Elementi matrične algebre u statistici

Master rad

Student:

Dubravka Punišić

Mentor:

Prof.dr Vesna Jevremović

Članovi komisije:

Prof. dr Vesna Jevremović, mentor

Prof. dr Aleksandar Lipkovski

Prof. dr Danko Jocić

Beograd

Septembar 2014.

Sadržaj

Uvod

Matrična algebra ima veoma važnu ulogu u statistici. U mnogim oblastima statistike uobičajeno je da se matrična algebra koristi za prikazivanje, izvođenje i potvrđivanje rezultata. Dve statističke oblasti u kojima je matrična algebra najzastupljenija su linearni statistički modeli i multivarijaciona analiza. To proizilazi iz činjenice da se ove oblasti bave proučavanjem velikog broja promenljivih koje se iz praktičnih razloga lakše zapisuju u matričnom obliku.

Sam rad predstavlja interdisciplinarni pregled različitih tema sa kojima smo upoznati tokom osnovnih studija, kao i neka dodatna saznanja u oblasti linearne algebre i različitih disciplina statistike. Od pomenutih statističkih disciplina to su prvenstveno linearni statistički modeli i vremenske serije. U vremenskim serijama matrice se primenjuju pri rešavanju sistema Džul-Vokerovih jednačina vektorskih autoregresionih modela, kao i prilikom ocenjivanja parcijalne korelacione funkcije. Kroskovarijacione matrice imaju primenu kod vektorskih vremenskih serija, a karakteristične vrednosti matrice parametara su bitne da bi se ustanovilo da li je vektorska vremenska serija stacionarna ili ne. Sopstvene vrednosti i sopstveni vektori koriste se i u analizi glavnih komponenata, koja je jedna od metoda multivarijacione analize. Tačnije, sopstvene vrednosti matrice kovarijacije su disperzije glavnih komponenata posmatranog sistema. Pored toga, matrični račun ima veliku primenu i u oblastima kao što su sistemi masovnog opsluživanja, procesi rađanja i umiranja, teorija informacija…, ali oni neće biti izloženi u ovom radu.

U prva tri poglavlja ovog rada biće predstavljene osnove matrične algebra. U poglavlju "Sistemi linearni jednačina" biće izložene definicije i osnovne osobine linearne jednačine i skupa linearnih jednačina. U poglavljima "Matrice" i "Determinante" upoznajemo se sa značenjem ovih termina kao i njihovom primenom radi lakšeg zapisivanja i rešavanja sistema linearnih jednačina. Ova tri poglavlja su neophodna da bismo na razumljiv i jasan način mogli da predstavimo upotrebu matričnog računa u statistici.

Četvrto poglavlje posvećeno je linearnim modelima kod kojih je broj jednačina manji od broja promenljivih, odnosno u situaciji kada nemamo dovoljan broj podataka na osnovu kojih određujemo nepoznate koeficijente u linearnom modelu. Tada koristimo uopštenu inverznu matricu.

U poglavlju "Matrica kovarijacije" prvi put u ovom radu se susrećemo sa pojmovima iz statistike kao sto su kovarijacija, korelacija i slučajni vektor. Dalje uopštavanje ovih pojmova dovodi nas do matrice kovarijacije i pojedinih njenih osobina koje su nam od značaja. U ovom poglavlju videćemo i kako se slučajna veličina može predstvaiti pomoću kvadratne forme, kao i ulogu matričnog računa u analizi glavnih komponenata.

Poglavlje "Matrice u metodi najmanjih kvadrata" pored same teoreme Gaus-Markova, njenog dokaza i primene, sadrži nekoliko tema sa kojima se moramo upoznati u cilju potpunog razumevanja teoreme. Ova teorema primenljiva je u slučaju kada je matrica linearnog modela potpunog ranga, odnosno kada imamo jednak ili veći broj jednačina nego li promenljivih.

Koliko god da matrice olakšavaju zapis linearnih statističkih modela, operacije koje vršimo nad matricama i dalje predstavljaju obiman računski zadatak. Pogotovo ako je broj promenljivih veliki. Zato je u ovom radu predstavljen softverski programi Matlab, koji je namenjen za odgovarajuća izračunavanja. Pored Matlab-a, izdvajaju se i statistički softver R, program SPSS, ali i Microsoft-ov program Excel. Ipak, za pravilno korišćenje takvih programa neophodno je poznavanje teorije i procesa koji se prilikom njihove upotrebe odvijaju.

O značaju matričnog računa u statistici govori i to da su cele knjige posvećene upravo toj temi. U toku pisanja ovog rada izdvojila se knjiga „Matrix Algebra From a Statistician's Perspective" koja detaljno i sveobuhvatno pokriva ovu temu.

1 Sistemi linearnih jednačina

Definicija 1.1. Linearna jednačina sa $\boldsymbol{n}$ nepoznatih nad poljem $\mathbb{F}$ predstavlja jednačinu oblika

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

pri čemu su $a_i, b \in \mathbb{F}(i = 1, \dots, n)$, a $x_i(i = 1, \dots, n)$ su nepoznate (promenljive).

Sistem linearnih jednačina (linearni sistem) sa m jednačina i n nepoznatih nad poljem $\mathbb{F}$ je skup od m takvih jednačina oblika

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

pri čemu su $a_{ij} \in \mathbb{F}(i = 1, \dots, m,\ j = 1, \dots, n)$ koeficjenti sistema, $b_i \in \mathbb{F}\ (i = 1, \dots, m)$ su slobodni članovi sistema, a $x_j(j = 1, \dots, n)$ su nepoznate (promenljive) sistema.

Rešenje sistema je uređena n –torka $(c_1, \dots, c_n) \in \mathbb{F}^n$ elemenata iz polja $\mathbb{F}$ za koje važi

$$a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1$$

$$a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2$$

$$\dots$$

$$a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m$$

tj. zamenom ovih elemenata u sistem dobija se m jednakosti u polju $\mathbb{F}$. Skup rešenja linearnog sistema označimo sa S. Važi da je $S \subset \mathbb{F}^n$.

Definicija 1.2. Sistem je saglasan, ako je skup rešenja neprazan. Sistem je nesaglasan, ako je skup rešenja prazan. Saglasan sistem je određen, ako je skup rešenja jednočlan. Saglasan sistem je neodređen, ako skup rešenja ima više od jednog elementa. Ako su svi slobodni članovi sistema jednaki 0, onda je sistem homogen. Ako postoji bar jedan slobodan član koji nije 0, onda je sistem nehomogen.

Svaki homogen sistem je saglasan, jer je u tom slučaju n –torka $(0, \dots, 0)$ sigurno rešenje sistema. Ovo rešenje se naziva trivijalno.

Na osnovu prethodne definicije, zaključujemo da je moguće sledeće:

(1) Sistem ima beskonačno mnogo rešenja tj. sistem je saglasan i neodređen.
(2) Sistem ima jedno jedinstveno rešenje tj. sistem je saglasan i određen.

(3) Sistem nema rešenje tj. sistem je nesaglasan.

Može se desiti da različiti sistemi imaju isti skup rešenja. Za takve sisteme kažemo da su ekvivalentni. Pojam ekvivalentnosti sistema je bitan, jer nam pomaže u određivanju rešenja sistema jednačina. Ako postupnim transformacijama jedan sistem svedemo na drugi tako da u svakom koraku imamo ekvivalentne sisteme, onda određivanjem rešenja poslednjeg sistema ujedno određujemo rešenje i polaznog sistema.

Definicija 1.3. Elementarna transformacija sistema $\boldsymbol{\mathcal{A}}$ prvog tipa $\boldsymbol{p_{ij}(i,j=1,\dots,m, i \neq j)}$ je zamena mesta $\boldsymbol{i}$ –te i $\boldsymbol{j}$ –te jednačine u sistemu.

$$\boldsymbol{p_{ij}: \mathcal{A} \mapsto p_{ij}(\mathcal{A})}$$

Elementarna transformacija sistema $\boldsymbol{\mathcal{A}}$ drugog tipa $\boldsymbol{q_{ij}(\lambda)(i,j=1,\dots,m, i \neq j, \lambda} \in \mathbb{F})$ je dodavanje $\boldsymbol{j}$ –toj jednačini u sistemu $\boldsymbol{i}$ –tu jednačinu pomnoženu brojem $\boldsymbol{\lambda}$.

$$\boldsymbol{q_{ij}(\lambda): \mathcal{A} \mapsto q_{ij}(\lambda)(\mathcal{A})}$$

Elementarna transformacija sistema $\boldsymbol{\mathcal{A}}$ trećeg tipa $\boldsymbol{r_i(\lambda)(i=1,\dots,m, \lambda} \in \mathbb{F} \setminus \{\boldsymbol{0}\})$ je množenje $\boldsymbol{i}$ –te jednačine u sistemu brojem $\boldsymbol{\lambda \neq 0}$.

$$\boldsymbol{r_i(\lambda): \mathcal{A} \mapsto r_i(\lambda)(\mathcal{A})}$$

Ako je sistem $\mathcal{A}'$ dobijen primenom konačnog niza elementarnih transformacija na sistem $\mathcal{A}$, onda kažemo da je $\mathcal{A}'$ dobijen ekvivalentnom transformacijom sistema $\mathcal{A}$ i to obeležavamo sa $\mathcal{A} \rightsquigarrow \mathcal{A}'$.

Stav 1.1. Ako je $\mathcal{A} \rightsquigarrow \mathcal{A}'$, onda je i $\mathcal{A}' \rightsquigarrow \mathcal{A}$ i važi $S(\mathcal{A}) = S(\mathcal{A}')$ tj. sistemi su ekvivalentni.

Dokaz se može naći u literaturi [5] na strani 61.

Teorema 1.2. (Gausova[1] metoda eliminacije):

(1) Svaki sistem oblika

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

ekvivalentan je sistemu oblika

$$a'_{1j_1}x_{j_1} + \cdots + a'_{1n}x_n = b'_1$$

[1]Karl Friedrich Gaus (1777-1855), nemački matematičar, fizičar i astronom

$$a'_{2j_2}x_{j_2} + \cdots + a'_{2n}x_n = b'_2$$

$$\dots$$

$$a'_{rj_r}x_{j_r} + \cdots + a'_{rn}x_n = b'_r$$

$$0 = b'_{r+1}$$

$$\dots$$

$$0 = b'_m$$

koji se naziva stepenasti sistem. Pri tome $1 \le j_1 < j_2 < \cdots < j_r \le n$ i $a'_{1j_1},\ a'_{2j_2}, \dots, a'_{rj_r} \neq 0$ se nazivaju glavni koeficijenti. Nepoznate $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ su vezane promenljive, a ostale promenljive su slobodne.

(2) Stepenasti sistem je saglasan tj. ima rešenje ako i samo ako ne sadrži jednačine oblika $0 = b'_k$ gde je $b'_k \neq 0$.

(3) Neka je stepenasti sistem saglasan. Tada:

(3.1) Za $r = n$ sistem je određen tj. ima jedinstveno rešenje. U tom slučaju mora da bude $j_i = i, n \le m$ i sistem ima oblik

$$a'_{11}x_1 + \cdots + a'_{1n}x_n = b'_1$$

$$a'_{22}x_2 + \cdots + a'_{2n}x_n = b'_2$$

$$\dots$$

$$a'_{nn}x_n = b'_n.$$

Rešenje se dobija hodom unazad.

$$x_k = \frac{1}{a'_{kk}}\left[b'_k - \sum_{j>k} a'_{kj}x_j\right](k = n, n-1, \dots, 1).$$

(3.2) Za $r < n$ sistem je neodređen. Skup rešenja se dobija tako što slobodnim promenljivima dajemo proizvoljne vrednosti iz polja $\mathbb{F}$, a vezane promenljive određujemo hodom unazad

$$x_{j_k} = \frac{1}{a'_{kj_k}}\left[b'_k - \sum_{j>j_k} a'_{kj}x_j\right](k = r, r-1, \dots, 1).$$

Dokaz se može naći u literaturi [5] na strani 63.

Stav 1.3. Broj vezanih promenljivih r kao i same vezane promenljive $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ sa $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ ne zavise od načina svođenja tj. redosleda elementarnih transformacija. Jednoznačno određeni broj r naziva se rang sistema.

Dokaz se može naći u literaturi [5] na strani 66.

Definicija 1.4. Ako je $r = n$ onda je to sistem sa potpunim rangom. A ako je $r < n$ onda je u pitanju sistem sa nepotpunim rangom.

2 Matrice

Definicija 2.1. Ako su $\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}$ dva prirodna broja, matrica nad poljem $\mathbb{F}$ tipa $\boldsymbol{m} \times \boldsymbol{n}$ je tablica oblika

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

čiji su članovi (elementi) $a_{ij} \in \mathbb{F}$ $(i = 1, \dots, m, j = 1, \dots, n)$.

Definicija 2.2. Matrica tipa $1 \times n$ se naziva vrsta, matrica tipa $m \times 1$ naziva se kolona, a matrica tipa $n \times n$ se naziva kvadratna matrica tipa n.

$M(m \times n, \mathbb{F}) = \left\{A = [a_{ij}]_{m \times n} | \forall a_{ij} \in \mathbb{F}\right\}$ je skup svih matrica tipa $m \times n$ nad poljem $\mathbb{F}$ u oznaci $A = [a_{ij}]$ gde je i redni broj vrste i j redni broj kolone. Posebna oznaka za kvadratne matrice tipa n je $M(n, \mathbb{F})$.

Definicija 2.3. Kvadratna matrica $A = [a_{ij}] \in M(n, \mathbb{F})$ je simetrična matrica ako je $a_{ij} = a_{ji}$ za sve $(i, j = 1, \dots, n)$.

Definicija 2.4. Kvadratna matrica $A = [a_{ij}] \in M(n, \mathbb{F})$ je dijagonalna matrica ako je $a_{ij} = 0$ za $i \neq j$. Pri tome se elementi $a_{ii} (i = 1, \dots, n)$ nazivaju dijagonalnim elementima.

Ako je polje $\mathbb{F} = \mathbb{R}$, onda se matrice nazivaju realnim, a ako je $\mathbb{F} = \mathbb{C}$ matrice se nazivaju kompleksnim.

2.1 Osnovni pojmovi i operacije u matričnom računu

Definicija 2.5. Ako su $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$ matrice $\boldsymbol{A} = [\boldsymbol{a_{ij}}]$, $\boldsymbol{B} = [\boldsymbol{b_{ij}}]$ istog tipa $\boldsymbol{m} \times \boldsymbol{n}$, zbir te dve matrice $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C}$ je matrica $\boldsymbol{C} = [\boldsymbol{c_{ij}}]$ tipa $\boldsymbol{m} \times \boldsymbol{n}$ takva da je $\boldsymbol{c_{ij}} = \boldsymbol{a_{ij}} + \boldsymbol{b_{ij}}$ za sve $\boldsymbol{i} = \boldsymbol{1}, \dots, \boldsymbol{m},\ \boldsymbol{j} = \boldsymbol{1}, \dots, \boldsymbol{n}$.

Stav 2.1. Skup $\boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$ svih matrica tipa $\boldsymbol{m} \times \boldsymbol{n}$ nad poljem $\mathbb{F}$ je komutativna grupa u odnosu na operaciju sabiranja matrica.

Dokaz se može naći u literaturi [5] na strani 69.

Neutral je matrica $0 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in M(m \times n, \mathbb{F})$ pri čemu je $0 \in \mathbb{F}$ neutral polja $\mathbb{F}$. Ovu matricu nazivamo nula-matrica tipa $m \times n$. Za svaki tip postoji posebna nula-matrica.

Suprotan element matrice $A = [a_{ij}] \in M(m \times n, \mathbb{F})$ je matrica $-A = [-a_{ij}] \in M(m \times n, \mathbb{F})$.

Definicija 2.6. Ako je $\boldsymbol{A} = [\boldsymbol{a_{ij}}] \in \boldsymbol{M}(\boldsymbol{m} \text{ x } \boldsymbol{n}, \mathbb{F})$ matrica tipa $\boldsymbol{m} \text{ x } \boldsymbol{n}$ i $\boldsymbol{\lambda} \in \mathbb{F}$ element polja $\mathbb{F}$ (skalar), proizvod matrice $\boldsymbol{A}$ i skalara $\boldsymbol{\lambda}$ je matrica $\boldsymbol{B} = \boldsymbol{\lambda A} = [\boldsymbol{b_{ij}}]$ tipa $\boldsymbol{m} \text{ x } \boldsymbol{n}$ čiji su elementi $\boldsymbol{b_{ij}} = [\boldsymbol{\lambda a_{ij}}]$ za sve $\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{m}$, $\boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{n}$.

Množenje matrice skalarom je spoljna operacija, jer je to operacija između elemenata dva različita skupa tj. $\mathbb{F} \times M(m \text{ x } n, \mathbb{F}) \to M(m \text{ x } n, \mathbb{F})$.

Stav 2.2. Ako su $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{m} \text{ x } \boldsymbol{n}, \mathbb{F})$ matrice $\boldsymbol{A} = [\boldsymbol{a_{ij}}]$, $\boldsymbol{B} = [\boldsymbol{b_{ij}}]$ i $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{F}$ skalari, tada važe jednakosti:

(1) $\lambda(A + B) = \lambda A + \lambda B$;
(2) $(\lambda + \mu)A = \lambda A + \mu A$;
(3) $(\lambda\mu)A = \lambda(\mu A)$;
(4) $1A = A$.

Dokaz se može naći u literaturi [5] na strani 70.

Definicija 2.7. Ako su $\boldsymbol{A} = [\boldsymbol{a_{ij}}] \in \boldsymbol{M}(\boldsymbol{m} \text{ x } \boldsymbol{n}, \mathbb{F})$ i $\boldsymbol{B} = [\boldsymbol{b_{jp}}] \in \boldsymbol{M}(\boldsymbol{n} \text{ x } \boldsymbol{k}, \mathbb{F})$, proizvod te dve matrice $\boldsymbol{AB} = \boldsymbol{C}$ je matrica $\boldsymbol{C} = [\boldsymbol{c_{ip}}]$ tipa $\boldsymbol{m} \text{ x } \boldsymbol{k}$ takva da je

$$c_{ip} = \sum_{j=1}^{n} a_{ij} b_{jp}$$

za sve $i = 1, \dots, m$, $p = 1, \dots, k$.

Množenje matrica je parcijalna operacija na skupu svih matrica. Ne mogu se množiti bilo koje dve matrice. Neophodan i dovoljan uslov da se matrice A i B mogu pomnožiti, i to baš tim redom, je da broj kolona matrice A bude jednak broju vrsta matrice B.

I kada postoje oba proizvoda AB i BA, što je moguće ako i samo ako je $A \in M(m \text{ x } n, \mathbb{F})$ i $B \in M(n \text{ x } m, \mathbb{F})$, rezultati su matrice različitog tipa. Čak i kada su rezultati matrice istog tipa, što je moguće ako i samo ako su $A, B \in M(n, \mathbb{F})$, rezultati su različite matrice. Tako da množenje matrica nije komutativna operacija.

Stav 2.3. Neka je $\boldsymbol{A} = [\boldsymbol{a_{ij}}]$, $\boldsymbol{B} = [\boldsymbol{b_{jk}}]$, $\boldsymbol{C} = [\boldsymbol{c_{kp}}]$. Uz uslov da svi zbirovi i proizvodi postoje, važi sledeće:

(1) asocijativnost množenja matrica $A(BC) = (AB)C$;
(2) desna distributivnost $A(B + C) = AB + AC$;
(3) leva distributivnost $(A + B)C = AC + BC$;
(4) U svakom skupu $M(n, \mathbb{F})$ postoji matrica E reda n koja se naziva jedinična matrica i oblika je

$$E = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

takva da za bilo koje dve matrice $A, B \in M(n, \mathbb{F})$, kada god su proizvodi definisani važi $AE = A$ i $EB = B$.

Dokaz se može naći u literaturi [5] na strani 73.

Kao i za nula-matricu, tako i za jediničnu matricu važi da za svaki tip postoji posebna jedinična matrica.

Ako posmatramo skup svih kvadratnih matrica reda n, onda množenje matrica postaje binarna operacija, jer se svake dve matrice tog reda mogu množiti.

Posledica 2.4. Skup svih kvadratnih matrica $\boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ sa operacijama sabiranja i množenja matrica je nekomutativan prsten sa jedinicom $\boldsymbol{E}$.

Ovaj prsten ima delitelje nule, što znači da proizvod dve nenula-matrice može biti nula-matrica.

Primer:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Definicija 2.8. Kvadratna matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ je regularna (nesingularna, inverzibilna) ako postoji matrica $\boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ takva da je $\boldsymbol{AB} = \boldsymbol{E}$ i $\boldsymbol{BA} = \boldsymbol{E}$. U protivnom, matrica $\boldsymbol{A}$ je singularna (neregularna, neinverzibilna).

Ako postoji matrica B iz definicije 2.8. onda je ona jedinstveno određena.

Definicija 2.9. Jedinstvena matrica $\boldsymbol{B}$ za koju važi $\boldsymbol{AB} = \boldsymbol{E}$ i $\boldsymbol{BA} = \boldsymbol{E}$ se obeležava sa $\boldsymbol{A^{-1}}$ i naziva se inverzna matrica matrice $\boldsymbol{A}$. Skup $\boldsymbol{M}(\boldsymbol{n}, \mathbb{F})^*$ svih inverzibilnih tj. regularnih matrica čini grupu u odnosu na operaciju množenja matrica. Ta grupa se obeležava sa $\boldsymbol{GL}(\boldsymbol{n}, \mathbb{F})$ i naziva se opšta linearna grupa nad poljem $\mathbb{F}$.

Množenje matrica nam omogućava da skraćeno zapisujemo linearne sisteme jednačina. Posmatrajmo sistem linearnih jednačina

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = c_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = c_2$$

$$\cdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = c_m.$$

Uvedimo oznake:

$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ je matrica sistema,

$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ je kolona nepoznatih,

$c = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$ je kolona slobodnih članova.

Tada se linearni sistem može predstaviti u matričnom obliku $Ax = c$.

Neka su rešenja prethodnog linearnog sistema slobodni članovi drugog linearnog sistema

$$b_{11}y_1 + b_{12}y_2 + \cdots + b_{1k}y_k = x_1$$

$$b_{21}y_1 + b_{22}y_2 + \cdots + b_{2k}y_k = x_2$$

…

$$b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nk}y_k = x_n.$$

Odnosno $By = x$. Zamenom promenljivih x preko promenljivih y, dobijamo

$$a_{11}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1k}y_k) + \cdots + a_{1n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nk}y_k) = c_1$$

$$a_{21}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1k}y_k) + \cdots + a_{2n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nk}y_k) = c_2$$

…

$$a_{m1}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1k}y_k) + \cdots + a_{mn}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nk}y_k) = c_m.$$

Posle sređivanja, sistem izgleda ovako

$$(a_{11}b_{11} + \cdots + a_{1n}b_{n1})y_1 + \cdots + (a_{11}b_{1k} + \cdots + a_{1n}b_{nk})y_k = c_1$$

$$(a_{21}b_{11} + \cdots + a_{2n}b_{n1})y_1 + \cdots + (a_{21}b_{1k} + \cdots + a_{2n}b_{nk})y_k = c_2$$

…

$$(a_{m1}b_{11} + \cdots + a_{mn}b_{n1})y_1 + \cdots + (a_{m1}b_{1k} + \cdots + a_{mn}b_{nk})y_k = c_m.$$

Matrica

$$\begin{bmatrix} a_{11}b_{11}+\cdots+a_{1n}b_{n1} & a_{11}b_{12}+\cdots+a_{1n}b_{n2} & \cdots & a_{11}b_{1k}+\cdots+a_{1n}b_{nk} \\ a_{21}b_{11}+\cdots+a_{2n}b_{n1} & a_{21}b_{12}+\cdots+a_{2n}b_{n2} & \cdots & a_{21}b_{1k}+\cdots+a_{2n}b_{nk} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11}+\cdots+a_{mn}b_{n1} & a_{m1}b_{12}+\cdots+a_{mn}b_{n2} & \cdots & a_{m1}b_{1k}+\cdots+a_{mn}b_{nk} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{bmatrix}$$

u skraćenom zapisu $Ax = c$ i $By = x$, odnosno $Ax = A(By) = (AB)y = c$. Ovde vidimo da se množenje matrica, operacija koja na prvi pogled deluje veštački, zapravo dobija na osnovu rada sa sistemimam linearnih jednačina i da otuda ima veliku i neposrednu primenu.

Definicija 2.10. Neka je $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a_{11}} & \cdots & \boldsymbol{a_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{a_{m1}} & \cdots & \boldsymbol{a_{mn}} \end{bmatrix} \in \boldsymbol{M(m \text{ x } n}, \mathbb{F})$ matrica tipa $\boldsymbol{m \text{ x } n}$. Njoj transponovana matrica je $\boldsymbol{A^T} = \begin{bmatrix} \boldsymbol{a_{11}} & \cdots & \boldsymbol{a_{m1}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{a_{1n}} & \cdots & \boldsymbol{a_{mn}} \end{bmatrix} \in \boldsymbol{M(n \text{ x } m}, \mathbb{F})$ čije su vrste kolone matrice $\boldsymbol{A}$, a kolone su vrste matrice $\boldsymbol{A}$.

Stav 2.5. (Osobine transponovanja):

Ako su $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \in M(m \text{ x } n, \mathbb{F})$ matrice i $\lambda \in \mathbb{F}$ skalar, tada važi:

(1) $(A^T)^T = A$;
(2) AA^T je kvadratna matrica;
(3) $(A+B)^T = A^T + B^T$;
(4) $(\lambda A)^T = \lambda A^T$;
(5) $(AB)^T = B^T A^T$.

Dokaz:

(1) $(A^T)^T = \left(\begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} \right)^T = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = A$

(2) $A \in M(m \text{ x } n, \mathbb{F})$ i $A^T \in M(n \text{ x } m, \mathbb{F}) \Longrightarrow A\, A^T \in M(m \text{ x } m, \mathbb{F})$ odnosno ovaj proizvod je zaista kvadratna matrica.

(3) $(A+B)^T = \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}\right)^T =$

$\left(\begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}\right)^T = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{m1}+b_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n}+b_{1n} & \cdots & a_{mn}+b_{mn} \end{bmatrix} =$

$\begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{mn} \end{bmatrix} = A^T + B^T$

(4) $(\lambda A)^T = \left(\lambda \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\right)^T = \left(\begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}\right)^T = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{m1} \\ \vdots & \ddots & \vdots \\ \lambda a_{1n} & \cdots & \lambda a_{mn} \end{bmatrix} =$

$\lambda \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} = \lambda A^T$

(5) Prvo proverimo dimenzije.

$A \in M(m \text{ x } n, \mathbb{F})$ i $B \in M(n \text{ x } p, \mathbb{F}) \Longrightarrow AB \in M(m \text{ x } p, \mathbb{F}) \Longrightarrow (AB)^T \in M(p \text{ x } m, \mathbb{F})$

$B^T \in M(p \text{ x } n, \mathbb{F})$ i $A^T \in M(n \text{ x } m, \mathbb{F}) \Longrightarrow B^T A^T \in M(p \text{ x } m, \mathbb{F})$

Formati ove dve matrice su zaista isti.

$$AB = [c_{ij}] \quad i = 1, \dots, m \quad j = 1, \dots, p$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$(AB)^T = [c_{ji}] \quad j = 1, \dots, p \quad i = 1, \dots, m$$

$$c_{ji} = \sum_{k=1}^{n} b_{jk} a_{ki} \quad (1)$$

$$B^T A^T = [d_{ji}] \quad j = 1, \dots, p \quad i = 1, \dots, m$$

$$d_{ji} = \sum_{k=1}^{n} b_{jk} a_{ki} \quad (2)$$

Upoređivanjem jednakosti (1) i (2) zaključujemo da je $(AB)^T = B^T A^T$.■

Definicija 2.11. Neka je $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M(n, \mathbb{F})$. Trag kvadratne matrice A, u oznaci trA je suma elemenata na dijagonali matrice tj.

$$trA = \sum_{i=1}^{n} a_{ii}.$$

Definicija 2.12. Neka je $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M(n, \mathbb{F})$ simetrična kvadratna matrica. Matrica A je idempotentna ako je $A^2 = A$.

Definicija 2.13. Matrica dobijena od matrice $\boldsymbol{A} = [\boldsymbol{a_{ij}}] \in \boldsymbol{M}(\boldsymbol{m} \, \mathbf{x} \, \boldsymbol{n}, \mathbb{F})$ istovremenim izbacivanjem nekoliko vrsta i kolona naziva se podmatrica matrice $\boldsymbol{A}$**.** Ako je izbačeno $\boldsymbol{k}$ vrsta $(\boldsymbol{0} \leq \boldsymbol{k} \leq \boldsymbol{m})$ i $\boldsymbol{l}$ kolona $(\boldsymbol{0} \leq \boldsymbol{l} \leq \boldsymbol{n})$ dobijena je podmatrica tipa $(\boldsymbol{m} - \boldsymbol{k}) \, \mathbf{x} \, (\boldsymbol{n} - \boldsymbol{l})$. Podmatrica dobijena izbacivanjem $\boldsymbol{i}$ −te vrste i $\boldsymbol{j}$ −te kolone matrice $\boldsymbol{A}$ obeležava se sa $\boldsymbol{A_{(ij)}}$ i ona je tipa $(\boldsymbol{m} - \mathbf{1}) \, \mathbf{x} \, (\boldsymbol{n} - \mathbf{1})$.

Definicija 2.14. Ako su date matrice $\boldsymbol{A_{ij}}$ dimenzija $\boldsymbol{m_i} \, \mathbf{x} \, \boldsymbol{n_j}$ $(\boldsymbol{i} = \mathbf{1}, \ldots, \boldsymbol{p}, \ \boldsymbol{j} = \mathbf{1}, \ldots, \boldsymbol{q})$ onda od njih možemo sastaviti blok-matricu $\boldsymbol{A}$ dimenzija $\boldsymbol{m} \, \mathbf{x} \, \boldsymbol{n}$, gde je $\boldsymbol{m} = \boldsymbol{m_1} + \cdots + \boldsymbol{m_p}$ i $\boldsymbol{n} = \boldsymbol{n_1} + \cdots + \boldsymbol{n_q}$ i svaka matrica $\boldsymbol{A_{ij}}$ je na mestu $\boldsymbol{ij}$. Matrica $\boldsymbol{A}$ je oblika

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pq} \end{bmatrix}.$$

Matrica A se može shvatiti i kao matrica čiji su elementi podmatrice A_{ij} – njenih blokovi. Ukoliko su matrice A_{ij} kvadratne, onda je i matrica A kvadratna. Pri tome, ako su sve vandijagonalne matrice $A_{ij} = 0 \ (i \neq j)$, matrica A je blok-dijagonalna tj.

$$A = \begin{bmatrix} A_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{nn} \end{bmatrix}.$$

Ako su sve A_{ii} dijagonalne matrice sa elementima λ_i na dijagonali, onda je i matrica A dijagonalna matrica oblika

$$A = \begin{bmatrix} \lambda_1 & \ldots & 0 & & & & \\ \vdots & \ddots & \vdots & \cdots & & \mathbf{0} & \\ 0 & \ldots & \lambda_1 & & & & \\ & \vdots & & \ddots & & \vdots & \\ & & & & \lambda_n & \ldots & 0 \\ & \mathbf{0} & & \cdots & \vdots & \ddots & \vdots \\ & & & & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Ako su pri tome svi λ_i međusobno jednaki, onda je to skalarna matrica.

$$A = \begin{bmatrix} \begin{matrix} \lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda \end{matrix} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{matrix} \lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda \end{matrix} \end{bmatrix} = \lambda \begin{bmatrix} \begin{matrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{matrix} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \begin{matrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{matrix} \end{bmatrix} = \lambda E$$

Definicija 2.15. Neka su date blok-matrice $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{Q} \\ \boldsymbol{R} & \boldsymbol{S} \end{bmatrix}$ i $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{G} & \boldsymbol{H} \end{bmatrix}$ tako da su matrice $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S}, \boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}$ odgovarajućih dimenzija kako bi sledeće opercije bile definisane. Tada važi:

(1) Sabiranje blok-matrica $A + B = \begin{bmatrix} P+E & Q+F \\ R+G & S+H \end{bmatrix}$;

(2) Transponovanje blok-matrice $A^T = \begin{bmatrix} P^T & R^T \\ Q^T & S^T \end{bmatrix}$;

(3) Množenje matrica $AB = \begin{bmatrix} PE+QG & PF+QH \\ RE+SG & RF+SH \end{bmatrix}$.

2.2 Pozitivno definitne matrice, ortogonalne matrice, sopstvene vrednosti i sopstveni vrektori

Pored toga što su nam ovi pojmovi bitni za pojedine dokaze koji slede, oni imaju veliku primenu i u drugim oblastima statistike.

Za ispitivanje stacionarnosti vektorske vremenske serije koristimo karakteristične vrednosti matrice parametara.

Kod analize glavnih komponenata, među posmatranim slučajnim veličinama izdvajaju se one koje su bitnije tj. koje nose više informacija nego ostale. Te slučajne veličine nazivamo glavnim komponentama i njihove disperzije su sopstvene vrednosti matrice kovarijacije tih slučanih veličina.

Definicija 2.16. Simetrična matrica $\boldsymbol{B} = [\boldsymbol{b_{ij}}] \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{R})$ je nenegativno definitna (definisana) ako za bilo koji nenula-vektor $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta_1} \\ \vdots \\ \boldsymbol{\beta_n} \end{bmatrix} \in \mathbb{R}^{\boldsymbol{n}}$ važi sledeća nejednakost

$$\beta^T B \beta = \sum_{i,j=1}^{n} \beta_i \beta_j b_{ij} \geq 0.$$

Izraz

$$\beta^T B\beta = \sum_{i,j=1}^{n} \beta_i\beta_j b_{ij}$$

se naziva kvadratna forma matrice B.

Lema 2.6. Neka je $D = [d_{ij}] \in M(n,\mathbb{R})$ dijagonalna matrica. Matrica D je nenegativno definitna ako i samo ako je su dijagonalni elementi $d_{11}, \dots, d_{nn}$ nenegativni.

Dokaz: Pođimo od toga da je dijagonalna matrica $\boldsymbol{D}$ nenegativno definitna. Po definici to znači da za bilo koji nenula-vektor $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta_1} \\ \vdots \\ \boldsymbol{\beta_n} \end{bmatrix} \in \mathbb{R}^{\boldsymbol{n}}$ važi da je

$$\beta^T B\beta = \sum_{i,j=1}^{n} \beta_i\beta_j d_{ij} \geq 0.$$

$$\sum_{i,j=1}^{n} \beta_i\beta_j d_{ij} = \sum_{\substack{i,j=1 \\ i\neq j}}^{n} \beta_i\beta_j d_{ij} + \sum_{i=1}^{n} \beta_i\beta_i d_{ii} = 0 + \sum_{i=1}^{n} {\beta_i}^2 d_{ii} \geq 0$$

Kao je je ${\beta_i}^2 \geq 0$, zaključujemo da je i $d_{ii} \geq 0$ za svako $i = 1, \dots, n$.■

Lema 2.7. Za bilo koju matricu $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{R})$, važi da je $\boldsymbol{A^T A}$ nenegativno definitna matrica.

Dokaz: $A \in M(m \times n, \mathbb{R}) \Rightarrow A^T \in M(n \times m, \mathbb{R}) \Rightarrow A^T A \in M(n \times n, \mathbb{R})$ tako da je matrica kvadratna. Direktnom proverom može se zaključiti i da je simetrična. Treba još dokazati da je kvadratna forma nenegativna.

Za bilo koji nenula-vektor $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{R}^n$ važi sledeće

$$z^T(A^T A)z = (z^T A^T)(Az) = (Az)^T(Az).$$

Uvedimo oznaku $Az = x$. Kako je $A \in M(m \times n, \mathbb{R})$ i $z \in M(n \times 1, \mathbb{R})$, sledi da je $x \in M(m \times 1, \mathbb{R})$.

$$z^T(A^T A)z = x^T x = [x_1 \quad \cdots \quad x_m] \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = {x_1}^2 + \cdots {x_m}^2 \geq 0 ■$$

Ova lema će se koristiti u dokazu Gaus-Markove teoreme.

Definicija 2.17. Matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ je ortogonalna matrica nad poljem $\mathbb{F}$ ako važi jedno od ekvivalentnih svojstava:

$$AA^T = E;$$

$$A^T A = E;$$

$$A^T = A^{-1}.$$

Definicija 2.18. Skalar $\lambda \in \mathbb{F}$ je sopstvena vrednost matrice $A \in M(n, \mathbb{F})$ ako je postoji nenula vektor $v \in \mathbb{F}^n$ tako da je $Av = \lambda v$. Svaki nenula vektor za koji ovo važi se naziva sopstveni vektor matrice A koji odgovara sopstvenoj vrednosti λ.

Stav 2.8. Simetrična matrica je nenegativno definitna ako i samo ako su joj sve sopstvene vrednosti nenegativni brojevi.

Dokaz se može naći u literaturi [3] na strani 549.

Stav 2.9. Neka je $B \in M(n, \mathbb{F})$ simetrična matrica i neka su $d_1, \dots, d_n$ sopstvene vrednosti matrice B. Tada postoji ortogonalna matrica $Q \in M(n, \mathbb{F})$ tako da je $Q^T B Q = D$, gde je D dijagonalna matrica sa elementima $d_1, \dots, d_n$na dijagonali.

Dokaz se može naći u literaturi [3] na strani 541.

Sledeća iskaz je direktna posledica stavova 2.7. i 2.8.

Posledica 2.10. Za svaku nenegativno definitnu matricu $\boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ postoji ortogonalna matrica $\boldsymbol{Q} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$, takva da važi $\boldsymbol{Q}^{\boldsymbol{T}} \boldsymbol{B} \boldsymbol{Q} = \boldsymbol{D}$, gde je $\boldsymbol{D}$ dijagonalna matrica sa nenegativnim elementima na dijagonali.

2.3 Elementarne transformacije matrica

Definicija 2.19. Neka je $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{m} \,\mathbf{x}\, \boldsymbol{n}, \mathbb{F})$ matrica. Elementarna transformacija vrsta prvog tipa $\boldsymbol{p_{ij}}$ $(\boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{m}, \boldsymbol{i} \neq \boldsymbol{j})$ jeste zamena mesta $\boldsymbol{i}$ −te i $\boldsymbol{j}$ − te vrste matrice $\boldsymbol{A}$.

$$p_{ij}: A \to p_{ij}(A)$$

Elementarna transformacija vrsta drugog tipa $\boldsymbol{q_{ij}(\lambda)}$ $(\boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{m}, \boldsymbol{i} \neq \boldsymbol{j}, \boldsymbol{\lambda} \in \mathbb{F})$ jeste dodavanje $\boldsymbol{j}$ − toj vrsti matrice $\boldsymbol{A}$ $\boldsymbol{i}$ −tu vrstu matrice $\boldsymbol{A}$ pomnoženu brojem $\boldsymbol{\lambda}$.

$$q_{ij}(\lambda): A \to q_{ij}(\lambda)(A)$$

Elementarna transformacija vrsta trećeg tipa $\boldsymbol{r_i(\lambda)}$ $(\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{m}, \boldsymbol{\lambda} \in \mathbb{F} \backslash \{\mathbf{0}\})$ jeste množenje $\boldsymbol{i}$ −te vrste matrice $\boldsymbol{A}$ brojem $\boldsymbol{\lambda} \neq \mathbf{0}$.

$$r_i(\lambda): A \to r_i(\lambda)(A)$$

Ako je matrica $B \in M(m \,\mathrm{x}\, n, \mathbb{F})$ dobijena primenom konačno mnogo elementarnih transformacija vrsta matrice A, onda kažemo da je matrica B dobijena ekvivalentnim transformacijama vrsta matrice A i pišemo $A \sim_v B$.

Stav 2.11. Relacija $\sim_v$ je relacija ekvivalencije na skupu matrica $\boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$.

Dokaz se može naći u literaturi [5] na strani 76.

Definicija 2.20.

$$\boldsymbol{P_{ij}} = \begin{bmatrix} \mathbf{1} & & \cdots & & \mathbf{0} \\ & \mathbf{0} & \ldots & \mathbf{1} & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & \mathbf{1} & \ldots & \mathbf{0} & \\ \mathbf{0} & & \cdots & & \mathbf{1} \end{bmatrix} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$$ je elementarna matrica prvog tipa, jer su zamenjene $\boldsymbol{i}$ −ta i $\boldsymbol{j}$ − ta vrsta matrice $\boldsymbol{E}$.

$$Q_{ij}(\lambda) = \begin{bmatrix} 1 & & \cdots & & 0 \\ & 1 & \ldots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & \lambda & \ldots & 1 & \\ 0 & & \cdots & & 1 \end{bmatrix} \in M(n, \mathbb{F})$$ je elementarna matrica drugog tipa, jer je j − toj vrsti matrice Edodata i −ta vrsta matrice E pomnožena brojem $\lambda \neq 0$.

$$R_i(\lambda) = \begin{bmatrix} 1 & & \cdots & & 0 \\ & \lambda & \ldots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & 0 & \ldots & 1 & \\ 0 & & \cdots & & 1 \end{bmatrix} \in M(n, \mathbb{F})$$ je elementarna matrica trećeg tipa, jer je i −ta vrsta matrice E pomnožena brojem λ.

Stav 2.12. Elementarne transformacije vrsta matrice $\boldsymbol{A}$ dobijaju se množenjem matrice $\boldsymbol{A}$ elementarnim matricama s leve strane tj.

$p_{ij}(A) = P_{ij}A;$ $\qquad$ $q_{ij}(\lambda)(A) = Q_{ij}(\lambda)A;$ $\qquad$ $r_i(\lambda)(A) = R_i(\lambda)A.$

Dokaz: Neposrednim izračunavanjem.

Posledica 2.13. $A \sim_v B$ ako i samo ako je $B = PA$ gde je $P \in M(n, \mathbb{F})$ proizvod elementarnih matrica.

Stav 2.14. Sve elementarne matrice su regularne:

${P_{ij}}^{-1} = P_{ij};$ $\qquad$ $Q_{ij}(\lambda)^{-1} = Q_{ij}(-\lambda);$ $\qquad$ $R_i(\lambda)^{-1} = R_i\left(\frac{1}{\lambda}\right).$

Dokaz se može naći u literaturi [5] na strani 78.

Sve pojmove koje smo uveli za vrste mogu se potpuno analogno uvesti i za kolone, pa tako postoje elementarne transformacije kolona matrice $p^{ij}, q^{ij}(\lambda), r^i(\lambda)$ i ekvivalentna transformacija kolona matrice $\sim_k$.

$$p^{ij}: A \to p^{ij}(A)$$

$$q^{ij}(\lambda): A \to q^{ij}(\lambda)(A)$$

$$r^i(\lambda): A \to r^i(\lambda)(A)$$

Ako je matrica $B \in M(m \times n, \mathbb{F})$ dobijena primenom konačno mnogo elementarnih transformacija kolona matrice A, onda kažemo da je matrica B dobijena ekvivalentnim transformacijama kolona matrice A i pišemo $A \sim_k B$.

Stav 2.15. Relacija $\sim_k$ je relacija ekvivalencije na skupu matrica $\boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$.

Dokaz: Slično kao dokaz stava 2.8.

Stav 2.16. Elementarne transformacije kolona matrice $\boldsymbol{A}$ dobijaju se množenjem matrice $\boldsymbol{A}$ elementarnim matricama s desne strane tj.

$p^{ij}(A) = AP_{ij};$ $\quad q^{ij}(\lambda)(A) = AQ_{ij}(\lambda);$ $\quad r^i(\lambda)(A) = AR_i(\lambda).$

Dokaz: Neposrednim izračunavanjem.

Posledica 2.17. $A \sim_k B$ ako i samo ako je $B = AQ$ gde je $Q \in M(n, \mathbb{F})$ proizvod elementarnih matrica.

Definicija 2.21. Matrice $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$ su elementarno ekvivalentne ako se matrica $\boldsymbol{B}$ dobija od matrice $\boldsymbol{A}$ konačnim nizom elementarnih transformacija vrsta i/ili kolona. Tada pišemo $\boldsymbol{A} \sim \boldsymbol{B}$.

Stav 2.18. Relacija $\sim$ je relacija ekvivalencije na skupu $\boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$.

Dokaz: Slično kao dokaz stava 2.8.

Posledica 2.19. $A \sim B$ ako i samo ako je $B = PAQ$, gde su P i Q proizvodi elementarnih matrica.

2.4 Vektorski prostori

Definicija 2.22. Vektorski prostor nad poljem $\mathbb{F}$ ($\mathbb{F}$ −vektorski prostor) predstavlja skup $\boldsymbol{V}$ sa jednom binarnom operacijom $\boldsymbol{V} \times \boldsymbol{V} \to \boldsymbol{V}$ (sabiranje u aditivnoj oznaci $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} + \boldsymbol{v}$) i jednom spoljnom operaciojm $\mathbb{F} \times \boldsymbol{V} \to \boldsymbol{V}$ (množenje elementom polja $\mathbb{F}$, u oznaci $(\boldsymbol{\lambda}, \boldsymbol{u}) \mapsto \boldsymbol{\lambda u}$), pri čemu su zadovoljene sledeće osobine:

(B1) $(V, +)$ je komutativna grupa;

(B2) za sve $u, v \in V$ i sve $\lambda, \mu \in \mathbb{F}$ važi:

(B2.1) $\lambda(u + v) = \lambda u + \lambda v;$

(B2.2)$(\lambda + \mu)u = \lambda u + \mu u;$

(B2.3) $\lambda(\mu u) = (\lambda\mu)u$;

(B2.4) $1u = u$ 1 je jedinica $\mathbb{F}$.

Elementi iz V se nazivaju vektori, a elementi iz polja $\mathbb{F}$ skalari. Nula element $o \in V$ naziva se nula-vektor. Ako je $\mathbb{F} = \mathbb{R}$, onda govorimo o realnom vektorskom prostoru, a ako je $\mathbb{F} = \mathbb{C}$, o kompleksnom prostoru.

Primer: Na osnovu stava 2.1. i stava 2.2. skup matrica $M(m \,\mathrm{x}\, n, \mathbb{F})$ određenog tipa nad poljem $\mathbb{F}$ je $\mathbb{F}$ −vektorski prostor u odnosu na operacije sabiranje matrica i množenje matrica skalarom.

Primer: Specijalni slučaj prostora matrica je prostor $\mathbb{F}^n = M(1 \,\mathrm{x}\, n, \mathbb{F})$ matrica sa jednom vrstom i n kolona tj. prostor uređenih n −torki elemenata polja $\mathbb{F}$. To je $\mathbb{F}$ −vektorski prostor sa operacijama koje se eksplicitno zapisuju na sledeći način, ako su $x = (x_1, \dots, x_n)$ $y = (y_1, \dots, y_n)$ iz $\mathbb{F}^n$ i $\lambda \in \mathbb{F}$, tada je

$$x + y = (x_1 + y_1, \dots, x_n + y_n);$$

$$\lambda x = (\lambda x_1, \dots, \lambda x_n).$$

Ovo je standardni ili aritmetički prostor nad poljem $\mathbb{F}$. Oznaku $\mathbb{F}^n$ koristimo i za prostor kolona $M(n \,\mathrm{x}\, 1, \mathbb{F})$.

Definicija 2.23. Neka je V vektorski prostor nad poljem $\mathbb{F}$, gde $\mathbb{F}$ označava polje $\mathbb{R}$ ili polje $\mathbb{C}$. Funkcija $\|\cdot\|: V \to [0, \infty)$ koja ima osobine

(1) $\|u\| = 0 \Leftrightarrow u = 0$;
(2) $\|\lambda \cdot u\| = |\lambda|\|u\|$ za svaki skalar $\lambda \in \mathbb{F}$ i svaki vektor $u \in V$;
(3) $\|u + v\| \leq \|u\| + \|v\|$ za svaka dva vektora $u, v \in V$;

naziva se normom na prostoru V, a prostor V zajedno sa datom normom naziva se normiran vektorski prostor ili kraće normiran prostor.

2.5 Rang matrice

Stav 2.20. Svaka matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{m}\,\mathbf{x}\,\boldsymbol{n}, \mathbb{F})$ je ekvivalentna blok-matrici oblika $\boldsymbol{E}_r = \begin{bmatrix} \boldsymbol{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \boldsymbol{M}(\boldsymbol{m}\,\mathbf{x}\,\boldsymbol{n}, \mathbb{F})$ pri čemu je $\boldsymbol{E} = \begin{bmatrix} \mathbf{1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{1} \end{bmatrix} \in \boldsymbol{M}(\boldsymbol{r}, \mathbb{F})$. Drugim rečima, uvek postoje matrice $\boldsymbol{P}$ **i** $\boldsymbol{Q}$ tako da je $\boldsymbol{E}_r = \boldsymbol{PAQ}$.

Dokaz: Množenje matrice A matricama P I Q sa leve i desne strane je zapravo primena elementarnih transformacija na vrste i kolone matrice A. Pa tako primenom elementarnih transformacija na vrste matrice A, matricu možemo svesti na stepenasti oblik, pri čemu možemo smatrati da su svi glavni elementi tj. oni koji nam služe za eliminaciju jednaki 1. Zatim primenimo elementarne transformacije na kolone matrice A koristeći već izabrane glavne elemente dobijemo da su svi elementi sem glavnih jednaki 0, a glavni elementi jednaki 1. Na kraju permutacijom vrsta ili kolona možemo dovesti sve glavne elemente tj. sve jedinice na pozicije 11,22,... rr, čime smo dobili traženi oblik E_r.■

Definicija 2.24. Broj $\boldsymbol{r}$ je jedinstveno određen i naziva se rang matrice $\boldsymbol{A}$. Obeležavamo ga i sa $\boldsymbol{r = rangA}$.

Stav 2.21. Neka su matrice $\boldsymbol{A, B} \in \boldsymbol{M(m \text{ x } n}, \mathbb{F})$. Tada je $\boldsymbol{A \sim B}$ ako i samo ako je $\boldsymbol{rangA} = \boldsymbol{rangB}$.

Dokaz: Neka je $A \sim B$ i $rangA = r$. Iz činjenice da je $rangA = r$, važi da je $E_r \sim A$. $E_r \sim A$ i $A \sim B$, a relacija $\sim$ je tranzitivna, pa je i $E_r \sim B$. Što znači da je i $rangB = r$.

Neka je $rangA = rangB = r$. Iz činjenice da je $rangA = r$, važi da je $E_r \sim A$. Iz činjenice da je $rangB = r$, važi da je $E_r \sim B$. Zbog simetričnosti i tranzitivnosti relacije $\sim$, važi da je $A \sim B$.■

Skup klasa $M(m \text{ x } n, \mathbb{F})/_\sim$ je konačan, ima $k+1$ klasu ekvivalencije, po jednu za svako $r = 0,1, \dots k$ gde je $k = min\{m, n\}$. Sve matrice iste klase imaju isti rang i u svakoj klasi ekvivalencije postoji jedinstvena matrica oblika E_r koje se može uzeti za predstavnika klase.

Rang matrice se može izračunati svođenjem matrice na oblik E_r primenom elementarnih transformacija. Pri tome se ne mora ići do kraja, već se možemo zaustaviti na obliku

$$A \sim \begin{bmatrix} b_{11} & b_{12} & & \cdots & & b_{1n} \\ & & & \cdots & & \\ 0 & b_{22} & & \cdots & & \vdots \\ 0 & 0 & \ddots & \cdots & & \\ 0 & 0 & 0 & b_{rr} & \cdots & b_{rn} \end{bmatrix}.$$

Ako su svi $b_{ii} \neq 0$, onda je $rangA = r$.

Teorema 2.22. Za svaku kvadratnu matricu $\boldsymbol{A} \in \boldsymbol{M(n}, \mathbb{F})$ sledeća tvrđenja su ekvivalentna:

(1) A je proizvod elementarnih matrica;
(2) A je regularna matrica;
(3) Homogeni linerani sistem $Ax = 0$ ima samo trivijalno rešenje;
(4) Linearni sistem $Ax = b$ sa proizvoljnim slobodnim članom je određen;
(5) $A \sim E$;
(6) $A \sim_v E$;
(7) $A \sim_k E$.

Dokaz se može naći u literaturi [5] na strani 80.

Posledica 2.23. Kvadratna matrica $A \in M(n, \mathbb{F})$ ima rang n ako i samo ako je regularna.

Posledica 2.24. Množenje matrice regularnom matricom, bilo sleva ili zdesna, prevodi polaznu matricu u njoj ekvivalentu, te ne menja njen rang.

Rang matrice se može definisati i na drugi način.

Definicija 2.25. Neka je $\boldsymbol{V}$ $\mathbb{F}$ −vektorski prostor i neka su $\boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_k} \in \boldsymbol{V}$. Rang $\boldsymbol{r} = \boldsymbol{rang}\{\boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_k}\}$ ovog konačnog skupa vektora je maksimalan broj linerano nezavisnih vektora tog skupa. Očigledno je $\boldsymbol{0} \leq \boldsymbol{r} \leq \boldsymbol{k}$.

Definicija 2.26. Ako je $\boldsymbol{A} = [\boldsymbol{a_{ij}}] \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$ matrica i

$$a_i = (a_{i1}, a_{i2}, \dots, a_{in}) \; i = 1, \dots, m$$

su njene vrste posmatrane kao vektori u $\mathbb{F}^n$ i

$$a^j = (a_{1j}, a_{2j}, \dots, a_{mj})^T \; j = 1, \dots, n$$

su njene kolone posmatrane kao vektori u $\mathbb{F}^m$. Tada je $rang_v(A) = rang\{a_1, a_2, \dots, a_m\}$ rang vrsta matrice A i $rang_k(A) = rang\{a^1, a^2, \dots, a^n\}$ je rang kolona matrice A.

Stav 2.25. Za svaku matricu $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{F})$ je rang vrsta jednak rangu kolona.

Dokaz se može naći u literaturi [5] na strani 137.

Definicija 2.27. Jedinstven ceo broj određen ovom definicijom naziva se rang matrice i obeležava se sa $\boldsymbol{rangA}$.

Posledica 2.26. Ako je $A \in M(m \times n, \mathbb{F})$, tada je $0 \leq rangA \leq min\{m, n\}$.

Na prvi pogled nije jasno da su termini definisani u definicijama 2.24. i 2.27. isti brojevi. Sledeći stav govori upravo o tome.

Stav 2.27. Elementarne transformacije vrsta ne menjaju linearne veze između kolona, a elementarne transformacije kolona ne menjaju linearne veze između vrsta.

Dokaz se može naći u literaturi [5] na strani 138.

Posledica 2.28. Elementarne transformacije vrsta (kolona) ne menjaju rang kolona (vrsta) matrice, odnosno ne menjaju rang matrice. Elementarnim transformacijama vrsta (kolona) možemo dobiti

$$A \sim_v \begin{bmatrix} E & B \\ 0 & 0 \end{bmatrix}, rang_k(A) = r;$$

$$A \sim_k \begin{bmatrix} E & 0 \\ C & 0 \end{bmatrix}, rang_v(A) = r.$$

Teorema 2.29. (o baznom minoru): Rang matrice $\boldsymbol{A}$ jednak je maksimumu dimenzija njenih kvadratnih inverzibilnih podmatrica. Odnosno, $\boldsymbol{rangA} = \boldsymbol{r}$ ako i samo ako postoji inverzibilna podmatrica ranga $\boldsymbol{r}$ tj. formata $\boldsymbol{r} \times \boldsymbol{r}$ i nema inverzibilnih matrica većeg formata.

Dokaz se može naći u literaturi [5] na strani 139.

Ovo znači da se rang matrice može odrediti tako što nađemo regularnu podmatricu najvećeg ranga tj. minor koji je različit od nule.

Teorema 2.30. (Kroneker[2]-Kapelijeva[3] teorema): Linearni sistem

$$Ax = b$$

sa $A \in M(m \times n, \mathbb{F})$, $x \in M(n \times 1, \mathbb{F})$ i $b \in M(m \times 1, \mathbb{F})$ je saglasan ako i samo ako je $rangA = rang(A, b)$, gde je (A, b) proširena matrica sistema.

Dokaz se može naći u literaturi [5] na strani 140.

2.6 Linearna preslikavanja

Definicija 2.28. Neka su $\boldsymbol{V}$ i $\boldsymbol{W}$ $\mathbb{F}$ −vektorski prostori. Preslikavanje $\boldsymbol{f}: \boldsymbol{V} \longrightarrow \boldsymbol{W}$ je linearno preslikavanje $\mathbb{F}$ −vektorskih prostora (linearni operator, linearna transformacija) ako je

(1) $f(u + v) = f(u) + f(v)$ za svaka dva vektora $u, v \in V$ (aditivnost);

(2) $f(\lambda u) = \lambda f(u)$ za svaki vektor $u \in V$ i svaki scalar $\lambda \in \mathbb{F}$ (homogenost).

Ova dva svojstva iz definicije mogu se zameniti jednim svojstvom koje se naziva linearnost. Odnosno, preslikavanje $f: V \longrightarrow W$ je $\mathbb{F}$ −linearno ako i samo ako za svaka dva vektora $u, v \in V$ i za svaka dva skalara $\lambda, \mu \in \mathbb{F}$ važi

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Ako je $W = \mathbb{F}$ onda je u pitanju linearni funkcional $f^*: V \longrightarrow \mathbb{F}$.

Primer: Neka su $V = M(1 \times n, \mathbb{F})$ i $W = M(1 \times m, \mathbb{F})$ matrice koje nad poljem $\mathbb{F}$ koje imaju samo 1 vrstu i n, odnosno m kolona. Ovi skupovi su ujedno i vektorski prostori na osnovu primera iz poglavlja 2.4. Data je i matrica $A = [a_{ij}] \in M(m \times n, \mathbb{F})$. Za svaki vektor $x = (x_1, \dots, x_n) \in V$ definišemo vektor $f(x) = y = (y_1, \dots, y_m) \in W$ na sledeći način

$$y_1 = a_{11}x_1 + \dots + a_{1n}x_n$$

$$\dots$$

[2] Leopold Kronecker (1823-1891), nemački matematičar
[3] Alfredo Capelli (1855-1910), italijanski matematičar

$$y_m = a_{m1}x_1 + \cdots + a_{mn}x_n.$$

Tako dobijamo preslikavanje $f: V \to W$ određeno matricom A, koje ćemo obeležavati sa f_A. Proverimo da je ovo preslikavanje linearno tj. da za svaka dva vektora $x, x' \in V$ i svaka dva skalara $\lambda, \lambda' \in \mathbb{F}$ važi sledeće

$$f(\lambda x + \lambda' x') = \lambda f(x) + \lambda' f(x') = \lambda y + \lambda' y'$$

gde je $f(x') = y' = \left({y'}_1, \dots, {y'}_m\right)$.

$$\begin{aligned} f(\lambda x + \lambda' x') &= f\left(\lambda(x_1, \dots, x_n) + \lambda'({x'}_1, \dots, {x'}_n)\right) = f\left((\lambda x_1, \dots, \lambda x_n) + (\lambda' {x'}_1, \dots, \lambda' {x'}_n)\right) \\ &= f(\lambda x_1 + \lambda' {x'}_1, \dots, \lambda x_n + \lambda' {x'}_n) = (z_1, \dots, z_m) = z \end{aligned}$$

Pri čemu je za svako $(i = 1, \dots, m)$ važi

$$\begin{aligned} z_i &= a_{i1}(\lambda x_1 + \lambda' {x'}_1) + \cdots + a_{in}(\lambda x_n + \lambda' {x'}_n) \\ &= \lambda(a_{i1}x_1 + \cdots + a_{in}x_n) + \lambda'(a_{i1}{x'}_1 + \cdots + a_{in}{x'}_n) = \lambda y_i + \lambda' {y'}_i. \end{aligned}$$

Tako da je zaista $f(\lambda x + \lambda' x') = \lambda y + \lambda' y'$.

Ako umesto vrsta posmatramo kolone tj. ako je $V = M(n \times 1, \mathbb{F})$ i $W = M(m \times 1, \mathbb{F})$, tada se formule

$$y_1 = a_{11}x_1 + \cdots + a_{1n}x_n$$

$$\dots$$

$$y_m = a_{m1}x_1 + \cdots + a_{mn}x_n$$

mogu napisati u obliku

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Odnosno,

$$y = Ax.$$

Svaka matrica $A \in M(m \times n, \mathbb{F})$ definiše jedno linearno preslikavanje vektorskih prostora $f_A: M(n \times 1, \mathbb{F}) \to M(m \times 1, \mathbb{F})$ formulom

$$f_A = Ax.$$

Pokažimo i obrnuti smer.

Svako linearno preslikavanje $L: M(n \times 1, \mathbb{F}) \to M(m \times 1, \mathbb{F})$ može se zapisati u obliku

$$L(x) = Ax$$

sa jedinstveno određenom matricom $A \in M(m \times n, \mathbb{F})$. Kolone tražene matrice A moraju biti slike standardnih baznih vektora tj.

$$L(e_i) = Ae_i = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}.$$

Bilo koji vektor $x \in M(n \times 1, \mathbb{F})$ može se zapisati kao linearna kombinacija vektora iz baze tj.

$$x = x_1 e_1 + \cdots + x_n e_n$$

Kada na njega primenimo linearno preslikavanje, dobijamo

$$\begin{aligned} L(x) = L(x_1 e_1 + \cdots + x_n e_n) = x_1 L(e_1) + \cdots + x_n L(e_n) = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ = \begin{bmatrix} a_{11}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} \\ = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax. \end{aligned}$$

Definicija 2.29. Linearno preslikavanje $\boldsymbol{f}: \boldsymbol{V} \longrightarrow \boldsymbol{W}$ je ograničeno ako postoji broj $\boldsymbol{M} > \boldsymbol{0}$ tako da za svako $\boldsymbol{u} \in \boldsymbol{V}$ važi

$$\|\boldsymbol{f}\boldsymbol{u}\| \leq \boldsymbol{M}\|\boldsymbol{u}\|.$$

Infimum brojeva $\boldsymbol{M}$ za koje važi prethodna nejednakost naziva se normom linearnog preslikavanja $\boldsymbol{f}$ i označava se sa $\|\boldsymbol{f}\|$.

Skup svih ograničenih linearnih preslikavanja $\boldsymbol{f}: \boldsymbol{V} \longrightarrow \boldsymbol{W}$ obrazuje vektorski prostor koji obeležavamo sa $\boldsymbol{\mathcal{B}}(\boldsymbol{V}, \boldsymbol{W})$.

Stav 2.31. Ako je $\boldsymbol{f}$ linearno preslikavanje, tada važi

$$\|\boldsymbol{f}\| = \sup_{\boldsymbol{u} \neq \boldsymbol{0}} \frac{\|\boldsymbol{f}\boldsymbol{u}\|}{\|\boldsymbol{u}\|} = \sup_{\|\boldsymbol{u}\| \leq \boldsymbol{1}} \|\boldsymbol{f}\boldsymbol{u}\| = \sup_{\|\boldsymbol{u}\| = \boldsymbol{1}} \|\boldsymbol{f}\boldsymbol{u}\|.$$

Dokaz: Kako je $\|\boldsymbol{f}\|$ infimum brojeva $\boldsymbol{M}$ za koje važi nejednakost $\frac{\|\boldsymbol{f}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \boldsymbol{M}$, to je onda $\|\boldsymbol{f}\|$ supremum brojeva na levoj strani nejednakosti, odnosno

$$\|\boldsymbol{f}\| = \sup_{\boldsymbol{u} \neq \boldsymbol{0}} \frac{\|\boldsymbol{f}\boldsymbol{u}\|}{\|\boldsymbol{u}\|}.$$

Zbog homogenosti je $\frac{\|\boldsymbol{f}\boldsymbol{u}\|}{\|\boldsymbol{u}\|} = \left\|\frac{\boldsymbol{f}\boldsymbol{u}}{\|\boldsymbol{u}\|}\right\| = \left\|\boldsymbol{f}\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\right\|$, pa je $\|\boldsymbol{f}\| = \mathbf{sup}_{\boldsymbol{u} \neq \boldsymbol{0}} \left\|\boldsymbol{f}\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\right\|$.

Kako $\frac{u}{\|u\|}$ ima normu 1 za sve vrednosti $u \neq 0$, važi da je $\|f\| = \sup_{\|u\|=1}\|fu\|$.

$$\|f\| = \sup_{\|u\|=1}\|fu\| \leq \sup_{\|u\|\leq 1}\|fu\|$$

Ova nejednakost važi jer je supremum po manjem skupu manji ili jednak supremumu po većem skupu. Sa druge strane, iz definicije ograničenosti linearnog preslikavanja važi

$$\|fu\| \leq \|f\|\|u\|.$$

A odatle i $\sup_{\|u\|\leq 1}\|fu\| \leq \|f\|$. Na kraju, možemo zaključiti da je

$$\|f\| = \sup_{\|u\|=1}\|fu\| = \sup_{\|u\|\leq 1}\|fu\|. \blacksquare$$

Definicija 2.30. Funkcija $f: X \longrightarrow Y$ između metričkih prostora (X, d) i (Y, ρ) je neprekidna u tački $a \in X$ ako za svako $\varepsilon > 0$ postoji $\delta > 0$ tako tako da je $\rho(f(x), f(a)) < \varepsilon$ uvek kada je $d(x, a) < \delta$. Funkcija je neprekidna na skupu X ako je neprekidna u svakoj tački tog skupa.

Definicija 2.31. Funkcija $f: X \longrightarrow Y$ između metričkih prostora (X, d) i (Y, ρ) je ravnomerno (uniformno) neprekidna ako za svako $\varepsilon > 0$ postoji $\delta > 0$ tako tako da je $\rho(f(x), f(y)) < \varepsilon$ uvek kada je $d(x, y) < \delta$.

Definicija 2.32. Funkcija $f: X \longrightarrow Y$ između metričkih prostora (X, d) i (Y, ρ) je Lipšic[4] neprekidna ako postoji konstanta L takva da je $\rho(f(x), f(y)) \leq Ld(x, y)$ za sve $x, y \in X$.

Svaka Lipšic neprekidna funkcija je i ravnomerno neprekidna, a svaka ravnomerno neprekidna funkcija je i neprekidna.

Stav 2.32. Linearno preslikavanje $f: V \longrightarrow W$ je neprekidno ako i samo ako je ograničeno.

Dokaz: Neka je f ograničeno preslikavanje. Treba dokazati da je neprekidno.

$$\|fu_n - fu\| = \|f(u_n - u)\| \leq \|f\|\|u_n - u\|$$

Vidimo da je f Lipšic neprekidno preslikavanje, sa konstantom $\|f\|$. To znači da je i neprekidno.

Neka je f neprekidno preslikavanje. Treba dokazati da je ograničeno. Pretpostavimo suprotno, tj. pretpostavimo da nije ograničeno preslikavanje. To na osnovu stava 2.21. znači da je $\sup_{\|u\|=1}\|fu\| = +\infty$. Odatle sledi da postoji niz $(u_n)_{n=1}^{\infty}$ takav da je $\|u_n\| = 1$ i $\|fu_n\| = c_n \to +\infty$. Uočimo drugi niz, $v_n = \frac{u_n}{c_n}$ za koji onda važi $\|v_n\| = \frac{1}{c_n} \to 0$ i $\|fv_n\| = \frac{c_n}{c_n} = 1$ tj. f nije neprekidno preslikavanje u tački $u = 0$, pa samim tim ni u jednoj tački skupa V, a to je kontradikcija. $\blacksquare$

[4]Rudolph Otto Sigismund Lipschity (1832-1903)

Definicija 2.33. Za neprekidno linearno preslikavanje $\boldsymbol{f}: \boldsymbol{V} \longrightarrow \boldsymbol{W}$ normiranih vektorskih prostora $\boldsymbol{V}$ i $\boldsymbol{W}$ kaže se da je odozdo ograničeno ako postoji konstanta $\boldsymbol{c} > \boldsymbol{0}$ tako da je $\|\boldsymbol{fu}\| \geq \boldsymbol{c}\|\boldsymbol{u}\|$ za svako $\boldsymbol{u} \in \boldsymbol{V}$.

Stav 2.33. Linearno preslikavanje $\boldsymbol{f}$ koje preslikava normiran vektorski prostor $\boldsymbol{V}$ na normiran vektorski prostor $\boldsymbol{W}$ ima inverz koji je linearan i ograničen ako i samo ako je ograničen odozgo.

Dokaz: Pretpostavimo da je linearno preslikavanje $\boldsymbol{f}$ ograničeno odozdo. Treba pokazati da postoji inverzno preslikavanje $\boldsymbol{f}^{-1}$ koje je linearno i ograničeno. Kako je preslikavanje $\boldsymbol{f}$ "na", onda će inverz postojti ako je $\boldsymbol{f}$ i "1-1".

$$\boldsymbol{fu} = \boldsymbol{fv}$$

$$\boldsymbol{fu} - \boldsymbol{fv} = \boldsymbol{0}$$

$$\|\boldsymbol{fu} - \boldsymbol{fv}\| = \boldsymbol{0}$$

$$\|\boldsymbol{f}(\boldsymbol{u} - \boldsymbol{v})\| = \boldsymbol{0}$$

Zbog ograničenosti preslikavanja $\boldsymbol{f}$, važi $\|\boldsymbol{f}(\boldsymbol{u} - \boldsymbol{v})\| \geq \boldsymbol{c}\|\boldsymbol{u} - \boldsymbol{v}\|$.

Tako da je $\boldsymbol{c}\|\boldsymbol{u} - \boldsymbol{v}\| \leq \|\boldsymbol{f}(\boldsymbol{u} - \boldsymbol{v})\| = \boldsymbol{0}$, odnosno $\|\boldsymbol{u} - \boldsymbol{v}\| = \boldsymbol{0}$. Odatle sledi da je $\boldsymbol{u} = \boldsymbol{v}$, a samim tim i da je preslikavanje $\boldsymbol{f}$ "1-1".

Neka je $\boldsymbol{w} \in \boldsymbol{W}$. Tada je $\|\boldsymbol{f}(\boldsymbol{f}^{-1}\boldsymbol{u})\| \geq \boldsymbol{c}\|\boldsymbol{f}^{-1}\boldsymbol{u}\|$ jer je preslikavanje $\boldsymbol{f}$ ograničeno odozdo.

$$\|\boldsymbol{f}(\boldsymbol{f}^{-1}\boldsymbol{u})\| \geq \boldsymbol{c}\|\boldsymbol{f}^{-1}\boldsymbol{u}\|$$

$$\|\boldsymbol{u}\| \geq \boldsymbol{c}\|\boldsymbol{f}^{-1}\boldsymbol{u}\|$$

$$\frac{\boldsymbol{1}}{\boldsymbol{c}}\|\boldsymbol{u}\| \geq \|\boldsymbol{f}^{-1}\boldsymbol{u}\|$$

Odnosno preslikavanje $\boldsymbol{f}^{-1}$ je ograničeno i $\|\boldsymbol{f}^{-1}\| \leq \frac{\boldsymbol{1}}{\boldsymbol{c}}$.

Pretpostavimo da preslikavanje $\boldsymbol{f}$ ima ograničen inverz. Treba dokazati da je preslikavanje $\boldsymbol{f}$ ograničeno odozdo.

$\|\boldsymbol{u}\| = \|\boldsymbol{f}^{-1}\boldsymbol{fu}\| \leq \|\boldsymbol{f}^{-1}\|\|\boldsymbol{fu}\|$ jer je $\boldsymbol{f}^{-1}$ ograničeno preslikavanje. Odavde sledi da je $\|\boldsymbol{fu}\| \geq \frac{\boldsymbol{1}}{\|\boldsymbol{f}^{-1}\|}\|\boldsymbol{u}\|$. Ako izaberemo da je $\frac{\boldsymbol{1}}{\|\boldsymbol{f}^{-1}\|} = \boldsymbol{c}$, onda prethodna nejednakost postaje $\|\boldsymbol{fu}\| \geq \boldsymbol{c}\|\boldsymbol{u}\|$ tj. preslikavanje $\boldsymbol{f}$ je odozdo ograničeno. ■

Definicija 2.34. Preslikavanje $\boldsymbol{f}: \boldsymbol{U} \times \boldsymbol{V} \longrightarrow \boldsymbol{W}$ je bilinearno ako je linearno po obe promenljive tj. ako za svako $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{U}$, svako $\boldsymbol{w}, \boldsymbol{y} \in \boldsymbol{V}$ i svako $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{F}$ važi

(1) $\boldsymbol{f}(\boldsymbol{\lambda u} + \boldsymbol{\mu v}, \boldsymbol{w}) = \boldsymbol{\lambda f}(\boldsymbol{u}, \boldsymbol{w}) + \boldsymbol{\mu f}(\boldsymbol{v}, \boldsymbol{w})$;

(2) $\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{\lambda w} + \boldsymbol{\mu y}) = \boldsymbol{\lambda f}(\boldsymbol{u}, \boldsymbol{w}) + \boldsymbol{\mu f}(\boldsymbol{u}, \boldsymbol{y})$.

Definicija 2.35. Bilinearno preslikavanje $\boldsymbol{f}: \boldsymbol{U} \times \boldsymbol{V} \longrightarrow \boldsymbol{W}$ je ograničeno ako postoji broj $\boldsymbol{M} > \boldsymbol{0}$ tako da za svako $\boldsymbol{u} \in \boldsymbol{U}$ i svako $\boldsymbol{v} \in \boldsymbol{V}$ važi

$$\|\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})\| \leq \boldsymbol{M}\|\boldsymbol{u}\|\|\boldsymbol{v}\|.$$

Infimum brojeva $\boldsymbol{M}$ za koje važi prethodna nejednakost naziva se normom bilinearnog preslikavanja $\boldsymbol{f}$ i označava se sa $\|\boldsymbol{f}\|$.

Skup svih ograničenih bilinearnih preslikavanja $\boldsymbol{f}: \boldsymbol{U} \times \boldsymbol{V} \longrightarrow \boldsymbol{W}$ obrazuje vektorski prostor koji obeležavamo sa $\boldsymbol{\mathcal{B}}(\boldsymbol{U} \times \boldsymbol{V}, \boldsymbol{W})$.

Definicija 2.36. Dva normirana prostora $\boldsymbol{V}$ i $\boldsymbol{W}$ su izomorfna ako postoji neprekidno linearno preslikavanje $\boldsymbol{f}: \boldsymbol{V} \rightarrow \boldsymbol{W}$ koje je "1-1" i "na" i čiji je inverz $\boldsymbol{f}^{-1}: \boldsymbol{W} \rightarrow \boldsymbol{V}$ neprekidno preslikavanje.

Ako pri tome još važi i da je $\|\boldsymbol{fu}\| = \|\boldsymbol{u}\|$ za svako $\boldsymbol{u} \in \boldsymbol{V}$ onda je preslikavanje $\boldsymbol{f}$ izometrički izomorfizam, a prostori $\boldsymbol{V}$ i $\boldsymbol{W}$ su izometrički izomorfni.

Definicija 2.37. Kvadratni funkcional bilinearnog preslikavanja $f: V^2 \longrightarrow W$ je jako pozitivan ako postoji $c > 0$ tako da je

$$f(u, u) \geq c\|u\|^2$$

za svako $\boldsymbol{u} \in \boldsymbol{V}$.

3 Determinante

Definicija 3.1. Ako je $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{bmatrix} \in \boldsymbol{M}(\boldsymbol{2}, \mathbb{F})$ kvadratna matrica reda 2, njena determinanta (reda 2) je broj $\boldsymbol{detA} = \begin{vmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{vmatrix} = \boldsymbol{ad} - \boldsymbol{bc} \in \mathbb{F}$.

Definicija 3.2. Ako je $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{12}} & \boldsymbol{a_{13}} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{32}} & \boldsymbol{a_{33}} \end{bmatrix} \in \boldsymbol{M}(\boldsymbol{3}, \mathbb{F})$ kvadratna matrica reda 3, njena determinanta (reda 3) je broj

$$detA = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \in \mathbb{F}.$$

Ovde vidimo da se determinanta trećeg reda izražava preko determinante drugog reda. Ovo se naziva razvoj determinante po prvoj koloni.

$$detA = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}$$

Definicija 3.3. Neka je $\boldsymbol{A} = [\boldsymbol{a_{ij}}] \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ kvadratna matrica reda $\boldsymbol{n}$, i neka su $\boldsymbol{A_{(11)}}, \boldsymbol{A_{(21)}}, \dots, \boldsymbol{A_{(n1)}}$ podmatrice matrice $\boldsymbol{A}$ dobijene izbacivanjem prve kolone i prve, druge, … , $\boldsymbol{n} -$ te vrste respektivno. Tada je determinanta matrice $\boldsymbol{A}$ ($\boldsymbol{n}$ −tog reda) jednaka

$$detA = a_{11}detA_{(11)} - a_{21}detA_{(21)} + \cdots + (-1)^{n+1}a_{n1}detA_{(n1)} \in \mathbb{F}.$$

Za $n = 1$ biće $detA = a_{11}$. Broj $A_{ij} = (-1)^{i+j}detA_{(ij)} \in \mathbb{F}$ naziva se algebarski komplement ili algebarski kofaktor elementa a_{ij} i to je determinanta za 1 manjeg reda. Sa ovako uvedenim oznakama, formula ima oblik

$$detA = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}.$$

Ovo je izračunavanje determinante n −tog reda razvojem po prvoj koloni.

Izračunavanje determinante n −tog reda je preslikavanje tzv. determinantno preslikavanje koje svakoj kvadratnoj matrici reda n dodeljuje neki broj iz polja $\mathbb{F}$.

$$det: M(n, \mathbb{F}) \to \mathbb{F}: A \longmapsto detA$$

Ako je $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$, tada je $detA = a_{11}a_{22} \dots a_{nn}$ tj. determinanta gornje trougaone matrice jednaka je proizvodu elemenata na dijagonali.

$$detA = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} + 0,$$

jer su svi $a_{i1} = 0$ $(i = 2, \dots, n)$. Ako nastavimo dalje ovaj postupak, dobijamo $detA = a_{11}a_{22} \dots a_{nn}$.

Vidimo da je izračunavanje determinante trougaone matrice jako jednostavno. Elementarnim transformacijama vrsta možemo svaku matricu svesti na trougaoni oblik, pa je bitno utvrditi koje su osobine determinanti prilikom elementarnih transformacija vrsta.

Stav 3.1. (Osobine determinanti): Funkcija $\boldsymbol{det}: \boldsymbol{M}(\boldsymbol{n}, \mathbb{F}) \to \mathbb{F}$ može se posmatrati kao funkcija $\boldsymbol{n}$ promenljivih vrsta i ima sledeća svojstva:

(1) aditivnost : k −ta vrsta je zbir dve vrste

$$det[\cdots \quad a+b \quad \cdots] = det[\cdots \quad a \quad \cdots] + det[\cdots \quad b \quad \cdots];$$

(2) homogenost : k −ta vrsta je proizvod skalara i vrste

$$det[\cdots \quad \lambda a \quad \cdots] = \lambda det[\cdots \quad a \quad \cdots];$$

(3) antisimetričnost :

$$det[\cdots \quad a \quad \cdots \quad b \quad \cdots] = -det[\cdots \quad b \quad \cdots \quad a \quad \cdots];$$

(4) normiranost :

$$detE = \begin{vmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{vmatrix} = 1.$$

(5) Linearnost ⇔ aditivnost + homogenost tj.

$$det[\cdots \quad \lambda b + \mu c \quad \cdots] = \lambda det[\cdots \quad b \quad \cdots] + \mu det[\cdots \quad c \quad \cdots].$$

Determinanta linearna funkcija svih svojih vrsta.

Dokaz: Dokaze za aditivnost i homogenost objedinićemo u dokazu za linearnost. Neka je $k-$ta vrsta matrice A reda n, oblika $a_k = \lambda b + \mu c$, gde su $\lambda, \mu \in \mathbb{F}$. Ovo svojstvo dokazaćemo indukcijom po n.

Za $n = 1$, matrica se svodi na broj iz polja $\mathbb{F}$, pa linearnost važi na osnovu osobina polja $\mathbb{F}$.

Pretpostavimo da tvrđenje važi za $n - 1$.

Dokazujemo da važi za n. Razvijamo po prvoj koloni determinantu reda n matrice A. Po definiciji dobijamo

$$detA = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}.$$

Uočimo k −tu vrstu i u sumi izdvojimo$k$ −ti sabirak

$$detA = \sum_{i\neq k} a_{i1}A_{i1} + a_{k1}A_{k1}.$$

Svi kofaktori prvog dela zbira sadrže k −tu vrstu bez njenog prvog elementa, a kofaktor A_{k1} ne sadrži k −tu vrstu uopšte. k −ta vrsta matrice A je oblika $a_k = \lambda b + \mu c$. Uvedimo matrice B i C koje su iste matrici A, osim što je k −ta vrsta matrice B jednaka b i k −ta vrsta matrice C jednaka c. Tada je u svakom od kofaktora A_{i1} $i \neq k$ jedna vrsta jednaka $\lambda b + \mu c$ jer se ti kofaktori dobijaju precrtavanjem prve kolone i ne k −te vrste. Na osnovu indukcijske pretpostavke je $A_{i1} = \lambda B_{i1} + \mu C_{i1}$ $i \neq k$, dok je $A_{k1} = B_{k1} = C_{k1}$ jer su ove matrice izuzimajući k −tu vrstu iste.

$$\begin{aligned} detA &= \sum_{i\neq k} a_{i1}A_{i1} + a_{k1}A_{k1} = \sum_{i\neq k} a_{i1}(\lambda B_{i1} + \mu C_{i1}) + (\lambda b_{k1} + \mu c_{k1})A_{k1} \\ &= \sum_{i\neq k} a_{i1}\lambda B_{i1} + \lambda b_{k1}A_{k1} + \sum_{i\neq k} a_{i1}\mu C_{i1} + \mu c_{k1}A_{k1} \\ &= \lambda\left(\sum_{i\neq k} a_{i1}B_{i1} + b_{k1}B_{k1}\right) + \mu\left(\sum_{i\neq k} a_{i1}C_{i1} + c_{k1}C_{k1}\right) = \lambda detB + \mu detC \end{aligned}$$

Time smo dokazali linearnost funkcije det. Za $\lambda = \mu = 1$ izvodi se osobina aditivnosti, a za $\lambda \in \mathbb{F}$ i $\mu = 0$, dobija se osobina homogenosti.

Dokažimo sada antisimetričnost. Dovoljno je dokazati za parove susednih elemenata. Neka je matrica B dobijena od matrice A, permutacijom k − te i $k+1$ − e vrste. Prilikom te permutacije u svim kofaktorima A_{i1} gde je $i \neq k, k+1$ dolazi do permutacije te dve vrste, pa je po indukcijskoj pretpostavci $B_{i1} = -A_{i1}$. Za $i = k, k+1$ je

$$B_{k1} = -A_{(k+1)1};$$

$$B_{(k+1)1} = -A_{k1}.$$

$$\begin{aligned} detA &= \sum_{i\neq k,k+1} a_{i1}A_{i1} + a_{k1}A_{k1} + a_{(k+1)1}A_{(k+1)1} \\ &= -\sum_{i\neq k,k+1} b_{i1}B_{i1} - b_{(k+1)1}B_{(k+1)1} - b_{k1}B_{k1} \\ &= -\left(\sum_{i\neq k,k+1} b_{i1}B_{i1} + b_{k1}B_{k1} + b_{(k+1)1}B_{(k+1)1}\right) = -detB \end{aligned}$$

Kod permutacija bilo koje 2 vrste potrebno je napraviti neparan broj permutacija susednih vrsta. Znak se menja pri svakoj takvoj permutaciji, a kako ih ima neparan broj, važi da će na kraju biti $detA = -detB$. Time je svojstvo antisimetričnosti u potpunosti dokazano.

Osobina normiranosti je očigledna na osnovu primera o determinanti gornje trougaone matrice. Kako su svi elementi na dijagonali jedinične matrice jednaki 1, sledi da je $detE = 1$.

Stav 3.2.

(1) $detp_{ij}(A) = -detA$
(2) $detr_i(\lambda)(A) = \lambda detA$
(3) Determinanta sa vrstom ili kolonom 0 je jednaka 0.
(4) Determinanta sa dve iste vrste ili kolone je jednaka 0.
(5) $detq_{ij}(\lambda)(A) = detA$

Dokaz:

(1) Na osnovu osobine antisimetričnosti.

(2) Na osnovu osobine homogenosti.

(3) Ako je $i-$ta vrsta matrice A jednaka 0, onda se matrica A može napisati u obliku $A = r_i(0)(A)$. Tada je

$$detA = det\big(r_i(0)(A)\big) = 0detA = 0.$$

(4) Ako su $i-$ta i $j-$ta vrsta matrice A jednake, onda je $p_{ij}(A) = A$. $detA = detp_{ij}(A)$ i na osnovu osobina antisimetričnosti važi $detp_{ij}(A) = -detA$, tako da je $detA = -detA$, odnosno $detA = 0$.

(5) $detq_{ij}(\lambda)(A) = \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_j + \lambda a_i \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{vmatrix} + \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ \lambda a_i \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{vmatrix} + \lambda \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{vmatrix} + \lambda 0 = \begin{vmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{vmatrix} =$

$detA$ ■

Teorema 3.3. (Jedinstvenost determinante) Ako je $\boldsymbol{F}: \boldsymbol{M}(\boldsymbol{n}, \mathbb{F}) \to \mathbb{F}$ funkcija matrice, odnosno funkcija $\boldsymbol{n}$ vrsta matrice, koja je linearna i antisimetrična, tada je $\boldsymbol{F}(\boldsymbol{A}) = \boldsymbol{F}(\boldsymbol{E})\boldsymbol{detA}$. Ako je pri tome funkcija $\boldsymbol{F}$ još i normirana tj. $\boldsymbol{F}(\boldsymbol{E}) = \mathbf{1}$, tada je $\boldsymbol{F} \equiv \boldsymbol{det}$.

Dokaz: $F: M(n, \mathbb{F}) \to \mathbb{F}$ i F je linearna i antisimetrična funkcija vrsta matrice. Treba dokazati da je $F(A) = F(E)detA$.

Elementarnim transformacijama vrsta prvog i drugog tipa matricu A možemo svesti na trougaonu matricu T. k je broj puta koliko smo izvršili elementarnu transformaciju prvog tipa. Zbog antisimetričnosti funkcije F pojavljuje se koeficjent $(-1)^k$. Elementarne transformacije drugog tipa ne utiču na funkciju F. Tako da je

$F(A) = (-1)^k F(T).$ (1)

Takođe,

$det(A) = (-1)^k det(T) = (-1)^k a_{11}a_{22} \dots a_{nn}.$ (2)

F je linearna i antisimetrična funkcija vrsta matrice pa se na osnovu toga korišćenjem Gausovog metoda može zaključiti da je

$F(T) = a_{11}a_{22} \dots a_{nn}F(E).$ (3)

Kada ovo zamenimo u jednačinu (1), dobijamo

$F(A) = (-1)^k a_{11}a_{22} \dots a_{nn}F(E).$ (4)

Upoređivanjem jednačina (2) i (4), vidimo da je

$$F(A) = F(E)detA.$$

Ako još važi i da je $F(E) = 1$, onda je

$$F(A) = detA.$$

Pošto ovo važi za bilo koju kvadratnu matricu A, možemo zaključiti da je

$$F \equiv det. \blacksquare$$

Na osnovu prethodne teoreme možemo reći da je determinanta jedinstvena funkcija kvadratne matrice koja je linearna, antisimetrična i normirana kao funkcija njenih vrsta.

Do sada smo matricu posmatrali kao uređenu n –torku vrsta. Možemo je posmatrati i kao uređenu n –torku kolona. Tada determinanta postaje funkcija kolona matrice.

Stav 3.4. Kao funkcija $\boldsymbol{n}$ promenljivih kolona, determinanta je linearna, antisimetrična i normirana funkcija.

Dokaz se može naći u literaturi [5] na strani 92.

Stav 3.5. $\boldsymbol{detA^T = detA}$

Dokaz: Primenimo teoremu o jedinstvenosti na funkciju $F(A) = detA^T$. $detA^T$ je linearna, antisimetrična funkcija kolona matrice A^T, a kako su kolone matrice A^T ujedno i vrste matrice A, možemo reći da je F linearna, antisimetrična funkcija vrsta matrice A. Na osnovu teoreme o jedinstvenosti važi da je $F(A) = F(E)detA$. Očigledno je $F(E) = detE^T = detE = 1$, tako da je funkcija F i normirana. To znači da je $F(A) = detA$, odnosno $detA^T = detA$. $\blacksquare$

Posledica 3.6. Sve što je rečeno za vrste determinante, važi i za kolone.

Stav 3.7. Determinantu možemo razviti po bilo kojoj vrsti ili koloni.

Stav 3.8. (Determinanta blok-matrice): Ako imamo blok-matricu $\boldsymbol{A}$ formata 2x2 sa nula-matricom ispod dijagonale, tada je

$$detA = det\begin{bmatrix} B & D \\ 0 & C \end{bmatrix} = detBdetC.$$

Dokaz: Neka su C i D fiksirane, neka je $\bar{B} = \begin{bmatrix} B & D \\ 0 & C \end{bmatrix}$ i posmatrajmo funkciju $F(B) = det\bar{B}$ kao funkciju kolona matrice B. Pretpostavimo da je $b_i = \lambda p_i + \mu q_i$ kolona matrice B. Isto važi i za matricu $\bar{B}$, jer su njene kolone jednake kolonama matrice B i dopunjene su nulama. Tako da je $\overline{b_\iota} = \lambda p_i + \mu q_i$. Odatle dobijamo

$$\begin{aligned} F(B) = det\bar{B} = det[\cdots \quad \overline{b_\iota} \quad \cdots] = det[\cdots \quad \lambda p_i + \mu q_i \quad \cdots] \\ = \lambda det[\cdots \quad p_i \quad \cdots] + \mu det[\cdots \quad q_i \quad \cdots] = \lambda F(P) + \mu F(Q). \end{aligned}$$

Ovo dokazuje linearnost funkcije F. Pri tome matrice P i Q su dobijene od matrice B zamenom kolone b_i kolonama p_i i q_i. Ako i –ta i j –ta kolona matrice B zamene mesta, isto će se desiti i sa matricom $\bar{B}$. To znači da će $\det(\bar{B})$ promeniti znak, a isto važi i za $F(B)$, čime je dokazana antisimetričnost funkcije F. Za normiranost važi $F(E) = \det\bar{E} = 1 \ldots 1detC = detC$. Sada na funkciju F primenimo teoremu o jedinstvenosti determinante, jer je ona linearna i antisimetrična funkcija kolona matrice B.

$$det\bar{B} = F(B) = detBF(E) = detBdetC \ \blacksquare$$

Teorema 3.9. (Bine[5]-Košijeva[6] teorema) Ako su $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$, tada je

$$det(AB) = detAdetB.$$

Dokaz: Neka je matrica B fiksirana i posmatrajmo funkciju $F(A) = \det(AB)$ kao funkciju vrsta matrice A. Ako su a_i $(i = 1, \ldots, n)$ vrste matrice A, onda su $c_i = a_i B (i = 1, \ldots, n)$ vrste matrice AB. Pretpostavimo da je $a_i = \lambda p_i + \mu q_i$. Onda je $c_i = a_i B = \lambda p_i B + \mu q_i B$.

$$\begin{aligned} F(A) = det(AB) = det[\cdots \quad c_i \quad \cdots] = det[\cdots \quad \lambda p_i B + \mu q_i B \quad \cdots] \\ = \lambda det[\cdots \quad p_i B \quad \cdots] + \mu det[\cdots \quad q_i B \quad \cdots] = \lambda(P) + \mu F(Q) \end{aligned}$$

Ovo dokazuje linearnost funkcije F. Pri tome matrice P i Q su dobijene od matrice A zamenom vrste a_i vrstama p_i i q_i. Ako i –ta i j –ta vrsta matrice A zamene mesta, isto će se desiti i sa matricom AB. To znači da će $\det(AB)$ promeniti znak, a isto važi i za$F(A)$, čime je dokazana antisimetričnost funkcije F. Za normiranost važi $F(E) = \det(EB) = detB$. Sada na funkciju F primenimo teoremu o jedinstvenosti determinante, jer je ona linearna i antisimetrična funkcija vrsta matrice A.

$$\det(AB) = F(A) = detAF(E) = detAdetB \blacksquare$$

Teorema 3.10. Matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{F})$ je regularna ako i samo ako je $\boldsymbol{detA \neq 0}$. U tom slučaju je

$$detA^{-1} = \frac{1}{detA} \quad \text{i} \quad A^{-1} = \frac{1}{detA}\bar{A}.$$

[5]Jacques Philippe Marie Binet (1786-1856), francuski matematičar
[6]Augustin-Louis Cauchy (1789-1857), francuski matematičar

Pri čemu se matrica $\bar{A}$ naziva adjungovana matrica i dobijena je od matrice A na sledeći način svaki element a_{ij} je zamenjen kofaktoromA_{ij}, a potom je dobijena matrica transponovana. Ovaj postupak naziva se još i hermitijansko transponovanje.

Dokaz:

(⇒) Matrica $\bar{A}$ se može izračunati za svaku matricu A. Važe sledeće jednakosti

$$A\bar{A} = [a_{ij}][A_{ji}] = \begin{bmatrix} detA & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & detA \end{bmatrix} = detAE \text{ i}$$

$$\bar{A}A = [A_{ji}][a_{ij}] = \begin{bmatrix} detA & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & detA \end{bmatrix} = detAE.$$

Tako da je $A\bar{A} = \bar{A}A = detAE$.

Ako je $detA \neq 0$, onda je $A[(detA)^{-1}\bar{A}] = [(detA)^{-1}\bar{A}]A = E$, što znači da je matrica A regularna i da je njen inverz $A^{-1} = (detA)^{-1}\bar{A}$.

(⇐) Ukoliko je pak matrica A regularna, onda postoji njoj inverzna matrica A^{-1}, tako da je $AA^{-1} = E$.

$\det(AA^{-1}) = detE$ i primenom Bine-Košijeve teoreme dobijamo $detAdetA^{-1} = 1$, odavde sledi da je $detA \neq 0$. ■

Za regularnost matrice A, dovoljan je jedan od uslova $AB = E$ i $BA = E$. Jer je

$$AB = E \Rightarrow \det(AB) = detE \Rightarrow detAdetB = 1 \Rightarrow detA \neq 0 \Rightarrow A \text{ je regularna.}$$

3.1 Upotreba Matlab-a u matričnom računu

Matlab sadrži tri velika paketa funkcija koje su primenljive samo za matrice. To su:

- matlab\elmat – elementarne matrice i obrada matrica;
- matlab\matfun – matrične funkcije iz numeričke linearne algebre;
- matlab\sparfun – slabo popunjene matrice tj. matrice koje imaju sadrže veliki broj nula.

Elementarne matrice:

Funkcija	Opis funkcije	Primer	Rezultat primera
$zeros()$	Svi elementi matrice su 0. Važi za	$zeros(2)$	$ans = \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}$
		$zeros(2{,}3)$	$ans = \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$

	kvadratne i pravougaone matrice.		
$ones()$	Svi elementi matrice su 1. Važi za kvadratne i pravougaone matrice.	$ones(2)$	$ans = \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$
		$ones(3,2)$	$ans = \begin{matrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{matrix}$
$eye()$	Matrica identiteta. U slučaju pravougaone matrice, kvadratna matrica je matrica identiteta, ostali elementi su nule.	$eye(2)$	$ans = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$
		$eye(2,3)$	$ans = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix}$
		$eye(3,2)$	$ans = \begin{matrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{matrix}$
$diag()$	Vraća elemente na dijagonali u obliku vektora. U slučaju pravougaone matrice, vraća elemente na dijagonali kvadratne podmatrice.	$A = [5\ 7\ 9\ ;\ 4\ 5\ 2\ ; 1\ 2\ 5]$ $diag(A)$	$A = \begin{matrix} 5 & 7 & 9 \\ 4 & 5 & 2 \\ 1 & 2 & 5 \end{matrix}$ $ans = \begin{matrix} 5 \\ 5 \\ 5 \end{matrix}$
		$B\ =\ [1\ 1\ 1\ ;\ 2\ 2\ 2\ ;\ 3\ 3\ 3\ ;\ 4\ 4\ 4]$ $diag(B)$	$B = \begin{matrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \end{matrix}$ $ans = \begin{matrix} 1 \\ 2 \\ 3 \end{matrix}$
		$C\ =\ [5\ 5\ 5\ 5\ ;\ 6\ 6\ 6\ 6]$ $diag(C)$	$C = \begin{matrix} 5 & 5 & 5 & 5 \\ 6 & 6 & 6 & 6 \end{matrix}$ $ans = \begin{matrix} 5 \\ 6 \end{matrix}$
$rank()$	Rang matrice. Funkcija je primenljiva na kvadratne i pravougaone matrice.	$A = [4\ 4\ 4\ ;\ 0\ 8\ 9\ ; 0\ 8\ 9]$ $rank(A)$	$A = \begin{matrix} 4 & 4 & 4 \\ 0 & 8 & 9 \\ 0 & 8 & 9 \end{matrix}$ $ans = 2$
$det()$	Determinanta matrice. Funkcija je primenljiva na kvadratne matrice.	$A = [1\ 2\ 3\ ; 4\ 5\ 6\ ; 7\ 8\ 9]$ $det(A)$	$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}$ $ans = 0$

$trace()$	Trag matrice tj. zbir elemenata na dijagonali. U slučaju pravougaone matrice, vraća trag kvadratne podmatrice.	$A = [1\ 2\ 3\ ;4\ 5\ 6\ ;7\ 8\ 9]$ $trace(A)$	$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}$ $ans = 15$
$rref()$	Primena Gaus-Žordanove eliminacije sa pivotiranjem glavnih koeficijenata. Funkcija je primenljiva na kvadratne i pravougaone matrice.	$A = [6\ 2\ 3\ 13\ ;5\ 11\ 10\ 8\ ;9\ 7\ 6\ 12\ ;4\ 14\ 15\ 1]$ $rref(A)$	$A = \begin{matrix} 6 & 2 & 3 & 13 \\ 5 & 11 & 10 & 8 \\ 9 & 7 & 6 & 12 \\ 4 & 14 & 15 & 1 \end{matrix}$ $ans = \begin{matrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{matrix}$
$adj()$	Adjungovana matrica. Funkcija je primenljiva na kvadratne matrice.	$A = [1\ 2\ 3\ ;4\ 5\ 6\ ;7\ 7\ 8]$ $adj(A)$	$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 7 & 8 \end{matrix}$ $ans = \begin{matrix} -2 & 5 & -3 \\ 10 & -13 & 6 \\ -7 & 7 & -3 \end{matrix}$
$transpose()$	Transponovana matrica. Funkcija je primenljiva na kvadratne i pravougaone matrice.	$A = [1\ 2\ 3\ ;4\ 5\ 6\ ;7\ 7\ 8]$ $transpose(A)$	$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 7 & 8 \end{matrix}$ $ans = \begin{matrix} 1 & 4 & 7 \\ 2 & 5 & 7 \\ 3 & 6 & 8 \end{matrix}$

Posebnu pažnju treba posvetiti rešavanju linearnih sistema jednačina. Razmotrimo sistem

$$Ax = b$$

gde je matrica sistema A formata $n \times n$, x je matrica nepoznatih formata $n \times 1$ i b je matrica slobodnih članova formata $n \times 1$. Ako je matrica A regularna tj. ako je $detA \neq 0$, odnosno $rangA = n$, rešenje sistema je

$$x = A^{-1}b$$

i za to se u Matlab-u koristi jedna od naredbi:

- Funkcija $inv()$ u sledećem obliku

$$x = inv(A) * transpose(b);$$

- Levo deljenje pomoću operatora "\" u sledećem obliku

$$x = A\backslash transpose(b).$$

Primer: Rešiti sistem linearnih jednačina.

$$2x + y - z = 2$$

$$-x + 2y + 3z = 4$$

$$x + y + z = 3$$

Matrica sistema je $A = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}$, a matrica slobodnih članova $b = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$.

$$A \ = \ [2\ 1\ \ -1\ ;\ -1\ 2\ 3\ ;\ 1\ 1\ 1]$$

$$b \ = \ [2\ 4\ 3]$$

Rešenje sistema možemo naći na dva načina.

1. $inv(A) * transpose(b)$
   $ans = \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}$
2. $A\backslash transpose(b)$
   $ans = \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}$

4 Uopštena inverzna matrica pravougaone matrice

Pri rešavanju sistema linearnih jednačina koji nemaju potpuni rang tj. u slučaju kada je broj jednačina manji od broja nepoznatih koristi se inverzna matrica pravougaone matrice (G-inverzna matrica, uopštena inverzna matrica). To je na primer situacija kada treba da ocenimo n nepoznatih parametara, a raspolažemo sa k opservacija, pri čemu je $k < n$.

Ovde razmatramo linearne modele $X = A\theta + \varepsilon$ za koje važi

$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ je $n-$ dimenzioni slučajni vektor tj. $X_1, \dots X_n$ su slučajne veličine, $A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{bmatrix}$ je matrica formata $n \times p$, gde je $p > n$, $\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix}$ je p-dimenzioni nepoznati parametar tj. $\theta_1, \dots, \theta_p$ su nepoznati parametri koje treba oceniti i $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ je n-dimenzioni slučajni vektor, gde su $\varepsilon_1, \dots \varepsilon_n$ su slučajne veličine.

Definicija 4.1. Uopštena inverzna matrica matrice $\boldsymbol{A}$ formata $\boldsymbol{m} \times \boldsymbol{n}$ je bilo koja matrica $\boldsymbol{G}$ formata $\boldsymbol{n} \times \boldsymbol{m}$ za koju važi

$$AGA = A.$$

Stav 4.1. Inverzna matrica $\boldsymbol{A}^{-1}$ regularne matrice $\boldsymbol{A}$ je ujedno i uopšteni inverz te matrice. Regularna matrica $\boldsymbol{A}$ nema drugih uopštenih inverza sem tog jednog.

Dokaz: Neka je G jedan uopšteni inverz matrice A. Tada važi

$$G = EGE = (A^{-1}A)G(AA^{-1}) = A^{-1}(AGA)A^{-1} = A^{-1}AA^{-1} = A^{-1}. \blacksquare$$

Svaka matrica ima bar jedan uopšteni inverz. Skup svih uopštenih inverznih matrica matrice A obeležava se sa $GI(A)$.

Teorema 4.2. Neka je $\boldsymbol{G} \in \boldsymbol{GI}(\boldsymbol{A})$. Onda važi sledeće:

(1) GA je idempotentna matrica tj. $(GA)^2 = GA$;

(2) Opšte rešenje homogenog sistema $AX = 0$ je

$$X = (GA - E)Z$$

gde je Z proizvoljan vektor;

(3) Opštenje rešenje saglasnog nehomogenog sistema $AX = Y$ je

$$X = GY + (GA - E)Z$$

gde je Z proizvoljan vektor.

Dokaz:

(1) $(GA)^2 = (GA)(GA) = G(AGA) = GA$

(2) Proverimo da je $X = (GA - E)Z$ zaista rešenje sistema $AX = 0$.

$$AX = A\big((GA - E)Z\big) = A(GAZ - EZ) = AGAZ - AZ = AZ - AZ = 0$$

(3) Proverimo da je $X = GY + (GA - E)Z$ zaista rešenje sistema $AX = Y$.

$$\begin{aligned} AX = A(GY + (GA - E)Z) = A(GY + GAZ - EZ) = AGY + AGAZ - AZ \\ = AGY + AZ - AZ = AGY + 0 = AGY \end{aligned}$$

$$Y = AX = (AGA)X = AG(AX) = AGY \blacksquare$$

Stav 4.3. Neka je $\boldsymbol{B}$ regularna matrica formata $\boldsymbol{r} \times \boldsymbol{r}$, $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \boldsymbol{DB^{-1}C} \end{bmatrix}$ je blok-matrica formata $\boldsymbol{m} \times \boldsymbol{n}$ i $\boldsymbol{rangA} = \boldsymbol{r}$. Onda je

$$Q \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} P \in GI(A).$$

Dokaz: Kao i obično obeležimo uopštenu inverznu matricu sa G tj. $G = Q \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} P$.

Pretpostavimo da bazni minor B matrice A nije u gornjem levom uglu. U tom slučaju primenjujemo zamenu mesta vrsta i kolona da bi to postigli.

Kako je matrica B formata $r \times r$, a $rangA = r$ to znači da su kolone matrice C linearno zavisne od kolona matrice B i vrste matrice D su linearno zavisne od vrsta matrice B. Elementarne transformacije kojima se kolone matrice C anuliraju pomoću kolona matrice B poklapaju se sa elementarnim transformacijama putem kojih se kolone matrice $DB^{-1}C$ anuliraju pomoću kolona matrice D. Slično važi i za vrste.

Već smo istakli da se elementarne transformacije kolona neke matrice dobijaju množenjem te matrice sa matricom Q sa leve strane, gde je Q proizvod elementarnih matrica. To znači da je

$$C = BQ.$$

Odnosno,

$$Q = B^{-1}C.$$

Kada to primenimo na matricu $DB^{-1}C$ dobijamo

$$DB^{-1}C = DQ.$$

Kako su matrice C i $DB^{-1}C$ dobijene množenjem matrica B i D istom matricom tj. matricom Q, možemo zaključiti da se transformacije za anuliranje kolona matrica C i $DB^{-1}C$ zaista poklapaju.

Slično tome, elementarne transformacije vrsta neke matrice dobijaju se množenjem te matrice sa matricom P sa desne strane, gde je P proizvod elementarnih matrica. To znači da je

$$D = PB.$$

Odnosno,

$$P = DB^{-1}.$$

Kada to primenimo na matricu $DB^{-1}C$ dobijamo

$$DB^{-1}C = PC.$$

Kako su matrice D i $DB^{-1}C$ dobijene množenjem matrica B i C istom matricom tj. matricom P, možemo zaključiti da se transformacije za anuliranje vrsta matrica D i $DB^{-1}C$ zaista poklapaju.

To znači da je $PAQ = \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix}$ gde je matrica B zaista bazni minor. Odavde sledi da je

$$A = P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1}.$$

Direktnom proverom videćemo da je $G = Q \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} P$ zaista uopštena inverzna matrica.

$$\begin{aligned}
AGA &= \left(P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1}\right)\left(Q \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} P\right)\left(P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1}\right) \\
&= P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} (Q^{-1}Q) \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} (PP^{-1}) \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} \\
&= P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} E \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} E \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} \\
&= P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} \begin{bmatrix} B^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} \\
&= P^{-1} \begin{bmatrix} BB^{-1} & 0 \\ DB^{-1} & 0 \end{bmatrix} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} = P^{-1} \begin{bmatrix} E & 0 \\ DB^{-1} & 0 \end{bmatrix} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} \\
&= P^{-1} \begin{bmatrix} B & C \\ DB^{-1}B & DB^{-1}C \end{bmatrix} Q^{-1} = P^{-1} \begin{bmatrix} B & C \\ D & DB^{-1}C \end{bmatrix} Q^{-1} = A \blacksquare
\end{aligned}$$

4.1 Upotreba Matlab-a za uopštene inverzne matrice

Razmotrimo sistem

$$Ax = b$$

gde je matrica sistema A formata $m \times n$, pri čemu je $m < n$. Takođe, x je matrica nepoznatih formata $n \times 1$ i b je matrica slobodnih članova formata $m \times 1$. Ovakav sistem ima više nepoznatih nego jednačina i rešenje sistema je zapravo uopštena inverzna matrica.

U Matlab-u postoji ugrađena funkcija $pinv()$ koja se koristi u ovakvoj situaciji i to u formi $x = pinv(A) * transpose(b)$.

Primer: Rešiti sistem linearnih jednačina.

$$x + 2y + 3z = 1$$

$$4x + 5y + 6z = 1$$

Matrica sistema je $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, a matrica slobodnih članova $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$A \ = [1\ 2\ 3\ ;\ 4\ 5\ 6]$$

$$b \ = \ [1\ 1]$$

Rešenje sistema se dobija pozivanjem sledeće naredbe: $pinv(A) * transpose(b)$.

$$ans = \begin{matrix} -0.5 \\ 0 \\ 0.5 \end{matrix}$$

5 Matrica kovarijacije

Pre nego što damo prve primere upotrebe matričnog računa u statistici, neophodno je definisati i neke osnovne pojmove iz statistike.

Ako su $X_1, \dots, X_n$ slučajne veličine, onda se $n-$dimenzionim slučajnim vektorom naziva njihov uređeni skup $X = (X_1, \dots, X_n)$.

Zapisan u matričnom obliku slučajni vektor je $X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$.

Matematičko očekivanje slučajnog vektora $X = (X_1, \dots, X_n)$ je vektor matematičkih očekivanja njegovih komponenti tj. $E(X) = (E(X_1), \dots, E(X_n))$ ili matrično

$$E(X) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{bmatrix}.$$

5.1 Kovarijacija, korelacija

Kovarijacija dve slučajne veličine daje jednu moguću meru zavisnosti tih slučajnih veličina. Ako je kovarijacija velika i pozitivna, onda obe slučajne veličine imaju veliko odstupanje od svojih očekivanih vrednosti i to na istu stranu, odnosno njihove vrednosti su znatno iznad ili znatno ispod očekivanih. Ako je kovarijacija velika i negativna, onda su odstupanja takođe velika, ali jedna slučajna veličina ima vrednosti znatno iznat očekivane, a druga znatno ispod očekivane.

Koeficijent korelacije takođe meri stepen zavisnosti dve slučajne veličine. Slučajne veličine X i Y su pozitivno (negativno) korelisane ako je njihov koeficijent korelacije $\rho(X, Y)$ pozitivan (negativan). Slučajne veličine X i Y su nekorelisane ako je njihov koeficijent korelacije $\rho(X, Y) = 0$. Pri tome uvek važi

$$|\rho(X, Y)| \leq 1.$$

Ako je $\rho(X, Y) = \pm 1$ onda je sa verovatnoćom 1 veza između slučajnih veličina X i Y linearna tj. $Y = \pm aX + b$, gde je $a \in \mathbb{R}^+$ i $b \in \mathbb{R}$. Ako je pak $\rho(X, Y) = 0$, onda možemo odbaciti mogućnost da je veza između ovih slučajnih veličina linearna, što ne znači da ne postoji neka druga veza.

Definicija 5.1. Neka su $\boldsymbol{X}$ i $\boldsymbol{Y}$ slučajne veličine sa konačnim i strogo većim od nula disperzijama. Kovarijacija (kovarijansa) slučajnih veličina $\boldsymbol{X}$ i $\boldsymbol{Y}$ u oznaci $\boldsymbol{cov}(\boldsymbol{X}, \boldsymbol{Y})$ je broj zadat sa

$$cov(X, Y) = E(X - EX)(Y - EY).$$

Koeficijent korelacije slučajnih veličina X i Y je broj

$$\rho(X,Y) = \frac{E(XY) - EYEX}{\sqrt{DX}\sqrt{DY}}.$$

$$\begin{aligned} cov(X,Y) &= E(X - EX)(Y - EY) = E(XY - XEY - YEX + EXEY) \\ &= E(XY) - E(XEY) - E(YEX) + E(EXEY) \\ &= E(XY) - EYEX - EXEY + EXEY = E(XY) - EXEY \end{aligned}$$

Ovo važi zbog svojstva matematičkog očekivanja:

- $E(X) \in \mathbb{R}$;
- $E(c) = c$ gde je $c \in \mathbb{R}$;
- $E(cX) = cE(X)$ gde je $c \in \mathbb{R}$.

$$\rho(X,Y) = \frac{E(XY) - EYEX}{\sqrt{DX}\sqrt{DY}} = \frac{cov(X,Y)}{\sigma X \sigma Y}$$

gde je σ_X standardna devijacija slučajne veličine X.

Kovarijacija slučajne veličine X sa njom samom se naziva autokovarijacija i predstavlja disperziju slučajne veličine X.

$$cov(X,X) = E(X - EX)(X - EX) = E(X - EX)^2 = DX$$

Koeficjent korelacije slučajne veličine X sa njom samom je uvek 1. Lako se može pokazati da je zaista tako.

$$\rho(X,X) = \frac{cov(X,X)}{\sigma X \sigma X} = \frac{DX}{DX} = 1$$

5.2 Matrica kovarijacije, korelacije i kroskovarijacije

Definicija 5.2. Neka je $\boldsymbol{X} = (\boldsymbol{X_1}, \dots, \boldsymbol{X_n})$ slučajni vektor, takav da za sve $\boldsymbol{i} = \boldsymbol{1}, \dots, \boldsymbol{n}$ važi $\boldsymbol{EX_i^2} < +\infty$. Za ovaj slučajni vektor definiše se matrica kovarijacije (matrica autokovarijacije, matrica kovarijanse) $\boldsymbol{cov(X)} \in \boldsymbol{M(n}, \mathbb{R})$ na sledeći način

$$cov(X) = E\big((X - EX)(X - EX)^T\big).$$

Razmotrimo malo detaljnije ovu jednakost

$$cov(X) = E\left(\begin{bmatrix} X_1 - EX_1 \\ \vdots \\ X_n - EX_n \end{bmatrix}\begin{bmatrix} X_1 - EX_1 \\ \vdots \\ X_n - EX_n \end{bmatrix}^T\right) = E\left(\begin{bmatrix} X_1 - EX_1 \\ \vdots \\ X_n - EX_n \end{bmatrix}\begin{bmatrix} X_1 - EX_1 & \dots & X_n - EX_n \end{bmatrix}\right)$$

$$= E\begin{bmatrix} (X_1-EX_1)(X_1-EX_1) & (X_2-EX_2)(X_1-EX_1) & \cdots & (X_n-EX_n)(X_1-EX_1) \\ (X_1-EX_1)(X_2-EX_2) & (X_2-EX_2)(X_2-EX_2) & \cdots & (X_n-EX_n)(X_2-EX_2) \\ \vdots & \vdots & \ddots & \vdots \\ (X_1-EX_1)(X_n-EX_n) & (X_2-EX_2)(X_n-EX_n) & \cdots & (X_n-EX_n)(X_n-EX_n) \end{bmatrix}$$

$$= \begin{bmatrix} E(X_1-EX_1)^2 & E(X_2-EX_2)(X_1-EX_1) & \cdots & E(X_n-EX_n)(X_1-EX_1) \\ E(X_1-EX_1)(X_2-EX_2) & E(X_2-EX_2)^2 & \cdots & E(X_n-EX_n)(X_2-EX_2) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_1-EX_1)(X_n-EX_n) & E(X_2-EX_2)(X_n-EX_n) & \cdots & E(X_n-EX_n)^2 \end{bmatrix}$$

$$= \begin{bmatrix} cov(X_1,X_1) & cov(X_2,X_1) & \cdots & cov(X_n,X_1) \\ cov(X_1,X_2) & cov(X_2,X_2) & \cdots & cov(X_n,X_2) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_1,X_n) & cov(X_2,X_n) & \cdots & cov(X_n,X_n) \end{bmatrix}$$

jer je

$$E(X_i - EX_i)\big(X_j - EX_j\big) = cov\big(X_i, X_j\big) \text{ za } i,j = 1,\dots,n.$$

Tako da je

$$cov(X) = \big[cov\big(X_i, X_j\big)\big]_{i,j=1}^{n}.$$

Kako je

$$cov(X_i, X_i) = D(X_i)$$

na dijagonali su disperzije slučajnih veličina X_i za $i = 1,\dots,n,$ a vandijagonalni elementi predstavljaju kovarijacije između slučajnih veličina X_i i X_j za $i,j = 1,\dots,n$ i $i \neq j$.

Neka su X_i, X_j nezavisne slučajne veličine $(i,j = 1,\dots,n)$ i $i \neq j$, tada je

$$cov\big(X_i, X_j\big) = E\big(X_iX_j\big) - EX_iEX_j = EX_iEX_j - EX_iEX_j = 0.$$

Odnosno, kovarijaciona matrica je dijagonalna, a elementi na dijagonali su disperzije slučajnih veličina $X_i (i = 1,\dots,n)$.

$$cov(X) = \begin{bmatrix} DX_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & DX_n \end{bmatrix}$$

Ako su još disperzije svih slučajnih veličina jednake σ^2, onda je kovarijaciona matrica skalarna.

$$cov(X) = \begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \sigma^2 E$$

Matrica kovarijacije se može predstaviti i pomoću matrica

$$\Gamma = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{bmatrix} \quad \text{i} \quad \mathrm{H} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{bmatrix}$$

gde je $\sigma_i (i = 1, \dots, n)$ standardna devijacija slučajne veličine X_i i $\rho_{ij} (i, j = 1, \dots, n)$ koeficijent korelacije između slučajnih veličina X_i i $X_j (i, j = 1, \dots, n)$.

Tada je

$$cov(X) = \Gamma \mathrm{H} \Gamma,$$

što se dokazuje direktnom proverom, jer je $cov(X_i, X_j) = \sigma_i \sigma_j \rho_{ij} (i, j = 1, \dots, n)$.

Matrica H se naziva matrica koeficijenata korelacije slučajnog vektora X.

$$\mathrm{H} = \rho(X) = \left[\rho(X_i, X_j)\right]_{i,j=1}^{n} = \rho_{ij}$$

$$\rho(X_i, X_j) = \frac{E(X_i X_j) - EX_i EX_j}{\sqrt{DX_i}\sqrt{DX_j}} = \frac{cov(X_i, X_j)}{\sigma X_i \sigma X_j}$$

Dijagonalni elementi ove matrice su u jedinice, jer u slučaju kada je $i = j$ tražimo koeficijent korelacije slučajne veličine X_i sa samom sobom, a to je 1.

Sledeća teorema povezuje važne pojmove statistike (odnosno verovatnoće) i matrične algebre.

Teorema 5.1. (Nenegativna definitnost matrice kovarijacije) Matrica $\boldsymbol{B} = [\boldsymbol{b_{ij}}] \in \boldsymbol{M}(\boldsymbol{n}, \mathbb{R})$ je kovarijaciona matrica nekog slučajnog vektora $\boldsymbol{X} = (\boldsymbol{X_1}, \dots, \boldsymbol{X_n})$ ako i samo ako je $\boldsymbol{B}$ simetrična i nenegativno definitna matrica.

Dokaz:

($\Rightarrow$) $B = [b_{ij}] \in M(n, \mathbb{R})$ je kovarijaciona matrica. To znači da je $b_{ij} = cov(X_i, X_j)$ za $i, j = 1, \dots, n$.

(1) $b_{ij} = cov(X_i, X_j) = EX_i X_j - EX_i EX_j = EX_j X_i - EX_j EX_i = cov(X_j, X_i) = b_{ji}$

Matrica B je simetrična.

(2) Za svaki vektor $(\beta_1, \dots, \beta_n) \in \mathbb{R}^n$ važi

$$\sum_{i,j=1}^{n} \beta_i\beta_j b_{ij} = \sum_{i,j=1}^{n} \beta_i\beta_j cov(X_i, X_j)$$
$$= \sum_{i,j=1}^{n} \beta_i\beta_j E(X_i - EX_i)(X_j - EX_j)$$
$$= E\sum_{i,j=1}^{n} \beta_i\beta_j (X_i - EX_i)(X_j - EX_j) = E\left[\sum_{i=1}^{n} \beta_i (X_i - EX_i)\right]^2 \geq 0.$$

Matrica B je nenegativno definitna.

($\Leftarrow$) Neka je matrica B simetrična i nenegativno definitna. Tada na osnovu posledice 2.10. postoji ortogonalna matrica M takva da važi $M^T BM = D$, gde je D dijagonalna matrica sa nenegativnim elementima na dijagonali.

$$D = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{bmatrix} \text{i } d_1, \dots d_n \geq 0$$

$$M^T BM = D$$

$(M^T)^{-1} M^T BMM^{-1} = (M^T)^{-1} DM^{-1}$ (1)

Kako je M ortogonalna matrica, to važi da je $M^T = M^{-1}$.

Kada se vratimo u jednakost (1), dobijemo

$B = MDM^T$. (2)

Neka je $D_1 = \begin{bmatrix} \sqrt{d_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{d_n} \end{bmatrix}$. Direktnom proverom se uočava da je $D_1 {D_1}^T = D$. Tada je jednačina (2) oblika $B = M(D_1 {D_1}^T)M^T$.

Zbog asocijativnosti množenja matrica, važi

$B = (MD_1)({D_1}^T M^T)$. (3)

Neka je $A = MD_1$. Tada je $A^T = (MD_1)^T = {D_1}^T M^T$, što znači da je jednačina (3)

$B = AA^T$.

Neka su $Y_1, \dots, Y_n \in N(0,1)$ nezavisnesločajne veličine i neka je $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = A \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$.

$$YY^T = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} [Y_1 \quad \cdots \quad Y_n] = \begin{bmatrix} {Y_1}^2 & \cdots & Y_n Y_1 \\ \vdots & \ddots & \vdots \\ Y_1 Y_n & \cdots & {Y_n}^2 \end{bmatrix}$$

$$E(YY^T) = E \begin{bmatrix} {Y_1}^2 & \cdots & Y_1 Y_n \\ \vdots & \ddots & \vdots \\ Y_1 Y_n & \cdots & {Y_n}^2 \end{bmatrix} = \begin{bmatrix} E{Y_1}^2 & \cdots & E(Y_1 Y_n) \\ \vdots & \ddots & \vdots \\ E(Y_1 Y_n) & \cdots & E{Y_n}^2 \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \boldsymbol{E}$$

Jer je $Y_i \in N(0,1)(i = 1, \dots, n)$, pa je $E{Y_i}^2 = DY_i = 1$ i $Y_i(i = 1, \dots, n)$ su nezavisne, pa je $E\left(Y_i Y_j\right) = EY_i EY_j = 0$.

$E(XX^T) = E((X - 0)(X - 0)^T) = E((X - EX)(X - EX)^T) = cov(X)$ (1)

Jer je $EX = E(AY) = AE(Y) = A0 = 0$.

$E(XX^T) = E\left((AY)(AY)^T\right) = E(AYY^T A^T) = AE(YY^T)A^T = A\boldsymbol{E}A^T = AA^T = B$ (2)

Kada uporedimo jednačine (1) i (2), vidimo da je B kovarijaciona matrica slučajnog vektora X.■

Neka je $X = (X_1, \dots, X_n)$ slučajni vektor i $A \in M(n, \mathbb{F})$ matrica čiji su elementi skalari iz polja $\mathbb{F}$. Linearna transformacija slučajnog vektora X je slučajni vektor$Y$, definisan na sledeći način

$$Y = AX.$$

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \text{ i } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$$Y = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} a_{11}X_1 + \cdots + a_{1n}X_n \\ \vdots \\ a_{n1}X_1 + \cdots + a_{nn}X_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$$

Matematičko očekivanje linearne transformacije slučajnog vektora je $EY = E(AX) = AEX$.

Kovarijaciona matrica linearne transformacije slučajnog vektora je

$$\begin{aligned} cov(Y) &= E\left((Y - EY)(Y - EY)^T\right) = E\left((AX - AEX)(AX - AEX)^T\right) \\ &= E\left(A(X - EX)\left(A(X - EX)\right)^T\right) = E(A(X - EX)(X - EX)^T A^T) \\ &= AE\left((X - EX)(X - EX)^T\right)A^T = Acov(X)A^T. \end{aligned}$$

Matrica kovarijacije se lako generalizuje za slučaj dva različita slučajna vektora istih dimenzija.

Definicija 5.3. Neka su $\boldsymbol{X} = (\boldsymbol{X_1}, \dots, \boldsymbol{X_n})$ i $\boldsymbol{Y} = (\boldsymbol{Y_1}, \dots, \boldsymbol{Y_n})$ $\boldsymbol{n}$ −dimenzioni slučajni vektori. Tada je matrica kroskovarijacije (kroskovarijanse) definisana na sledeći način

$$cov(X, Y) = E\left((X - EX)(Y - EY)^T\right).$$

Neka je $B \in M(n, \mathbb{F})$ matrica čiji su elementi skalari iz polja $\mathbb{F}$. Primenom linearne transformacije na slučajni vektor X, dobijen je slučajni vektor $Z = BX$.

$$\begin{aligned} cov(Y,Z) &= E\big((Y-EY)(Z-EZ)^T\big) = E\big((Y-EY)(Z^T-(EZ)^T)\big) \\ &= E\Big(\big(AX-E(AX)\big)\Big((BX)^T-\big(E(BX)\big)^T\Big)\Big) \\ &= E\Big(\big(AX-A(EX)\big)\Big(X^TB^T-\big(B(EX)\big)^T\Big)\Big) \\ &= E\big(A(X-EX)(X^TB^T-(EX)^TB^T)\big) = E\big(A(X-EX)(X^T-(EX)^T)B^T\big) \\ &= E\big(A(X-EX)(X-EX)^TB^T\big) = AE\big((X-EX)(X-EX)^T\big)B^T \\ &= Acov(X)B^T. \end{aligned}$$

5.3 Upotreba Matlab-a u statistici

Matlab ima programske pakete (eng. toolboxes) koji predstavljaju zaokruženu programsku celinu neke oblaski nauke ili tehinike. Tako i za statistiku postoji programski paket koji se naziva toolbox\stats.

Pre toga, napomenimo da se u Matlab-u izmerene vrednosti predstavljaju u obliku vektora, a izmerene vrednosti slučajnog vektora predstavljene su u obliku matrice. Pri tome, kolone matrice su pojedinačne slučajne veličine, a redovi matrice su pojedinačne observacije.

Funkcija	Opis funkcije	Primer	Rezultat primera
$mean()$	Ukoliko je argument funkcije vektor, onda je rezultat srednja vrednost elemenata vektora tj. uzoračko matematičko očekivanje slučajne veličine.	$A = [0 \quad 1 \quad 1]$ $mean(A)$	$A = 0 \quad 1 \quad 1$ $ans = 0.66667$
	Ukoliko je argument funkcije matrica, onda je rezultat vektor čiji su elementi	A $= [0\ 1\ 1;\ 2\ 3\ 2;\ 1\ 3\ 2;\ 4\ 2\ 2]$ $mean(A)$	$A = \begin{matrix} 0 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 3 & 2 \\ 4 & 2 & 2 \end{matrix}$ ans $= 1.7500 \quad 2.2500 \quad 1.7500$

	srednje vrednosti kolona te matrice tj. svaki element vektora je uzoračko matematičko očekivanje slučajnih veličina čije su observacije pojedinačne kolone.		
$var()$	Ukoliko je argument funkcije vektor, onda je rezultat popravljena disperzija elemenata vektora tj.popravljena uzoračka dispercija slučajne veličine.	$A = [1 \quad 2 \quad 3]$ $var(A)$	$A = 1 \quad 2 \quad 3$ $ans = 1$
	Ukoliko je argument funkcije matrica, onda je rezultat vektor čiji su elementi popravljene disperzije kolona te matrice tj. svaki element vektora je popravljena uzoračka disperzija slučajnih veličina čije su observacije	$A = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $var(A)$	$A = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ ans $= 12.500 \quad 24.500 \quad 18.000$

	pojedinačne kolone.		
	Drugi mogući argument ove funkcije je indikator koji može imati vrednost 0 ili 1. Ako je vrednost 0, onda se disperzija računa na način 1. Ako je vrednost indikatora 1, onda se disperzija računa na način 2.	$A = [1 \quad 2 \quad 3]$ $var(A, 0)$ $var(A, 1)$ $B = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $var(B, 0)$ $var(B, 1)$	$A = 1 \quad 2 \quad 3$ $ans = 1$ $ans = 0.66667$ $B = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = 12.500 \quad 24.500 \quad 18.000$ $ans = 6.250 \quad 12.250 \quad 9.00$
$std()$	Ukoliko je argument funkcije vektor, onda je rezultat popravljena standardna devijacija elementata vektora tj. popravljena uzoračka standardna devijacija slučajne veličine.	$A = [1 \quad 2 \quad 3]$ $std(A)$	$A = 1 \quad 2 \quad 3$ $ans = 1$
	Ukoliko je argument funkcije matrica, onda je rezultat vektor čiji su elementi popravljene standardne devijacije kolona te	$A = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $std(A)$	$A = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = 3.5355 \quad 4.9497 \quad 4.2426$

	matrice tj. svaki element vektora je popravljena standardna devijacija slučajnih veličina čije su observacije pojedinačne kolone.		
	Drugi mogući argument ove funkcije je indikator koji može imati vrednost 0 ili 1. Ako je vrednost 0, onda se standardna devijacija računa na način 1. Ako je vrednost indikatora 1, onda se standardna devijacija računa na način 2.	$A = [1 \quad 2 \quad 3]$ $std(A, 0)$ $std(A, 1)$ $B = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $std(B, 0)$ $std(B, 1)$	$A = 1 \quad 2 \quad 3$ $ans = 1$ $ans = 0.8165$ $B = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = 3.5355 \quad 4.9497 \quad 4.2426$ $ans = 2.50 \quad 3.50 \quad 3.00$
$cov()$	Ukoliko su argumenti funkcije dva vektora, onda je rezultat kovarijacija među slučajnim veličimana koje predstavljaju ti vektori. Neophodan je isti broj observacija za obe slučajne	$A = [1 \quad 2 \quad 3]$ $C = [7 \quad -3 \quad 1.5]$ $cov(A, C)$	$A = 1 \quad 2 \quad 3$ $C = 7 \quad -3 \quad 1.5$ $ans = -2.7500$

	veličine.		
	Ukoliko je argument funkcije matrica, onda su kolone slučajne veličine, a redovi observacije tih slučajnih veličina. Rezultat je matrica kovarijacije slučajnog vektora.	$B = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $cov(B)$	$B = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = \begin{matrix} 12.50 & 17.50 & 15.00 \\ 17.50 & 24.50 & 21.00 \\ 15.00 & 21.00 & 18.00 \end{matrix}$
	Drugi mogući argument ove funkcije je indikator koji može imati vrednost 0 ili 1. Ako je vrednost 0, formule pomoću kojih se računa kovarijacija su normalizovane sa $n-1$. Ako je vrednost 1, formule pomoću kojih se računa kovarijacija su normalizovane sa n.	$A = [1 \quad 2 \quad 3]$ $C = [7 \quad -3 \quad 1.5]$ $cov(A, C, 0)$ $cov(A, C, 1)$ $B = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $cov(B, 0)$ $cov(B, 1)$	$A = 1 \quad 2 \quad 3$ $C = 7 \quad -3 \quad 1.5$ $ans = -2.7500$ $ans = -1.8333$ $B = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = \begin{matrix} 12.50 & 17.50 & 15.00 \\ 17.50 & 24.50 & 21.00 \\ 15.00 & 21.00 & 18.00 \end{matrix}$ $ans = \begin{matrix} 6.250 & 8.750 & 7.500 \\ 8.750 & 12.250 & 10.500 \\ 7.500 & 10.500 & 9.000 \end{matrix}$
$corrcoef()$	Ukoliko su argumenti funkcije dva vektora, onda je rezultat koeficijent korelacije među	$A = [1 \quad 2 \quad 3]$ $C = [7 \quad -3 \quad 1.5]$ $corrcoef(A, C)$	$A = 1 \quad 2 \quad 3$ $C = 7 \quad -3 \quad 1.5$ $ans = -0.54909$

$\bar{x}$.

	slučajnim veličinama koje predstavljaju ti vektori. Neophodan je isti broj observacija za obe slučajne veličine.		
	Ukoliko je argument funkcije matrica, onda su kolone slučajne veličine, a redovi observacije tih slučajnih veličina. Rezultat je matrica korelacije slučajnog vektora.	$B = [4 \quad -2 \quad 1; 9 \quad 5 \quad 7]$ $corrcoef(B)$	$B = \begin{matrix} 4 & -2 & 1 \\ 9 & 5 & 7 \end{matrix}$ $ans = \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$

Napomena 1: Uzoračka disperzija var može da se računa na dva načina.

Način 1: popravljena uzoračka disperzija.

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

Način 2: nepopravljena uzoračaka disperzija.

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

gde je

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

Napomena 2: Uzoračka standardna devijacija σ može da se računa na dva načina.

Način 1: popravljena uzoračka standardna devijacija.

$$\sigma = \left(\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{\frac{1}{2}}$$

Način 2: nepopravljena uzoračka standardna devijacija.

$$\sigma = \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{\frac{1}{2}},$$

gde je

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

5.4 Slučajni vektori i kvadratne forme

Već je bilo reči o kvadratnim formama, ali sada ćemo ih koristiti na još jedan način.

Definicija 5.4. Neka je $X = (X_1, \dots, X_n)$ slučajni vektor i $A \in M(n, \mathbb{R})$ simetrična kvadratna matrica. Slučajna veličina $Q = X^T A X$ je kvadratna forma slučajnog vrektora X.

Zbog simetričnosti matrice A, postoji nekoliko načina kako možemo zapisati ovu kvadratnu formu:

$$Q = X^T A X = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} X_i X_j = \sum_{i=1}^{n} a_{ii} X_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_{ij} X_i X_j = \sum_{i=1}^{n} a_{ii} X_i^2 + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} a_{ij} X_i X_j$$

Sledeći stavovi tiču se raspodela i nezavisnosti kvadratnih formi slučajnih vektora. U dokazima ovih stavova koriste se pojedine osobine traga matrice, kao na primer da je trag matrice linerni operator i da je

$$tr(ABC) = tr(BCA) = tr(CAB).$$

Stav 5.2. Pretpostavimo da slučajni vektor $X = (X_1, \dots, X_n)$ ima očekivanje μ i matricu kovarijacije Σ. Neka je $Q = X^T A X$, gde je $A \in M(n, \mathbb{R})$ simetrična kvadratna matrica. Tada važi sledeće:

$$E(Q) = trA\Sigma + \mu^T A \mu.$$

Dokaz se može naći u literaturi [8] na strani 502.

Takođe je bitna i slecijalna dekompozicija realne simerične matice. Naime ako je A relana simetrična matrica, onda se ona može napisati u obliku

$$A = \Gamma^T \Lambda \Gamma,$$

gde je Λ dijagonalana matrica $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$,i $\lambda_1 \geq \cdots \geq \lambda_n$ su sopstvene vrednosti matrice A, a kolone $\Gamma^T = [v_1 \quad \cdots \quad v_n]$ su odgovarajući ortonormirani sopstveni vektri, što znači da je Γ ortogonalna matrica. Rang matrice A je broj sopstvenih vrednosti različitih od nule.

Stav 5.3. Neka slučajni vektor $X = (X_1, \dots, X_n)$ ima $n-$ dimenzionalnu normalnu raspodelu sa očekivanjem μ i matricom kovarijacije Σ koja je pozitivno definitna. Tada slučajna veličina

$$Q = (X-\mu)^T (\Sigma)^{-1} (X-\mu)$$

ima $\chi^2 -$ raspodelu sa n stepeni slobode.

Dokaz se može naći u literaturi [8] na strani 504.

Idempotentne matrice imaju primenu u ovoj oblasti. Pretpostavimo da je λ sopstvena vrednost idempotentne matrice A, a v odgovarajući sopstveni vektor. Tada važi

$$\lambda v = Av = A^2 v = \lambda A v = \lambda^2 v.$$

$$\lambda v - \lambda^2 v = 0$$

$$\lambda(1-\lambda)v = 0$$

Kako je $v \neq 0$, to znači da je $\lambda = 0$ ili $\lambda = 1$. Ovo znači da je rang idempotentne matrice jednak broju sopstvenih vrednosti koje imaju vrednost 1. Što se tiče specijalne dekompozicije dijagonalna matrica Λ na dijagonali ima samo jedinice i nule, a trag te matrice je jednak broju jedinica. Na osnovu toga imamo da je

$$trA = tr\Gamma^T \Lambda \Gamma = tr\Lambda\Gamma\Gamma^T = tr\Lambda \mathrm{E} = tr\Lambda = \mathrm{rangA}.$$

Stav 5.4. Neka je $X = (X_1, \dots, X_n)$ $n-$ dimenzioni slučajni vektor, gde su $X_1, \dots, X_n$ nezavisne slučajne veličine iz raspodele $N(0, \sigma^2)$. Neka je $Q = \sigma^{-2} X^T A X$ kvadratna forma sa simetričnom matricom A čiji je rang r. Tada Q ima $\chi^2 -$ raspodelu sa r stepeni slobode ako i samo ako je matrice A idempotentna.

Dokaz se može naći u literaturi [8] na strani 505.

Stav 5.5. Neka je $X = (X_1, \dots, X_n)$ $n-$ dimenzioni slučajni vektor, gde su $X_1, \dots, X_n$ nezavisne slučajne veličine iz raspodele $N(0, \sigma^2)$. Neka su $Q_1 = \sigma^{-2} X^T A X$ i $Q_2 = \sigma^{-2} X^T B X$ kvadratne forme sa simetričnim matricama A i B. Slučajne veličine Q_1 i Q_2 su nezavisne ako i samo ako je $AB = 0$.

Dokaz se može naći u literaturi [8] na strani 508.

5.5 Analiza glavnih komponenata i matrica kovarijacije

Analiza glavnih komponenata proučava matricu kovarijacije skupa slučajnih veličina na osnovu njihovih linearnih kombinacija. Iako je p slučajnih veličina izabrano kako bi se opisala varijabilnost celog sistema, često je slučaj da manji broj k slučajnih veličina opisauje veliki deo tog varibiliteta. Te slučajne veličine nazivaju se glavnim komponentama i one zapravo nose istu količinu informacija o sistemu kao i početnih p posmatranih slučajnih veličina.

Ovo nam pomaže da redukujemo broj podataka, jer se početni skup koji se sastoji od n merenja na p slučajnih veličina, može sveti na n merenja na k slučajnih veličina, pri čemu je naravno $k < p$. Takođe, ova analiza proučava odnos između slučajnih veličina, pa nam tako omogućava interpretacije do kojih inače ne bismo dosli.

Glavne komponente su linearne kombinacije slučajnih veličina, pa su i one same slučajne veličine.

Neka slučajni vektor $X = (X_1, \dots, X_p)$ ima matricu kovarijacije Σ sa sopstvenim vrednostima $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$. Posmatrajmo linearne kombinacije:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$

$$\dots$$

$$Y_p = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p.$$

Ako uvedemo sledeće oznake: $a_i = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{ip} \end{bmatrix}$ i $X = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}$, onda je $Y_i = {a_i}^T X$.

Na osnovu osobina matrice kovarijacije pretstavljenih u poglavlju 5.2, važi da je

$$D(Y_i) = cov({a_i}^T X) = {a_i}^T cov(X) a_i = {a_i}^T \Sigma a_i \quad i = 1, \dots, p$$

$$cov(Y_i, Y_j) = cov\left({a_i}^T X, {a_j}^T X\right) = {a_i}^T cov(X) a_j = {a_i}^T \Sigma a_j \quad i, j = 1, \dots, p$$

Glavne komponente su slučajne veličine $Y_1, Y_2, \dots, Y_p$ čije su disperzije što je moguće veće. Obzirom da se disperzija vektora $D(Y_i) = {a_i}^T \Sigma a_i$ može povećati množenjem vektora sa konstantom, ograničavamo se na vektore dužine 1.

Prva glavna komponenta je slučajna veličina $Y_1 = {a_1}^T X$ koja maksimizira vrednost $D(Y_1) = {a_1}^T \Sigma a_1$, uz uslov da je ${a_1}^T a_1 = 1$. Druga glavna komponenta je slučajna veličina $Y_2 = {a_2}^T X$ koja maksimizira vrednost $D(Y_2) = {a_2}^T \Sigma a_2$, uz uslove da je ${a_2}^T a_2 = 1$ i $cov({a_1}^T X, {a_2}^T X) = 0$. Shodno tome, $i -$ta glavna komponenta je slučajna veličina

$Y_i = a_i{}^T X$ koja maksimizira vrednost $D(Y_i) = a_i{}^T \Sigma a_i$, uz uslove da je $a_i{}^T a_i = 1$ i $cov\left(a_i{}^T X, a_j{}^T X\right) = 0$ za $j < i$.

Stav 5.6. Neka je $X = \left(X_1, \dots, X_p\right)$ $p -$ dimenzioni slučajni vektor sa matricom kovarijacije Σ. Parovi sopstvenih vrednosti i sopstvenih vektora matrice Σ su $(\lambda_1, e_1), \dots, \left(\lambda_p, e_p\right)$, pri čemu je $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$. Tada je i $-$ta glavna komponenta data sa

$$Y_i = e_i{}^T X = e_{i1} X_1 + e_{i2} X_2 + \cdots + e_{ip} X_p \quad i = 1, \dots, p,$$

uz takav izbor da je

$$D(Y_i) = \ e_i{}^T \Sigma e_i = \lambda_i \quad i = 1, \dots, p,$$

$$cov\left(Y_i, Y_j\right) = e_i{}^T \Sigma e_j = 0 \quad i \neq j \,.$$

Ako su neke sopstvene vrednosti međusobno jednake, onda izbor slučajnih vreličina Y_i nije jednoznačan.

Dokaz se može naći u literaturi [10] na strani 17.

Stav 5.7. Neka je $X = \left(X_1, \dots, X_p\right)$ $p -$ dimenzioni slučajni vektor sa matricom kovarijacije Σ. Parovi sopstvenih vrednosti i sopstvenih vektora matrice Σ su $(\lambda_1, e_1), \dots, \left(\lambda_p, e_p\right)$, pri čemu je $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$. Neka su $Y_i = e_i{}^T X = e_{i1} X_1 + e_{i2} X_2 + \cdots + e_{ip} X_p \quad i = 1, \dots, p$, glavne komponente. Tada važi:

$$\sigma_{11} + \cdots + \sigma_{pp} = \sum_{i=1}^{p} D(X_i) = \lambda_1 + \cdots + \lambda_p = \sum_{i=1}^{p} D(Y_i).$$

Dokaz se može naći u literaturi [10] na strani 17.

6 Matrice u metodi najmanjih kvadrata

Jedan od osnovnih zadataka statistike je ocenjivanje parametara u raspodeli posmatranog obeležja. Metoda najmanjih kvadrata je jedna od opštih metoda zajedno sa metodom maksimalne verodostojnosti i metodom momenata. Kada dođemo do neke ocene, onda se postavlja pitanje kvaliteta te ocene, te će u ovom poglavlju biti reči i o tome.

Teorema Gaus-Markova[7] se odnosi na linearne statističke modele koji se pomoću matrica mogu zapisati u obliku

$$X = A\theta + \varepsilon,$$

gde je θ matrica parametara koje treba oceniti.

Diferenciranje je sastavni deo metode najmanjih kvadrata, a kako je ovde parameter matrica, jedna od neophodnih tema je i diferenciranje matrica.

Teorema Gaus-Markova je veoma važna u primenama. Već na osnovu same postavke problema

$$X = A\theta + \varepsilon$$

pod uslovom da raspolažemo dovoljnim brojem informacija (tj. da je rang matrice modela potpun) možemo zaključiti koja je linearna ocena nepoznatog parametra $\theta = (\theta_1, \theta_2, \dots \theta_n)$ najefikasnija tj. koja linearna ocena ima najmanju disperziju. To je ona ocena koja se dobije metodom najmanjih kvadrata i oblika je

$$\hat{\theta} = (A^T A)^{-1} A^T X.$$

6.1 Metoda najmanjih kvadrata

Metoda najmanjih kvadrata koristi se u mnogo različitih situacija. Princip metode najmanjih kvadrata možemo objasniti na primeru linearne regresije.

Pretpostavimo da imamo dve veličine x i y. Postoje dve mogućnosti: te veličine su nezavisne ili zavisne. Ako su nezavisne onda bilo kojoj vrednosti veličine x odgovara bilo koja vrednost veličine y. Ako su pak zavisne, onda vrednost veličine x na neki način definiše koje su to moguće vrednosti veličine y. Oblici zavisnosti mogu biti različiti. Kada god primetimo da za jednake promene veličine x dolazi do jednakih promena veličine y, možemo posumnjati na linearnu zavisnost, koja se zapisuje u obliku $y = f(x) = ax + b$. Ona je potpuno određena ako su nam poznati brojevi a i b.

[7] Андре́й Андре́евич Ма́рков (1856-1922), ruski matematičar

Brojeve a i b određujemo po principu najmanjih kvadrata. Načelo ovog principa je da su najbolji oni parametri a i b za koje je suma kvadrata razlika između stvarnih (izmerenih) vrednosti $y_i (i = 1, \dots, n)$ i izračunatih vrednosti $f(x_i) = ax_i + b \ (i = 1, \dots, n)$ minimalna. To znači da parametre a i b određujemo iz uslova da izraz

$$\sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} \big(y_i - (ax_i + b)\big)^2$$

ima najmanju moguću vrednost.

Malo je drugačije ako metodu najmanjih kvadrata želimo da primenimo na matrice. Posmatramo sistem linearnih jednačina $Ax = b$ u kojem imamo više jednačina nego promenljivih, a pritom jednačine nisu jednake niti proporcionalne. U tom slučaju sistem nema rešenje, ali možemo napraviti neku ocenu približnog rešenja. Ako bi sistem bio rešiv onda bi važilo $Ax - b = 0$. Međutim, sistem nije rešiv, ali je prirodno zahtevati da razlika $Ax - b$ bude što manja tj. da norma vektora $\|Ax - b\|$ bude što manja.

$$\|Ax - b\| \longrightarrow min \ \Longrightarrow \|Ax - b\|^2 \longrightarrow min$$

Ako je vektor x predstavljen u matričnom obliku tj. $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, onda se kvadrat norme vektora može izraziti u oblik $\|x\|^2 = x^T x$. Pokažimo da je zaista tako,

$x = (x_1, \dots, x_n)$,

$\|x\| = \sqrt{{x_1}^2 + \cdots + {x_n}^2}$,

$\|x\|^2 = {x_1}^2 + \cdots + {x_n}^2$.

Sa druge strane imamo,

$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$,

$x^T x = [x_1 \quad \cdots \quad x_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = {x_1}^2 + \cdots + {x_n}^2$,

tako da je zaista $\|x\|^2 = x^T x$.

Koristeći ovu činjenicu, uslov metode najmanjih kvadrata se svodi na

$$(Ax - b)^T (Ax - b) \longrightarrow min.$$

6.2 Metoda najmanjih kvadrata u Matlab-u

Matlab ima ugrađenu funkciju koja izračunava polinom za aproksimiranje podataka metodom najmanjih kvadrata. U pitanju je funkcija $polyfit(x, y, n)$ gde su $x = (x_1, x_2, \dots, x_m)$ i $y = (y_1, y_2, \dots, y_m)$ vektori poznatih vrednosti, a n je stepen polinoma kojim se vrši aproksimacija. Rezultat funkcije je vektor koeficijenata $a = (a_n, a_{n-1}, \dots, a_0)$ tako da važi sledeća veza

$$y = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + a_n x^n.$$

Ako pretpostavimo da među promenljivima x i y postoji linearna veza, onda je traženi polinom oblika

$$y = a_0 + a_1 x,$$

a koeficijenti a_0 i a_1 se dobijaju pozivanjem funkcije $polyfit(x, y, 1)$.

Primer: U prvoj kolini su vrednosti nezavisno promenljive, a u drugoj koloni izmerene vrednosti zavisno promenljive.

1	12.9699	8.79636
2	12.9806	8.81486
3	13.0033	9.00868
4	13.0265	9.10698
5	13.0497	9.18665
6	13.0693	9.28436
7	13.0918	9.39767
8	13.1266	9.58123
9	13.1459	9.69405
10	13.1664	9.77536

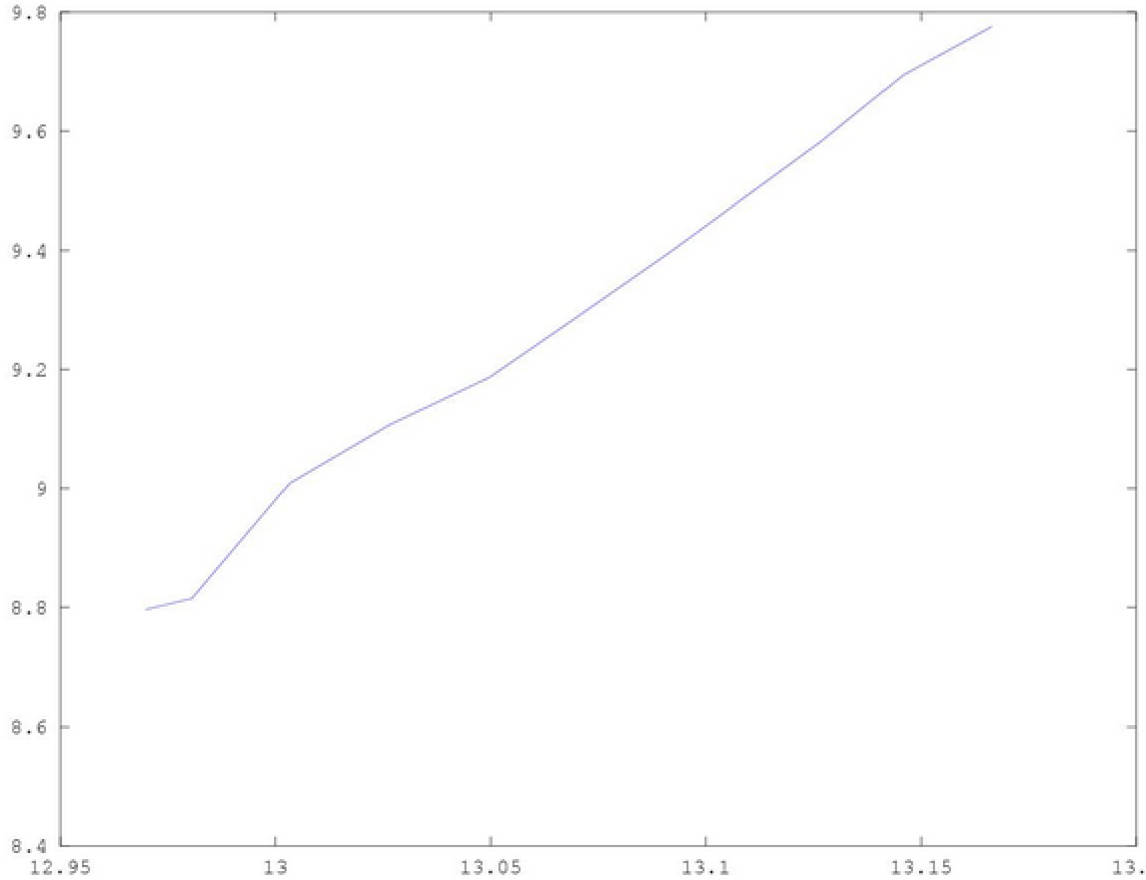
Smestimo ove vrednosti u dva vektora x i y.

x = [12.9699 12.9806 13.0033 13.0265 13.0497 13.0693 13.0918 13.1266 13.1459 13.1664]

y = [8.79636 8.81486 9.00868 9.10698 9.18665 9.28436 9.39767 9.58123 9.69405 9.77536]

Za grafičko prikazivanje podataka koristimo ugrađenu funkciju plot().

$$plot(x, y)$$



Na osnovu izgleda grafika, možemo pretpostaviti da postoji linearna zavisnost između vrednosti x i y.

$$polyfit(x, y, 1)$$

$$ans\ = 5.0043 \quad -56.1067$$

Na osnovu rezultata zaključujemo da je $a_1 = 5.0043$ i $a_0 = -56.1067$, tako da se linearna zavisnost može izraziti jednačinom $y = -56.1067 + 5.0043x$.

Pokažimo na koji još način metoda najmanjih kvadrata ima upotrebu u Matlab-u. Razmotrimo sistem $Ax = b$ gde je matrica sistema A formata $m \times n$, pri čemu je $m > n$. Takođe, x je matrica nepoznatih formata $n \times 1$ i b je matrica slobodnih članova formata $m \times 1$. U tom slučaju ne postoji jedinstveno rešenje, jer imamo više jednačina nego promenljivih. Ako nasumično izaberemo n jednačina od mogućih m dobićemo jedno rešenje za vektor x. Ako izaberemo nekih drugih n jednačina dobićemo drugačije rešenje za vektor x. U Matlab-u upotrebom levog deljenje dobićemo ono rešenje koje zadovoljava uslov najmanjih kvadrata.

Primer: Rešiti sistem linearnih jednačina.

$$2x - y = 1$$

$$x + 10y = 1$$

$$x + 2y = 1$$

Matrica sistema je $A = \begin{bmatrix} 2 & -1 \\ 1 & 10 \\ 1 & 2 \end{bmatrix}$, a matrica slobodnih članova $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

$$A \ = [2 - 1; 1\ 10; 1\ 2]$$

$$b \ = \ [1\ 1\ 1]$$

Rešenje sistema se dobija pozivanjem sledeće naredbe: $A * transpose(b)$.

$$ans = \begin{matrix} 0.584906 \\ 0.049057 \end{matrix}$$

6.3 Kvalitet ocene parametara

Neka je $(X_1, \dots, X_n)$ uzorak iz familije dopustivih verovatnoća $\mathcal{P} = \{F(x, \theta) : \theta \in \Theta\}$ i neka je statistika $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ ocena napoznatog parametra θ.

Definicija 6.1. Statistika $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ je postojana (konzistentna u verovatnoći) ocena parametra θ ako $\hat{\theta} \to \theta$ u verovatnoći kada $n \to \infty$, tj.

$$(\forall \varepsilon > 0) \lim_{n \to \infty} P\{|\hat{\theta} - \theta| > \varepsilon\} = 0.$$

Definicija 6.2. Statistika $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ je jako postojana (jako konzistentna) ocena parametra θ ako $\hat{\theta} \to \theta$ skoro sigurno kada $n \to \infty$, tj.

$$P\left\{\lim_{n \to \infty} \hat{\theta} = \theta\right\} = 1.$$

Definicija 6.3. Statistika $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ je postojana (konzistentna) u srednjekvadratnom smislu ocena parametra θ ako $\hat{\theta} \to \theta$ u srednje kvadratnom kada $n \to \infty$, tj.

$$\lim_{n \to \infty} E\left((\hat{\theta} - \theta)^2\right) = 0.$$

Definicija 6.4. Statistika $\boldsymbol{\hat{\theta}} = \boldsymbol{\hat{\theta}}(\boldsymbol{X_1}, \dots, \boldsymbol{X_n})$ je nepristrasna (centrirana) ocena parametra $\boldsymbol{\theta}$ ako za svako $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ važi jednakost

$$E(\hat{\theta}) = \theta.$$

Veličina $\delta(\theta) = E(\hat{\theta}) - \theta$ zove se sistematsko odstupanje (pristrasnost) ocene $\hat{\theta}$.

Definicija 6.5. Statistika $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}(\boldsymbol{X_1}, \dots, \boldsymbol{X_n})$ je asimptotski nepristrasna (asimptotski centrirana) ocena parametra $\boldsymbol{\theta}$ ako za svako $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ važi

$$\lim_{n\to\infty} \boldsymbol{E}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{\theta}.$$

Od dve nepristrasne ocene efikasnija je ona koja ima manju disperziju.

6.4 Matrično diferenciranje

U metodi najmanjih kvadrata tražimo minimum funkcije, tako da je diferenciranje neophodan deo postupka. Međutim, na primeru koji razmatramo u ovom radu nije moguće primeniti diferenciranje funkcije jedne ili vise promenljivih, već je neophodno dati uopštenje pojma diferenciranja. Zato ćemo se u ovom poglavlju baviti matričnim diferenciranjem.

Definicija 6.6. Neka je $\boldsymbol{U} \subset \mathbb{R}^{\boldsymbol{n}}$ otvoren skup i $\boldsymbol{a} = (\boldsymbol{a_1}, \dots, \boldsymbol{a_n}) \in \boldsymbol{U}$ fiksirana tačka. Za funkciju $\boldsymbol{f}: \boldsymbol{U} \to \mathbb{R}^{\boldsymbol{m}}$ kažemo da je jako diferencijabilna u tački $\boldsymbol{a}$ ako postoji ograničeno linearno preslikavanje $\boldsymbol{A}: \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ tako da za $\boldsymbol{h} = (\boldsymbol{h_1}, \dots, \boldsymbol{h_n}) \in \mathbb{R}^{\boldsymbol{n}}$ važi

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} = 0. \qquad (1)$$

Stav 6.1. Ograničeno linearno preslikavanje $\boldsymbol{A}$ je jedinstveno uvek kada postoji.

Dokaz: Pretpostavimo da postoji još neko ograničeno linearno preslikavanje $B: \mathbb{R}^n \to \mathbb{R}^m$ za koje važi

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Bh\|}{\|h\|} = 0. \qquad (2)$$

$$\begin{aligned} \lim_{h\to 0} \frac{\|(A-B)h\|}{\|h\|} &= \lim_{h\to 0} \frac{\|f(a+h) - f(a) - Bh - f(a+h) + f(a) + Ah\|}{\|h\|} \\ &= \lim_{h\to 0} \frac{\|f(a+h) - f(a) - Bh - (f(a+h) - f(a) - Ah)\|}{\|h\|} \\ &\leq \lim_{h\to 0} \frac{\|f(a+h) - f(a) - Bh\|}{\|h\|} + \lim_{h\to 0} \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} = 0 \qquad (3) \end{aligned}$$

Na osnovu jednakosti (1) i (2).

$$\|(A-B)h\| = \|h\| \frac{\|(A-B)th\|}{\|th\|} \quad \text{za } t > 0, h \neq 0$$

Ako $t \to 0$, onda i $\|th\| \to 0$, pa je

$$\|(A-B)h\| = \lim_{t\to 0} \|h\| \frac{\|(A-B)th\|}{\|th\|} = \|h\| \lim_{\|th\|\to 0} \frac{\|(A-B)th\|}{\|th\|} = 0$$

na osnovu jednakosti (3). To znači da je $\|(A-B)h\| = 0$ za svako h iz $\mathbb{R}^n$, odnosno $Ah = Bh$, a time i $A = B$.■

Definicija 6.8. Ograničeno linearno preslikavanje $\boldsymbol{A}$ nazivamo jakim (Frešeovim[8]) izvodom funkcije $\boldsymbol{f}$ u tački $\boldsymbol{a}$ i označavamo sa $\boldsymbol{f}'(\boldsymbol{a})$. Sam vektor $\boldsymbol{f}'(\boldsymbol{a})\boldsymbol{h}$ nazivamo vrednošću jakog (Frešeovog) diferencijala funkcije $\boldsymbol{f}$ u tački $\boldsymbol{a}$ u pravcu vektora $\boldsymbol{h}$.

Jednakost (1) se može napisati i na drugi način

$$f(a+h) - f(a) = f'(a)h + o(h) \quad \text{kada} \quad h \to 0 \qquad (3)$$

ili

$$(\forall \varepsilon > 0)(\exists \delta > 0)\|f(a+h) - f(a) - f'(a)h\| < \varepsilon\|h\| \quad \text{za sve } \|h\| < \delta.$$

Stav 6.2. Neka je dato preslikavanje

$$f: \mathbb{R}^n \to \mathbb{R}^m : x \longmapsto c.$$

Tada je f jako diferencijabilno i važi $f'(a) = 0$ za svako $a \in \mathbb{R}^n$.

Dokaz:

$f(a+h) - f(a) = c - c = 0$

$f'(a)h + o(h) = 0$ kada $h \to 0$

Odavde sledi da je $f'(a) = 0$ za svako $a \in \mathbb{R}^n$. ■

Stav 6.3. Neka je data matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{m} \times \boldsymbol{n}, \mathbb{R})$ i preslikavanje

$$f: \mathbb{R}^n \to \mathbb{R}^m : x \longmapsto Ax.$$

Tada je f jako diferencijabilno za svako $a \in \mathbb{R}^n$ i važi $f'(a) = A$ za svako $a \in \mathbb{R}^n$.

Dokaz:

$f(a+h) - f(a) = A(a+h) - Aa = Aa + Ah - Aa = Ah$

$f'(a)h + o(h) = Ah$ kada $h \to 0$

Odavde sledi da je $f'(a) = A$ za svako $a \in \mathbb{R}^n$.■

Definicija 6.9. Neka je $\boldsymbol{U} \subset \mathbb{R}^{\boldsymbol{n}}$ otvoren skup i $\boldsymbol{a} = (\boldsymbol{a_1}, \dots, \boldsymbol{a_n}) \in \boldsymbol{U}$ fiksirana tačka. Za funkciju $\boldsymbol{f}: \boldsymbol{U} \to \mathbb{R}^{\boldsymbol{m}}$ kažemo da je slabo (Gato[9]) diferencijabilna u tački $\boldsymbol{a}$ u pravcu vektora $\boldsymbol{h} = (\boldsymbol{h_1}, \dots, \boldsymbol{h_n}) \in \mathbb{R}^{\boldsymbol{n}}$, ako postoji limes

[8] Maurice Rene Frechet (1878-1973), francuski matematičar
[9] Rene Gateaux (1889-1914), francuski matematičar

$$Df(a,h) \overset{\text{def}}{=} \lim_{t\to 0}\frac{f(a+th)-f(a)}{t}, \quad (4)$$

tj. ako važi

$$\lim_{t\to 0}\left\|\frac{f(a+th)-f(a)}{t} - Df(a,h)\right\| = 0. \quad (5)$$

Izraz $Df(a,h)$ naziva se slabi tj. Gatoov diferencijal.

Izraz (4) može se napisati na još jedan način

$$f(a+th)-f(a) = Df(a,h)t + o(t) \quad \text{kada } t\to 0 \quad (6)$$

Slabi diferencijal $Df(a,h)$ ne mora biti linearan po h. Ali ako jeste, onda važi sledeća definicija.

Definicija 6.10. Ako je $\boldsymbol{Df(a,h) = Ah}$ gde je $\boldsymbol{A}$ ograničeno linearno preslikavanje skupa $\mathbb{R}^{\boldsymbol{n}}$ u skup $\mathbb{R}^{\boldsymbol{m}}$, onda je to $\boldsymbol{A}$ slabi (Gatoov) izvod funkcije $\boldsymbol{f}$ u tački $\boldsymbol{a}$ i označava se sa $\boldsymbol{f'_w(a)}$.

Drugim rečima, ograničeno linearno preslikavanje $f'_w(a)$ jednoznačno je određeno sa

$$\lim_{t\to 0}\left\|\frac{f(a+th)-f(a)}{t} - f'_w(a)h\right\| = 0$$

za sve $h \in \mathbb{R}^n$.

Stav 6.4. Ako preslikavanje f ima jak izvod u tački a, onda to preslikavanje ima i slabi izvod u istoj tački i ti izvodi se poklapaju tj. $f'_w(a) = f'(a)$.

Dokaz: Iz definicije jake diferencijabilnosti sledi da $a+th \to a$ kada $t\to 0$ za svako $h\in\mathbb{R}^n$, pa je

$$\lim_{t\to 0}\left\|\frac{f(a+th)-f(a)}{t} - f'(a)h\right\| = \|h\|\lim_{t\to 0}\frac{\|f(a+th)-f(a)-f'(a)(a+th-a)\|}{\|a+th-a\|}$$
$$= 0$$

Kada ovo uporedimo sa definicijom slabe diferencijabilnosti, vidimo da je $f'_w(a)h = f'(a)h$, odnosno $f'_w(a) = f'(a)$.■

Obrnuto ne mora da važi. Ali da bi smo to dokazali, trebaće nam sledeća lema i njena posledica.

Lema 6.5. Ukoliko je $U\subset\mathbb{R}^n$ otvoren skup koji sadrži odsečak $[\![a,b]\!] \overset{\text{def}}{=} \{a+\alpha(b-a) : 0 \le \alpha \le 1\}$ i ako je funkcija $f: U\to\mathbb{R}^m$ slabo diferencijabilna u svakoj tački odsečka $[\![a,b]\!]$, tada je

$$\|f(b)-f(a)\| \le \sup_{x\in[\![a,b]\!]}\|f'_w(x)\|_{B(\mathbb{R}^n,\mathbb{R}^m)}\|b-a\|$$

gde je $B(\mathbb{R}^n, \mathbb{R}^m)$ prostor ograničenih linearnih preslikavanje iz $\mathbb{R}^n$ u $\mathbb{R}^m$.

Dokaz se može naći u literaturi [1] na strani 319.

Posledica 6.6. Pod uslovima leme 6.5. važi

$$\|f(b) - f(a) - f'_w(a)(b-a)\| \leq \sup_{x \in [\![a,b]\!]} \|f'_w(x) - f'_w(a)\|_{B(\mathbb{R}^n,\mathbb{R}^m)} \|b-a\|.$$

Stav 6.7. Ako je $U \subset \mathbb{R}^n$ otvoren skup, a funkcija $f: U \to \mathbb{R}^m$ slabo diferencijabilna na U tako da je njen slab izvod $f'_w: U \to B(\mathbb{R}^n, \mathbb{R}^m)$ neprekidan u tački a, tada je f i jako diferencijabilna u a i važi

$$f'(a) = f'_w(a).$$

Dokaz: Kako je f'_w neprekidno preslikavanje u tački a, to važi da za svako $\varepsilon > 0$ postoji $\delta > 0$ tako da je $\|f'_w(x) - f'_w(a)\| < \varepsilon$ uvek kada je $\|x-a\| < \delta$ i za svako $x \in \mathbb{R}^n$. Za svako $z \in [\![a,x]\!]$ je

$$\|z-a\| = \|a + \alpha(x-a) - a\| = \|\alpha(x-a)\| = |\alpha|\|x-a\| \leq \|x-a\| < \delta,$$

jer je $0 \leq \alpha \leq 1$.

Odatle sledi

$$\|f(x) - f(a) - f'_w(a)(x-a)\| \leq \sup_{z \in [\![a,x]\!]} \|f'_w(z) - f'_w(a)\|_{B(\mathbb{R}^n,\mathbb{R}^m)} \|x-a\| \leq \varepsilon\|x-a\|.$$

Ovo znači da je $\lim_{x\to a} \frac{\|f(x)-f(a)-f'_w(a)(x-a)\|}{\|x-a\|} = 0$, odnosno preslikavanje f je jako diferencijabilno u tački a i $f'(a) = f'_w(a)$. ■

Ako je $f: \mathbb{R}^n \to \mathbb{R}^m$ diferencijabilno preslikavanje, onda je njegov jak izvod f' preslikavanje prostora $\mathbb{R}^n$ na prostor $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$. Ako je preslikavanje $f': \mathbb{R}^n \to \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ diferencijabilno u nekoj tački $a \in \mathbb{R}^n$, tada njegov izvod nazivamo drugim izvodom preslikavanja f u tački a i obeležavamo sa $f''(a)$. To znači da je $f''(a) \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ uvek kada postoji. Sledeći stav pokazuje da se elementi prostora $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ mogu predstaviti preglednije pomoću bilinearnih preslikavanja.

Stav 6.8. Svakom elementu $A \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ može se dodeliti element iz $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$ formulom

$$B(x, x') \stackrel{\text{def}}{=} (Ax)x'$$

za sve $x, x' \in \mathbb{R}^n$, koji uspostavlja izometrički izomorfizam ova dva prostora.

Dokaz: Funkcija koja svakom elementu iz prostora $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ dodeljuje element iz prostora $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$ po formuli $B(x, x') \stackrel{\text{def}}{=} (Ax)x'$ je linearna. Kako je

$\|B(x,x')\| = \|(Ax)x'\| \leq \|Ax\|\|x'\| \leq \|A\|\|x\|\|x'\|$ za sve $x, x' \in \mathbb{R}^n$, možemo zaključiti da je $\|B\| \leq \|A\|$.

Kako je $\|Ax\| = \sup_{\|x'\|\leq 1}\|(Ax)x'\| = \sup_{\|x'\|\leq 1}\|B(x,x')\| \leq \|B\|\|x\|$ za sve $x \in \mathbb{R}^n$, tako da je $\|A\| \leq \|B\|$.

Na osnovu prethodne dve nejednakosti, možemo zaključiti da je $\|A\| = \|B\|$. Korespodencija između $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ i $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$ koja je zadata formulom $B(x,x') \stackrel{\text{def}}{=} (Ax)x'$ je linearna, izometrična i "na", pa je i uzajamno jednoznačna. ■

Na osnovu ovog stava element $f''(a) \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ možemo takođe smatrati elementom prostora $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$.

Stav 6.9. Neka je data matrica $A \in M(m \times n, \mathbb{R})$ i preslikavanje

$$f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}: (x,y) \longmapsto y^T Ax.$$

Tada je f jako diferencijabilno za svako $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$ i za svako $a, h \in \mathbb{R}^n$ i svako $b, k \in \mathbb{R}^m$ važi

$$f'(a,b)(h,k) = b^T Ah + k^T Aa.$$

Dokaz:

$$\begin{aligned}
&|f(a+h, b+k) - f(a,b) - b^T Ah - k^T Aa| \\
&\qquad = |(b+k)^T A(a+h) - b^T Aa - b^T Ah - k^T Aa| \\
&\qquad = |(b^T + k^T)A(a+h) - b^T Aa - b^T Ah - k^T Aa| \\
&\qquad = |(b^T A + k^T A)(a+h) - b^T Aa - b^T Ah - k^T Aa| \\
&\qquad = |b^T Aa + b^T Ah + k^T Aa + k^T Ah - b^T Aa - b^T Ah - k^T Aa| = |k^T Ah| \\
&\qquad \leq \|k^T\|\|A\|\|h\| \leq \frac{1}{2}\|A\|(\|k\|^2 + \|h\|^2) = \frac{1}{2}\|A\|\|(h,k)\|^2 = o(\|(h,k)\|)
\end{aligned}$$

kada $\|(h,k)\| \to 0$. ■

Stav 6.10. Neka je data matrica $\boldsymbol{A} \in \boldsymbol{M}(\boldsymbol{n} \times \boldsymbol{n}, \mathbb{R})$ i preslikavanje

$$f: \mathbb{R}^n \to \mathbb{R}: x \longmapsto x^T Ax.$$

Tada je f jako diferencijabilno za svako $a \in \mathbb{R}^n$ i za svako $a, h \in \mathbb{R}^n$ važi

$$f'(a)h = a^T(A + A^T)h.$$

Dokaz:

$$
\begin{aligned}
|f(a+h)-f(a)-a^T(A+A^T)h| &= |(a+h)^TA(a+h)-a^TAa-a^T(Ah+A^Th)| \\
&= |(a^T+h^T)A(a+h)-a^TAa-a^TAh-a^TA^Th| \\
&= |(a^TA+h^TA)(a+h)-a^TAa-a^TAh-a^TA^Th| \\
&= |a^TAa+a^TAh+h^TAa+h^TAh-a^TAa-a^TAh-a^TA^Th| = |h^TAh| \\
&\leq \|A\|\|h\|^2 = o(\|h\|)
\end{aligned}
$$

kada $\|h\| \to 0$.

Naime, kako je $h^TAa \in \mathbb{R}$, to je $h^TAa = (h^TAa)^T = (Aa)^Th = a^TA^Th$. ■

Posledica 6.11. Ako je matrica $A \in M(n \times n, \mathbb{R})$ iz stava 6.10. simetrična, onda za svako $a \in \mathbb{R}^n$ važi

$$f'(a)h = 2a^TAh.$$

Ako je matrica A simetrična onda je $A^T = A$. Kada to primenimo u prethodnoj teoremi dobijamo

$$f'(a)h = a^T(A^T + A)h = a^T(A + A)h = 2a^TAh.$$

Posledica 6.12. Neka je dato preslikavanje

$$f: \mathbb{R}^n \to \mathbb{R}: x \longmapsto x^Tx.$$

Tada je f jako diferencijabilno za svako $a \in \mathbb{R}^n$ I za svako $a, h \in \mathbb{R}^n$važi

$$f'(a)h = 2a^Th.$$

Ako u stavu 6.10. izaberemo da je matrica A jedinična, onda važi

$$f'(a)h = a^T(E^T + E)h = 2a^TEh = 2a^Th.$$

Jaki diferencijal preslikavanja $f: \mathbb{R}^n \to \mathbb{R}^m$ definisali smo kao sliku elementa $h \in \mathbb{R}^n$ pri dejstvu ograničenog linearnog preslikavanja $f'(a) \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ tj. $df \overset{\text{def}}{=} f'(a)h$. Slično tome, diferencijal drugog reda se definiše sa $d^2f \overset{\text{def}}{=} f''(a)(h,h)$ odnosno kao slika elementa $(h,h) \in \mathbb{R}^n \times \mathbb{R}^n$ pri dejstvu ograničenog bilinearnog preslikavanja $f''(a) \in \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$. Po analogiji definišu se i diferencijali višeg reda.

Teorema 6.13. (Tejlorova[10] formula) Neka je $U \subset \mathbb{R}^n$ otvoren skup i neka je funkcija $f: U \to \mathbb{R}^m$ diferencijabilna $n-1$ puta na U i n puta diferencijabilna u tački $a \in U$. Tada važi

$$
\begin{aligned}
f(x)-f(a)-f'(a)(x-a)-\frac{1}{2!}f''(a)(x-a,x-a)-\cdots-\frac{1}{n!}f^{(n)}(a)(x-a,\ldots,x-a) \\
= o(\|x-a\|^n)
\end{aligned}
$$

kada $x \to a$.

[10]Brook Taylor (1685-1731)

Dokaz se može naći u literaturi [1] na strani 322.

Definicija 6.11. Funkcional $f: \mathbb{R}^n \to \mathbb{R}$ dostiže lokalni minimum u tački $a \in \mathbb{R}^n$ ako postoji $\varepsilon > 0$ tako da je $f(x) \geq f(a)$ za svako $x \in \mathbb{R}^n$ za koje važi $\|x - a\| < \varepsilon$.

Stav 6.14. Da bi jako diferencijabilni funkcional $f: \mathbb{R}^n \to \mathbb{R}$ imao u tački a lokalni ekstremum, neophodno je da bude $f'(a) = 0$.

Dokaz: Uočimo funkciju $g: \mathbb{R} \to \mathbb{R}: t \mapsto f(a + th)$. Ako funkcional f ima lokalni ekstremum u tački a, onda funkcija g ima lokalni ekstremum u tački 0 za svako $h \in \mathbb{R}^n$. Funkcija g je diferencijabilna u tački 0, pa je

$$g'(0) = f'(a)h = 0.$$

Za svako $h \in \mathbb{R}^n$. Odavde sledi da je $f'(a) = 0$ u tački a lokalnog ekstremuma.■

Stav 6.15. Neka je funkcional $f: \mathbb{R}^n \to \mathbb{R}$ diferencijabilan u okolini tačke a i ima drugi izvod u tački a. Ako je a tačka lokalnog minimuma za f, tada je $d^2 f(a) \geq 0$.

Dokaz: Krenimo od Tejlorove formule

$f(x) - f(a) - f'(a)(x - a) - \frac{1}{2!} f''(a)(x - a, x - a) = o(\|x - a\|^2)$ kada $x \to a$.

Kako $a + th \to a$ kada $t \to 0$ za svako $h \in \mathbb{R}^n$, važi sledeće

$f(a + th) - f(a) - f'(a)(a + th - a) - \frac{1}{2!} f''(a)(a + th - a, a + th - a) = o(\|a + th - a\|^2)$ kada $t \to 0$

$f(a + th) - f(a) - tf'(a)h - \frac{t^2}{2!} f''(a)(h, h) = o(t^2 \|h\|^2)$ kada $t \to 0$

$$\lim_{t \to 0} \frac{\left\| f(a + th) - f(a) - tf'(a)h - \frac{t^2}{2!} f''(a)(h, h) \right\|}{t^2 \|h^2\|} = 0.$$

Kako je a tačka lokalnog minimuma, to je $f'(a) = 0$, pa sledi

$$\frac{1}{2!} f''(a)(h, h) = \lim_{t \to 0} \frac{f(a + th) - f(a)}{t^2}$$

$\lim_{t \to 0} \frac{f(a+th) - f(a)}{t^2} \geq 0$ jer je a tačka lokalnog minimuma, pa je time i

$f''(a)(h, h) = d^2 f(a) \geq 0.$ ■

Stav 6.16. Ako je $f: \mathbb{R}^n \to \mathbb{R}$ diferencijabilan funkcional tako da je $df(a) = 0$, a $d^2 f(a)$ jako pozitivan kvadratni funkcional, tada je a tačka lokalnog minimum funkcionala f.

Dokaz: Kako je $d^2 f(a)$ pozitivan kvadratni funkcional, to postoji konstanta $c > 0$ tako da je

$$d^2f(a) = f''(a)(x-a, x-a) \geq c\|x-a\|^2$$

za sve $x \in \mathbb{R}^n$. Primenom Tejlorove formule i uslova da je $df(a) = 0$, za neko $\delta > 0$ dobijamo

$$\left|f(x) - f(a) - \tfrac{1}{2}f''(a)(x-a, x-a)\right| < \tfrac{c}{4}\|x-a\|^2 \quad \text{za } \|x-a\| < \delta.$$

Koristeći prethodne dve nejednakosti, dobijamo

$$\begin{aligned} f(x) - f(a) &\geq \frac{1}{2}f''(a)(x-a, x-a) - \left|f(x) - f(a) - \frac{1}{2}f''(a)(x-a, x-a)\right| \\ &\geq \frac{c}{2}\|x-a\|^2 - \frac{c}{4}\|x-a\|^2 = \frac{c}{4}\|x-a\|^2 \geq 0, \end{aligned}$$

što dokazuje da je a tačka lokalnog minimum.■

6.5 Teorema Gaus-Markova

Teorema 6.17. U klasi linearnih nepristrasnih ocena nepoznatog parametra $\boldsymbol{\theta}$ u linearnom modelu sa potpunim rangom $\boldsymbol{X} = \boldsymbol{A\theta} + \boldsymbol{\varepsilon}$, ocena dobijena metodom najmanjih kvadrata ima najmanju disperziju.

Dokaz:

$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ je n-dimenzioni slučajni vektor tj. $X_1, \dots X_n$ su slučajne veličine i

$A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{bmatrix}$ je matrica formata $n \times p$, gde je $p < n$. Obzirom da je ovo model sa potpunim rangom, $rangA = p$.

$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix}$ je p-dimenzioni nepoznati parametar tj. $\theta_1, \dots, \theta_p$ su nepoznati parametri koje treba oceniti i

$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ je n-dimenzioni slučajni vektor, gde su $\varepsilon_1, \dots \varepsilon_n$ su slučajne veličine takve da važi:

- $E(\varepsilon_i) = 0$ za $i = 1, \dots, n$
- $D(\varepsilon_i) = \sigma^2$ za$i = 1, \dots, n$
- $E(\varepsilon_i \varepsilon_j) = 0$ za $i, j = 1, \dots, n$ i $i \neq j$

Na osnovu osobina vektora ε, možemo zaključiti da je $cov(\varepsilon) = \sigma^2 E$. Slučajne veličine $\varepsilon_1, \dots \varepsilon_n$ nazivaju se slučajnim greškama.

U dokazu je prvi korak određivanje ocene nepoznatog parametra θ metodom najmanjih kvadrata. Zatim treba pokazati da je ta ocena linearna i nepristrasna, a onda i odrediti disperziju ocene. Na kraju dobijenu ocenu upoređujemo sa nekom drugom linearnom i nepristrasnom ocenom nepoznatog parametra θ.

(1) Metoda najmanjih kvadrata podrazumeva određivanje takve ocene nepoznatog parametra θ da norma vektora slučajne greške $\|\varepsilon\|$ bude što manja.

$$\|\varepsilon\| \longrightarrow min \implies \|\varepsilon\|^2 \longrightarrow min$$
$$\|\varepsilon\|^2 = \varepsilon^T \varepsilon$$
$$\varepsilon^T \varepsilon = (X - A\theta)^T (X - A\theta)$$

Jer je $X = A\theta + \varepsilon \implies \varepsilon = X - A\theta$.

Posmatrajmo funkcional $Q: \mathbb{R}^p \to \mathbb{R}: \theta \mapsto (X - A\theta)^T (X - A\theta)$

$$\begin{aligned} Q(\theta) &= (X - A\theta)^T (X - A\theta) = (X^T - (A\theta)^T)(X - A\theta) = (X^T - \theta^T A^T)(X - A\theta) \\ &= X^T X - X^T A\theta - \theta^T A^T X + \theta^T A^T A\theta \end{aligned}$$

Na osnovu stava 6.16. ako je $dQ(\hat{\theta}) = 0$ i $d^2 Q(\hat{\theta})$ jako pozitivan kvadratni funkcional, onda je $\hat{\theta}$ tačka lokalnog minimuma funkcionala Q.

$$dQ(\hat{\theta}) = 0 - X^T A - X^T A + 2\hat{\theta}^T A^T A = -2X^T A + 2\hat{\theta}^T A^T A$$

Naime, kako je $\hat{\theta}^T A^T X \in \mathbb{R}$, to je

$$\hat{\theta}^T A^T X = \left(\hat{\theta}^T A^T X\right)^T = (A^T X)^T \hat{\theta} = X^T A \hat{\theta}.$$

Takođe, $A^T A$ je simetrična matrica i tu osobinu koristimo prilikom diferenciranja.

$$-2X^T A + 2\hat{\theta}^T A^T A = 0$$

$$\hat{\theta}^T A^T A = X^T A$$

$$(A^T A)^T \hat{\theta} = A^T X$$

$rangA = rang_v A$, a ono što su vrste matrice A, to su kolone matrice A^T, pa važi $rang_v A = rang_k A^T$. Na kraju $rang_k A^T = rangA^T$.

Kako je $rangA = p$ i $rangA^T = p$, sledi da je $rangA^T A = p$.

Matrica $A^T A$ je formata $p \times p$ i $rangA^T A = p$. Na osnovu posledice 2.23. možemo zaključiti da je matrica $A^T A$ regularna tj. postoji inverz $(A^T A)^{-1}$.

$$\hat{\theta} = (A^T A)^{-1} A^T X$$

Treba još proveriti da li je $d^2Q(\hat{\theta})$ jako pozitivan kvadratni funkcional tj. da li postoji $c > 0$ tako da je $d^2Q(\hat{\theta})(h,h) \geq c\|h\|^2$ za svako $h \in \mathbb{R}^p$.

$d^2Q(\hat{\theta}) = 2A^TA$

$d^2Q(\hat{\theta})(h,h) = 2h^TA^TAh$

Matrica A^TA je nenegativno definitna na osnovu leme 2.7. Na osnovu posledice 2.10. postoji ortogonalna matrica $Q \in M(n,\mathbb{R})$, takva da je $Q^TA^TAQ = D$, gde je D dijagonalna matrica sa nenegativnim elementima na dijagonali.

Iz jednakosti $Q^TA^TAQ = D$, dobijamo da je $A^TA = QDQ^T$.

$d^2Q(\hat{\theta})(h,h) = 2h^TQDQ^Th$

Neka je $Q = P^T$, tada prethodna jednakost postaje

$$d^2Q(\hat{\theta})(h,h) = 2h^TP^TDPh = 2(Ph)^TDPh \geq 2\lambda\|Ph\|^2 \qquad (1)$$

Gde je $\lambda > 0$ minimalni element na dijagonali matrice D.

$$\|Ph\|^2 = \|Ph\|\|Ph\| = \|Ph\|\|h^TP^T\| = \|h^TP^TPh\| = \|h^Th\| = \|h\|^2 \qquad (2)$$

Iz (1) i (2), sledi da je $d^2Q(\hat{\theta})(h,h) \geq 2\lambda\|h\|^2$ za svako $h \in \mathbb{R}^p$ i $\lambda > 0$.

Dakle,

$$\hat{\theta} = (A^TA)^{-1}A^TX$$

je zaista loklani minimum funkcionala $Q: \mathbb{R}^p \to \mathbb{R}: \theta \mapsto (X - A\theta)^T(X - A\theta)$.

(2) Iz samog zapisa se vidi da je ocena $\hat{\theta}$ linearna.
(3) Izračunajmo očekivanje ocene $\hat{\theta}$.

$$\begin{aligned} E(\hat{\theta}) = E((A^TA)^{-1}A^TX) &= (A^TA)^{-1}A^TE(X) = (A^TA)^{-1}A^TE(A\theta + \varepsilon) \\ &= (A^TA)^{-1}A^T\big(E(A\theta) + E(\varepsilon)\big) = (A^TA)^{-1}A^T(AE(\theta) + 0) \\ &= (A^TA)^{-1}A^TAE(\theta) = (A^TA)^{-1}(A^TA)E(\theta) = E(\theta) \end{aligned}$$

Tako da je ocena $\hat{\theta}$ zaista nepristrasna.

(4) Određivanje disperzije ocene.

$$D(\hat{\theta}) = D((A^TA)^{-1}A^TX) = (A^TA)^{-1}A^Tcov(X)((A^TA)^{-1}A^T)^T$$

Na osnovu svojstva kovarijacionih matrica izloženih u poglavlju 5, važi sledeće

$$D(X) = D(A\theta + \varepsilon) = D(A\theta) + D(\varepsilon) = 0 + \sigma^2E = \sigma^2E.$$

Tako da je

$$
\begin{aligned}
D(\hat{\theta}) &= (A^TA)^{-1}A^T\sigma^2E((A^TA)^{-1}A^T)^T = \sigma^2(A^TA)^{-1}A^TEA((A^TA)^{-1})^T \\
&= \sigma^2(A^TA)^{-1}(A^TA)((A^TA)^T)^{-1} = \sigma^2((A^TA)^{-1}A^TA)(A^TA)^{-1} \\
&= \sigma^2E(A^TA)^{-1} = \sigma^2(A^TA)^{-1}.
\end{aligned}
$$

(5) Neka je $\tilde{\theta}$ neka druga linearna i nepristrasna ocena nepoznatog parametra θ oblika

$$
\tilde{\theta} = G^TX + \hat{\theta},
$$

gde je matrica G^T formata $p \times n$ izabrana tako da važi $G^TA = 0$. Ocena je očigledno linearna, proverimo da li je i nepristrasna.

$$
\begin{aligned}
E(\tilde{\theta}) = E(G^TX + \hat{\theta}) &= E(G^TX) + E(\hat{\theta}) = E(G^T(A\theta + \varepsilon)) + E(\hat{\theta}) \\
&= E(G^TA\theta + G^T\varepsilon) + E(\hat{\theta}) = E(G^TA\theta) + E(G^T\varepsilon) + E(\hat{\theta}) \\
&= G^TAE(\theta) + G^TE(\varepsilon) + E(\theta) = 0E(\theta) + G^T0 + E(\theta) = E(\theta)
\end{aligned}
$$

Ocena $\tilde{\theta}$ je zaista nepristrasna.

Radi upoređivanja ocena $\hat{\theta}$ i $\tilde{\theta}$, potrebno je izračunati disperziju ocene $\tilde{\theta}$.

$$
\begin{aligned}
D(\tilde{\theta}) = D(G^TX + \hat{\theta}) &= D(G^TX + (A^TA)^{-1}A^TX) = D((G^T + (A^TA)^{-1}A^T)X) \\
&= (G^T + (A^TA)^{-1}A^T)cov(X)\left((G^T + (A^TA)^{-1}A^T)\right)^T \\
&= (G^T + (A^TA)^{-1}A^T)\sigma^2E(G^T + (A^TA)^{-1}A^T)^T \\
&= \sigma^2(G^T + (A^TA)^{-1}A^T)(G + ((A^TA)^{-1}A^T)^T) \\
&= \sigma^2(G^T + (A^TA)^{-1}A^T)(G + A((A^TA)^{-1})^T) \\
&= \sigma^2(G^T + (A^TA)^{-1}A^T)(G + A((A^TA)^T)^{-1}) \\
&= \sigma^2(G^T + (A^TA)^{-1}A^T)(G + A(A^TA)^{-1}) \\
&= \sigma^2(G^TG + (A^TA)^{-1}A^TG + G^TA(A^TA)^{-1} + (A^TA)^{-1}A^TA(A^TA)^{-1}) \\
&= \sigma^2(G^TG + (A^TA)^{-1}(G^TA)^T + 0(A^TA)^{-1} + ((A^TA)^{-1}A^TA)(A^TA)^{-1}) \\
&= \sigma^2(G^TG + (A^TA)^{-1}0 + E(A^TA)^{-1}) = \sigma^2(G^TG + (A^TA)^{-1}) \\
&= \sigma^2G^TG + \sigma^2(A^TA)^{-1} = \sigma^2G^TG + D(\hat{\theta})
\end{aligned}
$$

Matrica G^TG je nenegativno definitna na osnovu leme 2.7. tako da je $\sigma^2G^TG \geq 0$, pa važi $D(\tilde{\theta}) \geq D(\hat{\theta})$. Dakle među linearnim nepristrasnim ocenama nepoznatog parametra θ najefikasnija ocena (u smislu minimalne disperzije) je ocena dobijena po metodi najmanjih kvadrata. ■

Primer: Metoda najmanjih kvadrata se koristi u regresionom modelu, a u ovom radu primena teoreme Gaus-Markova biće ilustrovana na jednom primeru primene u okviru regresionih vremenskih serija. Posmatrajmo vremensku seriju $(X_1, \dots X_n)$ koja se sastoji od jedne determinističke komponente (trenda) $f(t)$ i jedne slučajne komponente ε_t

$$X_t = f(t) + \varepsilon_t \quad t = 1,2, \dots n.$$

Pri čemu je ε_t niz takav da važi:

- $E(\varepsilon_t) = 0 \quad t = 1,2, \dots n;$
- $D(\varepsilon_t) = \sigma^2 \quad t = 1,2, \dots n;$
- $cov(\varepsilon_s, \varepsilon_t) = 0 \quad t, s = 1,2, \dots n \qquad s \neq t.$

Trend možemo predstaviti kao linearnu kombinaciju nepoznatih parametara

$$f(t) = f_1(t)\beta_1 + f_2(t)\beta_2 + \cdots + f_p(t)\beta_p \qquad t = 1,2, \dots n \text{ i } p < n$$

pri čemu su $f_j(t)(j = 1,2, \dots p\,)$ određeni brojevi odnosno vrednosti dobijene merenjem, a $\beta_j(j = 1,2, \dots p\,)$ su nepoznati parametri. Tada imamo

$$\begin{gathered} X_1 = f(1) + \varepsilon_1 = f_1(1)\beta_1 + f_2(1)\beta_2 + \cdots + f_p(1)\beta_p + \varepsilon_1 \\ X_2 = f(2) + \varepsilon_2 = f_1(2)\beta_1 + f_2(2)\beta_2 + \cdots + f_p(2)\beta_p + \varepsilon_2 \\ \dots \\ X_n = f(n) + \varepsilon_n = f_1(n)\beta_1 + f_2(n)\beta_2 + \cdots + f_p(n)\beta_p + \varepsilon_n. \end{gathered}$$

Ako uvdemo sledeće oznake

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad F = \begin{bmatrix} f_1(1) & f_2(1) & \cdots & f_p(1) \\ f_1(2) & f_2(2) & \cdots & f_p(2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(n) & f_2(n) & \cdots & f_p(n) \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Onda se jednačine mogu napisati u matričnom obliku

$$X = F\beta + \varepsilon.$$

Ovaj vremenski model naziva se regresioni model. Kada na ovaj model primenimo teoremu Gaus-Markova, dobijamo da ocena

$$\hat{\beta} = (F^T F)^{-1} F^T X$$

dobijena po metodi najmanjih kvadrata ima najmanju disperziu u klasi svih lineranih nepristrasnih ocena.

Zaključak

Slučajni vektori mogu se zapisivati u matričnom obliku. Baš kao i linearni sistemi, tako se i linearni statistički modeli mogu predstaviti preko matrica. Videli smo da i opservacije vremenskih serija imaju svoj matrični zapis. Sve su to samo neki od primera na kojima je lako uočljivo da matrični račun ima veliku primenu u statistici. Međutim, za to je potrebno vladanje ne samo osnovnim, nego i sofisticiranijim aparatom matrične algebre i linearne algebre uopšte.

U ovom radu osvrnuli smo se na osnove linearne algebre kao što su sistemi linearnih jednačina, osnovne operacije i osobine matrica i determinanti. Kada je u pitanju nešto složeniji matrični račun predstavljene su teme matričnog diferenciranja i uopštene inverzne matrice. Sve to u cilju pojednostavljenja i boljeg razumevanja pojedinih delova statistike, kao što je na primer teorema Gaus-Markova. Značaj teoreme je veliki, jer već na osnovu same postavke problema, pod uslovom da raspolažemo dovoljnim brojem informacija možemo zaključiti koja je linearna ocena nepoznatog parametra najefikasnija.

Uvođenje uopštene inverzne matrice omogućava nam da rešavamo one sisteme koji imaju beskonačno mnogo rešenja. To je slučaj kada imamo više nepoznatih nego jednačina tj. kada treba da ocenimo vise nepoznatih parametara nego što imamo uzoračkih podataka. Ako je pak situacija takva da imamo više uzoračkih podataka nego nepoznatih parametara koje ocenjujemo, onda je metoda najmanjih kvadrata jedan od mogućih načina da dodjemo do ocene nepoznatih parametara.

Sopstveni vektori i sopstvene vrednosti su bitni u vektorskim vremenskim serijama, analizi glavnih komponenata, primeni kvadratnih formi u statistici, itd. Ovo su samo neki od primera, jer matrični račun ima veoma veliku i bitnu upotrebu u statistici.

Literatura

[1] M.Arsenović, M.Dostanić i D.Jocić - "Teorija mere, funkcionalna analiza, teorija operatora", Zavod za udžbenike, Beograd 2012.

[2] M.Esert, T.Žilić - "MATLAB – Matrični laboratorij", Zagreb 2004.

[3] D.A.Harville – „Matrix Algebra from a Statistician's Perspective", IBM T.J. Watson Resarch Center Yourktown Heights, NY 2012

[4] B.Kovačević – „Stohastički sistemi i estimacija", Beograd 1999.

[5] A.Lipkovski – „Linearna algebra i analitička geometrija", Beograd 2007.

[6] J.Mališić – „Vremenske serije", Beograd 2002.

[7] P.Mladenović – „Verovatnoća i statistika", Beograd 2005.

[8] R.V.Hogg, J.W.McKean, A.T.Craig - „Introduction to Mathematical Statistics", University of Iowa, Western Michigan University 2005. Sixth edition

[9] http://cran.r-project.org. sajt sa kojeg se može preuzeti statistički program R i svi potrebni paketi

[10] http://www.matf.bg.ac.rs/p/files/42-ANALIZA_GLAVNIH_KOMPONENATA_02.pdf