vektor pravih parametara* Θ i *vektor pravih pogrešaka* v. Primjenom metode najmanjih kvadrata iz prethodnog poglavlja, dobijemo vektor procijenjenih parametara $\hat{\Theta}$, vektor procijenjenih pogrešaka $\hat{v}$ i vektor prilagođenih ili popravljenih mjerenja $\hat{y}$.

Uređena n-torka $\mathbf{z} = [z_1\ z_2\ \dots\ z_n]^T$ slučajnih varijabli z_k nad istim vjerojatnosnim prostorom zove se *slučajni vektor.* Vektor $E(z) = [E(z_1) \dots E(z_n)]^T$ čije komponente $E(z_k)$ su očekivanja slučajnih varijabli z_k zove se *očekivanje slučajnog vektora* **z**. Sasvim analogno definiraju se slučajna matrica i njeno očekivanje.

Teorem 2. Pretpostavimo da je vektor pravih pogrešaka **v** slučajni vektor s očekivanjem

$$E(v) = 0. \tag{19}$$

Tada vrijede slijedeća svojstva:

1° $E(y) = A\Theta$

2° $E(\hat{v}) = 0$

3° $E(\hat{y}) = E(y)$

4° $E(\hat{\Theta}) = \Theta.$

Svojstva 1°—4° lako se dokazuju na osnovu linearnosti očekivanja, tj. poznatih svojstava

$$E(Cz) = CE(z) \tag{20}$$

$$E(z + u) = E(z) + E(u) \tag{21}$$

gdje su z i u bilo koja dva slučajna vektora nad istim vjerojatnosnim prostorom, a C zadana matrica.

Kažemo da je vektor $\hat{\mathbf{z}}$ *centrirana procjena slučajnog vektora* **z** (ili da je $\hat{\mathbf{z}}$ *nepristrasna procjena, nepomjerena procjena, procjena u skladu s očekivanjem* vektora **z**, ako je

$$E(\hat{z}) = z. \tag{22}$$

Dakle, prema svojstvu 4°, vektor procijenjenih parametara $\hat{\Theta}$ je centrirana procjena pravih parametara Θ.

Kovarijaciona matrica D(z) slučajnog vektora z definira se prema

$$D(z) = E[(z - E(z))(z - E(z))^T]. \tag{23}$$

Ako je z slučajni vektor s kovarijacionom matricom D(z), C zadana matrica i

$$u = Cz, \tag{24}$$

onda se na osnovu definicije (23) kovarijacione matrice lako pokaže da vrijedi

$$D(u) = CD(z)C^T. \tag{25}$$

Ovo svojstvo se zove *zakon o rasprostiranju kovarijanci* ili u geodeziji *zakon o rasprostiranju grešaka.*

Teorem 3. Pretpostavimo da je vektor pravih pogrešaka $v = A\Theta - y$ slučajni vektor s očekivanjem

$$E(v) = 0, \tag{19}$$

i da je kovarijaciona matrica vektora y oblika

$$D(y) = \sigma_0^2 I, \qquad \sigma_0^2 > 0, \tag{26}$$

gdje je I jedinična matrica tipa (n,n), a σ_0^2 tzv. *faktor varijance*. Tada vrijedi:

5° $D(v) = \sigma_0^2 I$

6° $D(\hat{\Theta}) = \sigma_0^2 (A^T A)^{-1}$

7° $D(\hat{y}) = \sigma_0^2 A (A^T A)^{-1} A^T$

8° $D(\hat{v}) = \sigma_0^2 (I - A(A^T A)^{-1} A^T)$

9° $E(\hat{\sigma}_0^2) = \sigma_0^2, \qquad \hat{\sigma}_0^2 = \dfrac{\hat{v}^T \hat{v}}{n - m}$

$$10^\circ \quad E(\hat{D}(v)) = D(v), \qquad \hat{D}(v) = \hat{\sigma}_0^2 I$$

$$11^\circ \quad E(\hat{D}(\hat{\Theta})) = D(\hat{\Theta}), \qquad \hat{D}(\hat{\Theta}) = \hat{\sigma}_0^2 (A^T A)^{-1}$$

$$12^\circ \quad E(\hat{D}(\hat{y})) = D(\hat{y}), \qquad \hat{D}(\hat{y}) = \hat{\sigma}_0^2 A (A^T A)^{-1} A^T$$

$$13^\circ \quad E(\hat{D}(\hat{v})) = D(\hat{v}), \qquad \hat{D}(\hat{v}) = \hat{\sigma}_0^2 (I - A (A^T A)^{-1} A^T).$$

Svojstva 5°—8° dokazuju se na sasvim jednostavan način primjenom zakona o rasprostiranju varijanci. Svojstva 10°—13° su očigledna nakon što dokažemo svojstvo 9° koje nije trivijalno.

Dokaz svojstva 9°:

U dokazu ovog svojstva koristit ćemo jednostavno svojstvo traga produkta dviju matrica

$$\operatorname{tr}(XY) = \operatorname{tr}(YX)$$

koje naravno vrijedi ukoliko su oba množenja XY i YX definirana. (Trag kvadratne matrice je zbroj njenih elemenata na glavnoj dijagonali.)

Pošto je

$$\hat{v}^T \hat{v} = \operatorname{tr}(\hat{v}\hat{v}^T),$$

to imamo

$$E(\hat{v}^T \hat{v}) = E(\operatorname{tr}(\hat{v}\hat{v}^T)) = \operatorname{tr}(E(\hat{v}\hat{v}^T)) = \operatorname{tr}(D(\hat{v})) = \sigma_0^2 \operatorname{tr}(I - A(A^T A)^{-1} A^T) =$$
$$= \sigma_0^2 (\operatorname{tr}(I) - \operatorname{tr}(A(A^T A)^{-1} A^T)) = \sigma_0^2 (n - \operatorname{tr}(A^T A (A^T A)^{-1})) = \sigma_0^2 (n - m)$$

Dakle,

$$E(\hat{\sigma}_0^2) = E\left(\frac{\hat{v}^T \hat{v}}{n - m}\right) = \sigma_0^2.$$

Time je dokaz završen. Kovarijacione matrice $D(\hat{\Theta})$, $D(v)$, $D(\hat{v})$ i $D(\hat{y})$ mogu se odrediti prije primjene metode najmanjih kvadrata (a priori) ukoliko je poznat faktor varijacije σ_0^2. Veličina $\hat{\sigma}_0^2$ je prema svojstvu 9° centrirana procjena faktora varijance σ_0^2 i u geodeziji se zove *srednja kvadratna pogreška jedinice težine*. Matrice $\hat{D}(\hat{\Theta})$, $\hat{D}(v)$, $\hat{D}(\hat{v})$ i $\hat{D}(\hat{y})$ su centrirane procjene odgovarajućih kovarijacionih matrica (svojstva 10°—13°) i određuju se nakon primjene metode najmanjih kvadrata (a posteriori) kad je poznata procjena $\hat{\sigma}_0^2$.

Za dva slučajna vektora u i z definira se *kovarijaciona matrica* K(u, z) sa

$$K(u, z) = E[(u - E(u))(z - E(z))^T]. \tag{27}$$

U specijalnom slučaju kad je **u** = **z**, vidimo da je

$$K(z, z) = D(z). \tag{28}$$

Kažemo da su vektori **u** i **z** *nekorelirani*, ako je

$$K(u, z) = 0. \tag{29}$$

Na osnovu definicije (27) možemo odrediti kovarijacione matrice za bilo koja dva slučajna vektora, npr $K(\hat{y}, \hat{y})$, $K(y, \hat{\Theta})$, itd.

Teorem 4. Ako su ispunjene pretpostavke (19) i (26) teorema 3, tada su vektori $\hat{y}$ *i* $\hat{v}$ nekorelirani, tj. vrijedi

$$K(\hat{y}, \hat{v}) = 0. \tag{30}$$

Dokaz:

$$\begin{aligned} K(\hat{y}, \hat{v}) &= E[\hat{y} - E(\hat{y}))(\hat{v} - E(\hat{v}))^T] = E(A(A^T A)^{-1} A^T (y - \\ &\quad - E(y))[(A(A^T A)^{-1} A^T - I)(y - E(y))]^T) = \\ &= A(A^T A)^{-1} A^T E[(y - E(y))(y - E(y))^T](A(A^T A)^{-1} A^T - I) = \\ &= A(A^T A)^{-1} A^T D(y)(A(A^T A)^{-1} A^T - I) = \\ &= \sigma_0^2 (A(A^T A)^{-1} A^T - A(A^T A)^{-1} A^T) = 0. \end{aligned}$$

Time je dokaz teorema 4 završen. Zamislimo li da kovarijaciona matrica K(**u**, **z**) predstavlja »skalarni produkt« slučajnih vektora **u** i **z**, onda stohastička nekoreliranost (30) znači ortogonalnost vektora $\hat{y}$ i $\hat{v}$, a relacija

$$D(y) = D(\hat{y}) + D(\hat{v}), \tag{31}$$

koja se odmah dobije iz svojstava 7° i 8°, je Pitagorin poučak (Meissl 1982).

Poznato je da se *norma kvadratne matrice* X može definirati npr. sa

$$\|X\| = \sqrt{tr(XX^T)}. \tag{32}$$

Iz relacije (31) lako se izvede

$$\begin{gathered} D(y) D^T(y) = D(\hat{y}) D^T(\hat{y}) + D(\hat{v}) D^T(\hat{v}), \\ tr(D(y) D^T(y)) = tr(D(\hat{y}) D^T(\hat{y})) + tr(D(\hat{v}) D^T(\hat{v})), \\ \|D(y)\|^2 = \|D(\hat{y})\|^2 + \|D(\hat{v})\|^2, \end{gathered} \tag{33}$$

što je opet jedan oblik poznatog Pitagorinog poučka.

Teorem 5. Neka su ispunjene pretpostavke (19) i (26) teorema 3, neka je C zadana matrica i neka je

$$z = C\Theta \quad \text{i} \quad \hat{z} = C\hat{\Theta}. \tag{34}$$

Tada vrijedi

14° $E(\hat{z}) = z$

15° $D(\hat{z}) = \sigma_0^2 \, C (A^T A)^{-1} C^T$

16° $E(\hat{D}(\hat{z})) = D(\hat{z}), \qquad \hat{D}(\hat{z}) = \hat{\sigma}_0^2 \, C (A^T A)^{-1} C^T.$

Dokaz svojstava 14°—16° je trivijalan pa ga ispuštamo. Ova svojstva koriste se kod prognoziranja, uz ocjenu točnosti, linearnih funkcija parametra Θ.

3.1. *Najbolja centrirana linearna procjena parametara*

Reći ćemo da za simetrične matrice istog tipa X i Y vrijedi

$$Y \geq X, \text{ odnosno } Y - X \geq 0 \tag{35}$$

ako i samo ako je matrica Y—X nenegativno definitna, tj. ako i samo ako za svaki vektor z vrijedi

$$z^T (Y - X) z \geq 0. \tag{36}$$

Teorem 6. Neka su ispunjene prepostavke (19) i (26) teorema 3, neka je B matrica sa svojstvom

$$BA = I \tag{37}$$

i neka je

$$\hat{\Theta}_1 = By. \tag{38}$$

Tada vrijedi

17° $E(\hat{\Theta}_1) = \Theta$

18° $D(\hat{\Theta}_1) \geq D(\hat{\Theta}) = \sigma_0^2 (A^T A)^{-1}$

19° $E[(\hat{\Theta}_1 - \Theta)^T (\hat{\Theta}_1 - \Theta)] \geq E[(\hat{\Theta} - \Theta)^T (\hat{\Theta} - \Theta)] = \sigma_0^2 \operatorname{tr} (A^T A)^{-1},$

s time da umjesto znaka $\geq$ u 18° i 19° vrijedi znak = samo kad je

$$B = (A^T A)^{-1} A^T. \tag{39}$$

Dokaz svojstva 17°:

$$\hat{\Theta}_1 = By = B(A\Theta - v) = \Theta - Bv \Rightarrow E(\hat{\Theta}_1) = \Theta, \text{ zbog } E(v) = 0.$$

Dokaz svojstva 18°:

$$\begin{aligned} D(\hat{\Theta}_1) &= BD(y)B^T = B(D(\hat{y}) + D(\hat{v})B^T = \sigma_0^2 BA(A^T A)^{-1} A^T B^T + \\ &+ \sigma_0^2 B(I - A(A^T A)^{-1} A^T) B^T = \sigma_0^2 (A^T A)^{-1} + \\ &+ \sigma_0^2 B(I - A(A^T A)^{-1} A^T)(I - A(A^T A)^{-1} A^T) B^T = \\ &= D(\hat{\Theta}) + \sigma_0^2 (B - (A^T A)^{-1} A^T)(B - (A^T A)^{-1} A^T)^T. \end{aligned}$$

Time je svojstvo 18° dokazano, jer se lako vidi da je za svaku matricu X,

$$X^T X \geq 0 \quad \text{i}$$
$$X^T X = 0 \Leftrightarrow X = 0.$$

Dokaz svojstva 19°:

$$\begin{aligned} E[(\hat{\Theta}_1 - \Theta)^T (\hat{\Theta}_1 - \Theta)] &= E(\mathrm{tr}[(\hat{\Theta}_1 - \Theta)^T (\hat{\Theta}_1 - \Theta)]) = \\ &= E(\mathrm{tr}[(\hat{\Theta}_1 - \Theta)(\hat{\Theta}_1 - \Theta)^T]) = \mathrm{tr}(E[(\hat{\Theta}_1 - \Theta)(\hat{\Theta}_1 - \Theta)^T]) = \\ &= \mathrm{tr}(D(\hat{\Theta}_1)) \end{aligned}$$

$$E[(\hat{\Theta} - \Theta)^T (\hat{\Theta} - \Theta)] = \mathrm{tr}(D(\hat{\Theta}))$$

Nadalje, iz dokaza svojstva 18° slijedi

$$\mathrm{tr}(D(\hat{\Theta}_1)) = \mathrm{tr}(D(\hat{\Theta})) + \sigma_0^2 \,\mathrm{tr}[(B - (A^T A)^{-1} A^T)(B - (A^T A)^{-1} A^T)^T],$$

a time je svojstvo 19° dokazano, jer se lako vidi da je za svaku matricu X

$$\mathrm{tr}(X^T X) \geq 0 \quad \text{i}$$
$$\mathrm{tr}(X^T X) = 0 \Leftrightarrow X = 0.$$

Time je ujedno završen i dokaz teorema 6. Na osnovu svojstva 17°, skup svih vektora $\hat{\Theta}_1$ oblika $\hat{\Theta}_1 = By$, uz $BA = I$, zove se *skup svih centriranih linearnih procjena parametra* Θ.

Svojstvo 18° iskazujemo riječima kad kažemo da u skupu svih linearnih procjena parametra Θ, najmanju kovarijacionu matricu (u smislu uređaja (35)) ima procjena $\hat{\Theta} = (A^T A)^{-1} A^T y$ dobivena metodom najmanjih kvadrata.

Svojstvo 19° znači da u skupu svih linearnih procjena parametra Θ kovarijacionu matricu s najmanjim tragom (ponekad se kaže s najmanjom varijancom) ima procjena $\hat{\Theta} = (A^T A)^{-1} A^T y$ dobivena metodom najmanjih kvadrata. Takva procjena zove se *najbolja centrirana linearna procjena.*

3.2. *Najvjerojatnija procjena parametara*

Teorem 7. Neka su ispunjene pretpostavke (19) i (26) teorema 3 i neka je vektor pravih pogrešaka opažanja

$$v = A\Theta - y$$

normalno distribuirani slučajni vektor s funkcijom gustoće vjerojatnosti

$$\Phi = (2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{\|A\Theta - y\|^2}{2\sigma_0^2}\right). \tag{40}$$

Funkcija $\Phi = \Phi(\Theta, \sigma_0^2)$ ima (globalni) maksimum $\Phi(\hat{\Theta}, \tilde{\sigma}_0^2)$, gdje su

$$\hat{\Theta} = (A^T A)^{-1} A^T y, \tag{6}$$

$$\tilde{\sigma}_0^2 = \frac{\|A\hat{\Theta} - y\|^2}{n}. \tag{41}$$

U dokazu teorema koristit ćemo svojstvo

$$e^{x-1} \geq x \tag{42}$$

za svaki realni broj x, s time da umjesto znaka $\geq$ vrijedi znak $=$ ako i samo ako je $x = 1$. Ovo svojstvo se može lako dokazati ako napišemo Taylorovu formulu za funkciju

$$f(x) = e^{x-1} \tag{43}$$

u proizvoljnoj okolini točke $x = 1$ u obliku

$$f(x) = f(1) + (x-1) f'(1) + \frac{(x-1)^2}{2} f''(1 + \vartheta(x-1)) \quad 0 < \vartheta < 1, \tag{44}$$

dakle

$$e^{x-1} = x + \frac{(x-1)^2}{2} e^{\vartheta(x-1)}, \quad 0 < \vartheta < 1. \tag{45}$$

Odatle odmah slijedi svojstvo (42).

Dokaz teorema 7:

Treba dokazati da za svaki Θ i svaki $\sigma_0^2 > 0$ vrijedi

$$\Phi(\hat{\Theta}, \tilde{\sigma}_0^2) \geq \Phi(\Theta, \sigma_0^2). \tag{46}$$

Dokaz ćemo provesti tako da najprije pokažemo da vrijedi

$$\Phi(\hat{\Theta}, \tilde{\sigma}_0^2) \geq \Phi(\hat{\Theta}, \sigma_0^2), \tag{47}$$

a zatim

$$\Phi(\hat{\Theta}, \sigma_0^2) \geq \Phi(\Theta, \sigma_0^2). \tag{48}$$

Izraz (47) ekvivalentan je s

$$(2\pi\tilde{\sigma}_0^2)^{-n/2} \exp\left(-\frac{\|A\hat{\Theta} - y\|^2}{2\tilde{\sigma}_0^2}\right) \geq (2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{\|A\hat{\Theta} - y\|^2}{2\sigma_0^2}\right)$$

$$\Leftrightarrow \left(\frac{\tilde{\sigma}_0^2}{\sigma_0^2}\right)^{-n/2} \geq \exp\left(-\frac{\|A\hat{\Theta} - y\|^2}{2\sigma_0^2} + \frac{\|A\hat{\Theta} - y\|^2}{2\tilde{\sigma}_0^2}\right) = \exp\left(-\frac{n}{2}\left(\frac{\tilde{\sigma}_0^2}{\sigma_0^2} - 1\right)\right)$$

$$\Leftrightarrow \exp\left(\frac{\tilde{\sigma}_0^2}{\sigma_0^2} - 1\right) \geq \frac{\tilde{\sigma}_0^2}{\sigma_0^2},$$

što je istina zbog relacije (42). Nadalje, izraz (48) ekvivalentan je s

$$(2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{\|A\hat{\Theta} - y\|^2}{2\sigma_0^2}\right) \geq (2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{\|A\Theta - y\|^2}{2\sigma_0^2}\right)$$

$$\Leftrightarrow -\frac{\|A\hat{\Theta} - y\|^2}{2\sigma_0^2} \geq -\frac{\|A\Theta - y\|^2}{2\sigma_0^2}$$

$$\Leftrightarrow \quad \|A\Theta - y\|^2 \geq \|A\hat{\Theta} - y\|^2$$

što je istina na osnovu teorema 1. Još samo treba uočiti da umjesto znaka $\geq$ u relacijama (46), odnosno (47)—(48), može stajati znak $=$ onda i samo onda kad je

$$\sigma_0^2 = \tilde{\sigma}_0^2 \quad \text{i} \quad \Theta = \hat{\Theta}$$

što se lako vidi na osnovu relacije (42) i teorema 1. Time je teorem 7 u potpunosti dokazan. Za procjenu $\hat{\Theta}$ dobivenu metodom najmanjih kvadrata, a uz pretpostavku o normalnoj distribuciji pogrešaka opažanja (teorem 7), kaže se da je *navjerojatnija procjena parametra* Θ, a za

$$\tilde{\sigma}_0^2 = \frac{\|A\hat{\Theta} - y\|^2}{n} = \frac{\hat{v}^T \hat{v}}{n}$$

da je *najvjerojatnija procjena koeficijenta varijance*, odakle vidimo da ova procjena nije centrirana, tj.

$$E(\tilde{\sigma}_0^2) \neq \sigma_0^2.$$

Međutim, $\tilde{\sigma}_0^2$ je *asimptotski centrirana* jer vrijedi

$$\lim_{n\to\infty} E(\tilde{\sigma}_0^2) = \lim_{n\to\infty} E(\hat{\sigma}_0^2) \frac{n-m}{n} = \sigma_0^2 \lim_{n\to\infty} \frac{n-m}{n} = \sigma_0^2.$$

4. GAUß-MARKOVLJEV MODEL I NEKA NJEGOVA POOPĆENJA

Neka je A zadana matrica tipa (n,m), rang(A) = m, Θ vektor nepoznatih parametara, y slučajni vektor opažanja i $D(y) = \sigma_0^2 I$ kvorijaciona matrica vektora y, $\sigma_0^2 > 0$ nepoznati faktor. Tada se sistem

$$A\Theta = E(y), \qquad D(y) = \sigma_0^2 I \tag{49}$$

zove *Gauß-Markovljev model s potpunim rangom*. Teoremi 2—6 govore o procjenama i njihovim osnovnim svojstvima u ovom modelu. Pretpostavka ovog modela je da se očekivanje opažanja može prikazati kao linearna kombinacija zadanih vektora (vektori stupci matrice A) s koeficijentima koji su nepoznati parametri. Ovakva zavisnost naziva se također i *linearna regresija*, a procjenjivanje u modelu (49) *regresiona analiza*. U računu izjednačenja spomenuto procjenjivanje naziva se najčešće *posredno izjednačenje opažanja* ili *parametarsko izjednačenje*.

(i) U slučaju da je očekivanje opažanja nelinearna funkcija, ona se obično linearizira razvojem u Taylorov red, i tako nastaje model (49).

(ii) Ako je kovarijaciona matrica oblika $D(y) = \sigma_0^2 P^{-1}$, gdje je P (a time i P^{-1}) pozitivno definitna matrica, onda se model

$$A\Theta = E(y), \qquad D(y) = \sigma_0^2 P^{-1} \tag{50}$$

može svesti na model (49) npr. supstitucijama

$$y' = G^T y, \quad A' = G^T A \tag{51}$$

gdje je G donja trokutasta matrica iz rastava matrice P prema Choleskom

$$P = GG^T. \tag{52}$$

(iii) Daljnja moguća poopćenja modela (49) nastaju npr. ako je rang (A) < m ili ako je kovarijaciona matrica D(y) singularna. Njihovo proučavanje prelazi okvire ovog rada (Rao, Mitra 1971), (Koch 1980), (Meissl 1982), (Perović 1986).

Zahvaljujem se recenzentima i kolegama koji su čitali rukopis ovog rada na korisnim primjedbama.

LITERATURA

Czuber, E.: Theorie der Beobachtungsfehler. B. G. Teubner, Leipzig, 1891.

Čubranić, N.: Račun izjednačenja. Stručni otsjek N.S.O-e zagrebačkog Sveučilišta, Zagreb, 1948.

Čubranić, N.: Račun izjednačenja. Tehnička knjiga, Zagreb, 1958. (drugo izdanje 1967.).

Čubranić, N.: Teorija pogrešaka s računalom izjednačenja. Sveučilišna naklada Liber, Zagreb, 1980.

Hadživuković, S., R. Zegnal, K. Čobanović: Regresiona analiza. Privredni pregled, Beograd, 1982.

Helmert, F., R.: Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. B. G. Teubner, Leipzig, 1872.

Horvat, S.: Geodetsko računanje I. Udruženje slušača Tehničkog fakulteta, Zagreb, 1937.

Junašević, M.: Geometrijska interpretacija uvjeta minimuma sume kvadrata odstupanja. Zbornik radova povodom 50-godišnjice Geodetskog fakulteta. Geodetski fakultet Sveučilišta u Zagrebu, 1970, 165—176.

Klak, S.: Teorija pogrešaka i račun izjednačenja. Sveučilišna naklada Liber, Zagreb, 1982.

Koch, K., R.: Parameterschätzung und Hypotensentests in linearen Modellen. Dümmler, Bonn, 1980.

Kolmogorov, A. N.: K obosnovaniju metoda naimenših kvadratov. Uspehi matematičeskih nauk, tom 1, vip. 1, 1946, 57—70.

Macarol, S.: Praktična geodezija (treće izdanje). Tehnička knjiga, Zagreb, 1978.

Meissl, P.: Least Squares Adjustment :A Modern Approach. Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 43, Graz, 1982

Mihailović, K.: Račun izravnanja I, Građevinski fakultet Univerziteta u Beogradu, 1980.

Muminagić, A., V. Jovanović: Račun izravnanja. Vojnogeografski institut, Beograd, 1965.

Pašalić, S.: Račun izravnanja. Građevinski fakultet, Sarajevo, 1984.

Perović, G.: Singularna izravnanja. Naučna knjiga i Građevinski fakultet, Beograd, 1986.

Rao, C., R., S. K. Mitra: Generalized Inverse of Matrices and its Applications. John Wiley & Sons, New York, London, Sydney, Toronto, 1971.

Vranić, V.: Vjerojatnost i statistika (treće izdanje). Tehnička knjiga, Zagreb, 1971.

SAŽETAK

Kod nas se u geodetskoj literaturi zasnivanju metode najmanjih kvadrata, gotovo redovito, prilazilo s aspekta teorije pogrešaka i normalnosti njihove razdiobe. U ovom radu pokazuje se da i mnogo općenitije pretpostavke vode k istim procjenama parametara. Najvažnije činjenice o geometrijskoj i stohastičkoj prirodi metode najmanjih kvadrata istaknute su u obliku sedam teorema. Posebna pažnja posvećena je što jednostavnijem i kraćem dokazivanju tih teorema.

ABSTRACT

In Yugoslav geodetic literature, the founding of the method of least squares has almost as a rule been approached from the aspect of the theory of errors and the normality of their distribution. The present paper demonstrates that more general suppositions lead to same estimates of the parameters. The most important facts concerning the geometric and the stochastic nature of the method of least squares have been expressed in the form of seven theorems. A special attention has been paid to the simplest and shortest possible proofs of those theorems.

Primljeno: 1989-01-17