MAATRIKSID STATISTIKAS

§1. Tähistused, tehted, põhimõisted

Maatriksalgebra annab vahendid mitmemõõtmelise statistika esitamiseks kompaktsel ja ülevaatlikul kujul. Üleminek maatrikskeelele pärast T. W. Andersoni monograafia ilmumist 1958. aastal kiirendas ja süvendas mitmemõõtmelise statistika arengut oluliselt. Samas on olnud tegu vastastikku kasuliku arenguga. Statistikaprobleemid on viinud ka maatriksite teooria uute arendusteni viimastel aastakümnetel. Järgnev lühiülevaade on kokkusurutud esitus olulisematest mõistetest ja tulemustest, mida selles kursuses esitame. Põhjalikuma käsitluse huvilistele olgu siinkohal antud mõned viited kirjandusele. Toome välja just statistikasuunitlusega maatriksalgebra raamatud. Suhteliselt lihtsama käsitluse võib leida raamatust Searle (1982). 1990-ndate lõpus ilmus mitu head ja põhjalikku statistikutele mõeldud maatriksite teooria esitust: Harville (1997), Schott (1997), Rao & Rao (1998) Magnus & Neudecker (1999). Põhjalik käsitlus on antud ka raamatutes Kollo (1991) ja Kollo & von Rosen (2005, 2010). Kokkuvõtlikult on olulisemad maatriksalgebra mõisted ja tulemused ära toodud klassikalistes mitmemõõtmelise analüüsi monograafiates: Anderson (2003), Srivastava & Khatri (1979), Rao (1973), Muirhead (1982), Siotani, Hayakawa & Fujikoshi (1985) jt.

Järgnev materjal on jagatud kaheks: esimeses osas on ülevaade maatriksalgebra põhimõistetest ja vajalikest tulemustest, teine osa sisaldab kokkuvõtte plokkmaatriksitest.

Põhitehted.

Maatriks $\mathbf{A}$ mõõtmetega $m \times n$ on ristkülikukujuline tabel elementidest $a_{ij},\ i = 1, \ldots, m; j = 1, \ldots, n$:

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

kus element a_{ij} paikneb tabeli i-ndas reas ja j-ndas veerus. Elemendi a_{ij} tähisena kasutame ka tähistust $(\mathbf{A})_{ij}$. Maatriksi elementideks võivad olla erinevad matemaatilised objektid: reaalarvud, kompleksarvud, funktsioonid jne. Tekstis kasutame paralleelselt järgmisi fraase:

- maatriks $\mathbf{A}$ mõõtmetega $m \times n$,
- $m \times n$-maatriks $\mathbf{A}$,
- $\mathbf{A} : m \times n$.

Kui $m = n$, siis on $\mathbf{A}$ ruutmaatriks. Kui $\mathbf{A}$ on $m \times 1$-maatriks, on tegemist vektoriga. Sel juhul jäetakse teine indeks kirjutamata ja kasutatakse kirjapilti

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Maatriksid $\mathbf{A}$ ja $\mathbf{B}$ on võrdsed, $\mathbf{A} = \mathbf{B}$, kui nende mõõtmed on võrdsed ja vastavad elemendid on võrdsed. Maatriksit, mille elemendid on võrdsed arvuga 1 tähistame $\mathbf{1}$, vajadusel $\mathbf{1}_{m \times n}$, analoogiliselt ka nullidest koosneva maatriksi jaoks kasutame tähistusi $\mathbf{0}$ või $\mathbf{0}_{m \times n}$.

Maatriksi $\mathbf{A} : m \times n$ korrutis skalaariga c on $m \times n$ maatriks $c\mathbf{A}$, kus

$$(c\mathbf{A})_{ij} = ca_{ij}.$$

Skalaari all mõistetakse maatriksi elementidega sama tüüpi suurust.

Maatriksite $\mathbf{A} : m \times n$ ja $\mathbf{B} : m \times n$ summa $\mathbf{A} + \mathbf{B}$ on $m \times n$-maatriks elementidega

$$(\mathbf{A} + \mathbf{B})_{ij} = a_{ij} + b_{ij}.$$

Maatriksite liitmise ja skalaariga korrutamise olulisemad omadused on järgmised:

$$\begin{aligned}
\mathbf{A}+\mathbf{B} &= \mathbf{B}+\mathbf{A};\\
(\mathbf{A}+\mathbf{B})+\mathbf{C} &= \mathbf{A}+(\mathbf{B}+\mathbf{C});\\
\mathbf{A}+(-1)\mathbf{A} &= \mathbf{0};\\
(c_1+c_2)\mathbf{A} &= c_1\mathbf{A}+c_2\mathbf{A};\\
c(\mathbf{A}+\mathbf{B}) &= c\mathbf{A}+c\mathbf{B};\\
c_1(c_2\mathbf{A}) &= (c_1c_2)\mathbf{A}.
\end{aligned}$$

Maatriksite korrutamine on võimalik, kui esimese maatriksi veergude arv võrdub teise maatriksi ridade arvuga. Maatriksite $\mathbf{A}: m\times n$ ja $\mathbf{B}: n\times r$ korrutis $\mathbf{C}=\mathbf{AB}$ on $m\times r$-maatriks, kus

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

Maatriksite korrutamine ei ole kommutatiivne, küll aga kehtivad seosed

$$\begin{aligned}
\mathbf{A}(\mathbf{BC}) &= (\mathbf{AB})\mathbf{C};\\
\mathbf{A}(\mathbf{B}+\mathbf{C}) &= \mathbf{AB}+\mathbf{AC});\\
(\mathbf{A}+\mathbf{B})\mathbf{C} &= \mathbf{AC}+\mathbf{BC}.
\end{aligned}$$

Maatriksi $\mathbf{A}: m\times n$ transponeeritud maatriks $\mathbf{A}'$ on $n\times m$-maatriks, kus

$$(\mathbf{A}')_{ji}=a_{ij},\quad i=1,\ldots,m,\ j=1,\ldots,n.$$

Maatriksi transponeerimine on seotud liitmise ja kor-rutamisega järgmiselt:

$$\begin{aligned}
(\mathbf{A}')' &= \mathbf{A};\\
(\mathbf{A}+\mathbf{B})' &= \mathbf{A}'+\mathbf{B}';\\
(\mathbf{AB})' &= \mathbf{B}'\mathbf{A}'.
\end{aligned}$$

Liitmisega sarnane tehe on maatriksite elemendiviisiline korrutamine; kirjanduses ka Hadamard'i või Schuri korrutamine.
$m \times n$ maatriksite $\mathbf{A} = (a_{ij})$ ja $\mathbf{B} = (b_{ij})$ elemendiviisiline korrutis on $m \times n$ maatriks

$$\mathbf{A} \circ \mathbf{B} = (a_{ij} b_{ij}), \quad i = 1, \ldots, m; \; j = 1, \ldots, n.$$

Omadused

$$\begin{aligned} \mathbf{A} \circ \mathbf{B} &= \mathbf{B} \circ \mathbf{A}; \\ \mathbf{A} \circ (\mathbf{B} \circ \mathbf{C}) &= (\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C}; \\ \mathbf{A} \circ (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} \circ \mathbf{B}) + \mathbf{A} \circ \mathbf{C}; \\ (\mathbf{A} \circ \mathbf{B})' &= \mathbf{A}' \circ \mathbf{B}'. \end{aligned}$$

Erikujulised maatriksid.
Ühiku rolli etendab maatriksite hulgas *ühikmaatriks* $\mathbf{I}_n$:

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij}), \quad i, j = 1, \ldots, n,$$

kus δ_{ij} on Kroneckeri delta:

$$\delta_{ij} = \begin{cases} 1, & kui \; i = j; \\ 0, & kui \; i \neq j. \end{cases}$$

ühikmaatriks rahuldab võrdusi $\mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$, kui $\mathbf{A}$ on $m \times n$-maatriks.
Ruutmaatriks $\mathbf{A}$ on *sümmeetriline*, kui

$$\mathbf{A}' = \mathbf{A}$$

ja *kaldsümmeetriline*, kui

$$\mathbf{A} = -\mathbf{A}'.$$

Reaalarvuline ruutmaatriks $\mathbf{A} : n \times n$ on *ortogonaalne*, kui

$$\mathbf{A}'\mathbf{A} = \mathbf{I}_n$$

ja *idempotentne* , kui

$$\mathbf{A}^2 = \mathbf{A}.$$

Ortogonaalsed ja idempotentsed maatriksid on statistikas olulised. Vaatame lähemalt, mida maatriksi ortogonaalsus endaga kaasa toob. Olgu meil $\mathbf{A}$ esitatud veeruvektorite kaudu

$$\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

Siis

$$\mathbf{A}'\mathbf{A} = \begin{pmatrix} \mathbf{a}_1'\mathbf{a}_1 & \mathbf{a}_1'\mathbf{a}_2 & \cdots & \mathbf{a}_1'\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n'\mathbf{a}_1 & \mathbf{a}_n'\mathbf{a}_2 & \cdots & \mathbf{a}_n'\mathbf{a}_n \end{pmatrix} = \mathbf{I}_n.$$

Seega on ortogonaalmaatriksi veeruvektorid $\mathbf{a}_i$ ühiku pikkusega ja risti, kui $i \neq j$. Vaatame, kuidas teisendab ortogonaalmaatriks $\mathbf{A} : n \times n$ ruumi $\mathbb{R}^n$ loomuliku baasi vektoreid $\mathbf{e}_i$

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} - i - \text{s element.}$$

Korrutamise tulemusena saame

$$\mathbf{A}\mathbf{e}_i = \mathbf{a}_i, \quad i = 1, \ldots, n.$$

Järelikult ortogonaalmaatriks teisendab loomuliku baasi vektorid oma veeruvektoriteks, s.t. ortogonaalmaatriks pöörab baasi ruumis $\mathbb{R}^n$ mingi nurga võrra koordinaattelgede alguspunkti ümber. Samas on lihtne näha, et ortogonaalmaatriks jätab teisendamisel suvalise vektori pikkuse muutmata. Olgu $\mathbf{y} = \mathbf{A}\mathbf{x}$. Siis

$$\mathbf{y}'\mathbf{y} = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{I}_n\mathbf{x} = \mathbf{x}'\mathbf{x}.$$

Ruutmaatriks $\mathbf{A} : n \times n$ on *ülemine kolmnurkmaatriks*, kui tema allpool peadiagonaali asetsevad elemendid võrduvad nulliga. Tähistame sel juhul maatriksi $\mathbf{A}^{\triangle}$:

$$\mathbf{A}^{\triangle} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Kui ruutmaatriksi $\mathbf{A} : n \times n$ ülalpool peadiagonaali paiknevad elemendid on võrdsed nulliga, on tegemist *alumise kolmnurkmaatriksiga* $\mathbf{A}_{\triangle}$.

Ruutmaatriksi diagonaliseerimine on operatsioon, mis asendab $\mathbf{A} : n \times n$ väljaspool peadiagonaali asetsevad elemendid nullidega. *Diagonaliseeritud maatriksi* tähistame $\mathbf{A}_d$:

$$\mathbf{A}_d = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Vektorist $\mathbf{a}$ moodustatud diagonaalmaatriks on sarnase tähistusega:

$$\mathbf{a}_d = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

Ruutmaatriks $\mathbf{A}$ on alati esitatav kolmnurkmaatiksite ja diagonaliseeritud maatriksi kaudu

$$\mathbf{A} = \mathbf{A}^{\triangle} + \mathbf{A}_{\triangle} - \mathbf{A}_d.$$

Kompleksarvulisi maatrikseid läheb statistikas harva vaja. Kõige levinumad neist on Hermite'i maatriks ja normaalmaatriks. Olgu ruutmaatriksi $\mathbf{A}$ elemendid kompleksarvud. Tähistame maatriksi $\mathbf{A}$ elementide kaaskompleksarvudest moodustatud transponeeritud maatriksi $\mathbf{A}^*$.
Kompleksarvulist maatriksit $\mathbf{A}$ nimetatakse *Hermite'i maatriksiks*, kui

$$\mathbf{A} = \mathbf{A}^*$$

ja *normaalmaatriksiks*, kui

$$\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}.$$

Reaalarvuliste maatriksite korral nimetatakse ruutmaatriksit $\mathbf{A}$ normaalmaatriksiks, kui

$$\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A}.$$

Seega on ortogonaalmaatriks ka normaalmaatriks.

§2. Determinant ja pöördmaatriks

Ruutmaatriksi $\mathbf{A} : n \times n$ korral on tähtsaks arvuliseks karakteristikuks tema *determinant* $|\mathbf{A}|$. Tänapäeva arusaama kohaselt mõnevõrra ootamatult võeti determinant kasutusele enne maatriksi mõistet. Järgnevalt toome ära mõne ajaloolise viite.

Väidetavalt võttis determinandi esimesena kasutusele
Gottfried Wilhelm Leibniz (ka Leibnitz) – 1646-1716.

Leibniz, 1678

Jaapanlane Takakazu Seki (1642-1708) uuris determinante üldisemal kujul.

Seki, 1683

Kümme aastat pärast Seki tööd kasutas Leibniz determinante lineaarvõrrandite süsteemi lahendamisel.

Leibniz, 1693

Sõna "determinant" võttis kasutusele Carl Friedrich Gauss (1777-1855) aastal 1801.

Gauss, 1801

Esimene raamat determinantidest ilmus 1851. *Elementary Theorems Relating to Determinants*, autoriks William H. Spottiswoode ((1825-1883)

Spottiswoode, 1851

Järgmise raamatu autoriks oli Charles Ludwig Dodgson (1832-1898). Kaks aastat varem, 1865, ilmus tal raamatu "Alice imedemaal" esmatrükk Lewis Carrolli nime all.

Dodgson, 1867

Ruutmaatriksi $\mathbf{A}$ *determinant* defineeritakse võrdusega

$$|\mathbf{A}| = \sum_{j_1,\cdots,j_n} (-1)^{N(j_1,\cdots,j_n)} \prod_{i=1}^{n} a_{ij_i}, \tag{2.1}$$

kus summeeritakse üle arvude $1, 2, \ldots, n$ kõigi erinevate permutatsioonide $(j_1, \ldots, j_n)$ ja $N(j_1, \ldots, j_n)$ on permutatsiooni $(j_1, \ldots, j_n)$ inversioonide arv jõudmaks järjestuseni $1, 2, \ldots, n$. Inversioon seisneb kahe arvu vahetamises permutatsioonis nii, et suurem on pärast väiksemat. Determinandi arvutamiseks on valemit (2.1) tülikas rakendada vähegi suuremate n väärtuste korral. Üks võimalus arvutuste lihtsustamiseks on *miinori* mõiste rakendamine.

Elemendi a_{ij} *täiendmiinoriks* nimetatakse $\mathbf{A} : n \times n$ elementidest moodustatud $(n-1) \times (n-1)$-maatriksi $\mathbf{A}_{(ij)}$ determinanti. Maatriks $\mathbf{A}_{(ij)}$ saadakse maatriksist $\mathbf{A}$ selle i-nda rea ja j-nda veeru eraldamise teel. Täiendmiinorite abil saab maatriksi $\mathbf{A}$ determinandi esitada arendusena i-nda rea kaudu

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} |\mathbf{A}_{(ij)}| \quad \text{iga } i \text{ korral}, \tag{2.2}$$

või j-nda veeru kaudu

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{A}_{(ij)}| \quad \text{iga } j \text{ korral.} \tag{2.3}$$

Maatriksi $\mathbf{A}$ esimesest k reast ja k veerust moodustatud $k \times k$-alammaatriksi determinanti nimetatakse k-järku *nurgamiinoriks.*

Maatriksi $\mathbf{A} : n \times n$ k-ndat järku *peamiinor* on selle maatriksi elementidest koosnev $k \times k$-maatriksi determinant, kus selle alammaatriksi peadiagonaali elementideks on maatriksi $\mathbf{A}$ peadiagonaali elemendid, s.t. alammaatriks on moodustatud samade ridade ja veergude elementidest.

Miinori mõistet kasutatakse ka $m \times n$ maatriksi $\mathbf{A}$ korral.
Maatriksi $\mathbf{A} : m \times n$ k-ndat järku miinoriks nimetatakse maatriksi $\mathbf{A}$ elementidest k rea ja k veeru väljaeraldamisel nende ühistest elementidest moodustatud $k \times k$-maatriksi determinanti.

Esitame järgnevalt mõned determinantide olulisemad omadused:

$$|\mathbf{A}'| = |\mathbf{A}|; \tag{2.4}$$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|; \tag{2.5}$$

$$|\mathbf{I}_m + \mathbf{AB}| = |\mathbf{I}_n + \mathbf{BA}|, \tag{2.6}$$

kus $\mathbf{A}$ on $m \times n$-maatriks ja $\mathbf{B}$ on $n \times m$-maatriks.
Kui $|\mathbf{A}| \neq 0$, siis öeldakse, et maatriks $\mathbf{A} : n \times n$ on *pööratav e. regulaarne*, vastasel juhul, kui $|\mathbf{A}| = 0$, on meil tegemist *singulaarse* maatriksiga. Kui $|\mathbf{A}| \neq 0$, siis leidub maatriksil $\mathbf{A}$ *pöördmaatriks* $\mathbf{A}^{-1}$, mis on määratud võrdusega

$$\mathbf{AA}^{-1} = \mathbf{I}_n.$$

Pöördmaatriksil on järgmised omadused.

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n;$$

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}; \tag{2.7}$$

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}; \tag{2.8}$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Pöördmaatriksi üldelement avaldub valemiga

$$(\mathbf{A}^{-1})_{ij} = \frac{(-1^{i+j}|\mathbf{A}_{(ji)}|}{|\mathbf{A}|}, \tag{2.9}$$

kus $|\mathbf{A}_{(ji)}|$ on elemendi a_{ji} täiendmiinor.

Kahe maatriksi summa pöördmaatriksi avaldis on keerulisem.

Teoreem 2.1 (binomiaalne pöördteoreem). *Olgu järgnevas võrduses maatriksid* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *ja* $\mathbf{D}$ *sobivate dimensioonidega ja eksisteerigu seal esinevad pöördmaatriksid. Siis*

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}.$$

§3. Maatriksi astak

Vektorid $\mathbf{x}_i,\ i = 1, \ldots, r$ on *lineaarselt sõltumatud*, kui

$$\sum_{i=1}^{r} c_i \mathbf{x}_i = \mathbf{0},$$

parajasti siis, kui kõik kordajad $c_i = 0,\ i = 1, \ldots, r$.
Definitsioon 3.1. *Maatriksi* $\mathbf{A} : m \times n$ *astakuks* $r(\mathbf{A})$ *nimetatakse tema lineaarselt sõltumatute veergude maksimaalset arvu.*
Loetleme alljärgnevalt $m \times n$-maatriksi $\mathbf{A}$ astaku olulisemad omadused:

(i) $r(\mathbf{A}) = r(\mathbf{A}')$;

(ii) astak $r(\mathbf{A})$ on võrdne suurimat järku mittenullilise miinori järguga;

(iii) $r(\mathbf{A}) \leq \min(m, n)$;

(iv) $r(\mathbf{AB}) \leq \min(r(\mathbf{A}), r(\mathbf{B}))$;

(v) $r(\mathbf{A} + \mathbf{B}) \leq r(\mathbf{A}) + r(\mathbf{B})$;

(vi) olgu $\mathbf{A} : m \times m$, $\mathbf{B} : m \times n$ ja $\mathbf{C} : n \times n$, kui $\mathbf{A}$ ja $\mathbf{C}$ on pööratavad, siis $r(\mathbf{ABC}) = r(\mathbf{B})$;

(vii) $r(\mathbf{A}) = r(\mathbf{AA}') = r(\mathbf{A}'\mathbf{A})$;

(viii) kui $\mathbf{A} : m \times n$ ja $\mathbf{B}$ rahuldavad võrdust $\mathbf{AB} = \mathbf{0}$, siis $r(\mathbf{B}) \leq n - r(\mathbf{A})$.

Maatriks $\mathbf{A} : m \times n$ on *täisastakuga*, kui $r(\mathbf{A}) = m$ või $r(\mathbf{A}) = n$.
Praktikas kasutatakse astaku leidmiseks definitsiooni väga harva. Selle asemel kasutatakse maatriksi $\mathbf{A}$ teisendamist elementaarteisendustega nn. kanoonilisele kujule.
Maatriksi *elementaarteisendusi* oli kolm:

1) $\mathbf{A}$ ridade (veergude) vahetamine;

2) $\mathbf{A}$ rea (veeru) kõigi elementide korrutamine sama mittenullilise skalaariga;

3) $\mathbf{A}$ reale (veerule) teise rea (veeru) liitmine, kusjuures see teine rida (veerg) võib olla korrutatud mittenullilise skalaariga.

Elementaarteisendusi realiseerivad maatriksi $\mathbf{A}$ korrutamised teatud maatriksitega, nn. *elementaarmaatriksitega.* Need on järgmised.
1) Ridade ja veergude vahetamist teostab maatriks $\mathbf{E}_{ij}$, mis saadakse ühikmaatriksist i-nda ja j-nda rea vahetamisel (ühtlasi vahetuvad ka vastavad veerud).
Näiteks kolmandat järku maatriksi korral

$$\mathbf{E}_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Üldjuhul

$$\mathbf{E}_{ij} = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

kus $e_{ij} = e_{ji} = 1$, $i \neq j$ ja ülejäänud elemendid väljaspool peadiagonaali on nullid, peadiagonaalil $e_{ii} = e_{jj} = 0$.
Korrutises $\mathbf{AE}_{ij}$ on maatriksi $\mathbf{A}$ i-s ja j-s veerg vahetatud, korrutises $\mathbf{E}_{ij}\mathbf{A}$ on maatriksi $\mathbf{A}$ i-s ja j-s rida vahetatud.

2) Rea või veeru korrutamist skalaariga λ teostab maatriks $\mathbf{R}_{ii}(\lambda)$, mis saadakse ühikmaatriksist, kus ii-s element on λ

$$\mathbf{R}_{ii}(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

Korrutises $\mathbf{A}\mathbf{R}_{ii}(\lambda)$ on $\mathbf{A}$ i-s veerg korrutatud skalaariga λ, korrutises $\mathbf{R}_{ii}(\lambda)\mathbf{A}$ on $\mathbf{A}$ i-s rida korrutatud skalaariga λ.
3) Ühele reale (veerule) teise mittenullilise skalaariga λ korrutatud rea (veeru) liitmine.
Selle teisenduse realiseerib maatriksiga $\mathbf{P}_{ij}(\lambda)$ korrutamine. Maatriksis $\mathbf{P}_{ij}(\lambda)$ on lisaks peadiagonaali elementidele ka element indeksiga ij võrdne ühega. Näiteks kolmandat järku maatriks $\mathbf{P}_{12}(\lambda)$ on kujul

$$\mathbf{P}_{12}(\lambda) = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Korrutises $\mathbf{P}_{ij}(\lambda)\mathbf{A}$ on maatriksi $\mathbf{A}$ i-ndale reale liidetud λ-kordne j-s rida.
Korrutises $\mathbf{A}\mathbf{P}_{ij}(\lambda)$ on maatriksi $\mathbf{A}$ j-ndale veerule liidetud λ-kordne i-s veerg.
Elementaarmaatriksite definitsioonidest nähtub, et nad on täisastakuga, seega ka pööratavad; veelgi enam, elementaarmaatriksite pöördmaatriksid määravad omakorda elementaarteisenduse. Nimelt

1) $(\mathbf{E}_{ij})^{-1} = \mathbf{E}_{ij}$;
2) $(\mathbf{R}_{ii}(\lambda))^{-1} = \mathbf{R}_{ii}(\frac{1}{\lambda})$;
3) $(\mathbf{P}_{ij}(\lambda))^{-1} = \mathbf{P}_{ij}(-\lambda)$.

Samuti saame elementaarmaatriksi transponeerimisel elementaarmaatriksi.
Astaku omadusest (vi) järeldub, et $\mathbf{A}$ astak ei muutu, kui seda maatriksit korrutada elementaarmaatriksiga.
Definitsioon 3.2 *Öeldakse, et maatriksid* $\mathbf{A}$ *ja* $\mathbf{B}$ *on ekvivalentsed,* $\mathbf{A} \cong \mathbf{B}$, *kui*

$$\mathbf{B} = \mathbf{PAQ},$$

kus $\mathbf{P}$ *ja* $\mathbf{Q}$ *on elementaarmaatriksite korrutised.*
Kuna $\mathbf{P}$ ja $\mathbf{Q}$ on pööratavad (pööratavate maatriksite korrutised), siis

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{Q}^{-1}$$

ning järelikult $r(\mathbf{A}) = r(\mathbf{B})$.
Teoreem 3.1 *Iga astakuga* $r(\mathbf{A})$ $m \times n$*-maatriksi* $\mathbf{A} \neq \mathbf{0}$ *korral leiduvad elementaarmaatriksite korrutised* $\mathbf{P}$ *ja* $\mathbf{Q}$ *nii, et*

$$\mathbf{A} \cong \mathbf{PAQ} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Tõestus (skeem). Eelduse kohaselt vähemalt üks maatriksi $\mathbf{A}$ element ei võrdu nulliga. Olgu see a_{pq}. Korrutades vasakult maatriksit $\mathbf{A}$ elementaarmaatriksiga $\mathbf{E}_{1p}$ ja paremalt maatriksiga $\mathbf{E}_{1q}$ saame elemendi a_{pq} viia üles vasakusse nurka. Tähistame

$$\mathbf{E}_{1p}\mathbf{A}\mathbf{E}_{1q} = \mathbf{B}_1; \quad \mathbf{B}_1 \cong \mathbf{A}, \quad b_{11}^1 = a_{pq} \neq 0.$$

Teisendame maatriksi $\mathbf{B}_1$ esimese rea ja veeru ülejäänud elemendid nullideks. Korrutame maatriksit $\mathbf{B}_1$ vasakult maatriksitega

$$\mathbf{P}_{i1}\Big(\frac{-b_{i1}^1}{b_{11}^1}\Big), \quad i = 2, \ldots, m.$$

Sellega teisendame nullideks kõik esimese veeru elemendid peale b_{11}^1. Analoogiliselt, korrutades paremalt maatriksit $\mathbf{B}_1$ maatriksitega

$$\mathbf{P}_{i1}\Big(\frac{-b_{1j}^1}{b_{11}^1}\Big), \quad j = 2, \ldots, n$$

teisendame nullideks kõik esimese rea elemendid peale b_{11}^1. Saadud maatriksi tähistame $\mathbf{B}_2$,

$$\mathbf{A} \cong \mathbf{B}_2 = \begin{pmatrix} b_{11}^1 & 0 & \cdots & 0 \\ 0 & b_{22}^2 & \cdots & b_{2n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & b_{m2}^2 & \cdots & b_{mn}^2 \end{pmatrix}.$$

Kordame sama protseduuri maatriksi $\mathbf{B}_2$ all paremal nurgas oleva $(m-1) \times (n-1)$ alammaatriksiga ja nullime $\mathbf{B}_2$ teise rea ja veeru elemendid välja arvatud b_{22}^2 kohal olev mittenulliline element. Saame teisenduste tulemusena maatriksi $\mathbf{B}_3 \cong \mathbf{A}$. Jätkame seda protseduuri kuni maatriksis $\mathbf{B}_i$ leidub mittenullilisi elemente. Tulemusena saame maatriksi, kus üleval vasakus nurgas on r nullist erinevat elementi b_{ii}^i tänu eeldusele $r(\mathbf{A}) = r$. Korrutades nüüd saadud maatriksit vasakult elementaarmaatriksitega $\mathbf{R}\left(\frac{1}{b_{ii}^i}\right)$, $i = 1, \ldots, r$ saame teisendatud maatriksi teoreemi väites esitatud kujul

$$\mathbf{A} \cong \mathbf{PAQ} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

kus $\mathbf{P}$ ja $\mathbf{Q}$ tähistavad elementaarmaatriksite korrutisi. Saadud plokkmaatriksit nimetatakse *maatriksiga* $\mathbf{A}$ *ekvivalentseks kanooniliseks kujuks* või lihtsalt $\mathbf{A}$ *kanooniliseks kujuks*.

See teoreem annab meile meetodi maatriksi astaku leidmiseks: teisendame $\mathbf{A}$ elementaarteisenduste abil kanoonilisele kujule ja saame sealt teada astaku väärtuse.

Järgmine tulemus annab meile võimaluse esitada maatriks $\mathbf{A}$ astakuga $r(\mathbf{A}) = r$ kahe maatriksi korrutisena, kus üks maatriks on veergude ja teine ridade täisastakuga.

Teoreem 3.2 *Maatriks* $\mathbf{A} : m \times n$ *astakuga* $r(\mathbf{A}) = r$ *on esitatav kujul*

$$\mathbf{A} = \mathbf{KL},$$

kus $\mathbf{K} : m \times r$ *ja* $\mathbf{L} : r \times n$ *on täisastakuga maatriksid,* $r(\mathbf{K}) = r$ *ja* $r(\mathbf{L}) = r$.

Tõestus. Vastavalt eelmisele teoreemile leiduvad pööratavad maatriksid $\mathbf{P}$ ja $\mathbf{Q}$ nii, et

$$\mathbf{PAQ} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

kust

$$\mathbf{A} = \mathbf{P}^{-1} \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{Q}^{-1}.$$

Esitame maatriksid $\mathbf{P}^{-1}$ ja $\mathbf{Q}^{-1}$ plokkidest koosnevatena

$$\mathbf{P}^{-1} = (\mathbf{P}^{-1}_{(1)} \vdots \mathbf{P}^{-1}_{(2)}),$$

$$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{Q}^{-1}_{(1)} \\ \cdots \\ \mathbf{Q}^{-1}_{(2)} \end{pmatrix},$$

kus plokk $\mathbf{P}^{-1}_{(1)}$ on $m \times r$-maatriks ja $\mathbf{Q}^{-1}_{(1)}$ on $r \times n$-maatriks. Siis

$$\mathbf{A} = (\mathbf{P}^{-1}_{(1)} \vdots \mathbf{P}^{-1}_{(2)}) \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{Q}^{-1}_{(1)} \\ \cdots \\ \mathbf{Q}^{-1}_{(2)} \end{pmatrix} = (\mathbf{P}^{-1}_{(1)} \vdots \mathbf{0}) \begin{pmatrix} \mathbf{Q}^{-1}_{(1)} \\ \cdots \\ \mathbf{Q}^{-1}_{(2)} \end{pmatrix},$$

kust

$$\mathbf{A} = (\mathbf{P}^{-1}_{(1)} \mathbf{Q}^{-1}_{(1)}).$$

Võttes $\mathbf{K} = \mathbf{P}^{-1}_{(1)}$ ja $\mathbf{L} = \mathbf{Q}^{-1}_{(1)}$ olemegi saanud soovitud esituse, kuna maatriksite $\mathbf{P}$ ja $\mathbf{Q}$ pööratavuse tõttu on nendes maatriksites kõik read ja veerud lineaarselt sõltumatud ja seega ka plokid $\mathbf{P}^{-1}_{(1)}$ ja $\mathbf{Q}^{-1}_{(1)}$ täisastakuga.

§4. Omaväärtused, omavektorid

Omaväärtused ja omavektorid on ühed olulisemad maatriksalgebra mõisted mitmemõõtmelise statistika jaoks. Peakomponentide

analüüs, faktoranalüüs, korrespondentsanalüüs ja mitmed teised meetodid põhinevad sümmeetriliste maatriksite omaväärtustel ja omavektoritel.
Olgu $\mathbf{A} : n \times n$ reaalarvuline maatriks. Vaatame seda kui lineaarteisendust ruumis $\mathbb{R}^n$, mis teisendab suvalise vektori $\mathbf{x}$ teiseks vektoriks $\mathbf{y}$,

$$\mathbf{Ax} = \mathbf{y}.$$

Millised on need vektorid $\mathbf{x}$, mille siht säilib selle teisendusega? Selle nõude saame kirja panna võrrandina

$$\mathbf{Ax} = \lambda\mathbf{x}.$$

Teisiti saame selle esitada võrdusega

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}.$$

Sel võrrandil on mittetriviaalne lahend, kui

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0. \tag{4.1}$$

Võrrandit (4.1) nimetatakse maatriksi $\mathbf{A}$ *karakteristlikuks võrrandiks* ja selle lahendeid λ_i maatriksi $\mathbf{A}$ *omaväärtusteks*. Vektorit $\mathbf{x}_i$, mis rahuldab võrdust

$$\mathbf{Ax}_i = \lambda_i\mathbf{x}_i$$

nimetatakse omaväärtusele λ_i vastavaks *omavektoriks*. Determinandi definitsioonist järeldub, et karakteristlik võrrand on n-astme polünoom λ suhtes, polünoomil on aga n juurt, mis võivad olla nii reaalsed kui kompleksed, kusjuures nende hulgas võib olla kordseid. Seega on reaalarvulisel maatriksil n omaväärtust, mis võivad olla nii reaalarvulised kui kompleksarvud. Maatriksi $\mathbf{A}$ omaväärtuste hulka nimetatakse *spektriks*.
Vaatame lihtsat näidet.

Olgu

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} \qquad \mathbf{A} - \lambda\mathbf{I}_n = \begin{pmatrix} 1-\lambda & 2 \\ 3 & -\lambda \end{pmatrix}.$$

Karakteristliku võrrandi saame kujul

$$|\mathbf{A} - \lambda \mathbf{I}_n| = 0 = \lambda^2 - \lambda - 6,$$

kust omaväärtused $\lambda_1 = 3$; $\lambda_2 = -2$.
Omaväärtustele vastavad omavektorid ei ole üheselt määratud, varieeruda võivad nende pikkus ja suund. Kui $\mathbf{x}_i$ on omaväärtusele λ_i vastav omavektor, siis rahuldab ka $c\mathbf{x}_i$ omavektori nõudeid, kus $c \in \mathbb{R}$. Omavektori pikkuse osas lepitakse tavaliselt kokku, et vaadeldakse normeeritud omavektoreid pikkusega üks. Suuna fikseerimine on vähem oluline, vajadusel määratakse selleks kindlaks ühe mittenullilise koordinaadi märk (+ või –).
Statistika jaoks on eriti olulised sümmeetrilise maatriksi omaväärtused ja omavektorid. Järgnev teoreem annab meile nende kohta olulist teavet.
Teoreem 4.1 *Reaalarvulise sümmeetrilise maatriksi* $\mathbf{A}$ *kõik omaväärtused on reaalarvud.*

Tõestus. Olgu λ maatriksi $\mathbf{A}$ omaväärtus ja vastav omavektor $\mathbf{x} = \mathbf{u} + i\mathbf{v}$, kus $\mathbf{u}$ ja $\mathbf{v}$ on reaalarvulised vektorid. Siis

$$\mathbf{A}(\mathbf{u} + i\mathbf{v}) = \lambda(\mathbf{u} + i\mathbf{v}).$$

Korrutame seda võrdust vasakult kaaskompleksarvudest moodustatud transponeeritud vektoriga $(\mathbf{u} - i\mathbf{v})'$.

$$(\mathbf{u} - i\mathbf{v})'\mathbf{A}(\mathbf{u} + i\mathbf{v}) = \lambda(\mathbf{u} - i\mathbf{v})'(\mathbf{u} + i\mathbf{v}).$$

$$\mathbf{u}'\mathbf{A}\mathbf{u} - i\mathbf{v}'\mathbf{A}\mathbf{u} + i\mathbf{u}'\mathbf{A}\mathbf{v} + \mathbf{v}'\mathbf{A}\mathbf{v} = \lambda(\mathbf{u}'\mathbf{u} + \mathbf{v}'\mathbf{v}).$$

Maatriksi $\mathbf{A}$ sümmeetria tõttu $\mathbf{v}'\mathbf{A}\mathbf{u} = \mathbf{u}'\mathbf{A}\mathbf{v}$ ja kordajaga i liidetavad taanduvad välja ning vasakul pool võrdust on kahe reaalarvu summa. Paremal poolel on sulgudes samuti reaalarv, seega peab ka λ olema reaalarv.
Teoreem on tõestatud.
Tänu sellele, et omaväärtused on reaalarvud, saab ka omavektorid valida reaalarvuliste koordinaatidega.

Definitsioon 4.1 *Maatriksid* $\mathbf{A} : n \times n$ *ja* $\mathbf{B} : n \times n$ *on sarnased, kui*

$$\mathbf{B} = \mathbf{CAC}^{-1}.$$

Esitame järgnevalt olulisemad omaväärtuste ja omavektorite omadused.

(i) $\mathbf{A}$ ja $\mathbf{A}'$ on samade omaväärtustega.

(ii) Sarnased maatriksid $\mathbf{A}$ ja $\mathbf{B}$ on samade omaväärtustega.

(iii) Kui $\mathbf{A}$ on sümmeetriline maatriks omaväärtustega $\lambda_i \neq \lambda_j$, siis neile omaväärtustele vastavad omavektorid on ortogonaalsed.

(iv) Maatriksi $\mathbf{A}$ erinevatele omaväärtustele vastavad omavektorid on lineaarselt sõltumatud.

(v) Olgu $\mathbf{A}$ ja $\mathbf{B}$ $n \times n$-maatriksid, kusjuures $\mathbf{A}$ on pööratav. Siis $\mathbf{AB}$ ja $\mathbf{BA}$ on samade omaväärtustega.

(vi) Olgu $\mathbf{A} : n \times n$ pööratav omaväärtustega $\lambda_1, \ldots, \lambda_n$. Siis $\mathbf{A}^{-1}$ omaväärtused on $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$.

(vii) Olgu maatriks $\mathbf{A} : n \times n$ ortogonaalne. Siis tema omaväärtused on moodulilt võrdsed ühega (reaalarvulised omaväärtused on $+1$ või -1).

(viii) Idempotentse maatriksi omaväärtused on võrdsed kas nulli või ühega.

(ix) Maatriksi $\mathbf{A} : m \times n$ korral on maatriksitel $\mathbf{AA}'$ ja $\mathbf{A}'\mathbf{A}$ ühed ja samad mittenullilised omaväärtused.

(x) Kolmnurkse maatriksi omaväärtusteks on tema peadiagonaali elemendid.

(xi) Vähemalt üks singulaarse maatriksi omaväärtustest võrdub nulliga.

(xii) Maatriksid $\mathbf{A}$ ja $\mathbf{A} + c\mathbf{I}_n$ on ühede ja samade omavektoritega iga $c \in \mathbb{R}$ korral.

(xiii) Olgu $\mathbf{A}$ sümmeetriline $n \times n$-maatriks omaväärtustega λ_i. Siis $\sum_{i=1}^n \lambda_i^2 = \sum_{i,j=1}^n a_{ij}^2$.

Tõestame järgnevalt loetletud omadustest (iii) ja (iv).
(iii) *Tõestus.* Olgu $\lambda_i \neq \lambda_j$ sümmeetrilise maatriksi $\mathbf{A}$ omaväärtused, kusjuures

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i,$$

$$\mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_j.$$

Korrutame esimest võrdust vasakult transponeeritud vektoriga $\mathbf{x}_j'$ ja teist $\mathbf{x}_i'$. Saame võrdused

$$\mathbf{x}_j'\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_j'\mathbf{x}_i,$$

$$\mathbf{x}_i'\mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_i'\mathbf{x}_j.$$

Lahutame ülemisest võrdusest alumise. $\mathbf{A}$ sümmeetria tõttu

$$0 = (\lambda_i - \lambda_j)\mathbf{x}_i'\mathbf{x}_j.$$

Kuna $\lambda_i - \lambda_j \neq 0$, siis $\mathbf{x}_i'\mathbf{x}_j = 0$, mis on vektorite ristseisu tunnus. Järelikult erinevatele omaväärtustele vastavad omavektorid on ortogonaalsed.
(iv) *Tõestus.* Näitame vastuväiteliselt, et kahele suvalisele erinevale $\mathbf{A}$ omaväärtusele vastavad omavektorid on lineaarselt sõltumatud. Olgu

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \qquad \mathbf{A}\mathbf{y} = \mu\mathbf{y}.$$

Oletame, et omavektorid $\mathbf{x}$ ja $\mathbf{y}$ on lineaarselt sõltuvad. Siis leidub konstant $\alpha \neq 0$ nii, et $\alpha\mathbf{x} = \mathbf{y}$. Saame võrduste ahela

$$\mu\mathbf{y} = \mu\alpha\mathbf{x} = \mathbf{A}(\alpha\mathbf{x}) = \alpha\mathbf{A}\mathbf{x} = \alpha\lambda\mathbf{x}.$$

Lahutame võrduste rea teisest avaldisest viimase

$$\alpha(\mu - \lambda)\mathbf{x} = \mathbf{0}.$$

Kuna omaväärtused on erinevad ja omavektor $\mathbf{x} \neq \mathbf{0}$, saame vastuolu eeldusega ja $\alpha = 0$.
Mitmemõõtmelise statistika jaoks ülioluline on järgmine tulemus.
Teoreem 4.2 *Olgu* $\mathbf{A} : n \times n$ *sümmeetriline reaalarvuline maatriks. Siis leidub ortogonaalmaatriks* $\mathbf{P}$ *nii et*

$$\begin{aligned} \mathbf{P}'\mathbf{A}\mathbf{P} &= \mathbf{\Lambda}; \\ \mathbf{A} &= \mathbf{P}\mathbf{\Lambda}\mathbf{P}', \end{aligned}$$

kus $\mathbf{\Lambda}$ *on maatriksi* $\mathbf{A}$ *omaväärtuste diagonaalmaatriks.*

Tõestus. Teeme esmalt tõestuse läbi erijuhul kui maatriksil $\mathbf{A}$ on n erinevat omaväärtust $\lambda_1, \ldots, \lambda_n$. Olgu neile vastavad normeeritud omavektorid $\mathbf{p}_1, \ldots, \mathbf{p}_n$. Siis

$$\mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i; \qquad i = 1, \ldots, n,$$

kusjuures vektorid $\mathbf{p}_i$ ja $\mathbf{p}_j$ on omaduse (iii) põhjal ortogonaalsed,

$$\mathbf{p}_i'\mathbf{p}_j = 0; \qquad i \neq j,$$

$$\mathbf{p}_i'\mathbf{p}_i = 1; \qquad i = 1, \ldots, n.$$

Ülalolevad võrdused saame esitada kompaktsemalt maatrikskujul

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{\Lambda}, \tag{4.1}$$

$$\mathbf{P}'\mathbf{P} = \mathbf{I}_n,$$

kus $\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$. Korrutades vasakult võrdust (4.1) maatriksiga $\mathbf{P}'$, saame

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{\Lambda}$$

ja korrutades võrdust (4.1) paremalt maatriksiga $\mathbf{P}'$ saame võrduse

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Oleme saanud teoreemi väites esitatud võrdused.
Üldjuhul on tõestus tehniliselt pikem, aga mitte oluliselt keerulisem. Esitame järgnevalt tõestuse üldidee.
Näidata on vaja, et leidub selline ortogonaalmaatriks $\mathbf{P}$, et

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{D},$$

kus $\mathbf{D}$ on diagonaalne, kusjuures peadiagonaali elemendid d_{ii} on $\mathbf{A}$ omaväärtused. Selles on kerge veenduda:

$$|\mathbf{D} - \lambda\mathbf{I}_n| = |\mathbf{P}'\mathbf{A}\mathbf{P} - \lambda\mathbf{I}_n| = |\mathbf{P}'\mathbf{A}\mathbf{P} - \lambda\mathbf{P}'\mathbf{P}| = |\mathbf{P}'||\mathbf{A} - \lambda\mathbf{I}_n||\mathbf{P}|$$

$$= |\mathbf{A} - \lambda\mathbf{I}_n||\mathbf{P}'\mathbf{P}| = |\mathbf{A} - \lambda\mathbf{I}_n||\mathbf{I}_n| = |\mathbf{A} - \lambda\mathbf{I}_n|.$$

Kuidas seda maatriksit $\mathbf{P}$ leida?
Erinevatele omaväärtustele vastavad omavektorid $\mathbf{p}_i$ on ortogonaalsed, olgu neid k tükki. Saame neist moodustada osa maatriksist $\mathbf{P}$:

$$(\mathbf{p}_1, \ldots, \mathbf{p}_k).$$

Edasi, kui λ on r-kordne omaväärtus, siis saab tõestada, et ruumis $\mathbb{R}^n$ leidub r lineaarselt sõltumatut vektorit $\mathbf{z}(\lambda)_1, \ldots, \mathbf{z}(\lambda)_r$, mis rahuldavad omaväärtusele λ vastavate omavektorite nõudeid. Need saab aga ortogonaliseerida korrutades maatriksit

$$(\mathbf{z}(\lambda)_1, \ldots, \mathbf{z}(\lambda)_r)$$

paremalt ülemise kolmnurkmaatriksiga $\mathbf{T}^{\triangle} : r \times r$ (vaatame pärast tõestuse lõppu näidet, kuidas seda kolmnurkmaatriksit leida)

$$(\mathbf{z}(\lambda)_1, \ldots, \mathbf{z}(\lambda)_r)\mathbf{T}^{\triangle} = (\mathbf{o}_1, \ldots, \mathbf{o}_r).$$

Üks neist vektoritest $\mathbf{z}(\lambda)_j$ on juba varem olemas maatriksis $(\mathbf{p}_1, \ldots, \mathbf{p}_k)$. Lisame sealsetele veergudele ka ülejäänud $r-1$ leitud omavektorit $\mathbf{o}_i$ ja talitame nii kõigi kordsete omaväärtuste korral. Tulemusena saame ortogonaalsetest normeeritud vektoritest koosneva $n \times n$ maatriksi $\mathbf{P}$, mis vastab teoreemis esitatud nõuetele.

Näide ortogonaliseerimisest. Olgu meil kaks sõltumatut normeeritud omavektorit $\mathbf{z}_1$ ja $\mathbf{z}_2$, mis vastavad omaväärtusele λ. Leiame teisenduse $\mathbf{T}^{\triangle}$, millega teisendame need vektorid ortogonaalseteks.

$$\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \\ z_{31} & z_{32} \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} \\ 0 & t_{22} \end{pmatrix} = \begin{pmatrix} z_{11}t_{11} & z_{11}t_{12} + z_{12}t_{22} \\ z_{21}t_{11} & z_{21}t_{12} + z_{22}t_{22} \\ z_{31}t_{11} & z_{31}t_{12} + z_{32}t_{22} \end{pmatrix}.$$

Tahame, et teisendatud veeruvektorid oleksid risti. Valime $t_{11} = 1$, sellega jätame esimese veeruvektori muutmata,

$$z_{11}(z_{11}t_{12}+z_{12}t_{22})+z_{21}(z_{21}t_{12}+z_{22}t_{22})+z_{31}(z_{31}t_{12}+z_{32}t_{22}) = 0.$$

Liidetavaid ümber grupeerides saame

$$t_{12}(z_{11}^2 + z_{21}^2 + z_{31}^2) + t_{22}(z_{11}z_{12} + z_{21}z_{22} + z_{31}z_{32}) = 0.$$

Eeldasime, et vektorid $\mathbf{z}_i$ on normeeritud, seega elemendi t_{12} kordaja võrdub ühega. Kui t_{22} kordaja võrdub nulliga, siis $t_{12} = 0$. Vastasel korral

$$t_{12} + t_{22}(z_{11}z_{12} + z_{21}z_{22} + z_{31}z_{32}) = t_{22}(\mathbf{z}_1'\mathbf{z}_2) = 0.$$

Võtame

$$t_{12} = -t_{22}\mathbf{z}_1'\mathbf{z}_2).$$

Siis saame võtta ka $t_{22} = 1$ ja $t_{12} = -\mathbf{z}_1'\mathbf{z}_2$.

Oleme saanud ortogonaalsed normeeritud vektorid

$$\begin{pmatrix} z_{11} & z_{12} - z_{11}\mathbf{z}_1'\mathbf{z}_2 \\ z_{21} & z_{22} - z_{21}\mathbf{z}_1'\mathbf{z}_2 \\ z_{31} & z_{32} - z_{31}\mathbf{z}_1'\mathbf{z}_2 \end{pmatrix}.$$

Sümmeetriline maatriks $\mathbf{A}$ on esitatav omaväärtuste ja normeeritud omavektorite kaudu ka nn. *spektraallahutusena*

Järeldus 4.2.1 (spektraallahutus). *Sümmeetriline maatriks* $\mathbf{A}$ *on esitatav omaväärtuste ja normeeritud omavektorite kaudu kujul*

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i',$$

kus $\mathbf{p}_i$ *on omaväärtusele* λ_i *vastav normeeritud omavektor.*

Tõestus. Teoreemi 4.2 põhjal teame, et

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Järelduse tõestamiseks peame näitama, et maatriksi $\mathbf{A}$ üldelement a_{ij} avaldub samal kujul nii teoreemis toodud esituse kui järelduse väite korral. Maatriksite korrutamiseeskirja kohaselt ja arvestades transponeerimise omadust saame teoreemi väitest võrduse

$$a_{ij} = \sum_{k=1}^{n} p_{ik} \lambda_k p_{jk}.$$

Järelduse väitest saame sama avaldise

$$a_{ij} = \sum_{k=1}^{n} \lambda_k p_{ik} p_{jk}.$$

Järeldus 4.2.2 *Kehtib järgmine võrdus*

$$\sum_{k=1}^{n} \mathbf{p}_k \mathbf{p}_k' = \mathbf{I}_n.$$

Tõestus. Kuna $\mathbf{I}_n = \mathbf{P}'\mathbf{P}$, siis analoogiliselt eelmise järelduse tõestusega saame näidata, et ühikmaatriksi kahe erineva esituse üldelemendid on võrdsed.

Kasulikuks osutub ka järgmine tulemus, mida tuntakse *Rayleigh suhte* nime all.

Teoreem 4.3 (Rayleigh suhe). *Olgu* $\mathbf{A} : n \times n$ *sümmeetriline maatriks omaväärtustega* $\lambda_i,\ i = 1 \ldots, n$. *Siis iga mittenullilise* $\mathbf{x} \in \mathbb{R}^n$ *korral*

$$\lambda_{min} \le \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} \le \lambda_{max}.$$

Tõestus. Teoreemi 4.2 põhjal leidub $\mathbf{A}$ omavektorite ortogonaalmaatriks $\mathbf{P}$ nii, et $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$, kus $\mathbf{\Lambda}$ on omaväärtuste diagonaalmaatriks. Siis

$$\frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\mathbf{x}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{x}}{\mathbf{x}'\mathbf{P}\mathbf{P}'\mathbf{x}}.$$

Kuna $|\mathbf{P}| \neq 0$, siis ka $\mathbf{y} = \mathbf{P}'\mathbf{x} \neq \mathbf{0}$. Saame

$$\frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\mathbf{y}'\mathbf{\Lambda}\mathbf{y}}{\mathbf{y}'\mathbf{y}} = \frac{\sum_{i=1}^{n} \lambda_i y_i^2}{\sum_{i=1}^{n} y_i^2}.$$

Asendades kõik omaväärtused vähimaga ja suurimaga saamegi nõutavad võrratused.

Mida saab öelda mittesümmeetrilise maatriksi omavektorite ortogonaalsuse kohta?
Sel juhul eristatakse vasakpoolseid ja parempoolseid (ehk tavalisi) omavektoreid.
Definitsioon 4.2 *Nimetame vektorit* $\mathbf{y}$ *maatriksi* $\mathbf{A} : n \times n$ *vasakpoolseks omavektoriks, mis vastab omaväärtusele* μ, *kui*

$$\mathbf{y}'\mathbf{A} = \mu\mathbf{y}'. \tag{4.2}$$

Lihtne on veenduda, et μ on maatriksi $\mathbf{A}$ omaväärtus. Definitsioonist saame

$$\mathbf{y}'\mathbf{A} - \mu\mathbf{y}'\mathbf{I}_n = \mathbf{0}_{1\times n} = \mathbf{y}'(\mathbf{A} - \mu\mathbf{I}_n).$$

Kuna $\mathbf{y} \neq \mathbf{0}$, siis $|\mathbf{A} - \mu\mathbf{I}_n| = 0$, kust järeldub, et μ on maatriksi $\mathbf{A}$ omaväärtus.
Olgu $\mathbf{x}$ parempoolne omavektor, mis vastab omaväärtusele λ

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Korrutame viimast võrdust vasakult transponeeritud vektoriga $\mathbf{y}'$

$$\mathbf{y}'\mathbf{A}\mathbf{x} = \lambda\mathbf{y}'\mathbf{x}.$$

Korrutades võrdust (4.2) paremalt vektoriga $\mathbf{x}$, saame

$$\mathbf{y}'\mathbf{A}\mathbf{x} = \mu\mathbf{y}'\mathbf{x}.$$

Lahutades eelmisest võrdusest viimase, saame

$$0 = (\lambda - \mu)\mathbf{y}'\mathbf{x}.$$

Kui $\lambda \neq \mu$, siis vektorid $\mathbf{x}$ ja $\mathbf{y}$ on ortogonaalsed.
Saame sõnastada tulemuse.
Teoreem 4.4 *Maatriksi* $\mathbf{A} : n \times n$ *erinevatele omaväärtustele vastavad vasakpoolsed ja parempoolsed omavektorid on ortogonaalsed.*

§5. Maatriksi positiivne määratus

Kaks väga olulist mõistet mitmemõõtmelises statistikas on kovariatsioonimaatriks ja korrelatsioonimaatriks, mis mõlemad on kas positiivselt või mittenegatiivselt määratud. Tuletame meelde nende definitsioonid.
Definitsioon 5.1 *Juhusliku* p*-vektori* $\mathbf{x} = (X_1, \ldots, X_p)'$ *kovariatsioonimaatriks* $D\mathbf{x} : p \times p$ *on defineeritud keskväärtusena*

$$D\mathbf{x} = E[(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})'],$$

kui maatriksi elementide keskväärtused on lõplikud.
Tähistame $D\mathbf{x} = \mathbf{\Sigma}$. Siis

$$\sigma_{ij} = Cov(X_i, X_j) = E[(X_i - EX_i)(X_j - EX_j)], \; i \neq j,$$

$$\sigma_{ii} = DX_i, \; i = 1, \ldots, p.$$

Korrelatsioonimaatriksi $\mathbf{R}$ saame kovariatsioonimaatriksist teisenduse teel

$$\mathbf{R} = \mathbf{\Sigma}_d^{-\frac{1}{2}}\mathbf{\Sigma}\mathbf{\Sigma}_d^{-\frac{1}{2}},$$

kus
$$r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}, \; i \neq j, \quad r_{ii} = 1.$$
(Kontrollige lihtsal erijuhul, kui tekib kahtlus nende võrduste osas).
Alustame positiivse määratuse definitsioonist.
Definitsioon 5.2 *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on positiivselt (negatiivselt) määratud, kui*
$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0, \; (< 0) \qquad \textit{iga } \mathbf{x} \neq \mathbf{0}$$
korral.
Tähistame $\mathbf{A} > 0$, $(\mathbf{A} < 0)$.
Definitsioon 5.3 *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on positiivselt (negatiivselt) poolmääratud, kui*
$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0, \; (\leq 0) \qquad \textit{iga } \mathbf{x} \in \mathbb{R}^n$$
korral ja leidub $\mathbf{x}_0 \neq \mathbf{0}$ *nii, et*
$$\mathbf{x}_0'\mathbf{A}\mathbf{x}_0 = 0.$$
Tähistame $\mathbf{A} \geq 0$, $(\mathbf{A} \leq 0)$.
Definitsioon 5.4 *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on mittenegatiivselt (mittepositiivselt) määratud, kui*
$$\mathbf{A} > 0 \quad \textit{või} \quad \mathbf{A} \geq 0,$$
$$(\mathbf{A} < 0 \quad \textit{või} \quad \mathbf{A} \leq 0).$$
Mitmed positiivse määratuse omadused tuginevad järgmisel teoreemil.
Teoreem 5.1 *Olgu* $n \times n$*-maatriks* $\mathbf{A} > 0$ *ja* $\mathbf{B} : n \times m, \; m \leq n, \; r(\mathbf{B}) = m$. *Siis*
$$\mathbf{B}'\mathbf{A}\mathbf{B} > 0.$$

Tõestus. Võtame suvalise m-vektori $\mathbf{y} \neq \mathbf{0}$. Kuna $r(\mathbf{B}) = m$, siis $\mathbf{B}\mathbf{y} = \mathbf{x} \neq \mathbf{0}$.

Siis

$$\mathbf{y}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{y} = (\mathbf{B}\mathbf{y})'\mathbf{A}(\mathbf{B}\mathbf{y}) = \mathbf{x}'\mathbf{A}\mathbf{x} > 0.$$

Kuna $\mathbf{B}'\mathbf{A}\mathbf{B}$ on sümmeetriline, on positiivse määratuse tingimused täidetud ja $\mathbf{B}'\mathbf{A}\mathbf{B} > 0$.
Järgnevad omadused on ühtlasi ka kovariatsioonimaatriksi ja korrelatsioonimaatriksi omadused.

(i) Maatriks $\mathbf{A} : n \times n$ on positiivselt määratud parajasti siis, kui kõik tema nurgamiinorid on positiivsed;

(ii) Kui $\mathbf{A} > 0$, siis $\mathbf{A}^{-1} > 0$;

(iii) Maatriks $\mathbf{A} > 0$ parajasti siis, kui kõik tema omaväärtused $\lambda_i > 0$.

(iv) Olgu $\mathbf{A} > 0$, $\mathbf{A}$ ja $\mathbf{B} : n \times n$ ning $\mathbf{B}$ pööratav maatriks. Siis

$$\mathbf{B}'\mathbf{A}\mathbf{B} > 0.$$

(v) Suvalise $\mathbf{A}$ korral on $\mathbf{A}\mathbf{A}'$ ja $\mathbf{A}'\mathbf{A}$ mittenegatiivselt määratud.

(vi) Kui $\mathbf{A}$ on mittenegatiivselt määratud, siis leidub $\mathbf{A}^{-1}$ parajasti siis, kui $\mathbf{A} > 0$.

(vii) Kui $\mathbf{A} > 0$, $\mathbf{A} : n \times n$ ja $\mathbf{B} : n \times m$, $m \leq n$ ja $r(\mathbf{B}) = r$, siis $\mathbf{B}'\mathbf{A}\mathbf{B} > 0$ parajasti siis, kui $r = m$. Maatriks $\mathbf{B}'\mathbf{A}\mathbf{B} \geq 0$, kui $r < m$.

(viii) Kui $\mathbf{A} > 0$, $\mathbf{B} > 0$ ja $\mathbf{A} - \mathbf{B} > 0$, siis $\mathbf{B}^{-1} - \mathbf{A}^{-1} > 0$ ja $|\mathbf{A}| > |\mathbf{B}|$.

(ix) Kui $\mathbf{A} > 0$, $\mathbf{B} > 0$, siis $|\mathbf{A} + \mathbf{B}| \geq |\mathbf{A}| + |\mathbf{B}|$.

Tõestame omaduse (iii)

Tarvilikkus. Kui $\mathbf{A} > 0$, siis iga $\mathbf{x} \neq \mathbf{0}$ korral $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$. Esitame $\mathbf{A}$ omavektorite ortogonaalmaatriksi $\mathbf{P}$ ja omaväärtuste diagonaalmaatriksi $\mathbf{\Lambda}$ kaudu

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Kuna $\mathbf{A} > 0$, siis ka $\mathbf{P}'\mathbf{A}\mathbf{P} > 0$ teoreemi 5.1 kohaselt.
Seega

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{P}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{P} = \mathbf{\Lambda}.$$

Omadusest (i) järeldub, et $\lambda_i > 0, \quad i = 1, \ldots, n$. Teisiti jõuame samale tulemusele võttes vektoriks $\mathbf{x}$ baasivektori $\mathbf{e}_i$. Tänu $\mathbf{\Lambda}$ positiivsele määratusele

$$\mathbf{e}_i'\Lambda\mathbf{e}_i = \lambda_i > 0$$

iga i korral.

Piisavus. Olgu sümmeetrilise maatriksi $\mathbf{A}$ kõik omaväärtused $\lambda_i > 0$. Siis omaväärtuste diagonaalmaatriks $\mathbf{\Lambda}$ on positiivselt määratud. Olgu $\mathbf{P}$ omaväärtustele λ_i vastavatest omavektoritest moodustatud ortogonaalmaatriks. Siis aga ka

$$\mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{A} > 0.$$

Omadus (iii) on tõestatud.

Maatriksite järjestamisest

Järjestussuhet on maatriksite hulgas defineeritud mitmeti, kusjuures tegemist on osalise järjestusega. Statistikas kõige enam kasutatav järjestussuhe on nn. Löwneri järjestus

Definitsioon 5.4 *Olgu sümmeetrilised maatriksid* $\mathbf{A}$, $\mathbf{B}$: $n \times n$. *Öeldakse, et* $\mathbf{A}$ *ja* $\mathbf{B}$ *on ranges Löwneri järjestuses, kui*

$$\mathbf{A} - \mathbf{B} > 0.$$

Tähistatakse seda suhet sel juhul $\mathbf{A} >_L \mathbf{B}$.

Mõnevõrra nõrgema järjestuse annab tavaline *Löwneri järjestus.*
Definitsioon 5.5 *Olgu sümmeetrilised maatriksid* $\mathbf{A}, \mathbf{B} : n \times n$. *Öeldakse, et* $\mathbf{A}$ *ja* $\mathbf{B}$ *on Löwneri järjestuses,* $\mathbf{A} \geq_L \mathbf{B}$*, kui*

$$\mathbf{A} - \mathbf{B} = \mathbf{C}\mathbf{C}'$$

mingi maatriksi $\mathbf{C}$ *korral.*
Seega tavalise Lõwneri järjestuse korral $\mathbf{A}-\mathbf{B} > 0$ või $\mathbf{A}-\mathbf{B} \geq 0$. Erinevaid järjestussuhteid on sisse toodud ka ristkülikmaatriksite korral, kuid statistikas ei ole need kasutamist leidnud.

§6. Maatriksi jälg

Definitsioon 6.1. *Maatriksi* $\mathbf{A} : n \times n$ *jäljeks* $\mathrm{tr}(\mathbf{A})$ *nimetatakse tema peadiagonaali elementide summat*

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Esitame jälje olulisemad omadused:

(i) $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, kus $\mathbf{A} : m \times n$ ja $\mathbf{B} : n \times m$;

(ii) $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{B}^{-1}\mathbf{AB})$, kui $\mathbf{B}$ on pööratav;

(iii) $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{C}'\mathbf{AC})$, kui $\mathbf{C}$ on ortogonaalne;

(iv) $\mathrm{tr}(c_1\mathbf{A} + c_2\mathbf{B}) = c_1\mathrm{tr}\mathbf{A} + c_2\mathrm{tr}\mathbf{B}$, kus c_1, c_2 – konstandid;

(v) $\mathrm{tr}(\mathbf{A}) = r(\mathbf{A})$, kui $\mathbf{A}$ on idempotentne;

(vi) $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = 0$ siis ja ainult siis, kui $\mathbf{A} = \mathbf{0}$;

(vii) $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}')$;

(viii) $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{AA}') = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2, \quad \mathbf{A} : m \times n$;

(ix) $\mathrm{tr}(\mathbf{AB})^2 \leq \mathrm{tr}(\mathbf{A}^2\mathbf{B}^2), \quad \mathbf{A}, \mathbf{B}$ –sümmeetrilised;

(x) $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathrm{tr}(\mathbf{A}\mathbf{x}\mathbf{x}')$;

(xi) $r(\mathbf{A}) \geq \frac{(\mathrm{tr}\mathbf{A})^2}{\mathrm{tr}(\mathbf{A}^2)}$, kui $\mathbf{A} \neq \mathbf{0}$ on sümmeetriline;

(xii) $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i, \quad \mathbf{A}$ sümmeetriline;

(xiii) $\mathrm{tr}(\mathbf{A}^k) = \sum_{i=1}^n \lambda_i^k, \quad \mathbf{A}$ sümmeetriline;

(xiv) $\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$.

Tõestame omaduse (xii).
Esitame maatriksi $\mathbf{A}$ omavektorite ortogonaalmaatriksi $\mathbf{P}$ ja omaväärtuste diagonaalmaatriksi $\mathbf{\Lambda}$ kaudu

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Siis

$$\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{P}\mathbf{\Lambda}\mathbf{P}') = \mathrm{tr}(\mathbf{\Lambda}\mathbf{P}'\mathbf{P}) = \mathrm{tr}(\mathbf{\Lambda}) = \sum_{i=1}^n \lambda_i.$$

Järgnev tulemus, mis on tuntud *Aitken'i integraalina* on näide sellest kuidas tõenäosusteooria ja statistika mõisted on mõnikord ootamatult seotud maatriksalgebraga.

Lause 6.1 (Aitkeni integraal). *Olgu* $\mathbf{A} : n \times n$ *pööratav maatriks. Siis*

$$|\mathbf{A}| = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathrm{tr}[(\mathbf{A}\mathbf{A}')^{-1}\mathbf{x}\mathbf{x}']} d\mathbf{x}.$$

Tõestus. Kuna $\mathbf{A}$ on pööratav, siis $\mathbf{A}\mathbf{A}' > 0$. Tähistame $\mathbf{A}\mathbf{A}' = \mathbf{\Sigma}$. Kirjutame välja normaaljaotuse $N_n(\mathbf{0}, \mathbf{\Sigma})$ tihedusfunktsiooni, kus $\mathbf{\Sigma}$ on kovariatsioonimaatriks

$$f(\mathbf{x}) = \frac{1}{\sqrt{2\pi}^n |\mathbf{\Sigma}|^{1/2}} e^{-\frac{1}{2}\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}}.$$

Kuna $\mathbf{\Sigma} > 0$, siis ka $\mathbf{\Sigma}^{-1} > 0, \quad \mathbf{\Sigma}^{-1} = (\mathbf{A}\mathbf{A}')^{-1}$. Siis

$$f(\mathbf{x}) = \frac{1}{\sqrt{2\pi}^n |(\mathbf{A}\mathbf{A}')|^{1/2}} e^{-\frac{1}{2}\mathbf{x}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{x}}. \tag{6.1}$$

Et $f(\mathbf{x})$ on tõenäosustihedus, siis

$$\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} = 1.$$

Integreerime võrduse (6.1) mõlemat poolt üle $\mathbb{R}^n$ ja võtame arvesse seose $|(\mathbf{AA}')|^{1/2} = |\mathbf{A}|^{1/2}|\mathbf{A}'|^{1/2} = |\mathbf{A}|$

$$\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} = 1 = \frac{1}{\sqrt{2\pi}^n|\mathbf{A}|}\int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathbf{x}'(\mathbf{AA}')^{-1}\mathbf{x}}d\mathbf{x}.$$

Siit

$$|\mathbf{A}| = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathbf{x}'(\mathbf{AA}')^{-1}\mathbf{x}}d\mathbf{x} = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathrm{tr}[(\mathbf{AA}')^{-1}\mathbf{xx}']}d\mathbf{x}.$$

Viimase võrduse saamisel rakendasime jälje omadust (x).

§7. Üldistatud pöördmaatriks

Üldistatud pöördmaatriksite temaatika on üks näide maatriksalgebrast, mis on välja arenenud statistika vajadustest ja panustajateks on olnud suuresti statistikud. Enimviidatavaks on kujunenud monograafia kahelt statistika klassikult
C.R. Rao, S.M. Mitra (1971). *Generalized Inverse of Matrices and Its Applications.* Wiley, New York.

Muide, C.R. Rao sajandat sünnipäeva tähistasid statistikud üle maailma 2020. aasta septembris ja 1994. aastal oli ta meie Tartu Mitmemõõtmelise statistika konverentsi peaesineja.
Kui $\mathbf{A}$ on pööratav maatriks, siis lineaarne võrrand tundmatu $\mathbf{x}$ suhtes

$$\mathbf{Ax} = \mathbf{y} \tag{7.1}$$

lahendatakse maatriksi $\mathbf{A}$ pööramisega

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

Tegelikult esitab võrrand (7.1) võrrandisüsteemi, kus on n tundmatut x_i

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n. \end{cases}$$

Iga võrrand tähendab tingimust muutujate x_i vahel ja alati ei pruugi sugugi õnnestuda leida sama palju tingimusi kui on muutujaid, seda eriti siis, kui n ei ole väike. Samas võib juhtuda ka, et on küll leitud n tingimust, aga osa neist on lineaarselt sõltuvad ja me ei saa maariksit $\mathbf{A}$ pöörata.

Esimesel juhul soovime lahendada võrrandit (7.1), kus $\mathbf{A}$ on $m \times n$-maatriks, $m \neq n$, teisel juhul $\mathbf{A} : n \times n$ kuid $|\mathbf{A}| = 0$.

Selleks, et lahendada süsteem (7.1) nendel tingimustel, on sisse toodud üldistatud pöördmaatriksi mõiste. Selle lahendamisest saame rääkida vaid juhul, kui lahend leidub ja süsteem ei ole vastuoluline

Öeldakse et *võrrand* $\mathbf{Ax} = \mathbf{y}$ *on lahenduv*, kui leidub vektor $\mathbf{x}_0$ nii, et $\mathbf{Ax}_0 = \mathbf{y}$.

Kui $\mathbf{A}$ on pööratav, on selliseid vektoreid üks ja ainult üks, lahenduva süsteemi korral on lahendeid üldjuhul rohkem kui üks.

Definitsioon 7.1 *Olgu* $\mathbf{A} : m \times n$. *Maatriksi* $\mathbf{A}$ *üldistatud pöördmaatriksiks nimetatakse* $n \times m$*-maatriksit* $\mathbf{A}^-$*, kui*

$$\mathbf{x} = \mathbf{A}^-\mathbf{y}$$

on võrrandi $\mathbf{Ax} = \mathbf{y}$ *lahendiks iga* $\mathbf{y}$ *korral, kui see võrrand on lahenduv*

Teoreem 7.1 *Maatriks* $\mathbf{A}^- : n \times m$ *on maatriksi* $\mathbf{A} : m \times n$ *üldistatud pöördmaatriks parajasti siis, kui*

$$\mathbf{AA}^-\mathbf{A} = \mathbf{A}. \tag{7.2}$$

Tõestus. Tarvilikkus. Näitame, et $\mathbf{A}^-$ olemasolust järeldub võrdus (7.2). Üldistatud pöördmaatriksi olemasolu tõttu võrrand (7.1) on lahenduv ja leidub $\mathbf{x}_0$ nii, et $\mathbf{y} = \mathbf{Ax}_0$.

Aga definitsioonist (7.1) järeldub, et siis

$$\mathbf{x} = \mathbf{A}^{-}(\mathbf{A}\mathbf{x}_0),$$

sealt

$$\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{A}^{-}\mathbf{A}\mathbf{x}_0.$$

Aga

$$\mathbf{A}\mathbf{x} = \mathbf{y} = \mathbf{A}\mathbf{x}_0 = \mathbf{A}\mathbf{A}^{-}\mathbf{A}\mathbf{x}_0,$$

kust

$$\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}.$$

Piisavus. Kehtigu (7.2) ja näitame, et siis $\mathbf{x} = \mathbf{A}^{-}\mathbf{y}$. Olgu võrrand (7.1) lahenduv. Siis leidub vektor $\mathbf{w}$ nii, et $\mathbf{A}\mathbf{w} = \mathbf{y}$. Kuna $\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}$, siis saame järgmise võrduste ahela

$$\mathbf{A}\mathbf{A}^{-}\mathbf{A}\mathbf{w} = \mathbf{A}\mathbf{w} = \mathbf{A}\mathbf{A}^{-}\mathbf{y} = \mathbf{y},$$

See tähendab, et vektor $\mathbf{A}^{-}\mathbf{y}$ on võrrandi (7.1) lahendiks iga $\mathbf{y}$ korral, kui $\mathbf{A}\mathbf{x} = \mathbf{y}$ on lahenduv, ehk

$$\mathbf{x} = \mathbf{A}^{-}\mathbf{y}.$$

Tarvilik ja piisav tingimus tähendab, et võiksime võrdust (7.2) kasutada ka definitsioonina. Tihti algebralises käsitluses nii tehaksegi.
Üldistatud pöördmaatriksil $\mathbf{A}^{-}$ on rida puudusi. Tal puuduvad mitmed pöördteisendusele tavapära-sed omadused. Näiteks ei pruugi $\mathbf{A}$ olla maatriksi $\mathbf{A}^{-}$ üldistatud pöördmaatriks.
Definitsioon 7.2 *Öeldakse, et maatriksi* $\mathbf{A}$ *üldistatud pöördmaatriks* $\mathbf{A}^{-}$ *on maatriksi* $\mathbf{A}$ *refleksiivne üldistatud pöördmaatriks, kui ta rahuldab järgmist võrdust*

$$\mathbf{A}^{-}\mathbf{A}\mathbf{A}^{-} = \mathbf{A}^{-}.$$

Seega on $\mathbf{A}$ ja $\mathbf{A}^{-}$ teineteise üldistatud pöördmaatriksid kui $\mathbf{A}^{-}$ on refleksiivne üldistatud pöördmaatriks. Samas ei ole refleksiivne üldistatud pöördmaatriks ühene ka ruutmaatriksi korral

nagu tavaline pöördmaatriks. Üldistatud pöördmaatriksi ühesuse tagamiseks on vaja veel tingimusi lisada. Ühesuse nõue on rahuldatud *Moore-Penrose'i üldistatud pöördmaatriksi* korral.
Definitsioon 7.3 *Maatriksi* $\mathbf{A} : m \times n$ *Moore-Penrose'i üldistatud pöördmaatriksiks nimetatakse* $n \times m$ *maatriksit* $\mathbf{A}^+$*, kui on rahuldatud järgmised neli võrdust:*

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A};$$

$$\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+;$$

$$(\mathbf{A}\mathbf{A}^+)' = \mathbf{A}\mathbf{A}^+;$$

$$(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}.$$

Teoreem 7.2 *Moore-Penrose'i üldistatud pöördmaatriks* $\mathbf{A}^+$ *on ühene.*

Tõestus. Oletame vastuväiteliselt, et $m \times n$-maatriksi $\mathbf{A}$ jaoks leiduvad maatriksid $\mathbf{B}$ ja $\mathbf{C}$, mis mõlemad rahuldavad Moore-Penrose'i üldistatud pöördmaatriksi tingimusi, kusjuures $\mathbf{B} \neq \mathbf{C}$. Näitame, et see eeldus viib vastuolule.
Saame järgmise võrduste rea

$$\begin{aligned}
\mathbf{B} &= \mathbf{BAB} = \mathbf{B}(\mathbf{AB})' = \mathbf{BB}'\mathbf{A}' = \mathbf{BB}'(\mathbf{ACA})' \\
&= \mathbf{B}(\mathbf{AB})'(\mathbf{AC})' = \mathbf{BABAC} = \mathbf{BAC} = (\mathbf{BA})'\mathbf{CAC} \\
&= (\mathbf{BA})'(\mathbf{CA})'\mathbf{C} = \mathbf{A}'\mathbf{B}'\mathbf{A}'\mathbf{C}'\mathbf{C} = \mathbf{A}'\mathbf{C}'\mathbf{C} = (\mathbf{CA})'\mathbf{C} \\
&= \mathbf{CAC} = \mathbf{C},
\end{aligned}$$

vastuolu eeldusega $\mathbf{B} \neq \mathbf{C}$.
Üldistatud pöördmaatriksite definitsioonid ei ole konstruktiivsed. Kuidas leida maatriksi $\mathbf{A} : m \times n$ üldistatud pöördmaatriksit? Olgu $\mathbf{A} : m \times n$ astakuga $r(\mathbf{A}) = r$. Siis ridade ja veergude ümbervahetamisega saame alati jõuda olukorrani, kus ülal vasakus nurgas olev $r \times r$-alammaatriks on pööratav. Eeldame, et maatriks $\mathbf{A}$ ongi sellisel kujul ja saame selle esitada plokkidena

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix},$$

kus $\mathbf{B}$ on pööratav.
Väidame, et siis

$$\mathbf{A}^{-} = \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Selle väite saame tõestada oma kursuses hiljem.
Teise võimaluse üldistatud pöördmaatriksi leidmiseks saame kasutades astakuga r maatriksi $\mathbf{A} : m \times n$ esitust täisastakuga maatriksite korrutisena

$$\mathbf{A} = \mathbf{KL},$$

kus $\mathbf{K}$ on veergude ja $\mathbf{L}$ ridade täisastakuga r maatriksid.
Lause 7.1 *Ülaltoodud eeldustel*

$$\mathbf{A}^{-} = \mathbf{L}'(\mathbf{K}'\mathbf{A}\mathbf{L}')^{-1}\mathbf{K}'.$$

Tõestus. Kontrollime, kas tarvilik ja piisav tingimus teoreemis 7.1 on täidetud.

$$\begin{aligned} \mathbf{A}\mathbf{A}^{-}\mathbf{A} &= \mathbf{KLL}'(\mathbf{K}'\mathbf{AL}')^{-1}\mathbf{K}'\mathbf{KL} = \mathbf{KL}(\mathbf{L}'(\mathbf{K}'\mathbf{KLL}')^{-1}\mathbf{K}'\mathbf{KL} \\ &= \mathbf{KL}(\mathbf{L}'(\mathbf{LL}')^{-}1(\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{KL} = \mathbf{KL} = \mathbf{A}. \end{aligned}$$

Statistika jaoks eriti olulisel kohal on sümmeetrilised maatriksid. Osutub, et sel korral on Moore-Penrose'i üldistatud pöördmaatriks lihtsalt leitav omaväärtuste diagonaalmaatriksi ja omavektorite ortogonaalmaatriksi kaudu.
Saame vahetult kontrollida, et diagonaalmaatriksi $\mathbf{D} : n \times n$ korral on kõik neli Moore-Penrose'i üldistatud pöördmaatriksi tingimust täidetud järgmises lauses esitatud maatriksi korral.

Lause 7.2 *Diagonaalmaatriksi* $\mathbf{D} : n \times n$ *Moore-Penrose'i üldistatud pöördmaatriks* $\mathbf{D}^{+}$ *on diagonaalmaatriks diagonaali elementidega*

$$(\mathbf{D}^+)_{ii} = \begin{cases} \frac{1}{d_{ii}}, & \text{kui } d_{ii} \neq 0, \\ 0 & \text{kui } d_{ii} = 0. \end{cases}$$

Teoreem 7.3 *Olgu* $\mathbf{A}$: $n \times n$ *sümmeetriline maatriks. Siis Moore-Penrose'i üldistatud pöördmaatriks*

$$\mathbf{A}^+ = \mathbf{P}\mathbf{\Lambda}^+\mathbf{P}',$$

kus $\mathbf{P}$ *on* $\mathbf{A}$ *omavektorite ortogonaalmaatriks ja* $\mathbf{\Lambda}^+$ *omaväärtuste diagonaalmaatriksi* $\mathbf{\Lambda}$ *Moore-Penrose'i üldistatud pöördmaatriks..*

Tõestus. Lause 7.2 kohaselt

$$(\mathbf{\Lambda}^+)_{ii} = \begin{cases} \frac{1}{\lambda_i}, & \text{kui } \lambda_i \neq 0, \\ 0 & \text{kui } \lambda_i = 0. \end{cases}$$

Kontrollime, kas kõik definitsiooni 7.3 neli tingimust on täidetud. Kasutame seejuures maatriksi $\mathbf{A}$ esitust $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$.

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{P}\mathbf{\Lambda}^+\mathbf{P}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{A};$$

$$\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{P}\mathbf{\Lambda}^+\mathbf{P}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{P}\mathbf{\Lambda}^+\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}^+\mathbf{P}' = \mathbf{A}^+;$$

$$(\mathbf{A}\mathbf{A}^+)' = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{P}\mathbf{\Lambda}^+\mathbf{P}')' = \mathbf{P}\mathbf{\Lambda}^+\mathbf{P}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{A}\mathbf{A}^+;$$

$$(\mathbf{A}^+\mathbf{A})' = (\mathbf{P}\mathbf{\Lambda}^+\mathbf{P}'\mathbf{P}\mathbf{\Lambda}\mathbf{P}')' = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'\mathbf{P}\mathbf{\Lambda}^+\mathbf{P}' = \mathbf{A}^+\mathbf{A}.$$

Kõik tingimused on rahuldatud. Seega sümmeetrilise maatriksi korral taandub Moore-Penrose'i üldistatud pöördmaatriksi leidmine omaväärtusülesande lahendamisele.
Tuleme tagasi võrrandi

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

juurde. Millal on see võrrand lahenduv ja kuidas avaldub üldlahend?

Teoreem 7.4. *Olgu* $\mathbf{A}$ $m \times n$*-maatriks,* $\mathbf{x}$ n*-vektor ja* $\mathbf{y}$ m*-vektor. Võrrand* $\mathbf{Ax} = \mathbf{y}$ *on lahenduv parajasti siis, kui*

$$\mathbf{AA}^+\mathbf{y} = \mathbf{y}$$

ning selle üldlahend on kujul

$$\mathbf{x} = \mathbf{A}^+\mathbf{y} + (\mathbf{I}_n - \mathbf{A}^+\mathbf{A})\mathbf{q}, \tag{7.3}$$

kus $\mathbf{q}$ *on suvaline* n*-vektor.*

Tõestus. Kontrollime esmalt lahenduvuse tingimuse.
Tarvilikkus. Olgu (7.1) lahenduv. Siis leidub $\mathbf{x}_0$ nii, et $\mathbf{Ax}_0 = \mathbf{y}$. Seega

$$\mathbf{y} = \mathbf{Ax}_0 = \mathbf{AA}^+\mathbf{Ax}_0 = \mathbf{AA}^+\mathbf{y}.$$

Piisavus. Kehtib võrdus

$$\mathbf{AA}^+\mathbf{y} = \mathbf{y}.$$

Võtame $\mathbf{x}* = \mathbf{A}^+\mathbf{y}$, siis

$$\mathbf{Ax}* = \mathbf{AA}^+\mathbf{y} = \mathbf{y}.$$

Seega $\mathbf{x}*$ on lahend ja võrrand $\mathbf{Ax} = \mathbf{y}$ on lahenduv.
Näitame, et üldlahend on kujul (7.3). Eeldame, et võrrand on lahenduv ja seega

$$\mathbf{AA}^+\mathbf{y} = \mathbf{y}.$$

Siis $\mathbf{x}^0 = \mathbf{A}^+\mathbf{y} + (\mathbf{I}_n - \mathbf{A}^+\mathbf{A})\mathbf{q}$ on (7.1) lahend iga vektori $\mathbf{q}$ korral, kuna

$$\mathbf{Ax}^0 = \mathbf{AA}^+\mathbf{y} + (\mathbf{A} - \mathbf{AA}^+\mathbf{A})\mathbf{q} = \mathbf{y}.$$

Peame veel näitama, et suvaline $\mathbf{Ax} = \mathbf{y}$ lahend $\tilde{\mathbf{x}}$ on esitatav kujul (7.3). Võtame

$$\mathbf{q} = \tilde{\mathbf{x}} - \mathbf{A}^+\mathbf{y},$$

siis

$$\mathbf{A}^+\mathbf{y}+(\mathbf{I}_n-\mathbf{A}^+\mathbf{A})(\tilde{\mathbf{x}}-\mathbf{A}^+\mathbf{y}) = \mathbf{A}^+\mathbf{y}+\tilde{\mathbf{x}}-\mathbf{A}^+\mathbf{y}-\mathbf{A}^+\mathbf{A}\tilde{\mathbf{x}}+\mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{y}.$$

Kuna $\mathbf{A}^+\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}^+\mathbf{y}$, siis

$$\mathbf{A}^+\mathbf{y} + (\mathbf{I}_n - \mathbf{A}^+\mathbf{A})(\tilde{\mathbf{x}} - \mathbf{A}^+\mathbf{y}) = \tilde{\mathbf{x}}.$$

Esitame alljärgnevalt loetelu üldistatud pöördmaatriksi omadustest.

(i) $\mathbf{A}^+ = \mathbf{A}^{-1}$, kui $\mathbf{A}$ on pööratav.

(ii) $(\mathbf{A}^+)^+ = \mathbf{A}$.

(iii) $(\mathbf{A}')^+ = (\mathbf{A}^+)'$.

(iv) $\mathbf{A}^+ = \mathbf{A}$, kui $\mathbf{A}$ on sümmeetriline ja idempotentne.

(v) $\mathbf{A}\mathbf{A}^+$ ja $\mathbf{A}^+\mathbf{A}$ on idempotentsed maatriksid.

(vi) $\mathbf{A}$, $\mathbf{A}^+$, $\mathbf{A}\mathbf{A}^+$ ja $\mathbf{A}^+\mathbf{A}$ on võrdse astakuga.

(vii) $\mathbf{A}'\mathbf{A}\mathbf{A}^+ = \mathbf{A}' = \mathbf{A}^+\mathbf{A}\mathbf{A}'$.

(viii) $\mathbf{A}'(\mathbf{A}^+)'\mathbf{A}^+ = \mathbf{A}^+ = \mathbf{A}^+(\mathbf{A}^+)'\mathbf{A}'$.

(ix) $(\mathbf{A}'\mathbf{A})^+ = \mathbf{A}^+(\mathbf{A}^+)'$

(x) $\mathbf{A}^+ = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}$, kui $\mathbf{A}$ on ridade täisastakuga.

(xi) $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$, kui $\mathbf{A}$ on veergude täisastakuga.

(xii) $\mathbf{A} = \mathbf{0} \Longleftrightarrow \mathbf{A}^+ = \mathbf{0}$.

(xiii) $\mathbf{A}\mathbf{B} = \mathbf{0} \Longleftrightarrow \mathbf{B}^+\mathbf{A}^+ = \mathbf{0}$.

(xiv) $r(\mathbf{A}^-) \geq r(\mathbf{A})$.

(xv) $r(\mathbf{A}\mathbf{A}^-) = r(\mathbf{A}^-\mathbf{A}) = r(\mathbf{A})$.

(xvi) $\mathbf{AA}^-$ ja $\mathbf{A}^-\mathbf{A}$ on idempotentsed.

(xvii) Kui $\mathbf{A} : m \times n$, siis $\mathbf{A}^-\mathbf{A} = \mathbf{I}_n$ parajasti siis, kui $r(\mathbf{A}) = n$.

§8. Maatriksite esitused

Vaatame võimalusi maatriksi $\mathbf{A}$ esitamiseks teatud (lihtsamate) maatriksite korrutisena. Selliseid korrutisi nimetataksegi esitusteks või lahutusteks või faktorisatsioonideks. Inglise keeles on kasutusel sõnad *decomposition* ja *factorization*.
Seni oleme loengutes vaadanud juba kaht sellist esitust.
Teoreemi 3.2 kohaselt on $m \times n$ maatriks $\mathbf{A}$ astakuga $r(\mathbf{A}) = r$ esitatav täisastakuga maatriksite $\mathbf{K} : m \times r$ ja $\mathbf{L} : r \times n$ korrutisena

$$\mathbf{A} = \mathbf{KL}.$$

Sümmeetrilise maatriksi $\mathbf{A}$ saime esitada normeeritud omavektorite ortogonaalmaatriksi $\mathbf{P}$ ja omaväärtuste diagonaalmaatriksi $\mathbf{\Lambda}$ kaudu

$$\mathbf{A} = \mathbf{P\Lambda P}'.$$

Mitmed ruutmaatriksi esitused põhinevad järgmisel teoreemil.
Teoreem 8.1 *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on positiivselt määratud parajasti siis kui* $\mathbf{A}$ *on esitatav korrutisena* $\mathbf{A} = \mathbf{BB}'$ *mingi* $\mathbf{B} : n \times n$ *korral,* $|\mathbf{B}| \neq 0$. *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on mittenegatiivselt määratud kui* $\mathbf{A}$ *on esitatav korrutisena* $\mathbf{A} = \mathbf{BB}'$ *mingi* $\mathbf{B} : n \times m$ *korral.*

Tõestus. Tõestame esmalt teoreemi esimese väite.
Tarvilikkus. Olgu $\mathbf{A} > 0$. Kuna $\mathbf{A} = \mathbf{A}'$, saame maatriksi $\mathbf{A}$ esitada omaväärtuste ja omavektorite kaudu

$$\mathbf{A} = \mathbf{P\Lambda P}'.$$

Võtame

$$\mathbf{B} = \mathbf{P\Lambda}^{\frac{1}{2}},$$

siis

$$\mathbf{A} = \mathbf{BB}',$$

kusjuures

$$|\mathbf{B}| = |\mathbf{P}| \prod_{i=1}^{n} \sqrt{\lambda_i}.$$

Kuna ortogonaalmaatriksi determinant $|\mathbf{P}| \neq 0$ ja $\mathbf{A}$ positiivse määratuse tõttu $\lambda_i > 0,\ i = 1, \ldots, n$, siis $|\mathbf{B}| \neq 0$. Esituse nõuded on täidetud.
Piisavus. Olgu

$$\mathbf{A} = \mathbf{B}\mathbf{B}',\ |\mathbf{B}| \neq 0$$

ja $\mathbf{y} \in \mathbb{R}^n,\ \mathbf{y} \neq \mathbf{0}$. Siis

$$\mathbf{y}'\mathbf{A}\mathbf{y} = \mathbf{y}'\mathbf{B}\mathbf{B}'\mathbf{y} = \mathbf{x}'\mathbf{x} > 0,$$

kus $\mathbf{x} = \mathbf{B}'\mathbf{y} \neq \mathbf{0}$, sest $\mathbf{B}$ on pööratav.
Teoreemi esimene väide on tõestatud.
Teise väite kohaselt $\mathbf{A} > 0$ või $\mathbf{A} \geq 0$. Lähtume jälle esitusest

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Kui $\lambda_i > 0,\ i = 1, \ldots, n$, siis ülalesitatud tõestuse kohaselt $\mathbf{A} > 0$ ja $\mathbf{B} = \mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}$.
Vaatame juhtu, kus $|\mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}| = 0$. Siis vähemalt üks omaväärtus $\lambda_i = 0$. Olgu mittenulliliste omaväärtuste arv $k < n$. Võtame

$$\mathbf{B} = \mathbf{P}^*\mathbf{\Lambda}^{*\frac{1}{2}},$$

kus $\mathbf{\Lambda}^*: k \times k$ on mittenullilistest omaväärtustest moodustatud diagonaalmaatriks ja $\mathbf{B}^*: n \times k$ koosneb neile vastavatest normeeritud omavektoritest. Siis

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{P}^*\mathbf{\Lambda}^*\mathbf{P}^{*'} = \mathbf{B}\mathbf{B}'.$$

Positiivselt poolmääratud maatriksi $\mathbf{A}$ korral peab leiduma $\mathbf{x}_0 \neq \mathbf{0}$ nii, et

$$\mathbf{x}_0'\mathbf{A}\mathbf{x}_0 = \mathbf{x}_0'\mathbf{B}\mathbf{B}'\mathbf{x}_0 = 0.$$

Võtame vektoriks $\mathbf{x}_0$ nullilisele omaväärtusele vastava omavektori. Siis erinevatele omaväärtustele vastavate omavektorite ortogonaalsuse tõttu $\mathbf{B}'\mathbf{x}_0 = \mathbf{0}$ ja $\mathbf{x}_0'\mathbf{B}\mathbf{B}'\mathbf{x}_0 = 0$.
Teoreem on tõestatud.
Järgmine esitus on tuntud Cholesky lahutusena (*decomposition*).
Teoreem 8.2 *Kui* $\mathbf{A} > 0$, *siis leidub ühene positiivsete diagonaalielementidega alumine kolmnurkmaatriks* $\mathbf{T}_\triangle$ *nii, et*

$$\mathbf{A} = \mathbf{T}_\triangle(\mathbf{T}_\triangle)'.$$

Tõestus. Esitame järgnevalt range tõestuse asemel idee kuidas maatriksi $\mathbf{T}_\triangle$ elemente määrata. Osutub, et elemendid $t_{ij},\ i \geq j$ saame leida järjestikku ridade kaupa. Esimese diagonaalielemendi saame lihtsalt

$$t_{11}^2 = a_{11}; \quad t_{11} = \sqrt{a_{11}}.$$

Teise rea korral lähtume võrdustest

$$t_{21}t_{11} = a_{21} \quad t_{21}^2 + t_{22}^2 = a_{22}.$$

Saame

$$t_{21} = \frac{a_{21}}{\sqrt{a_{11}}}; \quad t_{22} = \sqrt{a_{22} - \frac{a_{21}^2}{a_{11}}} = \sqrt{\frac{a_{22}a_{11} - a_{21}^2}{a_{11}}} > 0,$$

kuna teist järku nurgamiinor on positiivne.
Nii mööda ridu edasi liikudes saame järjest määrata kõik maatriksi $\mathbf{T}_\triangle$ elemendid.

Toome järgnevalt tõestuseta ära ühe võimaliku esituse ristkülikmaatriksi jaoks.
Teoreem 8.3 *Olgu* $\mathbf{A} : m \times n$ *maatriks astakuga* $r(\mathbf{A}) = r$. *Siis leiduvad alumine kolmnurkmaatriks* $\mathbf{T}_\triangle : m \times m$ *ja ortogonaalmaatriks* $\mathbf{P} : n \times n$ *nii, et*

$$\mathbf{A} = \mathbf{T}_\triangle \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P}.$$

Järeldus 8.3.1 *Maatriks* $\mathbf{A} : m \times n$ *on esitatav korrutisena*

$$\mathbf{A} = \mathbf{W}\mathbf{V},$$

kus $\mathbf{W}$ *koosneb alumise kolmnurkmaatriksi* $\mathbf{T}_\triangle$ *esimesest* r *veerust ja* $\mathbf{V}$ *ortogonaalmaatriksi* $\mathbf{P}$ *esimesest* r *reast.*
Üks enimkasutatavaid ristkülikmaatriksi esitusi on *singulaarväärtuslahutus (Singular Value Decomposition).*
(Michael Greenacre, statistikaprofessor ja muusikaharrastaja, on kirjutanud laulu - *It Had to Be U - the SVD Song*, mis on leitav Youtube's *Statistical Songs* alt.)
Teoreem 8.4 *Olgu* $\mathbf{A} : m \times n$ *astakuga* $r(\mathbf{A}) = r$. *Siis leiduvad ortogonaalmaatriksid* $\mathbf{P} : m \times m$ *ja* $\mathbf{Q} : n \times n$ *nii, et*

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{Q}' \quad ja \quad \mathbf{D} = \mathbf{P}'\mathbf{A}\mathbf{Q},$$

kus

(i) $\mathbf{D} = \boldsymbol{\Delta}$, *kui* $m = n = r$;

(ii) $\mathbf{D} = (\boldsymbol{\Delta} \vdots \mathbf{0})$, *kui* $m = r < n$;

(iii)
$$\mathbf{D} = \begin{pmatrix} \boldsymbol{\Delta} \\ \mathbf{0} \end{pmatrix}, \ kui\ r = n < m;$$

(iv)
$$\mathbf{D} = \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \ kui\ r < m,\ r < n;$$

ja $\boldsymbol{\Delta} : r \times r$ *on positiivsete elementidega diagonaalmaatriks ning maatriksi* $\boldsymbol{\Delta}^2$ *diagonaali elemendid on maatriksite* $\mathbf{A}\mathbf{A}'$ *ja* $\mathbf{A}'\mathbf{A}$ *ühised mittenullilised omaväärtused*
Maatriksi $\boldsymbol{\Delta}$ *diagonaali elemente nimetatakse maatriksi* $\mathbf{A}$ *singulaarväärtusteks.*

Tõestus. Tõestame teoreemi juhul (iv), teised on lihtsamad erijuhud. Olgu $\boldsymbol{\Delta}^2 : r \times r$ maatriksi $\mathbf{A}'\mathbf{A}$ mittenulliliste omaväärtuste diagonaalmaatriks ja $\boldsymbol{\Delta} : r \times r$ diagonaalmaatriks, mille diagonaali elemendid on positiivsed ruutjuured $\mathbf{\Delta}^2$ diagonaali elementidest. Kuna $\mathbf{A}'\mathbf{A}$ on sümmeetriline maatriks, siis leidub ortogonaalmaatriks $\mathbf{Q} : n \times n$ nii, et

$$\mathbf{Q}'\mathbf{A}'\mathbf{A}\mathbf{Q} = \begin{pmatrix} \mathbf{\Delta}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \; kui \; r < m, \; r < n;$$

Jagame $\mathbf{Q}$ kaheks osaks $\mathbf{Q} = (\mathbf{Q}_1 \vdots \mathbf{Q}_2)$, kus $\mathbf{Q}_1$ on $n \times r$-maatriks ja $\mathbf{Q}_2$ – $n \times (n-r)$-maatriks. Plokke korrutades saame kaks võrdust

$$\mathbf{Q}_1'\mathbf{A}'\mathbf{A}\mathbf{Q}_1 = \mathbf{\Delta}^2, \tag{8.1}$$

$$\mathbf{Q}_2'\mathbf{A}'\mathbf{A}\mathbf{Q}_2 = \mathbf{0}.$$

Siit järeldub võrdus

$$\mathbf{A}\mathbf{Q}_2 = \mathbf{0}. \tag{8.2}$$

Olgu $\mathbf{P} = (\mathbf{P}_1 \vdots \mathbf{P}_2) : m \times m$ maatriks, kus

$$\mathbf{P}_1 = \mathbf{A}\mathbf{Q}_1\mathbf{\Delta}^{-1} : m \times r$$

ja $\mathbf{P}_2 : m \times (m-r)$ selline, et tema veerud on ortogonaalsed ja ühiku pikkusega ning ortogonaalsed $\mathbf{P}_1$ veergudega. On kerge veenduda, et nii konstrueeritud $\mathbf{P}$ on ortogonaalmaatriks

$$\begin{pmatrix} \mathbf{P}_1' \\ \mathbf{P}_2' \end{pmatrix} (\mathbf{P}_1 \vdots \mathbf{P}_2) = \begin{pmatrix} \mathbf{\Delta}^{-1}\mathbf{Q}_1'\mathbf{A}' \\ \mathbf{P}_2' \end{pmatrix} (\mathbf{A}\mathbf{Q}_1\mathbf{\Delta}^{-1} \vdots \mathbf{P}_2)$$

$$= \begin{pmatrix} \mathbf{\Delta}^{-1}\mathbf{Q}_1'\mathbf{A}'\mathbf{A}\mathbf{Q}_1\mathbf{\Delta}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-r} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-r} \end{pmatrix}.$$

Viimase võrduse saamisel kasutasime seost (8.1). Ortogonaalsuse tõttu

$$\mathbf{P}_2'\mathbf{P}_1 = \mathbf{0} = \mathbf{P}_2'\mathbf{A}\mathbf{Q}_1\mathbf{\Delta}^{-1}.$$

Siit

$$\mathbf{P}_2'\mathbf{A}\mathbf{Q}_1 = \mathbf{0}_{(m-r)\times r} \tag{8.3}$$

ja korrutis

$$\mathbf{P}'\mathbf{A}\mathbf{Q} = \begin{pmatrix} \mathbf{P}_1'\mathbf{A}\mathbf{Q}_1 & \mathbf{P}_1'\mathbf{A}\mathbf{Q}_2 \\ \mathbf{P}_2'\mathbf{A}\mathbf{Q}_1 & \mathbf{P}_2'\mathbf{A}\mathbf{Q}_2 \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{\Delta}^{-1}\mathbf{Q}_1'\mathbf{A}'\mathbf{A}\mathbf{Q}_1 & \mathbf{P}_1'\mathbf{A}\mathbf{Q}_2 \\ \mathbf{P}_2'\mathbf{A}\mathbf{Q}_1 & \mathbf{P}_2'\mathbf{A}\mathbf{Q}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Võrduste (8.2) ja (8.3) tõttu on viimases maatriksis kolm plokki võrdsed nullmaatriksitega, $\mathbf{\Delta}$ saame tänu võrdusele (8.1).
Tõestuses on maatriks $\mathbf{Q}$ moodustatud maatriksi $\mathbf{A}'\mathbf{A}$ ortonormeeritud omavektoritest

$$\mathbf{A}'\mathbf{A} = \mathbf{Q}\mathbf{D}'\mathbf{P}'\mathbf{P}\mathbf{D}\mathbf{Q}' = \mathbf{Q}\mathbf{D}'\mathbf{D}\mathbf{Q}'.$$

Samas maatriksi $\mathbf{P}$ moodustavad maatriksi $\mathbf{A}\mathbf{A}'$ ortonormeeritud omavektorid, kuna

$$\mathbf{A}\mathbf{A}' = \mathbf{P}\mathbf{D}\mathbf{Q}'\mathbf{Q}\mathbf{D}'\mathbf{P}' = \mathbf{P}\mathbf{D}\mathbf{D}'\mathbf{P}'.$$

Siit tekib võimalus sõnastada singulaarväärtuslahutus ka järgmiselt.
Järeldus 8.4.1 *Olgu* $\mathbf{A} : m \times n$ *astakuga* $r(\mathbf{A}) = r$. *Siis leiduvad maatriksid* $\mathbf{P}_1 : m \times r$ *ja* $\mathbf{Q}_1 : n \times r$ *nii, et* $\mathbf{P}_1'\mathbf{P}_1 = \mathbf{Q}_1'\mathbf{Q}_1 = \mathbf{I}_r$ *ja*

$$\mathbf{A} = \mathbf{P}_1\mathbf{\Delta}\mathbf{Q}_1',$$

kus $\mathbf{\Delta}$ *on maatriksi* $\mathbf{A}$ *singulaarväärtustest moodustatud positiivsete diagonaalielementidega diagonaalmaatriks.*
Järelduse tõestuseks märgime vaid, et viimastest võrdustest saame seosed

$$\mathbf{P}_1'\mathbf{A}\mathbf{A}'\mathbf{P}_1 = \mathbf{\Delta}^2,$$

$$\mathbf{Q}_1'\mathbf{A}'\mathbf{A}\mathbf{Q}_1 = \mathbf{\Delta}^2.$$

Varasema konstruktsiooni tõttu $\mathbf{P}_1$ sõltus $\mathbf{Q}_1$ valikust

$$\mathbf{P}_1 = \mathbf{A}\mathbf{Q}_1\mathbf{\Delta}^{-1}.$$

Oleksime võinud alustada ka maatriksist $\mathbf{P}_1$ ning siis tulnuks valida

$$\mathbf{Q}_1 = \mathbf{A}'\mathbf{P}_1\mathbf{\Delta}^{-1}.$$

Kuidas on see lahutus kooskõlas sümmeetrilise maatriksi esitusega omaväärtuste ja omavektorite kaudu, kus diagonaalil võivad esineda ka negatiivsed arvud?
Sel juhul on singulaarväärtused maatriksi $\mathbf{A}$ omaväärtuste absoluutväärtused. Kui $\mathbf{P}$ on $\mathbf{A}$ omavektorite ortogonaalmaatriks, siis maatriksis $\mathbf{Q}$ on positiivsetele omaväärtustele λ_i vastavad $\mathbf{Q}$ veerud $\mathbf{q}_i = \mathbf{p}_i$. Negatiivsete omaväärtuste korral $\mathbf{q}_i = -\mathbf{p}_i$.

§9. Projektorid

Vaatame eukleidilist ruumi $\mathbb{R}^n$ ja selles kaht alamruumi $\mathbb{M}$ ja $\mathbb{N}$, mis on teineteise täiendruumid, s.t.

$$\mathbb{R}^n = \mathbb{M} \oplus \mathbb{N}.$$

Siis suvaline vektor $\mathbf{x} \in \mathbb{R}^n$ avaldub üheselt summana

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \tag{9.1}$$

kus $\mathbf{x}_1 \in \mathbb{M}$ ja $\mathbf{x}_2 \in \mathbb{N}$. Vektorit $\mathbf{x}_1$ nimetatakse vektori $\mathbf{x}$ *projektsiooniks ruumile* $\mathbb{M}$ *paralleelselt ruumiga* $\mathbb{N}$ ja vektorit $\mathbf{x}_2$ vektori $\mathbf{x}$ *projektsiooniks ruumile* $\mathbb{N}$ *paralleelselt ruumiga* $\mathbb{M}$.
Kui iga $\mathbf{x}_1 \in \mathbb{M}$ ja iga $\mathbf{x}_2 \in \mathbb{N}$ korral $\mathbf{x}_1'\mathbf{x}_2 = 0$, siis alamruumid $\mathbb{M}$ ja $\mathbb{N}$ on teineteise ortogonaalsed täiendid. Sel juhul on tegemist $\mathbf{x}$ ortogonaalsete projektsioonidega ruumidele $\mathbb{M}$ ja $\mathbb{N}$.

Näide. Vaatame kolmemõõtmelises ruumis $\mathbb{R}^3$, kus alamruumi $\mathbb{M}$ moodustavad x-teljel asuvad punktid $\mathbf{x}_1 = (x, 0, 0)'$ ja alamruumi $\mathbb{N}$ moodustavad yz-tasandi punktid $\mathbf{x}_2 = (0, y, z)'$. Siis suvaline punkt $\mathbf{x}$ ruumis $\mathbb{R}^3$ avaldub summana $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, kusjuures $\mathbf{x}_1'\mathbf{x}_2 = 0$, s.t. tegemist on ortogonaalsete projektsioonidega.

Lihtne on kontrollida, et alamruumidele projekteerimine on lineaarne operatsioon. Olgu meil lisaks vektorile $\mathbf{x}$ teine vektor

$$\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2,$$

kus $\mathbf{y}_1 \in \mathbb{M}$ ja $\mathbf{y}_2 \in \mathbb{N}$.
Siis $\alpha\mathbf{x}_1 + \beta\mathbf{y}_1$ on $\alpha\mathbf{x} + \beta\mathbf{y}$ projektsioon ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$ ja $\alpha\mathbf{x}_2 + \beta\mathbf{y}_2$ on $\alpha\mathbf{x} + \beta\mathbf{y}$ projektsioon ruumile $\mathbb{N}$ paralleelselt ruumiga $\mathbb{M}$.
Tähistame P lineaarse operaatori, mis projekteerib ruumi $\mathbb{R}^n$ vektorid alamruumi $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$,

$$P\mathbf{x} = \mathbf{x}_1.$$

Nimetame seda lineaarset operaatorit *projektoriks* ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$. Kui ruumid $\mathbb{M}$ ja $\mathbb{N}$ on teiseteise ortogonaalsed täiendid, siis on tegemist *ortogonaalse projektoriga.*
Kuna iga lineaarne operaator ruumis $\mathbb{R}^n$ on esitatav $n \times n$-maatriksina, siis kasutamegi edaspidi maatriksesitust ja vaatame projektorina maatriksit $\mathbf{P}$. Lihtne on kontrollida, et kui $\mathbf{P}$ on projektor ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$, siis $\mathbf{I}_n - \mathbf{P}$ on projektor ruumile $\mathbb{N}$ paralleelselt ruumiga $\mathbb{M}$.
Tõepoolest, rakendame maatriksit $\mathbf{I}_n - \mathbf{P}$ vektorile $\mathbf{x}$

$$(\mathbf{I}_n - \mathbf{P})\mathbf{x} = \mathbf{x} - \mathbf{x}_1.$$

Siis aga

$$\mathbf{x}_1 + (\mathbf{x} - \mathbf{x}_1) = \mathbf{x},$$

kust esituse (9.1) ühesuse tõttu $\mathbf{x} - \mathbf{x}_1 = \mathbf{x}_2$.
Olgu $\mathbf{P}$ projektor ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$. Võrdus

$$\mathbf{P}\mathbf{x} = \mathbf{x}$$

kehtib parajasti siis, kui $\mathbf{x} \in \mathbb{M}$ ja

$$\mathbf{P}\mathbf{x} = \mathbf{0}$$

parajasti siis, kui $\mathbf{x} \in \mathbb{N}$.
Projektorit iseloomustav omadus on idempotentsus.
Teoreem 9.1 *Lineaarne teisendus* $\mathbf{P}$ *ruumil* $\mathbb{R}^n$ *on projektor mingile tema alamruumile* $\mathbb{M}$ *paralleelselt täiendalamruumiga* $\mathbb{N}$ *parajasti siis, kui ta on idempotentne.*

Tõestus. Tarvilikkus. Olgu $\mathbf{P}$ projektor alamruumile $\mathbb{M}$ paralleelselt täiendalamruumiga $\mathbb{N}$. Näitame, et siis $\mathbf{P}^2 = \mathbf{P}$. Vaatame vektorit

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2.$$

Siis

$$\mathbf{P}^2\mathbf{x} = \mathbf{P}(\mathbf{P}\mathbf{x}) = \mathbf{P}\mathbf{x}_1 = \mathbf{x}_1.$$

Aga ka

$$\mathbf{P}\mathbf{x} = \mathbf{x}_1, \qquad \Longrightarrow \qquad \mathbf{P}^2 = \mathbf{P}.$$

Piisavus. Olgu maatriks $\mathbf{P}$ idempotentne, $\mathbf{P}^2 = \mathbf{P}$. Defineerime alamruumid

$$\mathbb{V}_1 = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{P}\mathbf{x} = \mathbf{x}\};$$

$$\mathbb{V}_2 = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{P}\mathbf{x} = \mathbf{0}\}.$$

Konstruktsiooni tõttu on $\mathbb{V}_1$ ja $\mathbb{V}_2$ teineteise täiendruumid. Esitame suvalise $\mathbf{x} \in \mathbb{R}^n$ summana

$$\mathbf{x} = \mathbf{P}\mathbf{x} + (\mathbf{I}_n - \mathbf{P})\mathbf{x}.$$

Kuna $\mathbf{P}$ on idempotentne, siis võrdusest

$$\mathbf{P}\mathbf{x} - \mathbf{P}^2\mathbf{x} = \mathbf{0}$$

järeldub, et $(\mathbf{I}_n - \mathbf{P})\mathbf{x} \in \mathbb{V}_2$, sest

$$\mathbf{P}((\mathbf{I}_n - \mathbf{P})\mathbf{x}) = \mathbf{0}.$$

Samas

$$\mathbf{P}\mathbf{x} = \mathbf{P}^2\mathbf{x} = \mathbf{P}(\mathbf{P}\mathbf{x}).$$

Seega $\mathbf{Px} \in \mathbb{V}_1$. Oleme saanud lahutuse ruumide $\mathbb{V}_1$ ja $\mathbb{V}_2$ suhtes, kus $\mathbf{P}$ on projektor alamruumile $\mathbb{V}_1$ paralleelselt alamruumiga $\mathbb{V}_2$.
Ühtlasi oleme näidanud, et iga idempotentne operaator ruumil $\mathbb{R}^n$ on teatav projektor.

Projektori mõiste saame üle kanda juhule, kus kahe alamruumi asemel on k alamruumi. Siis

$$\mathbb{R}^n = \oplus_{i=1}^k \mathbb{M}_i = \mathbb{M}_1 \oplus \ldots \oplus \mathbb{M}_k,$$

sealjuures iga vektor $\mathbf{x}$ esitub üheselt summana

$$\mathbf{x} = \sum_{i=1}^{k} \mathbf{x}_i, \quad \mathbf{x}_i \in \mathbb{M}_i. \tag{9.2}$$

Kujutus $\mathbf{P}_j$ ruumil $\mathbb{R}^n$, mis on määratud võrdusega

$$\mathbf{P}_j \mathbf{x} = \mathbf{x}_j$$

on projektor ruumile $\mathbb{M}_j$ paralleelselt ülejäänud alamruumide summaga $\oplus_{i=1}^k \mathbb{M}_i,\ i \neq j$. Seejuures

$$\sum_{i=1}^{k} \mathbf{P}_j = \mathbf{I}_n, \tag{9.3}$$

$$\mathbf{P}_j \mathbf{P}_l = \delta_{jl} \mathbf{P}_j,$$

kus Kroneckeri delta

$$\delta_{jl} = \begin{cases} 1, & kui\ j = l; \\ 0, & kui\ j \neq l. \end{cases}$$

Esimese neist võrdustest saame, kui rakendame võrduse (9.3) vasakut ja paremat poolt vektorile $\mathbf{x}$, kust saame võrduse (9.2). Teise võrduse korral, kui $j \neq l$

$$\mathbf{P}_j(\mathbf{P}_l \mathbf{x}) = \mathbf{P}_j \mathbf{x}_l = \mathbf{0} \Longrightarrow \mathbf{P}_j \mathbf{P}_l = \mathbf{0}.$$

Vaatame maatriksit $\mathbf{A} : m \times n$ kui teisendust ruumist $\mathbb{R}^n$ ruumi $\mathbb{R}^m$. Kujutised $\mathbf{Ax}$ moodustavad ruumis $\mathbb{R}^m$ alamruumi, mis erijuhul võib kokku langeda ka kogu ruumiga.
Kuidas leida projektor, mis projekteerib ruumi $\mathbb{R}^n$ vektorid maatriksi $\mathbf{A} : m \times n$ veeruvektorite poolt määratud alamruumi? Vaata-me esmalt, kuidas tekivad ortogonaalsed projektorid.
Olgu $\mathbf{z}_1, \ldots, \mathbf{z}_n$ ortonormeeritud baas ruumis $\mathbb{R}^n$ ja määraku vektorid $\mathbf{z}_1, \ldots, \mathbf{z}_r$; $r < n$ alamruumi $\mathbb{M}$ ruumis $\mathbb{R}^n$. Võtame suvalise vektori $\mathbf{x} \in \mathbb{R}^n$ ja olgu $\alpha_1, \ldots, \alpha_n$ sellised konstandid, et

$$\mathbf{x} = (\alpha_1 \mathbf{z}_1 + \cdots + \alpha_r \mathbf{z}_r) + (\alpha_{r+1} \mathbf{z}_{r+1} + \cdots + \alpha_n \mathbf{z}_n).$$

Tähistame $\alpha_1 \mathbf{z}_1 + \cdots + \alpha_r \mathbf{z}_r = \mathbf{x}_1$ ja $\alpha_{r+1} \mathbf{z}_{r+1} + \cdots + \alpha_n \mathbf{z}_n = \mathbf{x}_2$.
Olgu

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \\ \cdots\cdots \\ \alpha_{r+1} \\ \vdots \\ \alpha_r \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \cdots\cdots \\ \boldsymbol{\alpha}_2 \end{pmatrix}; \qquad \mathbf{Z} = (\mathbf{Z}_1 \vdots \mathbf{Z}_2),$$

kus

$$\mathbf{Z}_1 = (\mathbf{z}_1, \ldots, \mathbf{z}_r); \qquad \mathbf{Z}_2 = (\mathbf{z}_{r+1}, \ldots, \mathbf{z}_n).$$

Siis

$$\mathbf{x} = \mathbf{Z}\boldsymbol{\alpha} = \mathbf{Z}_1 \boldsymbol{\alpha}_1 + \mathbf{Z}_2 \boldsymbol{\alpha}_2.$$

Tänu ortogonaalsusele

$$\mathbf{Z}_1' \mathbf{Z}_1 = \mathbf{I}_r, \qquad \mathbf{Z}_1' \mathbf{Z}_2 = \mathbf{0}.$$

Siis

$$(\mathbf{Z}_1 \mathbf{Z}_1')\mathbf{x} = \mathbf{Z}_1 \mathbf{Z}_1' \mathbf{Z}\boldsymbol{\alpha} = \mathbf{Z}_1 \mathbf{Z}_1' (\mathbf{Z}_1 \vdots \mathbf{Z}_2)\boldsymbol{\alpha} = \mathbf{Z}_1 \mathbf{Z}_1' (\mathbf{Z}_1 \vdots \mathbf{Z}_2) \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \cdots\cdots \\ \boldsymbol{\alpha}_2 \end{pmatrix}$$

$$= \mathbf{Z}_1(\mathbf{I}_r \vdots \mathbf{0}) \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \dots\dots \\ \boldsymbol{\alpha}_2 \end{pmatrix} = (\mathbf{Z}_1 \vdots \mathbf{0}) \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \dots\dots \\ \boldsymbol{\alpha}_2 \end{pmatrix} = \mathbf{Z}_1 \boldsymbol{\alpha}_1 = \mathbf{x}_1.$$

Seega, kui $\mathbf{Z}_1 : n \times r$ ortonormeeritud veerud määravad alamruumi $\mathbb{M}$, siis ortogonaalne projektor sinna ruumi on maatriks $\mathbf{Z}_1\mathbf{Z}_1'$. Näeme, et tegemist on idempotentse maatriksiga. Projektor täiendruumi ($\mathbb{M}$ ortogonaalsesse täiendisse) on kujul $(\mathbf{I}_n - \mathbf{Z}_1\mathbf{Z}_1')$.
Kui me võtame mingi r lineaarselt sõltumatust veeruvektorist koosneva täisastakuga maatriksi $\mathbf{A} : n \times r$, siis kõikvõimalikud nende vektorite linaarkombinatsioonid määravad r-mõõtmelise alamruumi ruumis $\mathbb{R}^n$. Kuidas leida projektor, mis projekteerib sellesse alamruumi?
Korrutades maatriksit $\mathbf{A}$ paremalt pööratava maatriksiga $\mathbf{T}$: $r \times r$ saame need veeruvektorid ortonormeerida. Oleme saanud tulemuseks r ortonormeeritud baasivektorit. Tähistame neist moodustatud maatriksi taas $\mathbf{Z}_1$

$$\mathbf{Z}_1 = \mathbf{AT}.$$

Projektor vastavasse alamruumi

$$\mathbf{Z}_1\mathbf{Z}_1' = \mathbf{ATT}'\mathbf{A}'.$$

Samas

$$\mathbf{Z}_1'\mathbf{Z}_1 = \mathbf{T}'\mathbf{A}'\mathbf{AT} = \mathbf{I}_r.$$

Viimane võrdus on võimalik vaid siis, kui

$$\mathbf{A}'\mathbf{A} = (\mathbf{T}')^{-1}\mathbf{T}^{-1}$$

ehk

$$\mathbf{TT}' = (\mathbf{A}'\mathbf{A})^{-1}.$$

Seega projektor, mis projekteerib suvalise vektori ruumis $\mathbb{R}^n$ maatriksi $\mathbf{A}$ veergude poolt määratud alamruumi, on kujul

$$\mathbf{P_A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'. \tag{9.4}$$

Projektor maatriksi $\mathbf{A} : m \times r$ veergude poolt määratud alamruumi täiendruumile on kujul $\mathbf{I}_n - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$.
Teoreem 9.2. *Olgu* $\mathbf{A} : n \times r,\ r \leq n$ *täisastakuga* $r(\mathbf{A}) = r$ *maatriks veeruvektoritega* $\mathbf{a}_i,\ i = 1, \ldots, r$. *Siis projektor* $\mathbf{P_A}$ *maatriksi* $\mathbf{A}$ *veeruvektorite poolt määratud alamruumile on kujul*

$$\mathbf{P_A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'.$$

Projektor võrduses (9.4) on statistikule tuttav konstruktsioon. Kui meil on tegemist lineaarse mudeliga

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

kus vaatluste vektor $\mathbf{y}$ ja vigade vektor $\boldsymbol{\epsilon}$ on n-vektorid, $\mathbf{X}$ on $n \times k$-plaanimaatriks ja $\boldsymbol{\beta}$ – tundmatute parameetrite k-vektor, siis vähimruutude hinnang parameetervektorile $\boldsymbol{\beta}$ on kujul

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{9.5}$$

Korrutame selle võrduse mõlemaid pooli vasakult maatriksiga $\mathbf{X}$. Siis

$$\mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

Teisisõnu, parameetrite vähimruutude hinnangu kujutis plaanimaatriksiga on vaatlustulemuste $\mathbf{y}$ projektsioon plaanimaatriksi veeruvektorite poolt määratud alamruumile.
Vähimruutude hinnangu (9.5) tuletame kursuses edaspidi maatrikstuletist kasutades.

PLOKKMAATRIKSID

§10. Tehted ja tähistused, determinant ja pöördmaatriks.

Matemaatilises statistikas on tihti olukord, kus andmetes või neid kirjeldavates seosemaatriksites on sisemine struktuur, mida oleks vaja arvesse võtta analüüsi käigus. Tihti saame sisemist struktuuri arvesse võttes tunduvalt lihtsustada ka vajalikke arvutusi. Selles osas vaadeldavad mõisted kannavad statistikale orienteeritud maatriksalgebra esitustes tihti ka "uuema maatriksalgebra" (*newer matrix algebra*) nimetust (vt. näiteks Magnus & Neudecker (1999)). Selle all peetakse silmas eelkõige mõisteid otsekorrutis, kommuteerimismaatriks, vec-operaator ja maatrikstuletis ning nende omadusi ja vastastikuseid seoseid.

Definitsioon 10.1. *Maatriksit* $\mathbf{A} : m \times n$ *nimetatakse plokkmaatriksiks, kui ta koosneb plokkidest* $\mathbf{A}_{ij} : m_i \times n_j,\ i = 1, \ldots, u;\ j = 1, \ldots, v$ *nii et*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1v} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{u1} & \mathbf{A}_{u2} & \cdots & \mathbf{A}_{uv} \end{pmatrix}, \quad \sum_{i=1}^{u} m_i = m, \quad \sum_{j=1}^{v} n_j = n.$$

Plokkmaatriksi jaoks kasutame tähistusi

$$\mathbf{A} = [\mathbf{A}_{ij}], \quad i = 1, \ldots, u;\ j = 1, \ldots, v$$

ja

$$\mathbf{A} = [\mathbf{A}]_{ij}, \quad i = 1, \ldots, u;\ j = 1, \ldots, v.$$

Plokkmaatriksi elementide määramisel kasutatakse kahekordseid indekseid, mis võimaldavad arvesse võtta plokkstruktuuri.
Öeldakse, et plokkmaatriksi element asub (k, l)*-ndas reas ja* (g, h)*-ndas veerus, kui ta on* k*-nda plokkide rea* l*-ndas reas ja* g*-nda plokkide veeru* h*-ndas veerus. Sel juhul kasutatakse tähistust* $a_{(k,l)(g,h)}$ *või* $(\mathbf{A})_{(k,l)(g,h)}$.

Tavalist tähistust kasutades saame võrduse

$$a_{(k,l)(g,h)} = a_{\sum_{i=1}^{k-1} m_i + l, \sum_{j=1}^{g-1} n_j + h}.$$

Plokkmaatriks $\mathbf{A} = [\mathbf{A}_{ij}], \quad i = 1, \ldots, u; \; j = 1, \ldots, v$ on *plokk-diagonaalne*, kui $u = v$ ja $\mathbf{A}_{ij} = 0$, kui $i \neq j$.
Esitame järgnevalt tehted maatriksitega arvestades plokkstruktuuri. Plokkide dimensioonid jätame välja kirjutamata, kui need järgivad definitsiooni 10.1.

(i) $c\mathbf{A} = [c\mathbf{A}_{ij}], \quad i = 1, \ldots, u; \; j = 1, \ldots, v;$

(ii) $\mathbf{A} + \mathbf{B} = [\mathbf{A} + \mathbf{B}]_{ij} = [\mathbf{A}_{ij}] + [\mathbf{B}_{ij}], \quad i = 1, \ldots, u; \; j = 1, \ldots, v;$

(iii) Kui $\mathbf{A} = [\mathbf{A}_{ij}] : m \times n, \quad i = 1, \ldots, u; \; j = 1, \ldots, v$ ja $\mathbf{B} = [\mathbf{B}_{jl}] : n \times r, \quad j = 1, \ldots, v; \; l = 1, \ldots, w$, siis korrutis $\mathbf{AB}$ on plokkmaatriks, mis koosneb $m_i \times r_l$-plokkidest

$$[\mathbf{AB}]_{il} = \sum_{j=1}^{v} \mathbf{A}_{ij} \mathbf{B}_{jl},$$

kui $\mathbf{A}_{ij}$ on $m_i \times n_j$-plokid ja $\mathbf{B}_{jl}$ on $n_j \times r_l$-plokid.

Paljudes rakendustes on oluline nn. 2×2-plokkstruktuur. Olgu maatriksi $\mathbf{A}$ plokid $\mathbf{A}_{ij} : m_i \times n_j, \; i, j = 1, 2$ ja maatriksi $\mathbf{B}$ plokid $\mathbf{B}_{jl} : n_j \times r_l, \; j, l = 1, 2$. Siis

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix}.$$

Tegelikult oleme seni 2×2 plokkmaatriksite plokiviisilist korrutamist juba korduvalt kasutanud.

(iv) Transponeerimisel vahetavad rea- ja veeruindeksid koha

$$(\mathbf{A})_{(k,l)(g,h)} = (\mathbf{A}')_{(g,h)(k,l)}.$$

Ühtlasi vahetavad diagonaalivälised plokid transponeerituna kohad esialgse maatriksiga võrreldes.

2×2-plokkmaatriksite korral

$$\mathbf{A}' = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}' = \begin{pmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{21} \\ \mathbf{A}'_{12} & \mathbf{A}'_{22} \end{pmatrix}$$

(v) Plokkdiagonaalse ruutmaatriksi $\mathbf{A} = [\mathbf{A}_{ii}],\ i = 1, \ldots, u$ pöördmaatriks on plokkdiagonaalne, kui diagonaalplokid $[\mathbf{A}_{ii}] : n_i \times n_i$ on pööratavad

$$\mathbf{A}^{-1} = [\mathbf{A}_{ii}^{-1}],\ i = 1, \ldots, u.$$

2×2-plokkmaatriksi korral

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{pmatrix}.$$

(vi) Plokkdiagonaalse maatriksi $\mathbf{A} = [\mathbf{A}_{ii]}],\ i = 1, \ldots, u$ determinant on diagonaalplokkide $\mathbf{A}_{ii} : n_i \times n_i$ determinantide korrutis

$$|\mathbf{A}| = \prod_{i=1}^{u} |\mathbf{A}_{ii}|.$$

2×2-plokkmaatriksi korral

$$|\mathbf{A}| = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}||\mathbf{A}_{22}|.$$

Vaatame tõestuses 2×2 juhtu, üldjuht on põhimõtteliselt sarnane. Esitame maatriksi $\mathbf{A}$ kahe maatriksi korrutisena. Siis

$$\begin{vmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix} = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{vmatrix} \begin{vmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}||\mathbf{A}_{22}|.$$

(vii) Kui $\mathbf{A} : n \times n$ on kujul

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix},$$

kus $\mathbf{A}_{ii} : n_i \times n_i$, siis

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{22}|.$$

Tõestus. Esitame maatriksi $\mathbf{A}$ kahe plokkmaatriksi korrutisena ja leiame tegurite determinandid.

$$\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix} = \begin{vmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix} \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{vmatrix} = |\mathbf{A}_{11}||\mathbf{A}_{22}|.$$

(viii) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

ja plokk $\mathbf{A}_{22} : n_2 \times n_2$ pööratav. Siis

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|.$$

Tõestus. Korrutame maatriksit $\mathbf{A}$ vasakult maatriksiga

$$\mathbf{B} = \begin{pmatrix} \mathbf{I}_{n_1} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}$$

ja paremalt maatriksiga

$$\mathbf{C} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}_{n_2} \end{pmatrix}.$$

Siis korrutis

$$\mathbf{BA} = \begin{pmatrix} \mathbf{A}_{11} - \mathbf{A}12\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

ja

$$\mathbf{BAC} = \begin{pmatrix} \mathbf{A}_{11} - \mathbf{A}12\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix}.$$

Võtame korrutisest **BAC** determinandi. Kuna $|\mathbf{C}| = |\mathbf{B}| = 1$, siis

$$|\mathbf{BAC}| = |\mathbf{A}| = \begin{vmatrix} \mathbf{A}_{11} - \mathbf{A}12\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{vmatrix}.$$

Omaduse (vi) põhjal

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|.$$

(ix) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

ja plokk $\mathbf{A}_{11} : n_1 \times n_1$ pööratav. Siis

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|.$$

Tõestus on analoogne eelmise omaduse tõestusega.
Omadustes (viii) ja (ix) esinevad plokkide vahed kannavad Schuri täiendite (*Schur complements*) nime.

Definitsioon 10.2 *2×2-plokkmaatriksi* $\mathbf{A}$ *ploki* $\mathbf{A}_{ii}$ *Schuri täiendiks* $\mathbf{A}_{jj\cdot i}$ *nimetatakse maatriksit*

$$\mathbf{A}_{jj\cdot i} = \mathbf{A}_{jj} - \mathbf{A}_{ji}\mathbf{A}_{ii}^{-1}\mathbf{A}_{ij}, \quad i, j = 1, 2.$$

Järgnevalt vaatame 2×2 plokkmaatriksi $\mathbf{A} : n \times n$ pöördmaatrikseid.

(x) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}, \text{ siis } \mathbf{A}^{-1} = \begin{pmatrix} \mathbf{I}_{n_1} & -\mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}.$$

Kui

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{I}_{n_2} \end{pmatrix}, \text{ siis } \mathbf{A}^{-1} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ -\mathbf{A}_{21} & \mathbf{I}_{n_2} \end{pmatrix}.$$

Vahetu korrutamise tulemusena $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$ mõlemal juhul.

(xi) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

ja $\mathbf{A}_{11} : n_1 \times n_1$ ning $\mathbf{A}_{22} : n_2 \times n_2$ pööratavad. Siis

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22}^{-1} \end{pmatrix}.$$

Tõestus. Esimese sammuna esitame maatriksi $\mathbf{A}$ kolme maatriksi korrutisena

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}.$$

Kontrollige võrdus!
Kuna kõigi kolme teguri pöördmaatriksid on lihtsal kujul, on $\mathbf{A}$ pöördmaatriks

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_{n_2} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix}^{-1}.$$

Omaduste (v) ja (x) põhjal

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{pmatrix}.$$

Läbi korrutades saame

$$\begin{aligned} \mathbf{A}^{-1} &= \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_{n_2} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22}^{-1} \end{pmatrix}. \end{aligned}$$

Kontrollige korrutamist!

Üldjuhul on 2×2 plokkmaatriksi pöördmaatriksi plokkide avaldised tunduvalt keerulisemad, kuid tänu dimensiooni alanemisele on siiski tihti otstarbekas neid valemeid kasutada, eriti kui pööratava maatriksi determinant on nullilähedane.

(xii) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \text{ ja } \mathbf{A}_{11} : n_1 \times n_1 \text{ ning } \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$$

pööratavad , siis

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

kus

$$\begin{aligned}
\mathbf{B}_{11} &= \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}, \\
\mathbf{B}_{22} &= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}, \\
\mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}, \\
\mathbf{B}_{21} &= -\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}.
\end{aligned}$$

Tõestus. Peame tõestama, et

$$\mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \mathbf{I}_n.$$

Seega peame näitama järgmise nelja võrduse kehtivust:

$$\begin{aligned}
\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} &= \mathbf{I}_{n_1}, \\
\mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} &= \mathbf{I}_{n_2}, \\
\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} &= \mathbf{0}, \\
\mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} &= \mathbf{0}.
\end{aligned}$$

Alustame kolmandast võrdusest.

$$\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} = -\mathbf{A}_{11}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{A}_{12}\mathbf{B}_{22} = \mathbf{0}.$$

Kontrollime teist võrdust.

$$\mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} = -\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{A}_{22}\mathbf{B}_{22}$$

$$= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\mathbf{B}_{22} = \mathbf{B}_{22}^{-1}\mathbf{B}_{22} = \mathbf{I}_{n_1}.$$

Esimene võrdus on ka lihtsalt kontrollitav.

$$\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} = \mathbf{A}_{11}\mathbf{A}_{11}^{-1}+\mathbf{A}_{11}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}+\mathbf{A}_{12}\mathbf{B}_{21}$$

$$= \mathbf{I}_{n_1} - \mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{A}_{12}\mathbf{B}_{21} = \mathbf{I}_{n_1}.$$

Viimane võrdus on kõige keerukam kontrollida.

$$\mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1}+\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}+\mathbf{A}_{22}\mathbf{B}_{21}.$$

Kuna $\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} = -\mathbf{B}_{21}$, siis

$$\mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{A}_{22}\mathbf{B}_{21}$$

$$= \mathbf{A}_{21}\mathbf{A}_{11}^{-1} + (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\mathbf{B}_{21}$$

$$= \mathbf{A}_{21}\mathbf{A}_{11}^{-1} + \mathbf{B}_{22}^{-1}(-\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}) = \mathbf{A}_{21}\mathbf{A}_{11}^{-1} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1} = \mathbf{0}.$$

Oleme omaduse (xii) tõestanud.

(xiii) Olgu $n \times n$-maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

Kui $\mathbf{A}_{22} : n_2 \times n_2$ ja $\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ on pööratavad, siis

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

kus

$$
\begin{aligned}
\mathbf{B}_{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}, \\
\mathbf{B}_{22} &= \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}, \\
\mathbf{B}_{12} &= -\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}, \\
\mathbf{B}_{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11}.
\end{aligned}
$$

Tõestus on analoogiline eelmise omaduse tõestusega.

Märkus. Pöördmaatriksi elemendid on teatavasti üheselt määratud, aga nagu näeme, on plokiviisiliselt võimalik pöördmaatriksit erinevalt esitada. Tihti kasutatakse selle esitust kujul, mis on ära toodud järgmises teoreemis. Sel korral on tehtud rohkem eeldusi plokkide pööratavuse kohta.

Teoreem 10.2 *Olgu pööratav* 2×2 *plokkmaatriks*

$$
\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}
$$

ja diagonaalplokid $\mathbf{A}_{ii}$ *ning nende Schuri täiendid* $\mathbf{A}_{ii.j}$. $i, j = 1, 2$ *pööratavad. Siis*

$$
\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{22\cdot 1}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{11\cdot 2}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{22\cdot 1}^{-1} & \mathbf{A}_{11\cdot 2}^{-1} \end{pmatrix},
$$

kus Schuri täiendid

$$
\mathbf{A}_{11\cdot 2} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12},
$$

$$
\mathbf{A}_{22\cdot 1} = \mathbf{A}_{11} - \mathbf{A}_{2112}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}.
$$

Tõestuse idee on sama, mis omaduse (xii) korral, kuid siin on täiendavalt vaja kasutada ka binomiaalset pöördteoreemi (Teoreem 2.1 teises loengus).

Viimane tulemus selles paragrahvis on seotud 2×2 plokkmaatriksite astakuga.

Teoreem 10.3 *Olgu*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

kus $\mathbf{A} : m \times n$ *on astakuga* $r(\mathbf{A}) = r$ *ja* $\mathbf{A}_{11} : r \times r$. *Kui* $r(\mathbf{A}_{11}) = r$, *siis*

$$\mathbf{A}_{22} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}.$$

Tõestus. Olgu $r(\mathbf{A}_{11}) = r$. Näitame, et

$$r(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}) = 0.$$

Siis

$$\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} = \mathbf{0},$$

kust

$$\mathbf{A}_{22} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}.$$

Selleni jõudmiseks näitame, et

$$r(\mathbf{A}) = r\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = r + r(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}).$$

Moodustame korrutise, kus $\mathbf{A}$ on korrutatud pööratava maatriksiga, selle tulemusena astak ei muutu:

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}_{m-r} \end{pmatrix}\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{pmatrix}.$$

Kontrollige korrutamise tulemust!
Siit

$$r(\mathbf{A}) = r = r(\mathbf{A}_{11}) + r(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}).$$

Seega $r(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}) = 0$. Väide on tõestatud.

See tulemus annab meile võimaluse leida vastus ühele varasemale küsimusele. Väitsime üldistatud pöördmaatriksi paragrahvis, et kui maatriks

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix},$$

kus $r(\mathbf{A}) = r(\mathbf{B}) = r$ ja $\mathbf{B} : r \times r$, siis

$$\mathbf{A}^{-} = \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Selleks tuleb näidata, et $\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}$. Korrutamisel saame

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix} \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{D}\mathbf{B}^{-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{D}\mathbf{B}^{-1}\mathbf{C} \end{pmatrix}$$

Teoreemi 10.3 kohaselt antud eeldustel $\mathbf{D}\mathbf{B}^{-1}\mathbf{C} = \mathbf{E}$ ja

$$\begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

on maatriksi $\mathbf{A}$ üldistatud pöördmaatriks.

§11. Kommuteerimismaatriks.

Vaadeldav mõiste on kirjanduses tuntud ka kui permutatsiooni-maatriks *(permutation matrix)*, kuid enamus autoreid kasutab nimetust kommuteerimismaatriks *(commutation matrix)*.

Definitsioon 11.1A. *Plokkmaatriksit* $\mathbf{K}_{m,n} : mn \times mn$, *mis koosneb* $n \times m$*-plokkidest, nimetatakse kommuteerimismaatriksiks, kui*

$$(\mathbf{K}_{m,n})_{(i,j)(g,h)} = \begin{cases} 1, & g = j,\ i = h, \\ 0, & \textit{vastasel juhul} \end{cases},$$

kus $i,h = 1,\ldots,m$; $j,g = 1,\ldots,n$. Kirjutame näitena välja maatriksi $\mathbf{K}_{2,3}$

$$\mathbf{K}_{2,3} = \begin{pmatrix} 1 & 0 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 & \vdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 1 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 1 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 & \vdots & 0 & 1 \end{pmatrix}.$$

Näeme, et tulemuseks on permuteeritud ühikmaatriks, kus igas reas ja veerus on üks 1 ja ülejäänud elemendid on nullid ning sealjuures järgitakse teatud plokkstruktuuri.
Plokkide kaudu võime defineerida kommuteerimismaatriksi järgmiselt
Definitsioon 11.1B. *Kommuteerimismaatriks* $\mathbf{K}_{m,n}$ *on* $mn \times mn$*-plokkmaatriks, mis koosneb* $n \times m$*-plokkidest, kusjuures ig-ndas plokis on gi-s element 1 ja ülejäänud elemendid on nullid,* $i = 1\ldots,m$; $g = 1,\ldots,n$.
Toome ära mõned $\mathbf{K}_{m,n}$ omadused:

(i) $\mathbf{K}_{m,n} = \mathbf{K}'_{n,m}$.

Tõestus. Näitame, et maatriksite vastavad elemendid on võrdsed.

$$((\mathbf{K}_{n,m})')_{(i,j)(g,h)} = (\mathbf{K}_{n,m})_{(g,h)(i,j)} = \begin{cases} 1, & g = j,\ i = h, \\ 0, & \textit{vastasel juhul} \end{cases}$$

$$= (\mathbf{K}_{m,n})_{(i,j)(g,h)}.$$

Vastavad elemendid on võrdsed ja mõlema maatriksi plokid on $n \times m$.

(ii) $\mathbf{K}_{m,n}\mathbf{K}_{n,m} = \mathbf{I}_{mn}$.

Tõestus. Näitame, et maatriksite korrutise peadiagonaali elemendid võrduvad ühega. Korrutise üldelement

$$(\mathbf{K}_{m,n}\mathbf{K}_{n,m})_{(i,j)(g,h)} = \sum_{k=1}^{n}\sum_{l=1}^{m}(\mathbf{K}_{m,n})_{(i,j)(k,l)}(\mathbf{K}_{n,m})_{(k,l)(g,h)}$$

$$= \begin{cases} 1, & i = l = g,\ j = k = h, \\ 0, & vastasel\ juhul. \end{cases}$$

Korrutise element võrdub ühega, kui korrutismaatriksis reaindeks võrdub veeruindeksiga, s.t. element asub peadiagonaalil.

(iii) $\mathbf{K}_{m,1} = \mathbf{K}_{1,m} = \mathbf{I}_m$.

(iv) $\mathbf{K}_{m,n}$ on ortogonaalmaatriks.

(v) $|\mathbf{K}_{m,n}| = \pm 1$.

(vi) Olgu $\mathbf{A} : p \times q$ plokkmaatriks, mis koosneb $r \times s$ plokkidest

$$\mathbf{A} = [\mathbf{A}_{ig}],\ i = 1, \ldots, m;\ g = 1, \ldots, n;\ mr = p;\ ns = q.$$

Siis plokkmaatriks $\mathbf{AK}_{n,s}$ koosneb $r \times n$ plokkidest ja selle korrutismaatriksi (g,h)-s veerg on (h,g)-s maatriksi $\mathbf{A}$ veerg.

Tõestus. Näitame, et vastavad elemendid on võrdsed.

$$(\mathbf{AK})_{(i,j)(g,h)} = (\mathbf{A})_{(i,j)(h,g)}.$$

$$(\mathbf{AK})_{(i,j)(g,h)} = \sum_{k=1}^{n}\sum_{l=1}^{s}(\mathbf{A})_{(i,j)(k,l)}(\mathbf{K})_{(k,l)(g,h)}$$

$$= \sum_{k=1}^{n}\sum_{l=1}^{s}(\mathbf{A})_{(i,j)(k,l)} \times \begin{cases} 1 & kui\ \ k = h\ ja\ l = g; \\ 0 & mujal. \end{cases}$$

$$= (\mathbf{A})_{(i,j),(h,g)}.$$

Jälgige dimensioonide muutumist korrutamisel.

(vii) Olgu $\mathbf{A} : p \times q$ plokkmaatriks, mis koosneb $r \times s$ plokkidest

$$\mathbf{A} = [\mathbf{A}_{ig}],\ i = 1, \ldots, m;\ g = 1, \ldots, n;\ mr = p;\ ns = q.$$

Siis plokkmaatriks $\mathbf{K}_{r,m}\mathbf{A}$ koosneb $m \times s$ plokkidest ja (i, j)-s korrutismaatriksi rida on (j, i)-s maatriksi $\mathbf{A}$ rida.

Tõestus. Tõestus on analoogiline eelmise omaduse tõestusega.

(viii) Olgu $\mathbf{A} : p \times q$ plokkmaatriks, mis koosneb $r \times s$ plokkidest

$$\mathbf{A} = [\mathbf{A}_{ig}],\ i = 1, \ldots, m;\ g = 1, \ldots, n;\ mr = p;\ ns = q.$$

Siis plokkmaatriks $\mathbf{K}_{r,m}\mathbf{A}\mathbf{K}_{n,s}$ koosneb $m \times n$ plokkidest ja (i, j)-s korrutismaatriksi rida on (j, i)-s maatriksi $\mathbf{A}$ rida ning (g, h)-s korrutismaatriksi veerg on (h, g)-s maatriksi $\mathbf{A}$ veerg

$$(\mathbf{K}_{r,m}\mathbf{A}\mathbf{K}_{n,s})_{(i,j)(g,h)} = (\mathbf{A})_{(j,i),(h,g)}.$$

See omadus võtab kokku kaks eelmist.

§12. Otsekorrutis.

Inglise keeles on selle mõiste kõige levinum nimetus *Kronecker product*, aga kasutatakse ka nimetusi *direct product* ja *tensor product*. Eesti keeles eelistame *otsekorrutist.*

Definitsioon 12.1 *Maatriksite* $\mathbf{A} : m \times n$ *ja* $\mathbf{B} : r \times s$ *otsekorrutis* $\mathbf{A} \otimes \mathbf{B}$ *on* $mr \times ns$*-plokkmaatriks, mis koosneb* $r \times s$ *plokkidest*

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}], \quad i = 1, \ldots, m;\ j = 1, \ldots, n,$$

kus

$$a_{ij}\mathbf{B} = \begin{pmatrix} a_{ij}b_{11} & \cdots & a_{ij}b_{1s} \\ \vdots & \ddots & \vdots \\ a_{ij}b_{r1} & \cdots & a_{ij}b_{rs} \end{pmatrix}.$$

Kui me vaatame $\mathbf{A} \otimes \mathbf{B}$ kui tavalist maatriksit, saame tema üldelemendi jaoks võrduse

$$(\mathbf{A} \otimes \mathbf{B})_{(i,j)(g,h)} = (\mathbf{A} \otimes \mathbf{B})_{(i-1)r+j,(g-1)s+h}.$$

Otsekorrutise põhiomadused on järgmised.
(i) $(\mathbf{A} \otimes \mathbf{B})_{(i,j)(g,h)} = a_{ig}b_{jh}$.

Tõestus. Maatriksi $\mathbf{A}$ element määrab definitsiooni kohaselt ploki indeksid ja $\mathbf{B}$ indeksid määravad otsekorrutise elemendi paiknemise ploki sees, seega võrdus kehtib.

(ii) $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$.

Tõestus. Näitame, et vastavad elemendid on võrdsed.

$$(\mathbf{A} \otimes \mathbf{B})'_{(i,j)(g,h)} = (\mathbf{A} \otimes \mathbf{B})_{(g,h)(i,j)} = a_{gi}b_{hj} = (\mathbf{A})'_{ig}(\mathbf{B})'_{jh}$$

$$= (\mathbf{A}' \otimes \mathbf{B}')_{(i,j)(g,h)}.$$

(iii) Distributiivsus liitmise suhtes

$$(\mathbf{A} + \mathbf{B}) \otimes (\mathbf{C} + \mathbf{D}) = \mathbf{A} \otimes \mathbf{C} + \mathbf{A} \otimes \mathbf{D} + \mathbf{B} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{D}.$$

(iv) Assotsiatiivsus

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}).$$

(v) Maatriksite $\mathbf{A} : m \times n$, $\mathbf{B} : n \times w, \mathbf{C} : r \times s$ ja $\mathbf{D} : s \times t$ korral kehtib võrdus

$$(\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) = (\mathbf{AB}) \otimes (\mathbf{CD}). \tag{12.1}$$

Tõestus. Plokkmaatriksite korral on kahe maatriksi võrduse tõestamiseks kaks võimalust: kas näidata vastavate elementide võrdumine ja dimensioonide võrdsus või vastavate plokkide võrdsus ja dimensioonide võrdsus. Kasutame selle omaduse tõestamisel viimast võimalust. Kontrollime esmalt dimensioonide võrdsuse. Vasakul poolel on $(\mathbf{A} \otimes \mathbf{C}) : mr \times ns$ maatriks ja $\mathbf{B} \otimes \mathbf{D} : ns \times wt$ maatriks, korrutis on $mr \times wt$ maatriks.

Paremal poolel $\mathbf{AB}$ on $m \times w$ maatriks ja $\mathbf{CD} : r \times t$-maatriks. Otsekorrutis $(\mathbf{AB}) \otimes (\mathbf{CD})$ on aga $mr \times wt$ maatriks, seega dimensioonid on võrdsed.
Näitame vastavate plokkide võrduse.
Paremal poolel

$$[(\mathbf{AB}) \otimes (\mathbf{CD})]_{ij} = [(\mathbf{AB})_{ij}(\mathbf{CD})] = \Big[\sum_{k=1}^{n} a_{ik}b_{kj}(\mathbf{CD})\Big],$$

plokkide dimensioonid on $r \times t$.
Vasakul poolel

$$[(\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})]_{ij} = \sum_{k=1}^{n} [\mathbf{E}]_{ik}[\mathbf{F}]_{kj},$$

kus $[\mathbf{E}]_{ik}$ on $\mathbf{A} \otimes \mathbf{C}$ i-nda plokkide rea k-s plokk ja $[\mathbf{F}]_{kj}$ on j-ndas $\mathbf{B} \otimes \mathbf{D}$ plokkide veerus k-s plokk, plokkide dimensioonid on vastavalt $r \times s$ ja $s \times t$, korrutise plokk $r \times t$. Plokid on paremal ja vasakul sama dimensiooniga. Siis

$$\sum_{k=1}^{n} [\mathbf{E}]_{ik}[\mathbf{F}]_{kj} = \Big[\sum_{k=1}^{n} a_{ik}\mathbf{C}b_{kj}\mathbf{D}\Big] = \Big[\sum_{k=1}^{n} a_{ik}b_{kj}(\mathbf{CD})\Big].$$

Plokk vasakul poolel võrdub plokiga paremal poolel, väide on tõestatud.
(vi) Olgu $\mathbf{A} : m \times m$ ja $\mathbf{B} : r \times r$ pööratavad, siis

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}.$$

Tõestus. Eelmise omaduse põhjal

$$(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1})(\mathbf{A} \otimes \mathbf{B}) = \mathbf{I}_{mr}.$$

(vii) Olgu $\mathbf{A} : m \times n$, $\mathbf{B} : r \times s$ siis

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{K}_{m,r}(\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{s,n}.$$

Tõestus. Kontrollime dimensioone. Vasakul $\mathbf{A} \otimes \mathbf{B}$ on $mr \times ns$ maatriks, plokid $r \times s$. Paremal $\mathbf{K}_{m,r}$ plokid on $r \times m$, $\mathbf{B} \otimes \mathbf{A}$ plokid on $m \times n$ ja $\mathbf{K}_{s,n}$ plokid on $n \times s$. Korrutise plokid on $r \times s$, mõlemal poolel on plokid sama dimensiooniga. Parema poole maatriks on ka $mr \times ns$, maatriksid on samade dimensioonidega. Näitame, et vastavad elemendid on võrdsed. Vasak pool:

$$(\mathbf{A} \otimes \mathbf{B})_{(i,j)(g,h)} = a_{ig} b_{jh}.$$

Parem pool:

$$(\mathbf{K}_{m,r}(\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{s,n})_{(i,j)(g,h)} = (\mathbf{B} \otimes \mathbf{A})_{(j,i)(h,g)}$$

kommuteerimismaatriksi omaduse (viii) põhjal. Siis omaduse (i) tõttu

$$(\mathbf{B} \otimes \mathbf{A})_{(j,i)(h,g)} = b_{jh} a_{ig}.$$

Vastavad elemendid on võrdsed.
Järeldus 1. Omaduse (vii) tähistustes

$$(\mathbf{B} \otimes \mathbf{A}) = \mathbf{K}_{r,m}(\mathbf{A} \otimes \mathbf{B})\mathbf{K}_{n,s}.$$

Järeldus 2. Omaduse (vii) tähistustes

$$\mathbf{K}_{r,m}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{s,n}.$$

(viii) Kui $\mathbf{A} : m \times m$ ja $\mathbf{B} : r \times r$, siis

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^r |\mathbf{B}|^m;$$

Tõestus.

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A} \otimes \mathbf{I}_r||\mathbf{I}_m \otimes \mathbf{B}| = |\mathbf{B}|^m |\mathbf{K}_{m,r}(\mathbf{I}_r \otimes \mathbf{A})\mathbf{K}_{r,m}|$$

kuna maatriks $\mathbf{I}_m \otimes \mathbf{B}$ on plokkdiagonaalne. Siis

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{B}|^m |\mathbf{K}_{m,r}||\mathbf{I}_r \otimes \mathbf{A}||\mathbf{K}_{r,m}| = |\mathbf{B}|^m |\mathbf{I}_r \otimes \mathbf{A}| = |\mathbf{B}|^m ||\mathbf{A}|^r,$$

kuna $|\mathbf{K}_{m,r}||\mathbf{K}_{r,m}| = |\mathbf{K}_{m,r}\mathbf{K}_{r,m}|$ ja $\mathbf{K}_{m,r}\mathbf{K}_{r,m}$ on ühikmaatriks determinandiga 1 ning $\mathbf{I}_r \otimes \mathbf{A}$ on plokkdiagonaalne.

(ix) vektorite $\mathbf{a}$ ja $\mathbf{b}$ korral

$$\mathbf{a} \otimes \mathbf{b}' = \mathbf{a}\mathbf{b}' = \mathbf{b}' \otimes \mathbf{a}.$$

Kirjapildi lihtsustamiseks kasutame tähistust

$$\mathbf{a}^{\otimes k} = \underbrace{\mathbf{a} \otimes \cdots \otimes \mathbf{a}}_{k \text{ tegurit}}.$$

§13. vec-operaator

Definitsioon 13.1 *Olgu* $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ $m \times n$*-maatriks. Vektoriseerimisoperaator* $\mathrm{vec}(\cdot)$ *on operaator maatriksite ruumist* $\mathbb{R}^{m\times n}$ *ruumi* $\mathbb{R}^{mn}$ *kujul*

$$\mathrm{vec}\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

vec-operaator toodi matemaatikasse ligi 140 aastat tagasi (Sylvester, 1884), mitmemõõtmelise statistika jaoks "avastati" see circa 50 aastat tagasi ja praegu ei ole seda tähistust artiklites enam vaja selgitada. Enne kui läheme selle operaatori omaduste juurde, tuleb rääkida tähistustest. Kui tegemist on plokkidest koosnevate rea- või veeruvektoritega, siis kasutatakse järgmist tähistust. Plokkidest koosnev veeruvektor $\mathbf{a}$ koosneb vaid ridade plokkidest (alamveergudest) ja vektori $\mathbf{a}$ elementi i-nda ploki j-ndas reas tähistame $\mathbf{a}_{(i,j)1}$. Kui tegemist on transponeeritud vektoriga $\mathbf{a}'$, siis on tegemist veergude plokkidega ja sel juhul tähistatakse g-nda ploki h-s element $(\mathbf{a})'_{1(g,h)}$. Nendes tähistustes maatriksi $\mathbf{A} : m \times n$ i-nda veeru j-s element on vektoris $\mathrm{vec}\mathbf{A}$ kohal $(\mathrm{vec}\mathbf{A})_{(i,j)1}$.

OMADUSED.

Järgnevad omadused seovad vec-operaatori teiste maatriks-operatsioonidega.

(i) Kui $\mathbf{A}$ on $m \times n$-maatriks, siis

$$(\mathrm{vec}\mathbf{A})_{(i,j)1} = a_{ji}.$$

(ii) Kui $\mathbf{A}$ on $m \times n$-maatriks, siis

$$\mathbf{K}_{m,n}\mathrm{vec}\mathbf{A} = \mathrm{vec}\mathbf{A}'.$$

Tõestus. Omadus on otsene järeldus kommuteerimismaatriksi omadusest (vii), mille kohaselt maatriksi $\mathbf{A}$ korrutamisel vasakult kommuteerimismaatriksiga on korrutise (ij)-s rida maatriksi $\mathbf{A}$ (ji)-s rida. Praegu koosneb see rida ühest elemendist.

$$(\mathbf{K}_{m,n}\mathrm{vec}\mathbf{A})_{(i,j),1} = \sum_{k=1}^{n}\sum_{l=1}^{m}\mathbf{K}_{(i,j)(k,l)}(\mathrm{vec}\mathbf{A})_{(k,l)1}$$

$$= (\mathrm{vec}\mathbf{A})_{(j,i)1} = (\mathbf{A})_{(i,j)} = (\mathbf{A}')_{(j,i)} = (\mathrm{vec}\mathbf{A}')_{(i,j)1}.$$

Vastavad elemendid on võrdsed.

(iii) Kui $\mathbf{A}$ on $m \times n$-maatriks, siis

$$\mathrm{vec}\mathbf{A} = \mathbf{K}_{n,m}\mathrm{vec}\mathbf{A}'.$$

Tõestus. Võrduse saame eelmisest omadusest, kui korrutame vasakult mõlemat poolt maatriksiga $\mathbf{K}_{n,m}$.

(iv) Kui $\mathbf{A}$ on $n \times n$ sümmeetriline maatriks, siis

$$\mathrm{vec}\mathbf{A} = \mathbf{K}_{n,n}\mathrm{vec}\mathbf{A}.$$

Seda võrdust kasutatakse mõnikord võrdsete indeksitega kommuteerimismaatriksi definitsioonina.

(v) Olgu maatriksid $\mathbf{A} : m \times n$, $\mathbf{B} : n \times r$, $\mathbf{C} : r \times s$, siis

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}. \tag{13.1}$$

Tõestus. See on rakendustes üks olulisemaid vec-operaatori omadusi. Näitame, et vastavad elemendid vasakul ja paremal pool on võrdsed. Vasak pool:

$$(\mathrm{vec}(\mathbf{ABC}))_{(i,j)1} = \begin{pmatrix} \mathbf{ABc}_1 \\ \dots \\ \mathbf{ABc}_s \end{pmatrix}_{(i,j)1}$$

$$= \sum_{l=1}^{r} (\mathbf{AB})_{jl} c_{li} = \sum_{k=1}^{n} \sum_{l=1}^{r} a_{jk} b_{kl} c_{li}.$$

Parem pool:

$$((\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B})_{(i,j)1} = \sum_{k=1}^{r} \sum_{l=1}^{n} (\mathbf{C}' \otimes \mathbf{A})_{(i,j)(k,l)} (\mathrm{vec}\mathbf{B})_{(k,l)1}$$

$$= \sum_{k=1}^{r} \sum_{l=1}^{n} (\mathbf{C}')_{ik} a_{jl} b_{lk} = \sum_{k=1}^{r} \sum_{l=1}^{n} c_{ki} a_{jl} b_{lk}.$$

Vasaku ja parema poole vastavad elemendid on võrdsed, omadus on tõestatud.

(vi) Kui $\mathbf{A}$ ja $\mathbf{B}$ on $m \times n$-maatriksid, siis

$$\mathrm{tr}(\mathbf{A}'\mathbf{B}) = (\mathrm{vec}\mathbf{A})'\mathrm{vec}\mathbf{B} = (\mathrm{vec}\mathbf{B})'\mathrm{vec}\mathbf{A}. \tag{13.2}$$

Tõestus. Olgu $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ ja $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$. Siis vasakul pool

$$\mathrm{tr}(\mathbf{A}'\mathbf{B}) = \sum_{i=1}^{n} (\mathbf{A}'\mathbf{B})_{ii} = \sum_{i=1}^{n} \mathbf{a}_i' \mathbf{b}_i.$$

Võrduse (13.2) esimese võrduse paremal poolel

$$(\mathrm{vec}\mathbf{A})'\mathrm{vec}\mathbf{B} = \sum_{i=1}^{n} \mathbf{a}_i' \mathbf{b}_i.$$

Võrduse (13.2) paremal poolel teine võrdus kehtib, kuna vektorite skalaarkorrutis on kommutatiivne. Vasak pool võrdub parema poolega, väide kehtib.

(vii) Olgu $\mathbf{A} : m \times n$, $\mathbf{B} : n \times r$, $\mathbf{C} : r \times s$ ja $\mathbf{D} : s \times m$ maatriksid, siis

$$\mathrm{tr}(\mathbf{ABCD}) = (\mathrm{vec}\mathbf{A}')'(\mathbf{D}' \otimes \mathbf{B})\mathrm{vec}\mathbf{C}.$$

Tõestus. Vaatame jälje all olevat korrutist kahe maatriksi korrutisena

$$\mathrm{tr}(\mathbf{ABCD}) = (\mathrm{vec}(\mathbf{AB})')'\mathrm{vec}(\mathbf{CD}) = (\mathrm{vec}(\mathbf{B}'\mathbf{A}'))'\mathrm{vec}(\mathbf{CD}).$$

Kirjutame välja vektorite avaldised

$$\mathrm{vec}(\mathbf{CD}) = \mathrm{vec}(\mathbf{I}_r\mathbf{CD}) = (\mathbf{D}' \otimes \mathbf{I}_r)\mathrm{vec}\mathbf{C};$$

$$\begin{aligned}(\mathrm{vec}(\mathbf{AB})')' &= (\mathrm{vec}(\mathbf{B}'\mathbf{A}'\mathbf{I}_m))' = ((\mathbf{I}_m \otimes \mathbf{B}')\mathrm{vec}\mathbf{A}')' \\ &= (\mathrm{vec}\mathbf{A}')'(\mathbf{I}_m \otimes \mathbf{B}).\end{aligned}$$

Siis

$$\begin{aligned}\mathrm{tr}(\mathbf{ABCD}) &= (\mathrm{vec}\mathbf{A}')'(\mathbf{I}_m \otimes \mathbf{B})(\mathbf{D}' \otimes \mathbf{I}_r)\mathrm{vec}\mathbf{C} \\ &= (\mathrm{vec}\mathbf{A}')'(\mathbf{D}' \otimes \mathbf{B})\mathrm{vec}\mathbf{C}.\end{aligned}$$

Viimase võrduse saamisel kasutasime otsekorrutise omadust (v), valem (12.1).

(viii) Olgu $\mathbf{A} : m \times n$ ja $\mathbf{B} : r \times s$, maatriksid,

$$\mathrm{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_n \otimes \mathbf{K}_{s,m} \otimes \mathbf{I}_r)(\mathrm{vec}\mathbf{A} \otimes \mathrm{vec}\mathbf{B}).$$

Tõestus. Selle omaduse tõestus on tehnilisem. Esitame maatriksid $\mathbf{A}$ ja $\mathbf{B}$ veergude kaudu Olgu

$$\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n), \qquad \mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_s)$$

ning tähistame ühikmaatriksite veerud $(\mathbf{I}_n)_i$ ja $(\mathbf{I}_s)_j$. Maatriksid $\mathbf{A}$ ja $\mathbf{B}$ saame esitada nende vektorite kaudu summana

$$\mathbf{A} = \sum_{i=1}^{n} \mathbf{a}_i(\mathbf{I}_n)_i', \qquad \mathbf{B} = \sum_{j=1}^{s} \mathbf{b}_j(\mathbf{I}_s)_j'.$$

Siis

$$\operatorname{vec}(\mathbf{A}\otimes\mathbf{B}) = \operatorname{vec}\Big[\Big(\sum_{i=1}^{n}\mathbf{a}_i(\mathbf{I}_n)'_i\Big)\otimes\Big(\sum_{j=1}^{s}\mathbf{b}_j(\mathbf{I}_s)'_j\Big)\Big]$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{s}\operatorname{vec}[(\mathbf{a}_i(\mathbf{I}_n)'_i)\otimes(\mathbf{b}_j(\mathbf{I}_s)'_j)]$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{s}\operatorname{vec}[(\mathbf{a}_i\otimes\mathbf{b}_j)((\mathbf{I}_n)_i\otimes(\mathbf{I}_s)_j)'].$$

Viimase võrduse saamisel kasutasime otsekorrutise omadust (v), valem (12.1). Kasutame nüüd vec-operaatori omadust (v), valem (13.1)

$$\operatorname{vec}(\mathbf{A}\otimes\mathbf{B}) = \sum_{i=1}^{n}\sum_{j=1}^{s}\operatorname{vec}[(\mathbf{a}_i\otimes\mathbf{b}_j)1((\mathbf{I}_n)_i\otimes(\mathbf{I}_s)_j)']$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{s}[((\mathbf{I}_n)_i\otimes(\mathbf{I}_s)_j)\otimes(\mathbf{a}_i\otimes\mathbf{b}_j)]\operatorname{vec}1.$$

Vahetame kommuteerimismaatriksite abil keskmise otsekorrutise tegurite järjekorra

$$\operatorname{vec}(\mathbf{A}\otimes\mathbf{B}) = \sum_{i=1}^{n}\sum_{j=1}^{s}[(\mathbf{I}_n)_i\otimes[\mathbf{K}_{s,m}(\mathbf{a}_i\otimes(\mathbf{I}_s)_j)\mathbf{K}_{1,1}]\otimes(\mathbf{b}_j)].$$

Korrutame esimest ja viimast maatriksit vasakult ühikmaatriksiga ja rakendame taaskord otsekorrutise omadust (v) kolme teguri korral, valem (12.1), samas $\mathbf{K}_{1,1}=1$.

$$\operatorname{vec}(\mathbf{A}\otimes\mathbf{B}) = \sum_{i=1}^{n}\sum_{j=1}^{s}\mathbf{I}_n(\mathbf{I}_n)_i\otimes[\mathbf{K}_{s,m}(\mathbf{a}_i\otimes(\mathbf{I}_s)_j)]\otimes\mathbf{I}_r\mathbf{b}_j$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{s}(\mathbf{I}_n\otimes\mathbf{K}_{s,m}\otimes\mathbf{I}_r)[(\mathbf{I}_n)_i\otimes\mathbf{a}_i\otimes(\mathbf{I}_s)_j\otimes\mathbf{b}_j]$$

$$= (\mathbf{I}_n \otimes \mathbf{K}_{s,m} \otimes \mathbf{I}_r)\Big[\sum_{i=1}^{n}((\mathbf{I}_n)_i \otimes \mathbf{a}_i) \otimes \sum_{j=1}^{s}((\mathbf{I}_s)_j) \otimes \mathbf{b}_j)\Big]$$

$$= (\mathbf{I}_n \otimes \mathbf{K}_{s,m} \otimes \mathbf{I}_r)\Big[\sum_{i=1}^{n}((\mathbf{I}_n)_i \otimes \mathbf{a}_i)\mathrm{vec}1 \otimes \sum_{j=1}^{s}((\mathbf{I}_s)_j) \otimes \mathbf{b}_j)\mathrm{vec}1\Big].$$

Rakendame vec-operaatori omadust (v), valem (13.1)

$$\mathrm{vec}(\mathbf{A}\otimes\mathbf{B}) = (\mathbf{I}_n\otimes\mathbf{K}_{s,m}\otimes\mathbf{I}_r)\Big[\sum_{i=1}^{n}\mathrm{vec}(\mathbf{a}_i(\mathbf{I}_n)'_i)\otimes\sum_{j=1}^{s}\mathrm{vec}(\mathbf{b}_j(\mathbf{I}_s)'_j)\Big]$$

$$= (\mathbf{I}_n \otimes \mathbf{K}_{s,m} \otimes \mathbf{I}_r)(\mathrm{vec}\mathbf{A} \otimes \mathrm{vec}\mathbf{B}).$$

Omadus on tõestatud.

Näide otsekorrutise ja vec-operaatori kasutamisest.
Olgu meil üldkogum kirjeldatud juhusliku p-vektoriga

$$\mathbf{x} = (X_1, \ldots, X_p)'$$

ja meil on sellest konkreetne valim mahuga n

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix}.$$

Arvud x_{ij} on juhuslike suuruste $X_{ij} \sim X_i$ väärtused. Konkreetset valimit kirjeldab teoreetiline valim – juhuslik maatriks

$$\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{p1} & X_{p2} & \cdots & X_{pn} \end{pmatrix} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n),$$

kus juhuslikud vektorid $\mathbf{x}_i = (X_{1i}, X_{2i}, \ldots, X_{pi})'$ on sõltumatud ja sama jaotusega, $\mathbf{x}_i \sim \mathbf{x}$.

Klassikalise mitmemõõtmelise statistika eelduseks on normaaljaotus $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kus

$$E\mathbf{x} = \boldsymbol{\mu} = (\mu_1, \dots, \mu_p)', \qquad D\mathbf{x} = \boldsymbol{\Sigma} = E[(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})'].$$

Kuidas panna kirja kogu teoreetilise valimi jaotus?
Esitame maatriksi $\mathbf{X}$ vektoriseerituna

$$\mathrm{vec}\mathbf{X} \sim N_{pn}(\boldsymbol{\mu}_{\mathrm{vec}\mathbf{X}}, \boldsymbol{\Sigma}_{\mathrm{vec}\mathbf{X}}).$$

Kuidas avalduvad $\boldsymbol{\mu}_{\mathrm{vec}\mathbf{X}}$ ja $\boldsymbol{\Sigma}_{\mathrm{vec}\mathbf{X}}$? Kuna kõik juhuslikud vektorid on sama jaotusega, $\mathbf{x}_i \sim \mathbf{x}$, siis

$$E(\mathrm{vec}\mathbf{X}) = \begin{pmatrix} \boldsymbol{\mu} \\ \dots \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} \otimes \boldsymbol{\mu} = \mathbf{1}_n \otimes \boldsymbol{\mu}.$$

Kovariatsioonimaatriks

$$D(\mathrm{vec}\mathbf{X}) = \begin{pmatrix} D\mathbf{x}_1 & Cov(\mathbf{x}_1, \mathbf{x}_2) & \cdots & Cov(\mathbf{x}_1, \mathbf{x}_n) \\ Cov(\mathbf{x}_2, \mathbf{x}_1) & D\mathbf{x}_2 & \cdots & Cov(\mathbf{x}_2, \mathbf{x}_n) \\ \dots & \dots & \dots & \dots \\ Cov(\mathbf{x}_n, \mathbf{x}_1) & Cov(\mathbf{x}_n, \mathbf{x}_2 & \cdots & D\mathbf{x}_n \end{pmatrix},$$

kus $Cov(\mathbf{x}_i, \mathbf{x}_j) = E[(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})']$. Tänu vektorite $\mathbf{x}_i$ sõltumatusele $Cov(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{0}$ ja

$$D(\mathrm{vec}\mathbf{X}) = \begin{pmatrix} D\mathbf{x}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & D\mathbf{x}_2 & \cdots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \cdots & D\mathbf{x}_n \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} & \cdots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma} = \end{pmatrix} = \mathbf{I}_n \otimes \boldsymbol{\Sigma}.$$

Seega

$$\mathrm{vec}\mathbf{X} \sim N_{pn}(\mathbf{1}_n \otimes \boldsymbol{\mu}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}).$$

§14. Frechet' tuletisest

Mitmemõõtmelises statistikas on sageli tegemist olukorraga, kus on vaja leida tuletisi mitmemuutuja- või maatriksfunktsioonidest olukorras, kus ka funktsiooni väärtused on vektorid või maatriksid. Selline on olukord näiteks suurima tõepära ja vähimruutude hinnangute leidmisel ja juhuslike vektorite ning maatriksite asümptootilise normaaljaotuse leidmisel. Funktsionaalanalüüsis on kasutusele võetud operaatorkujul mitmeid erinevaid tuletisi selliste olukordade jaoks, nimetame Gâteaux' ehk nõrka tuletist ja Frechet' ehk tugevat tuletist. Neist viimane sobib hästi statistikaprobleemide lahendamiseks.

Meie lähem eesmärk on sisse tuua maatrikstuletise mõiste, mis tähendab ühe maatriksi $\mathbf{Y}$ elementide diferentseerimist teise maatriksi $\mathbf{X}$ elementide järgi, s. t. meil on tegemist osatuletistega. Kui $\mathbf{Y} : r \times s$ ja $\mathbf{X} : p \times q$, siis meil on osatuletised

$$\frac{\partial y_{ij}}{\partial x_{kl}}, \quad i = 1, \dots, r;\ j = 1, \dots, s;\ k = 1, \dots, p;\ l = 1, \dots, q,$$

kus $y_{ij} = y_{ij}(\mathbf{X})$.

Kas igasugust osatuletiste kogumit sobib nimetada tuletiseks? Missuguseid nõudeid peaks rahuldama mõistlikult defineeritud tuletis? Siin saame kasutada matemaatilise analüüsi abi. Vaatame kaht eukleidilist ruumi $\mathbb{U}$ ja $\mathbb{V}$ ja kujutust

$$f : \mathbb{U} \to \mathbb{V}.$$

Meil $\mathbb{U}$ on $p \times q$ maatriksite ruum ja $\mathbb{V}$ on $r \times s$ maatriksite ruum, s. t. $\mathbf{X} \in \mathbb{R}^{p\times q}$ ja $f(\mathbf{X}) = \mathbf{Y} \in \mathbb{R}^{r\times s}$. Milline on olukord ühemõõtmelisel juhul? Tuletis punktis x toodi sisse piirväärtusena

$$f'(x) = \lim_{\delta\to 0} \frac{f(x+\delta) - f(x)}{\delta}.$$

Siis väikese δ korral

$$\frac{f(x+\delta) - f(x)}{\delta} = f'(x) + \frac{\varepsilon(x,\delta)}{\delta},$$

kus $\frac{\varepsilon(x,\delta)}{\delta} \to 0$, kui $\delta \to 0$. Siit saame võrduse

$$f(x+\delta) = f(x) + f'(x)\delta + \varepsilon(x,\delta). \tag{14.1}$$

Kui f on mitmemuutuja funktsioon, siis tuletis esitatakse osatuletiste kaudu täistuletisena.
Kui

$$f : \mathbb{U} \to \mathbb{V},$$

kus $\mathbb{U}$ ja $\mathbb{V}$ on eukleidilised ruumid, siis on tuletisi kasutusele võetud rohkem kui üks, näiteks nõrk ehk Gâteaux' tuletis ja tugev ehk Frechet' tuletis. Maatrikstuletise näol on tegemist Frechet' tuletise maatriksrealisatsiooniga. Väliselt on tegemist ühemõõtmelise võrduse (14.1) analoogiga.
Definitsioon 14.1. *Olgu $\mathbb{U}$ n-mõõtmeline ja $\mathbb{V}$ m-mõõtmeline eukleidiline ruum ning $f : \mathbb{U} \to \mathbb{V}$. Siis kujutus f on Frechet' mõttes diferentseeruv punktis $\mathbf{x}$, kui*

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + f'(\mathbf{x})\mathbf{h} + \varepsilon(\mathbf{x},\mathbf{h}),$$

kus $f'(\mathbf{x})$ on pidev lineaarne operaator ja ühtlaselt iga $\mathbf{h}$ korral

$$\frac{||\varepsilon(\mathbf{x},\mathbf{h})||}{||\mathbf{h}||} \to 0, \text{ kui } ||\mathbf{h}|| \to 0.$$

Siin norm

$$||\mathbf{h}|| = \sqrt{h_1^2 + \cdots + h_n^2}.$$

Pidevat lineaarset kujutust $f'(\mathbf{x})$ nimetatakse $f(\mathbf{x})$ *Frechet' tuletiseks* punktis $\mathbf{x}$ ja liidetavat $f'(\mathbf{x})\mathbf{h}$ *Frechet' diferentsiaaliks.*
Kujutus $f(\mathbf{x})$ on diferentseeruv hulgas $\mathcal{A}$, kui ta on diferentseeruv igas $\mathcal{A}$ punktis. Frechet' tuletise olulisemad omadused on esitatud järgmises teoreemis.
Teoreem 14.1. *Olgu $\mathbb{U}$ n-mõõtmeline ja $\mathbb{V}$ m-mõõtmeline eukleidiline ruum ning $f : \mathbb{U} \to \mathbb{V}$.*

(i) *Kui $f(\mathbf{x}) = const$, siis $f'(\mathbf{x}) = \mathbf{0}$.*

(ii) *Kui* $f(\mathbf{x})$ *on lineaarne kujutus, siis* $f'(\mathbf{x}) = f(\mathbf{x})$.

Tõepoolest, siis $f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + f(\mathbf{h})$.

(iii) *Kui* $f(\mathbf{x})$ *ja* $g(\mathbf{x})$ *on diferentseeruvad punktis* $\mathbf{x}$*, siis ka* $f(\mathbf{x}) + g(\mathbf{x})$ *ja* $cf(\mathbf{x})$ *on diferentseeruvad punktis* $\mathbf{x}$ *ja*

$$(f(\mathbf{x}) + g(\mathbf{x}))' = f'(\mathbf{x}) + g'(\mathbf{x}), \quad (cf(\mathbf{x}))' = cf'(\mathbf{x}).$$

(iv) *Olgu* $\mathbb{U}$, $\mathbb{V}$ *ja* $\mathbb{W}$ *eukleidilised ruumid ning*

$$f : \mathbb{U} \to \mathbb{V}, \qquad g : \mathbb{V} \to \mathbb{W}$$

nii, et f *on diferentseeruv punktis* $\mathbf{x}$ *ja* g *punktis* $\mathbf{y} = f(\mathbf{x})$.

Siis kujutus $h = f \odot g : \mathbb{U} \to \mathbb{W}$ *on diferentseeruv punktis* $\mathbf{x}$ *ja*

$$h'(\mathbf{x}) = g'(\mathbf{y}) \odot f'(\mathbf{x}),$$

kus $\odot$ *tähistab operaatorite kompositsiooni.*

Omadus (iv) kujutab endast liitfunktsiooni diferentseerimise eeskirja. Rakenduste jaoks on ülioluline, et saaksime liitfunktsioonist tuletist võtta. Nõrgal tuletisel omadus (iv) puudub.

Kuna kujutuse $f(\mathbf{x})$ Frechet' tuletis on lineaarne ja pidev kujutus, siis saab seda esitada maatriksina. Kui Frechet' tuletis eksisteerib punktis $\mathbf{x}$, siis eksisteerivad ka osatuletised $\frac{\partial f_j(\mathbf{x})}{\partial x_i}$, $\quad i = 1, \ldots, n;\ j = 1, \ldots, m$ ja maatriks, mis esitab Frechet' tuletist, koosneb kõikvõimalikest osatuletistest.

Vastupidine ei ole tõene – kõigi osatuletiste eksisteerimisest ei pruugi järelduda Frechet' tuletise olemasolu. On olemas tarvilik ja piisav tingimus selleks, et osatuletiste olemasolust järelduks Frechet' tuletise eksisteerimine, aga see on praktikas raskesti kont-rollitav. Õnneks leiduvad piisavad tingimused, mida on mugav kasutada. Need on ära toodud järgmises teoreemis.

Teoreem 14.2 *Olgu* $\mathbb{U}$ n*-mõõtmeline ja* $\mathbb{V}$ m*-mõõtmeline eukleidiline ruum ning* $f : \mathbb{U} \to \mathbb{V}$ *selline kujutus, et kõik osatuletised* $\frac{\partial f_j(\mathbf{x})}{\partial x_i}$ *eksisteerivad lahtises punkti* $\mathbf{x}_0$ *ümbruses ning on pidevad*

punktis $\mathbf{x}_0$. *Siis* $f(\mathbf{x})$ *on diferentseeruv Frechet' mõttes punktis* $\mathbf{x}_0$.
Edaspidi eeldame maatrikstuletise defineerimisel nende piisavate tingimuste täidetust.

§15. Maatrikstuletis

Kasutame osatuletiste järjestamiseks definitsiooni, mille võttis kasutusele Heinz Neudecker (Neudecker, 1969). Kirjapildi lihtsustamiseks kasutame edaspidi $(\mathrm{vec}\mathbf{A})'$ asemel tähistust $\mathrm{vec}'\mathbf{A}$. Maatrikstuletise omaduste juures läheb vaja nii kommuteerimismaatriksi kui vec-operaatori ja otsekorrutise omadusi.

Definitsioon 15.1 *Olgu* $r \times s$*-maatriksi* $\mathbf{Y}$ *elemendid funktsioonid* $p \times q$*-maatriksi* $\mathbf{X}$ *elementidest. Nimetame* $rs \times pq$*-maatriksit* $\frac{d\mathbf{Y}}{d\mathbf{X}}$ *maatriksi* $\mathbf{Y}$ *tuletiseks maatriksi* $\mathbf{X}$ *järgi lahtises piirkonnas* $\mathcal{D}$, *kui kõik osatuletised* $\frac{\partial y_{kl}}{\partial x_{gh}}$ *eksisteerivad ja on pidevad selles piirkonnas ning osatuletise* $\frac{\partial y_{kl}}{\partial x_{gh}}$ *paiknemine maatriksis on määratud otsekorrutisega*

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y}, \tag{15.1}$$

kus

$$\frac{d}{d\mathrm{vec}'\mathbf{X}} = \left(\frac{\partial}{\partial x_{11}}, \ldots, \frac{\partial}{\partial x_{p1}}, \frac{\partial}{\partial x_{12}}, \ldots, \frac{\partial}{\partial x_{p2}}, \ldots, \frac{\partial}{\partial x_{1q}}, \ldots, \frac{\partial}{\partial x_{pq}}\right).$$

Märkus. Osatuletiste järjestamiseks on mitmeid võimalusi. Näiteks $\frac{d}{d\mathbf{X}} \otimes \mathbf{Y}$, kus säilib maatriksite plokkstruktuur ja ühes plokis on kõigi $\mathbf{Y}$ elementide osatuletised x_{ij} järgi. Võib defineerida tuletise vec-operaatorite kaudu ka kui otsekorrutise $\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \mathrm{vec}'\mathbf{Y}$. 1970- ja 80-ndatel aastatel toodigi artiklites sisse mitmeid erinevaid esitusi ja ei olnud üksmeelt maatrikstuletise mõiste osas. Valdavaks on jäänud ülaltoodud definitsioon ja seda põhjusel, et ainult selle korral on liitfunktsiooni diferentseerimise eeskiri väliselt samasugune ühemõõtmelise juhuga.

Kirjutame plokkmaatriksi (15.1) välja elementide kaudu:

$$
\mathbf{A} = \begin{pmatrix}
\frac{\partial y_{11}}{\partial x_{11}} & \cdots & \frac{\partial y_{11}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{11}}{\partial x_{1q}} & \cdots & \frac{\partial y_{11}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{r1}}{\partial x_{11}} & \cdots & \frac{\partial y_{r1}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{r1}}{\partial x_{1q}} & \cdots & \frac{\partial y_{r1}}{\partial x_{pq}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial y_{12}}{\partial x_{11}} & \cdots & \frac{\partial y_{12}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{12}}{\partial x_{1q}} & \cdots & \frac{\partial y_{12}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{r2}}{\partial x_{11}} & \cdots & \frac{\partial y_{r2}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{r2}}{\partial x_{1q}} & \cdots & \frac{\partial y_{r2}}{\partial x_{pq}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial y_{1s}}{\partial x_{11}} & \cdots & \frac{\partial y_{1s}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{1s}}{\partial x_{1q}} & \cdots & \frac{\partial y_{1s}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{rs}}{\partial x_{11}} & \cdots & \frac{\partial y_{rs}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{rs}}{\partial x_{1q}} & \cdots & \frac{\partial y_{rs}}{\partial x_{pq}}
\end{pmatrix},
$$

Paneme tähele, et tuletise definitsioonist tulenevad järgmised võrdused

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathrm{vec}\mathbf{Y}}{d\mathrm{vec}\mathbf{X}} = \frac{d\mathbf{Y}}{d\mathrm{vec}\mathbf{X}} = \frac{d\mathrm{vec}'\mathbf{Y}}{d\mathbf{X}}.$$

See tähendab, et maatriksite $\mathbf{X}$ ja $\mathbf{Y}$ vektoriseerimine ei muuda tuletise kuju.

Maatrikstuletise omadused esitame loeteluna, maatriksite mõõtmed toome ära, kui need erinevad definitsioonis esinevatest.

(i)

$$\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{I}_{pq}.$$

Omadus tuleneb vahetult definitsioonist.

(ii) Kui $\mathbf{Y} = c\mathbf{X},\ c \in \mathbb{R}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = c\mathbf{I}_{pq}.$$

Ka see omadus on järeldus definitsioonist.

(iii) Kui $\mathrm{vec}\mathbf{Y} = \mathbf{A}\mathrm{vec}\mathbf{X}$, kus $\mathbf{A}$ on konstantne maatriks, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \mathbf{A}.$$

Tõestus. Definitsiooni kohaselt

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes (\mathbf{A}\mathrm{vec}\mathbf{X}).$$

Rakendades otsekorrutise omadust

$$(\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) = (\mathbf{AB}) \otimes (\mathbf{CD}),$$

saame

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = 1 \cdot \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes (\mathbf{A}\mathrm{vec}\mathbf{X}) = (1 \otimes \mathbf{A})(\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{X})$$

$$= \mathbf{A}\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{A}.$$

(iv) kui $\mathbf{Y} = \mathbf{U} + \mathbf{V}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathbf{U}}{d\mathbf{X}} + \frac{d\mathbf{V}}{d\mathbf{X}}.$$

Omadus järeldub vahetult definitsioonist ja otsekorrutise omadustest.

(v) Liitfunktsiooni tuletis. Olgu $\mathbf{Z} = \mathbf{Z}(\mathbf{Y}) : m \times n$, $\mathbf{Y} = \mathbf{Y}(\mathbf{X}) : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}}\frac{d\mathbf{Y}}{d\mathbf{X}}.$$

Tõestus. Matemaatilisest analüüsist on teada mitmemuutuja liitfunktsiooni osatuletise leidmise eeskiri. Olgu meil p-muutuja funktsioon $f(y_1, \ldots, y_p)$,

kusjuures $y_i = y_i(x_1, \ldots, x_n)$, $i = 1, \ldots, p$, siis

$$\frac{\partial f(y_1, \ldots, y_p)}{\partial x_i} = \sum_{j=1}^{p} \frac{\partial f(y_1, \ldots, y_p)}{\partial y_j} \frac{\partial y_j}{\partial x_i}.$$

Meil on vektorite asemel maatriksid ja tuleb summeerida üle kõigi maatriksi elementide

$$\frac{\partial z_{kl}}{\partial x_{gh}} = \sum_{j=1}^{r} \sum_{j=1}^{s} \frac{\partial z_{kl}}{\partial y_{ij}} \frac{\partial y_{ij}}{\partial x_{gh}},$$

$k = 1, \ldots, m$; $l = 1, \ldots, n$; $i = 1, \ldots, r$; $j = 1, \ldots, s$;

$g = 1, \ldots, p$; $h = 1, \ldots, q$.

Maatrikstuletise definitsiooni kohaselt

$$\frac{dz_{kl}}{d\mathbf{Y}} = \frac{d}{d\mathrm{vec}'\mathbf{Y}} \otimes z_{kl} = \Big(\frac{\partial z_{kl}}{\partial y_{11}}, \cdots, \frac{\partial z_{kl}}{\partial y_{rs}}\Big).$$

Samuti

$$\frac{d\mathbf{Y}}{dx_{gh}} = \frac{d}{dx_{gh}} \otimes \mathrm{vec}\mathbf{Y} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{gh}} \\ \cdots \\ \frac{\partial y_{rs}}{\partial x_{gh}} \end{pmatrix}.$$

Nendes tähistustes

$$\frac{\partial z_{kl}}{\partial x_{gh}} = \frac{dz_{kl}}{d\mathbf{Y}} \frac{d\mathbf{Y}}{dx_{gh}}.$$

Kuna see võrdus kehtib iga indeksite väärtuste komplekti korral ja

$$\frac{\partial z_{kl}}{\partial x_{gh}} = \Big(\frac{d\mathbf{Z}}{d\mathbf{X}}\Big)_{(l,k)(h,g)}, \text{ siis } \frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}} \frac{d\mathbf{Y}}{d\mathbf{X}}.$$

(vi) Olgu $\mathbf{Y} = \mathbf{AXB}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A}).$$

Tõestus. Definitsiooni kohaselt

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}(\mathbf{AXB}) = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes [(\mathbf{B}' \otimes \mathbf{A})\mathrm{vec}\mathbf{X}].$$

Omaduse (iii) tõttu väide kehtib

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A}).$$

□

(vii) Olgu $\mathbf{Z} = \mathbf{AYB}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid ja $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A})\frac{d\mathbf{Y}}{d\mathbf{X}}.$$

Tõestus. Omaduse tõestamisel rakendame järjest omadusi (v) ja (vi).

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}}\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d(\mathbf{AYB})}{d\mathbf{Y}}\frac{d\mathbf{Y}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A})\frac{d\mathbf{Y}}{d\mathbf{X}}.$$

□

(viii) Olgu $\mathbf{X} : p \times q$, siis

$$\frac{d\mathbf{X}'}{d\mathbf{X}} = \mathbf{K}_{p,q}, \qquad \frac{d\mathbf{X}}{d\mathbf{X}'} = \mathbf{K}_{q,p}.$$

Tõestus. Kasutame vec-operaatori omadusi

$$\mathrm{vec}(\mathbf{A})' = \mathbf{K}_{p,q}\mathrm{vec}\mathbf{A};$$

$$\mathrm{vec}\mathbf{A} = \mathbf{K}_{q,p}\mathrm{vec}(\mathbf{A})'.$$

Siis
$$\frac{d\mathbf{X}'}{d\mathbf{X}} = \frac{d}{d\mathbf{X}} \otimes (\mathbf{K}_{p,q}\mathrm{vec}\mathbf{X}) = \mathbf{K}_{p,q}$$
omaduse (iii) kohaselt.

Teine väide
$$\frac{d\mathbf{X}}{d\mathbf{X}'} = \frac{d}{d(\mathrm{vec}\mathbf{X}')'} \otimes \mathrm{vec}\mathbf{X} = \frac{d}{d(\mathbf{K}_{p,q}\mathrm{vec}\mathbf{X})'} \otimes \mathrm{vec}\mathbf{X}$$
$$= \frac{d}{d(\mathrm{vec}'\mathbf{X})\mathbf{K}_{q,p}} \otimes \mathrm{vec}\mathbf{X} = \frac{d}{d\mathrm{vec}'\mathbf{X}}\mathbf{K}_{q,p} \otimes \mathrm{vec}\mathbf{X} \cdot 1$$
$$= \Big(\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{X}\Big)(\mathbf{K}_{q,p} \otimes 1) = \mathbf{K}_{q,p}.$$
□

(ix) Kui $\mathbf{X}$ on $p \times p$-maatriks, siis
$$\frac{d\mathbf{X}_d}{d\mathbf{X}} = (\mathbf{K}_{p,p})_d.$$
Võrduse saame vahetult definitsiooni rakendades.

(x) Kui $m \times n$-maatriks $\mathbf{W} = \mathbf{W}(\mathbf{Y}, \mathbf{Z})$ on $r \times s$-maatriksi $\mathbf{Y}$ ja $k \times l$-maatriksi $\mathbf{Z}$ funktsioon, kusjuures $\mathbf{Y}$ ja $\mathbf{Z}$ on $\mathbf{X} : p \times q$ funktsioonid, siis
$$\frac{d\mathbf{W}}{d\mathbf{X}} = \frac{d\mathbf{W}}{d\mathbf{Y}}\bigg|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}} + \frac{d\mathbf{W}}{d\mathbf{Z}}\bigg|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}}.$$
See omadus on kõige üldisem vaadeldavatest omadustest.

Tõestus. Näitame, et matemaatilisest analüüsist tuntud mitmemuutuja funktsiooni diferentseerimise eeskiri on sellisel maatrikskujul esitatav. Kõige lihtsamal kujul oli valem järgmine. Olgu meil funktsioon $w(y, z)$, kus $y = y(x)$ ja $z = z(x)$. Siis
$$\frac{dw}{dx} = \frac{dw}{dy}\bigg|_{z=const} \frac{dy}{dx} + \frac{dw}{dz}\bigg|_{y=const} \frac{dz}{dx}.$$

Kui $w = w(\mathbf{y}, \mathbf{z})$, kus $\mathbf{y} = (y_1, \dots, y_n)'$, $y_i = y_i(x)$ ja $\mathbf{z} = (z_1, \dots, z_m)'$, $z_j = z_j(x)$, siis

$$\frac{dw}{dx} = \sum_{i=1}^{n} \frac{\partial w}{\partial y_i}\Big|_{\mathbf{z}=const} \frac{\partial y_i}{\partial x} + \sum_{j=1}^{m} \frac{\partial w}{\partial z_j}\Big|_{\mathbf{y}=const} \frac{\partial z_j}{\partial x}.$$

Selle diferentseerimiseeskirja järgi osatuletis

$$\frac{\partial w_{ij}}{\partial x_{gh}} = \sum_{u=1}^{r}\sum_{v=1}^{s} \frac{\partial w_{ij}}{\partial y_{uv}}\Big|_{\mathbf{Z}=const} \frac{\partial y_{uv}}{\partial x_{gh}} + \sum_{u=1}^{k}\sum_{v=1}^{l} \frac{\partial w_{ij}}{\partial z_{uv}}\Big|_{\mathbf{Y}=const} \frac{\partial z_{uv}}{\partial x_{gh}}.$$

Maatrikstuletise definitsiooni kohaselt

$$\frac{dw_{ij}}{d\mathbf{Y}} = \left(\frac{\partial w_{ij}}{\partial y_{11}}, \cdots, \frac{\partial w_{ij}}{\partial y_{rs}}\right),$$

$$\frac{dw_{ij}}{d\mathbf{Z}} = \left(\frac{\partial w_{ij}}{\partial z_{11}}, \cdots, \frac{\partial w_{ij}}{\partial z_{kl}}\right)$$

ning

$$\frac{d\mathbf{Y}}{dx_{gh}} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{gh}} \\ \cdots \\ \frac{\partial y_{rs}}{\partial x_{gh}} \end{pmatrix}, \qquad \frac{d\mathbf{Z}}{dx_{gh}} = \begin{pmatrix} \frac{\partial z_{11}}{\partial x_{gh}} \\ \cdots \\ \frac{\partial z_{kl}}{\partial x_{gh}} \end{pmatrix}.$$

Seega

$$\frac{\partial w_{ij}}{\partial x_{gh}} = \frac{dw_{ij}}{d\mathbf{Y}}\Big|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{dx_{gh}} + \frac{dw_{ij}}{d\mathbf{Z}}\Big|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{dx_{gh}}.$$

Kuna see võrdus kehtib iga indeksite i, j, g, h kombinatsiooni korral ja

$$\left(\frac{d\mathbf{W}}{d\mathbf{X}}\right)_{j,i)(h,g)} = \frac{\partial w_{ij}}{\partial x_{gh}},$$

on väide tõestatud.

□

(xi) Korrutise tuletis. Kui $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$ ja $\mathbf{Z} = \mathbf{Z}(\mathbf{X})$, kus $\mathbf{Y} : r \times s$, $\mathbf{Z} : s \times n$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d(\mathbf{YZ})}{d\mathbf{X}} = (\mathbf{Z}' \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathbf{I}_n \otimes \mathbf{Y})\frac{d\mathbf{Z}}{d\mathbf{X}}.$$

Tõestus. Kasutame eelmist omadust

$$\frac{d(\mathbf{YZ})}{d\mathbf{X}} = \left.\frac{d(\mathbf{YZ})}{d\mathbf{Y}}\right|_{\mathbf{Z}=const}\frac{d\mathbf{Y}}{d\mathbf{X}} + \left.\frac{d(\mathbf{YZ})}{d\mathbf{Z}}\right|_{\mathbf{Y}=const}\frac{d\mathbf{Z}}{d\mathbf{X}}.$$

Omaduse (vii) abil saame väite lisades sobivalt ühikmaatriksid

$$\frac{d(\mathbf{YZ})}{d\mathbf{X}} = (\mathbf{Z}' \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathbf{I}_n \otimes \mathbf{Y})\frac{d\mathbf{Z}}{d\mathbf{X}}.$$

□

(xii) Kui $\mathbf{Y}$ on $r \times r$-maatriks, siis

$$\frac{d\mathbf{Y}^n}{d\mathbf{X}} = \Big(\sum_{i+j=n-1,\ i,j\geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j\Big)\frac{d\mathbf{Y}}{d\mathbf{X}},$$

kus $\mathbf{Y}^0 = \mathbf{I}_r,\ n \geq 1$.

Tõestus. Kasutame matemaatilise induktsiooni meetodit. Kui $n = 1$, siis ainus võimalus on $i = j = 0$,

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = (\mathbf{I}_r \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathbf{Y}}{d\mathbf{X}},$$

väide kehtib.

Kui $n = 2$, siis eelmise omaduse põhjal

$$\begin{aligned}\frac{d\mathbf{Y}^2}{d\mathbf{X}} = \frac{d(\mathbf{YY})}{d\mathbf{X}} &= (\mathbf{Y}' \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathbf{I}_r \otimes \mathbf{Y})\frac{d\mathbf{Y}}{d\mathbf{X}} \\ &= \Big[\sum_{i+j=1,\ i,j\geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j\Big]\frac{d\mathbf{Y}}{d\mathbf{X}}.\end{aligned}$$

Valem kehtib.

Oletame, et valem kehtib $n = k$ korral, näitame, et siis kehtib ka $k+1$ korral. Seega teame

$$\frac{d\mathbf{Y}^k}{d\mathbf{X}} = \Big(\sum_{i+j=k-1,\ i,j\geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j \Big) \frac{d\mathbf{Y}}{d\mathbf{X}}.$$

Siis rakendades omadust (x), saame

$$\frac{d\mathbf{Y}^{(k+1)}}{d\mathbf{X}} = \frac{d(\mathbf{Y}\mathbf{Y}^k)}{d\mathbf{X}} = ((\mathbf{Y}')^k \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathbf{I}_r \otimes \mathbf{Y})\frac{d\mathbf{Y}^k}{d\mathbf{X}} =$$

$$= ((\mathbf{Y}')^k \otimes \mathbf{I}_r)\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathbf{I}_r \otimes \mathbf{Y})\Big(\sum_{i+j=k-1,\ i,j\geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j \Big)\frac{d\mathbf{Y}}{d\mathbf{X}} =$$

$$= \Big(\sum_{i+j=k,\ i,j\geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j \Big)\frac{d\mathbf{Y}}{d\mathbf{X}}.$$

□

(xiii) Olgu $p \times p$-maatriks $\mathbf{X}$ pööratav, siis

$$\frac{d\mathbf{X}^{-1}}{d\mathbf{X}} = -(\mathbf{X}')^{-1} \otimes \mathbf{X}^{-1}.$$

Tõestus. Lähtume samasusest

$$\mathbf{X}^{-1}\mathbf{X} = \mathbf{I}_p.$$

Rakendame korrutise tuletist (omadus (xi)).

$$\frac{d(\mathbf{X}^{-1}\mathbf{X})}{d\mathbf{X}} = (\mathbf{X}' \otimes \mathbf{I}_p)\frac{d(\mathbf{X}^{-1})}{d\mathbf{X}} + (\mathbf{I}_p \otimes \mathbf{X}^{-1})\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{0}.$$

Viime teise liidetava paremale poole.

$$(\mathbf{X}' \otimes \mathbf{I}_p)\frac{d(\mathbf{X}^{-1})}{d\mathbf{X}} = -(\mathbf{I}_p \otimes \mathbf{X}^{-1}).$$

Korrutame vasakult mõlemaid pooli teguriga $(\mathbf{X}' \otimes \mathbf{I}_p)^{-1}$

$$\frac{d(\mathbf{X}^{-1})}{d\mathbf{X}} = -(\mathbf{X}' \otimes \mathbf{I}_p)^{-1}(\mathbf{I}_p \otimes \mathbf{X}^{-1}) = -(\mathbf{X}')^{-1} \otimes \mathbf{X}^{-1}.$$

Kasutasime viimase võrduse saamiseks otsekorrutise omadust (v).

□

(xiv) Kui $\mathbf{Y}$ on pööratav $r \times r$-maatriks, siis

$$\frac{d\mathbf{Y}^{-n}}{d\mathbf{X}} = -\Big(\sum_{i+j=n-1,\ i,j\geq 0} (\mathbf{Y}')^{-i-1} \otimes \mathbf{Y}^{-j-1}\Big)\frac{d\mathbf{Y}}{d\mathbf{X}},$$

kus $\mathbf{Y}^0 = \mathbf{I}_r,\ n \geq 1$.

Tõestus. Tõestus põhineb kahel eelmisel omadusel. Tähistame $\mathbf{Z} = \mathbf{Y}^{-1}$, siis

$$\frac{d\mathbf{Y}^{-n}}{d\mathbf{X}} = \frac{d\mathbf{Z}^n}{d\mathbf{X}} = \Big(\sum_{i+j=n-1,\ i,j\geq 0} (\mathbf{Z}')^{i} \otimes \mathbf{Z}^{j}\Big)\frac{d\mathbf{Y}^{-1}}{d\mathbf{X}}$$

$$= \Big(\sum_{i+j=n-1,\ i,j\geq 0} (\mathbf{Y}')^{-i} \otimes \mathbf{Y}^{-j}\Big)(-(\mathbf{Y}')^{-1} \otimes \mathbf{Y}^{-1})\frac{d\mathbf{Y}}{d\mathbf{X}}$$

$$= -\Big(\sum_{i+j=n-1,\ i,j\geq 0} (\mathbf{Y}')^{-i-1} \otimes \mathbf{Y}^{-j-1}\Big)\frac{d\mathbf{Y}}{d\mathbf{X}}.$$

□

(xv) olgu $\mathbf{Z} : m \times n$, $\mathbf{Y} : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = (\mathbf{I}_s \otimes \mathbf{K}_{n,r} \otimes \mathbf{I}_m)\Big[(\mathbf{I}_{rs} \otimes \mathrm{vec}\mathbf{Z})\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathrm{vec}\mathbf{Y} \otimes \mathbf{I}_{mn})\frac{d\mathbf{Z}}{d\mathbf{X}}\Big].$$

Tõestus. vec-operaatori omaduse (viii) kohaselt

$$\mathrm{vec}(\mathbf{Y} \otimes \mathbf{Z}) = (\mathbf{I}_s \otimes \mathbf{K}_{n,r} \otimes \mathbf{I}_m)(\mathrm{vec}\mathbf{Y} \otimes \mathrm{vec}\mathbf{Z}).$$

Tähistame kirjapildi lihtsustamiseks

$$\mathbf{M} = (\mathbf{I}_s \otimes \mathbf{K}_{n,r} \otimes \mathbf{I}_m).$$

$\mathrm{vec}(\mathbf{Y} \otimes \mathbf{Z})$ esitusest saame

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}(\mathbf{Y} \otimes \mathbf{Z})$$

$$= \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes (\mathbf{M}(\mathrm{vec}\mathbf{Y} \otimes \mathrm{vec}\mathbf{Z})) = (1 \otimes \mathbf{M})\Big[\frac{d}{d\mathbf{X}} \otimes (\mathrm{vec}\mathbf{Y} \otimes \mathrm{vec}\mathbf{Z})\Big].$$

Rakendame maatrikstuletise omadust (x)

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = \frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{Y}}\Big|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}} + \frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{Z}}\Big|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}}$$

$$= \mathbf{M}\Big\{\frac{d}{d\mathbf{Y}} \otimes (\mathrm{vec}\mathbf{Y} \otimes \mathrm{vec}\mathbf{Z})\Big|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}}$$

$$+ \frac{d}{d\mathbf{Z}} \otimes (\mathrm{vec}\mathbf{Y} \otimes \mathrm{vec}\mathbf{Z})\Big|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}}\Big\}$$

$$= \mathbf{M}\Big\{\frac{d}{d\mathbf{Y}} \otimes ((\mathbf{I}_{rs}\mathrm{vec}\mathbf{Y}) \otimes (\mathrm{vec}\mathbf{Z} \cdot 1))\Big|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}}$$

$$+ \frac{d}{d\mathbf{Z}} \otimes ((\mathrm{vec}\mathbf{Y} \cdot 1) \otimes (\mathbf{I}_{mn}\mathrm{vec}\mathbf{Z}))\Big|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}}\Big\}$$

$$= \mathbf{M}\Big\{1 \cdot \frac{d}{d\mathbf{Y}} \otimes [(\mathbf{I}_{rs} \otimes \mathrm{vec}\mathbf{Z})\mathrm{vec}\mathbf{Y}]\Big|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}}$$

$$+ 1 \cdot \frac{d}{d\mathbf{Z}} \otimes [(\mathrm{vec}\mathbf{Y} \otimes \mathbf{I}_{mn})\mathrm{vec}\mathbf{Z}]\Big|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}}\Big\}$$

$$= \mathbf{M}\Big\{(\mathbf{I}_{rs} \otimes \mathrm{vec}\mathbf{Z})\frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathrm{vec}\mathbf{Y} \otimes \mathbf{I}_{mn})\frac{d\mathbf{Z}}{d\mathbf{X}}\Big\}.$$

□

§16. Mõned tuletise rakendused

Determinandi tuletis.

Kui $\mathbf{X}$ on pööratav $p \times p$-maatriks, siis

$$\frac{d|\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}|\mathrm{vec}'(\mathbf{X}')^{-1}.$$

Tõestus. Kasutame maatriksi determinandi esitust i-nda rea järgi.

$$|\mathbf{X}| = \sum_{j=1}^{p} x_{ij}(-1)^{i+j}|\mathbf{X}_{(ij)}|,$$

kus $|\mathbf{X}_{(ij)}|$ on x_{ij} täiendmiinor. Teisalt teame, et pöördmaatriksi element

$$(\mathbf{X}^{-1})_{ij} = \frac{(-1)^{i+j}|\mathbf{X}_{(ji)}|}{|\mathbf{X}|}.$$

Seega

$$\frac{d|\mathbf{X}|}{dx_{ij}} = (-1)^{i+j}|\mathbf{X}_{(ij)}| = (\mathbf{X}^{-1})_{ji}|\mathbf{X}| = (\mathbf{X}')^{-1}_{ij}|\mathbf{X}|.$$

Et tulemus kehtib iga indeksipaari $i, j = 1, \ldots, p$ korral ja tuletis on $(1 \times p^2)$-vektor, siis saame võrduse kui transponeerime ka paremal pool

$$\frac{d|\mathbf{X}|}{d\mathrm{vec}'\mathbf{X}} = |\mathbf{X}|\mathrm{vec}'(\mathbf{X}')^{-1}.$$

□

Jälje tuletis.

Kui $\mathbf{X}$ on $p \times p$-maatriks, siis

$$\frac{d(\mathrm{tr}\mathbf{X})}{d\mathbf{X}} = \mathrm{vec}'\mathbf{I}_p.$$

Tõestus. Esitame jälje kujul

$$\mathrm{tr}\mathbf{X} = \mathrm{tr}(\mathbf{I}_p\mathbf{X}).$$

Siis vec-operaatori omaduse kohaselt

$$\mathrm{tr}\mathbf{X} = \mathrm{vec}'\mathbf{I}_p\mathrm{vec}\mathbf{X}$$

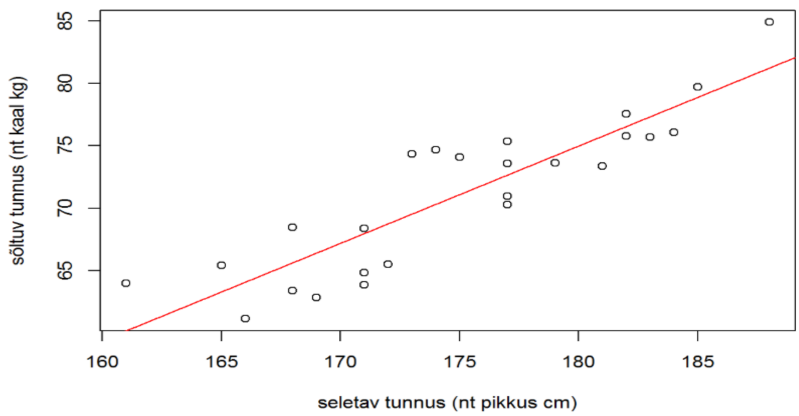
ja

$$\frac{d(\mathrm{tr}\mathbf{X})}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes (\mathrm{vec}'\mathbf{I}_p\mathrm{vec}\mathbf{X}) = \mathrm{vec}'\mathbf{I}_p$$

tuletise omaduse (iii) tõttu.

□

Vähimruutude hinnang.
Vaatame vaatlustulemuste punktiparve lähendamist sirgega. Meil on kaks tunnust, X ja Y, kusjuures meid huvitab tunnuse Y sõltuvus tunnusest X; näiteks uuritakse kaalu Y sõltuvust pikkusest X. Igale indiviidile vastab tasandil punkt (x_i, y_i), $i = 1, \ldots, n$. Tahame läbi punktiparve panna sirge, mis oleks parim vähimruutude (hälvete ruutude summa) mõttes.



Otsime sõltuvust kujul

$$Y = aX + b,$$

kusjuures vaatlustulemused saab esitada kujul

$$y_i = ax_i + b + \varepsilon_i, \;\; i = 1, \ldots, n,$$

kus ε_i on hälve sirgest. Vektorkujul saab selle lühemalt kirja panna

$$\mathbf{y} = a\mathbf{x} + b\mathbf{1}_n + \boldsymbol{\varepsilon}.$$

Vaja on leida a ja b. Eelmise mudeli saame kirja panna veel kompaktsemalt

$$\mathbf{y} = \mathbf{K}\begin{pmatrix} a \\ b \end{pmatrix} + \boldsymbol{\varepsilon},$$

kus

$$\mathbf{K} = \begin{pmatrix} x_1 & 1 \\ \cdots & \cdots \\ x_n & 1 \end{pmatrix} \qquad (16.1).$$

See on erijuht üldisemast *lineaarsest mudelist*, kus me otsime hinnanguid parameetritele $\theta_1, \ldots, \theta_k$, et avaldada uuritav tunnus Y teiste fikseeritud/seletavate tunnuste kaudu lineaarse funktsioonina (meie näites vaatame iga pikkuse korral kaalu väärtust). Siis mudel on kujul

$$\mathbf{y} = \mathbf{K}\boldsymbol{\theta} + \boldsymbol{\varepsilon},$$

kus $\mathbf{K}$ on teadaolev $n \times k$ maatriks, nn. plaanimaatriks ja $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)'$. Nõuame, et meie otsitav sirge oleks parim vähimruutude mõttes. Meie näite korral

$$D = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2 = \min$$

ehk maatrikskujul

$$D = \left(\mathbf{y} - \mathbf{K}\begin{pmatrix} a \\ b \end{pmatrix}\right)'\left(\mathbf{y} - \mathbf{K}\begin{pmatrix} a \\ b \end{pmatrix}\right) = \min,$$

Üldise lineaarse mudeli korral

$$D = (\mathbf{y} - \mathbf{K}\boldsymbol{\theta})'(\mathbf{y} - \mathbf{K}\boldsymbol{\theta}) = \min.$$

Ekstreemumi korral peab kehtima võrdus

$$\frac{dD}{d\boldsymbol{\theta}} = \mathbf{0}.$$

Leiame tuletise

$$\frac{dD}{d\boldsymbol{\theta}} = \frac{d[(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})]}{d\boldsymbol{\theta}}.$$

Rakendame korrutise tuletise omadust

$$\frac{dD}{d\boldsymbol{\theta}} = ((\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'\otimes 1)\frac{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'}{d\boldsymbol{\theta}} + (1\otimes(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})')\frac{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})}{d\boldsymbol{\theta}} =$$

$$= (\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'(-\mathbf{K}) + (\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'(-\mathbf{K}) = \mathbf{0}$$

kuna

$$\frac{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'}{d\boldsymbol{\theta}} = \frac{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'}{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})}\frac{d(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})}{d\boldsymbol{\theta}} = \mathbf{K}_{n,1}(-\mathbf{K}) = -\mathbf{K}.$$

Seega

$$(\mathbf{y}-\mathbf{K}\boldsymbol{\theta})'\mathbf{K} = \mathbf{0}, \quad \mathbf{y}'\mathbf{K} = \boldsymbol{\theta}'\mathbf{K}'\mathbf{K}.$$

Eeldame, et maatriks $\mathbf{K}$ on täisastakuga, siis $\mathbf{K}'\mathbf{K}$ on pööratav ja

$$\boldsymbol{\theta}' = \mathbf{y}'\mathbf{K}(\mathbf{K}'\mathbf{K})^{-1},$$

kust vähimruutude hinnang lineaarse mudeli parameetritele on kujul

$$\widehat{\boldsymbol{\theta}} = (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{y}.$$

Selle hinnangu leidsime projektorite paragrahvis.

§17. Kõrgemat järku tuletised.

Lisaks tuletisele on statistikas vaja ka kõrgemat järku tuletisi. Need defineeritakse rekursiivselt.

Definitsioon 17.1. *Olgu* $\mathbf{Y} = \mathbf{Y}(\mathbf{X}) : r \times s$ *ja* $\mathbf{X} : p \times q$. *Maatriksi* $\mathbf{Y}$ *k-ndat järku tuletiseks maatriksi* $\mathbf{X}$ *järgi lahtises* *piirkonnas $\mathcal{D}$ nimetatakse maatrikstuletist* $\mathbf{Y}$ $(k-1)$*-st järku tuletisest*

$$\frac{d^k\mathbf{Y}}{d\mathbf{X}^k} = \frac{d}{d\mathbf{X}}\Big(\frac{d^{k-1}\mathbf{Y}}{d\mathbf{X}^{k-1}}\Big), \quad k \geq 2,$$

kui eksisteerivad pidevad k-ndat järku osatuletised maatriksi $\mathbf{Y}$ *elementidest maatriksi* $\mathbf{X}$ *elementide järgi selles piirkonnas $\mathcal{D}$.*
Definitsiooni kohaselt k-ndat järku tuletis $\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}$ on $((pq)^{k-1} \times pq)$ maatriks. Selles on kerge veenduda järjestikku tuletisi võttes. Lähtudes definitsioonist saab tuletada $\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}$ jaoks avaldise ka ilma rekursiivse seoseta.

Lause 17.1. *Tuletise $\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}$ saab esitada kujul*

$$\frac{d^k\mathbf{Y}}{d\mathbf{X}^k} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}\mathbf{X}}}_{(k-1)\ \text{tegurit}} \otimes \mathrm{vec}\mathbf{Y}.$$

Tõestus. Tõestada saame väite matemaatilise induktsiooni teel. Kui $k = 2$, siis

$$\frac{d^2\mathbf{Y}}{d\mathbf{X}^2} = \frac{d}{d\mathbf{X}}\Big(\frac{d\mathbf{Y}}{d\mathbf{X}}\Big) = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big(\frac{d\mathbf{Y}}{d\mathbf{X}}\Big)$$
$$= \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big(\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y}\Big).$$

Rakendame otsekorrutise viimast omadust vektorite kohta

$$\mathbf{a}' \otimes \mathbf{b} = \mathbf{b}\mathbf{a}'$$

ja lisame vektorite korrutises nende vahele 1×1 maatriksi 1

$$\frac{d^2\mathbf{Y}}{d\mathbf{X}^2} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big(\mathrm{vec}\mathbf{Y} \cdot 1 \cdot \frac{d}{d\mathrm{vec}'\mathbf{X}}\Big).$$

Kasutame vec-operaatori omadust (v), $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}$

$$\frac{d^2\mathbf{Y}}{d\mathbf{X}^2} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y}.$$

Väide kehtib $k = 2$ korral.
Oletame, et kehtib ka $k = n - 1$ korral,

$$\frac{d^{n-1}\mathbf{Y}}{d\mathbf{X}^{n-1}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y},$$

kus $\frac{d}{d\mathrm{vec}\mathbf{X}}$ esineb $n-2$ korda. Näitame, et kehtib ka $k = n$ korral. Siis definitsiooni kohaselt

$$\frac{d^{n}\mathbf{Y}}{d\mathbf{X}^{n}} = \frac{d}{d\mathrm{vec}'\mathbf{X}}\Big(\frac{d^{n-1}\mathbf{Y}}{d\mathbf{X}^{n-1}}\Big) = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big(\frac{d^{n-1}\mathbf{Y}}{d\mathbf{X}^{n-1}}\Big)$$

$$= \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big(\underbrace{\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}\mathbf{X}}}_{(n-1)\text{ tegurit}} \otimes \mathrm{vec}\mathbf{Y}\Big) = \cdots$$

Kasutame jälle vektorite otsekorrutise omadust $\mathbf{a}' \otimes \mathbf{b} = \mathbf{b}\mathbf{a}'$ ja lisame vektorite vahele 1×1 maatriksi 1

$$\cdots = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\Big[\Big(\underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}\mathbf{X}}}_{(n-2)\text{ tegurit}}\Big) \otimes \mathrm{vec}\Big(\mathbf{Y} \cdot 1 \cdot \frac{d}{d\mathrm{vec}'\mathbf{X}}\Big)\Big]$$

$$= \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}\mathbf{X}}}_{(n-1)\text{ tegurit}} \otimes \mathrm{vec}\mathbf{Y}.$$

□

Kõrgemat järku tuletise puhul on oluline ka see, et tulemus ei tohi sõltuda tuletise võtmise järjekorrast, näiteks ühemõõtmelisel juhul teist järku tuletise kolmandat järku tuletis on viiendat järku tuletis. See omadus kehtib ka maatrikstuletise korral

Lause 17.2. *Olgu maatriksite dimensioonid nagu Lauses 1. Maatriksi* $\mathbf{Y}$ *k-ndat järku tuletise l-ndat järku tuletis on maatriksi* $\mathbf{Y}$ $(k+l)$*-ndat järku tuletis maatriksi* $\mathbf{X}$ *järgi.*

Tõestus. Tõestus taandub Lause 17.1 rakendamisele. Selle kohaselt

$$\frac{d^k\mathbf{Y}}{d\mathbf{X}^k} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}}}_{(k-1)\ \text{tegurit}} \otimes \mathrm{vec}\mathbf{Y}.$$

Siis

$$\frac{d^l\mathbf{Y}}{d\mathbf{X}^l}\Big(\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}\Big) = \underbrace{\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}}}_{l\ \text{tegurit}} \otimes \mathrm{vec}\Big(\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}\Big).$$

Kuna kasutades jälle vektorite otsekorrutise omadust

$$\mathrm{vec}\Big(\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}\Big) = \mathrm{vec}\Big(\frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}}}_{(k-1)\ \text{tegurit}} \otimes \mathrm{vec}\mathbf{Y}\Big)$$

$$= \mathrm{vec}\Big(\underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}}}_{(k-1)\ \text{tegurit}} \otimes \mathrm{vec}\Big(\mathbf{Y} \cdot 1 \cdot \frac{d}{d\mathrm{vec}'\mathbf{X}}\Big)\Big)$$

$$= \Big(\underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}'\mathbf{X}}}_{k\ \text{tegurit}} \otimes \mathrm{vec}\mathbf{Y}\Big)\mathrm{vec}1.$$

Saame tulemuseks

$$\frac{d^l\mathbf{Y}}{d\mathbf{X}^l}\Big(\frac{d^k\mathbf{Y}}{d\mathbf{X}^k}\Big) = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \underbrace{\frac{d}{d\mathrm{vec}\mathbf{X}} \otimes \cdots \otimes \frac{d}{d\mathrm{vec}\mathbf{X}}}_{(l-1)+k\ \text{tegurit}} \otimes \mathrm{vec}\mathbf{Y}.$$

□

Märkus. Erijuhul, kui $k = 2$ nimetatakse kirjanduses teist järku tuletist tihti Hessi maatriksiks. Statistikas kasutatakse seda informatsioonimaatriksi leidmisel. Seda läheb vaja hüpoteeside kontrollimisel ja regressioonanalüüsis. Informatsioonimaatriks leitakse pidevate jaotuste korral kui keskväärtus logaritmilise tõepärafunktsiooni teist järku maatrikstuletisest.

§18. Taylori rida

Matemaatilisest analüüsist on teile teada Taylori rea mõiste. Juhul kui funktsioon $f(x)$, $f : \mathbb{R} \to \mathbb{R}$, on $(m+1)$ korda diferentseeruv, saab selle funktsiooni kohal x_0 esitada reana

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{1}{2} f^{(2)}(x_0)(x - x_0)^2 + \cdots$$

$$+ \frac{1}{m!} f^{(m)}(x_0)(x - x_0)^m + R_m(x),$$

kus $f^{(k)}(x_0)$ on funktsiooni $f(x)$ k-s tuletis kohal x_0 ja jääkliige

$$R_m(x) = \frac{1}{(m+1)!} f^{(m+1)}(x^*)(x - x_0)^m; \quad x^* = x_0 + \theta(x - x_0),$$

kus $0 < \theta < 1$. Mitmemuutuja funktsiooni korral jääb põhimõtteliselt kõik samaks. Olgu

$$f : \mathbb{R}^p \to \mathbb{R}, \quad \mathbf{x} = (x_1, \ldots, x_p)', \ \mathbf{x}_0 = (x_1^0, \ldots, x_p^0)'.$$

Siis kui kõik $(m+1)$-st järku osatuletised on pidevad piirkonnas $\mathcal{D}$ ja $\mathbf{x}, \ \mathbf{x}_0 \in \mathcal{D}$

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^{p} \frac{\partial f(\mathbf{x})}{\partial x_i}\Big|_{\mathbf{x}=\mathbf{x}_0} (x_i - x_i^0)$$

$$+ \frac{1}{2!} \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}\Big|_{\mathbf{x}=\mathbf{x}_0} (x_i - x_i^0)(x_j - x_j^0) + \cdots$$

$$+ \frac{1}{m!} \sum_{i_1, \ldots, i_m = 1}^{p} \frac{\partial^m f(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_m}}\Big|_{\mathbf{x}=\mathbf{x}_0} (x_{i_1} - x_{i_1}^0) \cdots (x_{i_m} - x_{i_m}^0) + R_m(\mathbf{x}),$$

kus jääkliige

$$R_m(\mathbf{x}) = \frac{1}{(m+1)!} \sum_{i_1,\ldots,i_{m+1}=1}^{p} \frac{\partial^{(m+1)} f(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_{m+1}}}\Big|_{\mathbf{x}=\mathbf{x}^*} (x_{i_1} - x^0_{i_1}) \times \cdots$$

$$\times (x_{i_{m+1}} - x^0_{i_{m+1}})$$

ja $\mathbf{x}^* \in \mathcal{D}$. Kompaktsemalt saame selle kirja panna kujul

$$f(\mathbf{x}) = f(\mathbf{x}_0) +$$

$$+ \sum_{k=1}^{m} \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^{p} \frac{\partial^k f(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_k}}\Big|_{\mathbf{x}=\mathbf{x}_0} (x_{i_1} - x^0_{i_1}) \cdots (x_{i_k} - x^0_{i_k}) + R_m(\mathbf{x}). \tag{18.1}$$

Meil on vaja minna üks samm edasi. Mis saab siis, kui funktsioon $\mathbf{f} : \mathbb{R}^p \to \mathbb{R}^q$?
Siis saame esitada $\mathbf{f}$ koordinaatfunktsioonide vektorina

$$\mathbf{f} = \begin{pmatrix} f_1 \\ \cdots \\ f_q \end{pmatrix}, \qquad f_i : \mathbb{R}^p \to \mathbb{R},$$

kus iga koordinaatfunktsioon on esitatav Taylori reana (18.1).
Oleks hea, kui saaks kogu selle keerulise avaldise, mis sisaldab summasid ja indekseid, esitada kompaktsel maatrikskujul. See on võimalik. Esitame Taylori rea maatrikskujul järgmise teoreemina.
Teoreem 18.1. *Kui funktsioon* $\mathbf{f} : \mathbb{R}^p \to \mathbb{R}^q$ *omab pidevaid (m+1) järku osatuletisi punkti* $\mathbf{x}_0$ *ümbruses* $\mathcal{D}$*, siis funktsioon* $\mathbf{f}(\mathbf{x})$ *on arendatav Taylori ritta kohal* $\mathbf{x}_0$*,* $\mathbf{x} = (x_1, \ldots, x_p)'$*,* $\mathbf{x}_0 = (x^0_1, \ldots, x^0_p)'$

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \sum_{k=1}^{m} \frac{1}{k!} \left[(\mathbf{x}-\mathbf{x}_0)^{\otimes(k-1)} \otimes \mathbf{I}_q\right]' \frac{d^k \mathbf{f}(\mathbf{x})}{d\mathbf{x}^k}\Big|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x}-\mathbf{x}_0) + R_m(\mathbf{x}),$$

kus

$$R_m(\mathbf{x}) = \frac{1}{(m+1)!}\left[(\mathbf{x}-\mathbf{x}_0)^{\otimes m}\otimes\mathbf{I}_q\right]'\frac{d^{m+1}\mathbf{f}(\mathbf{x})}{d\mathbf{x}^{m+1}}\bigg|_{\mathbf{x}=\mathbf{x}^*}(\mathbf{x}-\mathbf{x}_0),\ \mathbf{x}^*\in\mathcal{D}.$$

Rakenduste jaoks on kõige olulisemad rea kaks esimest liiget.

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\bigg|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)$$

$$+\frac{1}{2}\left[(\mathbf{x}-\mathbf{x}_0)\otimes\mathbf{I}_q\right]'\frac{d^2\mathbf{f}(\mathbf{x})}{d\mathbf{x}^2}\bigg|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)+\cdots$$

Tõestus. Tõestuseks tuleks näidata, et funktsiooni $\mathbf{f}(\mathbf{x})$ i-nda koordinaadi $f_i(\mathbf{x})$ jaoks teoreemi väitest välja kirjutatav avaldis on kujul (18.1). Selleks peab k-nda liidetava korral kehtima võrdus, $k = 1, 2, \ldots$

$$\left\{\left[(\mathbf{x}-\mathbf{x}_0)^{\otimes(k-1)}\otimes\mathbf{I}_q\right]'\frac{d^k\mathbf{f}(\mathbf{x})}{d\mathbf{x}^k}\bigg|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)\right\}_i =$$

$$= \sum_{i_1,\ldots,i_k=1}^{p}\frac{\partial^k f_i(\mathbf{x})}{\partial x_{i_1}\cdots\partial x_{i_k}}\bigg|_{\mathbf{x}=\mathbf{x}_0}(x_{i_1}-x_{i_1}^0)\cdots(x_{i_k}-x_{i_k}^0).$$

Tehniliselt on selle näitamine tülikas, tuleb välja kirjutada maatriksid elementide kaupa ja näidata, et nende korrutamisel saadakse avaldis sellisel kujul nagu liidetav valemis (18.1). Me ei hakka seda üldkujul läbi tegema. (Koos vajaliku lemmaga võtab see enda alla neli lehekülge tihedaid teisendusi). Näitame, et see võrdus kehtib esimest tuletist sisaldava liidetava jaoks.

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\bigg|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_q(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_q(\mathbf{x})}{\partial x_p} \end{pmatrix}_{(\mathbf{x}-\mathbf{x}_0)} \begin{pmatrix} x_1 - x_1^0 \\ \cdots \\ x_p - x_p^0 \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^p \frac{\partial f_1(\mathbf{x})}{\partial x_i}\Big|_{(\mathbf{x}-\mathbf{x}_0)}(x_i - x_i^0) \\ \cdots \\ \sum_{i=1}^p \frac{\partial f_q(\mathbf{x})}{\partial x_i}\Big|_{(\mathbf{x}-\mathbf{x}_0)}(x_i - x_i^0) \end{pmatrix}.$$

Oleme saanud kõigi koordinaatide korral Taylori rea (18.1) esimest tuletist sisaldava liikme avaldise. Järgmiste liidetavate korral on vaja järgida põhimõtteliselt sama skeemi.

□

Kui me vektoriseerime vasakut ja paremat poolt, siis vasak pool ei muutu, paremal saame aga Taylori rea mõnevõrra teisel kujul.

Järeldus 18.1.1. *Teoreemi eeldustel on funktsiooni* $\mathbf{f} : \mathbb{R}^p \to \mathbb{R}^q$ *Taylori rida kohal* $\mathbf{x}_0$ *esitatav kujul*

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \sum_{k=1}^{m} \frac{1}{k!} \left[(\mathbf{x}-\mathbf{x}_0)^{\otimes k} \otimes \mathbf{I}_q\right]' \mathrm{vec} \frac{d^k \mathbf{f}(\mathbf{x})}{d\mathbf{x}^k}\Big|_{\mathbf{x}=\mathbf{x}_0} + R_m(\mathbf{x}).$$

Selle esituse saamiseks rakendame teoreemi 18.1 avaldises liidetavatele vec-operaatori omadust (v)

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}.$$

Tähtsal erijuhul, kui $f : \mathbb{R}^p \to \mathbb{R}$ Taylori rea kuju lihtsustub.

Järeldus 18.1.2. *Teoreemi eeldustel on funktsiooni* $f : \mathbb{R}^p \to \mathbb{R}$ *Taylori rida punktis* $\mathbf{x}_0$ *kujul*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{k=1}^{m} \frac{1}{k!} \left[(\mathbf{x}-\mathbf{x}_0)^{\otimes(k-1)}\right]' \frac{d^k f(\mathbf{x})}{d\mathbf{x}^k}\Big|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x}-\mathbf{x}_0) + R_m(\mathbf{x}),$$

kus

$$R_m(\mathbf{x}) = \frac{1}{(m+1)!} \left[(\mathbf{x}-\mathbf{x}_0)^{\otimes m}\right]' \frac{d^{m+1} f(\mathbf{x})}{d\mathbf{x}^{m+1}}\Big|_{\mathbf{x}=\mathbf{x}^*} (\mathbf{x}-\mathbf{x}_0),\ \mathbf{x}* \in \mathcal{D}.$$

Sel juhul väline kirjapilt ei ole oluliselt keerulisem Taylori rea kujust ühemõõtmelisel juhul. Rakendades vec-operaatorit saab Taylori rida sel korral ühemõõtmelisele juhule veelgi sarnasema kuju

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{k=1}^{m} \frac{1}{k!} \mathrm{vec}' \Big(\frac{d^k f(\mathbf{x})}{d\mathbf{x}^k}\Big|_{\mathbf{x}=\mathbf{x}_0} \Big) \left[(\mathbf{x}-\mathbf{x}_0)^{\otimes k}\right] + R_m(\mathbf{x}).$$

Pärast vektoriseerimist kasutasime siin veel vektorite korrutamisel kommutatiivsust, $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$.
Vaja oleks üle kanda Taylori rea maatriksesitus ka juhule, kui vektori $\mathbf{x}$ asemel on maatriks $\mathbf{X}$ ja vektorfunktsiooni $\mathbf{f}$ asemel maatriksfunktsioon $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$. Selle üldise juhu saame taandada vektoritele, kui vektoriseerime maatriksid tuletisi sisaldavates liikmetes ja siis saame rakendada teoreemi 18.1. Esitame üldjuhu järgmise teoreemina.
Teoreem 18.2. *Kui maatriksfunktsioon* $\mathbf{Y} : \mathbb{R}^{p\times q} \to \mathbb{R}^{r\times s}$ *omab pidevaid (m+1) järku osatuletisi maatriksi* $\mathbf{X}_0$ *ümbruses* $\mathcal{D}$*, siis funktsioon* $\mathbf{Y}(\mathbf{X})$ *on arendatav pärast vektoriseerimist Taylori ritta kohal* $\mathbf{X}_0$*,*

$$\mathbf{Y}(\mathbf{X}) = \mathbf{Y}(\mathbf{X}_0)$$

$$+\sum_{k=1}^{m} \frac{1}{k!}\left[(\mathrm{vec}(\mathbf{X}-\mathbf{X}_0))^{\otimes(k-1)} \otimes \mathbf{I}_{rs}\right]' \frac{d^k\mathbf{Y}(\mathbf{X})}{d\mathbf{X}^k}\bigg|_{\mathbf{X}=\mathbf{X}_0}$$

$$\times \mathrm{vec}(\mathbf{X}-\mathbf{X}_0) + R_m(\mathrm{vec}\mathbf{X}),$$

kus

$$R_m(\mathrm{vec}\mathbf{X}) = \frac{1}{(m+1)!}\left[(\mathrm{vec}(\mathbf{X}-\mathbf{X}_0))^{\otimes m} \otimes \mathbf{I}_{rs}\right]'$$

$$\times \frac{d^{m+1}\mathbf{Y}(\mathbf{X})}{d\mathbf{X}^{m+1}}\bigg|_{\mathbf{X}=\mathbf{X}^*} \mathrm{vec}(\mathbf{X}-\mathbf{X}_0), \ \ \mathrm{vec}\mathbf{X}^* \in \mathcal{D}.$$

Ka siin on kõige olulisem erijuht $Y : \mathbb{R}^{p\times q} \to \mathbb{R}$. Näiteks kui $p = q$ ja funktsiooniks Y on maatriksi $\mathbf{X}$ determinant või tema jälg. Sel juhul on Taylori rea kuju lihtsam.
Järeldus 18.2.1. *Funktsioon* $Y : \mathbb{R}^{p\times q} \to \mathbb{R}$ *on teoreemi eelduste täidetuse korral arendatav Taylori ritta kohal* $\mathbf{X}_0$ *kujul*

$$Y(\mathbf{X}) = Y(\mathbf{X}_0)+$$

$$+\sum_{k=1}^{m} \frac{1}{k!}\left[(\mathrm{vec}(\mathbf{X}-\mathbf{X}_0))^{\otimes(k-1)}\right]' \frac{d^kY(\mathbf{X})}{d\mathbf{X}^k}\bigg|_{\mathbf{X}=\mathbf{X}_0} \mathrm{vec}(\mathbf{X}-\mathbf{X}_0)+\cdots .$$

Miks on statistikas Taylori reaksarendusi vaja? Tavaliselt vajatakse vaid esimest või kaht esimest tuletist sisaldavat liiget selleks, et lähendada keerulise või tundmatu jaotusega valimi funktsiooni jaotust.

Näide 1. Arendada Taylori ritta kohal $\mathbf{X} = \mathbf{I}_p$ funktsioon $Y = \frac{1}{|\mathbf{X}|}$. Leida esimest ja teist järku tuletistega liikmed.
Lahendus.

$$Y = 1 + \frac{dY}{d\mathbf{X}}\Big|_{\mathbf{X}=\mathbf{I}_p} \mathrm{vec}(\mathbf{X}-\mathbf{I}_p) + \frac{1}{2}\mathrm{vec}'(\mathbf{X}-\mathbf{I}_p)\frac{d^2Y}{d\mathbf{X}^2}\Big|_{\mathbf{X}=\mathbf{I}_p}\mathrm{vec}(\mathbf{X}-\mathbf{I}_p) + \cdots$$

Leiame esimese tuletise

$$\frac{dY}{d\mathbf{X}} = \frac{d|\mathbf{X}|^{-1}}{d\mathbf{X}} = \frac{d|\mathbf{X}|^{-1}}{d|\mathbf{X}|}\frac{d|\mathbf{X}|}{d\mathbf{X}}$$

$$= -\frac{1}{|\mathbf{X}|^2}|\mathbf{X}|\mathrm{vec}'(\mathbf{X}^{-1})' = -\frac{1}{|\mathbf{X}|}\mathrm{vec}'(\mathbf{X}^{-1})'.$$

Leidsime determinandi tuletise eelmises loengus. Kohal $\mathbf{X} = \mathbf{I}_p$

$$\frac{dY}{d\mathbf{X}}\Big|_{\mathbf{X}=\mathbf{I}_p} = -\mathrm{vec}'\mathbf{I}_p.$$

Teist järku tuletise leidmisel rakendame korrutise tuletist

$$\frac{d^2Y}{d\mathbf{X}^2} = \frac{d}{d\mathbf{X}}\Big(\frac{dY}{d\mathbf{X}}\Big) = \frac{d}{d\mathbf{X}}(-\frac{1}{|\mathbf{X}|}\mathrm{vec}'(\mathbf{X}^{-1})')$$

$$= -(\mathrm{vec}(\mathbf{X}^{-1})' \otimes 1)\frac{d|\mathbf{X}|^{-1}}{d\mathbf{X}} - (\mathbf{I}_{p^2} \otimes \frac{1}{|\mathbf{X}|})\frac{d\mathrm{vec}'(\mathbf{X}^{-1})'}{d\mathbf{X}}.$$

Leiame eraldi viimase tuletise.

$$\frac{d\mathrm{vec}'(\mathbf{X}^{-1})'}{d\mathbf{X}} = \frac{d(\mathbf{X}^{-1})'}{d\mathbf{X}} = \frac{d(\mathbf{X}^{-1})'}{d\mathbf{X}^{-1}}\frac{d(\mathbf{X}^{-1})}{d\mathbf{X}} = -\mathbf{K}_{p,p}((\mathbf{X}^{-1})'\otimes\mathbf{X}^{-1}).$$

Seega teist järku tuletis on

$$\frac{d^2Y}{d\mathbf{X}^2} = -\mathrm{vec}(\mathbf{X}^{-1})'\frac{d|\mathbf{X}|^{-1}}{d\mathbf{X}} + \frac{1}{|\mathbf{X}|}\mathbf{K}_{p,p}((\mathbf{X}^{-1})' \otimes \mathbf{X}^{-1})$$

$$= \frac{1}{|\mathbf{X}|}[\mathrm{vec}(\mathbf{X}^{-1})'\mathrm{vec}'(\mathbf{X}^{-1})' + \mathbf{K}_{p,p}((\mathbf{X}^{-1})' \otimes \mathbf{X}^{-1})].$$

Viimase võrduse saamisel kasutasime ära varem leitud esimese tuletise avaldise. Kohal $\mathbf{X} = \mathbf{I}_p$

$$\left.\frac{d^2Y}{d\mathbf{X}^2}\right|_{\mathbf{X}=\mathbf{I}_p} = \mathrm{vec}\mathbf{I}_p\mathrm{vec}'\mathbf{I}_p + \mathbf{K}_{p,p}.$$

Kirjutame välja reaksarenduse asendades tuletiste avaldised

$$Y = 1 - \mathrm{vec}'\mathbf{I}_p\mathrm{vec}(\mathbf{X} - \mathbf{I}_p)$$

$$+\frac{1}{2}\mathrm{vec}'(\mathbf{X} - \mathbf{I}_p)(\mathrm{vec}\mathbf{I}_p\mathrm{vec}'\mathbf{I}_p + \mathbf{K}_{p,p})\mathrm{vec}(\mathbf{X} - \mathbf{I}_p) + \cdots$$

Lihtsustame kasutades maatriksi jälje ja vec-operaatori vahelist seost.

$$Y = 1 - \mathrm{tr}\mathbf{X} + p + \frac{1}{2}[\mathrm{vec}'(\mathbf{X} - \mathbf{I}_p)\mathrm{vec}\mathbf{I}_p\mathrm{vec}'\mathbf{I}_p\mathrm{vec}(\mathbf{X} - \mathbf{I}_p)$$

$$+\mathrm{vec}'(\mathbf{X} - \mathbf{I}_p)\mathbf{K}_{p,p}\mathrm{vec}(\mathbf{X} - \mathbf{I}_p)] + \cdots$$

$$= 1 + p - \mathrm{tr}\mathbf{X} + \frac{1}{2}[(\mathrm{tr}(\mathbf{X} - \mathbf{I}_p))^2 + \mathrm{tr}(\mathbf{X} - \mathbf{I}_p)^2] + \cdots .$$

Viimase võrduse saamisel saime kasutada maatriksi jälje omadust $\mathrm{tr}\mathbf{A} = \mathrm{tr}\mathbf{A}'$.

Märkus. Kui $\mathbf{X}$ elemendid on juhuslikud suurused ja Taylori rea liikmed kahanevad, siis rea esimeste liikmete abil saaksime lähendada pöördmaatriksi determinandi jaotust jälje jaotusega, mis on determinandist tunduvalt lihtsam funktsioon. Selline olukord ongi juhul, kui $\mathbf{X}$ on näiteks valimi kovariatsioonimaatriks.

Näide 2. Olgu meil valim $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ normaaljaotusest $\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$, $\mathbf{x} = (X_1, \ldots, X_p)$. Valimi kovariatsioonimaatriks

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})',$$

kus $\overline{\mathbf{x}}$ on valimi keskmine. Valimi üldistatud dispersioon on $|\mathbf{S}|$. Vaatleme hüpoteese

$$H_0 : \mathbf{\Sigma} = \mathbf{I}_p;$$

$$H_1 : \mathbf{\Sigma} \neq \mathbf{I}_p.$$

Nullhüpotees tähendab, et $\mathbf{x}$ koordinaadid X_i on sõltumatud ja ühikulise dispersiooniga. Oletame, et nullhüpotees kehtib. Determinandi jaotust me nullhüpoteesi kehtimisel ei tea, arendame $|\mathbf{S}|$ Taylori ritta kohal $\mathbf{S} = \mathbf{I}_p$. Valimi kovariatsioonimaatriksi funktsioonide korral on Taylori rea teist tuletist sisaldav liige kõrgemat järku väike võrreldes esimese liidetavaga ja järgmised liikmed on kahanevad. Seetõttu kasutame valimi üldistatud dispersiooni $|\mathbf{S}|$ lähendamiseks vaid esimest tuletist sisaldavat liiget.

$$|\mathbf{S}| = |\mathbf{I}_p| + \frac{d|\mathbf{S}|}{d\mathbf{S}}\bigg|_{\mathbf{S}=\mathbf{I}_p} \mathrm{vec}(\mathbf{S} - \mathbf{I}_p) + \cdots$$

$$= 1 + |\mathbf{S}|\mathrm{vec}(\mathbf{S}^{-1})'|_{\mathbf{S}=\mathbf{I}_p}\mathrm{vec}(\mathbf{S}-\mathbf{I}_p) + \cdots = 1 + \mathrm{vec}'\mathbf{I}_p\mathrm{vec}(\mathbf{S}-\mathbf{I}_p) + \cdots .$$

Siis

$$|\mathbf{S}| + p - 1 \approx \sum_{i=1}^{p} (\mathbf{S})_{ii}.$$

Nullhüpoteesi kehtimisel X_i dispersiooni hinnang $(\mathbf{S})_{ii} \sim \chi^2_{n-1}$ ja tänu sõltumatusele ja χ^2-jaotuse aditiivsusele on $|\mathbf{S}| + p - 1$ ligikaudu χ^2-jaotusega vabadusastmete arvuga $p(n-1)$.

§19. Karakteristlik funktsioon ja juhuslike vektorite momendid.

Kui meil on juhuslik suurus X, siis tema jaotus on alati määratud

– jaotusfunktsiooniga $F_X(x) = P(X \leq x)$;

– karakteristliku funktsiooniga $\varphi_X(t) = Ee^{itX}$.

Kui X on pidev juhuslik suurus, siis määrab jaotuse ka tihedusfunktsioon $f_X(x) = \frac{dF_X(x)}{dx}$. Juhuslike suuruste momendid on defineeritud kui keskväärtused:

– k-ndat järku moment $m_k(X) = E(X^k)$;

– k-ndat järku tsentraalne moment $\overline{m}_k(X) = E(X - EX)^k$.
Momente saab leida ka karakteristlikust funktsioonist tuletist võttes

$$m_k(X) = \frac{1}{i^k} \frac{d^k \varphi_X(t)}{dt^k}\Big|_{t=0}.$$

Kui momendid eksisteerivad, on karakteristliku funktsiooni abil neid tihti lihtsam leida kui tihedusfunktsiooni kasutades. Näiteks normaaljaotuse $N(\mu, \sigma^2)$ korral on tihedusfunktsioon kujul

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(- \frac{(x-\mu)^2}{2\sigma^2}\Big)$$

ja k-nda momendi leidmiseks tuleb leida integraal

$$\int_{-\infty}^{\infty} x^k f(x) dx,$$

mis tähendab ositi integreerimist. Karakteristlik funktsioon on aga kujul

$$\varphi(t) = \exp\Big(it\mu - \frac{t^2\sigma^2}{2}\Big),$$

milest tuletise võtmine on palju lihtsam. Juhusliku vektori $\mathbf{x} = (X_1, \ldots, X_p)'$ korral määravad jaotuse samuti jaotusfunktsioon ja karakteristlik funktsioon:

$$F_{\mathbf{x}}(x_1, \ldots, x_p) = P(\mathbf{x} \leq (x_1, \ldots, x_p)'),$$

$$\varphi_{\mathbf{x}}(\mathbf{t}) = Ee^{i\mathbf{t}'\mathbf{x}},$$

pideva juhusliku vektori korral ka tihedusfunktsioon

$$f_{\mathbf{x}}(x_1, \ldots, x_p) = \frac{\partial^k F_{\mathbf{x}}(x_1, \ldots, x_p)}{\partial x_1 \ldots, \partial x_p}.$$

Kõige olulisemad juhusliku vektori arvkarakteristikud on keskväärtus $E\mathbf{x}$ ja kovariatsioonimaatriks $D\mathbf{x} = E[(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})']$.

k-ndat järku segamomendid on kujul

$$E(X_1^{i_1} \times \cdots \times X_p^{i_p}), \qquad \sum_{j=1}^{p} i_j = k, \ i_j \geq 0.$$

Me ei saa üle kanda juhusliku suuruse momendi definitsiooni juhuslikule vektorile – vektorit astendada ei saa. Seetõttu defineeritakse momendid karakteristliku funktsiooni abil.

Definitsioon 19.1. *Eksisteerigu kõik juhusliku vektori* $\mathbf{x}$ *k-ndat järku absoluutsed segamomendid* $E|(X_1^{i_1} \times \cdots \times X_p^{i_p})|$, $\sum_{j=1}^{p} i_j = k$, $i_j \geq 0$. *Siis juhusliku vektori* $\mathbf{x}$ *k-ndat järku moment*

$$m_k(\mathbf{x}) = \frac{1}{i^k} \frac{d^k \varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^k}\Big|_{\mathbf{t}=\mathbf{0}}.$$

Tsentraalsed momendid tuuakse sisse momentide kaudu

Definitsioon 19.2. *Leidugu juhuslikul vektoril* $\mathbf{x}$ *k-ndat järku moment, siis vektori* $\mathbf{x}$ *k-ndat järku tsentraalne moment*

$$\overline{m}_k(\mathbf{x}) = m_k(\mathbf{x} - E\mathbf{x}).$$

Lisaks juhuslikele vektoritele on vaja ka juhuslikke maatrikseid. Juba andmemaatriks ise on tõenäosuslikult kirjeldatav juhuslik maatriks. Normaaljaotuse korral vaatasime seda enne tuletiste paragrahve. Nagu juhusliku suuruse ja juhusliku vektori korral, nii ka juhusliku maatriksi $\mathbf{X} : p \times q$ jaotuse määrab üheselt karakteristlik funktsioon ja selle momente saab leida karakteristliku funktsiooni diferentseerimisega. Maatriksi $\mathbf{X} : p \times q$ karakteristlik funktsioon ja selle abil momentide leidmine taandatakse vektoriseerimise teel juhuslike vektorite juhule

$$\mathbf{X} \longrightarrow \mathrm{vec}\mathbf{X}.$$

Definitsioon 19.3. *Juhusliku maatriksi* $\mathbf{X} : p \times q$ *karakteristlik funktsioon*

$$\varphi_{\mathbf{X}}(\mathbf{T}) \equiv \varphi_{\mathrm{vec}\mathbf{X}}(\mathrm{vec}\mathbf{T}) = Ee^{i\mathrm{vec}'\mathbf{T}\mathrm{vec}\mathbf{X}} = Ee^{i\mathrm{tr}(\mathbf{T}'\mathbf{X})}.$$

Siin argument $\mathbf{T}$ kuulub maatriksiga $\mathbf{X}$ samasse klassi – kui $\mathbf{X}$ on sümmeetriline, siis ka $\mathbf{T}$ on sümmeetriline; kui $\mathbf{X} > 0$, siis ka $\mathbf{T} > 0$, jne.

Definitsioon 19.4. *Eksisteerigu kõik juhusliku maatriksi* $\mathbf{X}$: $p \times q$ *k-ndat järku absoluutsed segamomendid* $E|(X_{11}^{i_{11}} \times \cdots \times X_{pq}^{i_{pq}})|$, $\sum_{j=1}^{p}\sum_{l=1}^{q} i_{jl} = k$, $i_{jl} \geq 0$. *Siis juhusliku maatriksi* $\mathbf{X}$ *k-ndat järku moment*

$$m_k(\mathbf{X}) = \frac{1}{i^k}\frac{d^k\varphi_{\mathbf{X}}(\mathbf{T})}{d\mathbf{T}^k}\Big|_{\mathbf{T}=\mathbf{0}}.$$

Arendame vektori $\mathbf{x}$ karakteristliku funktsiooni Taylori ritta kohal $\mathbf{t} = \mathbf{0}$

$$\varphi_{\mathbf{x}}(\mathbf{t}) = \varphi_{\mathbf{x}}(\mathbf{0}) + \sum_{j=1}^{m}\frac{1}{j!}[\mathbf{t}^{\otimes(j-1)}]'\frac{d^k\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^k}\Big|_{\mathbf{t}=\mathbf{0}}\mathbf{t} + \cdots .$$

Paneme tähele, et karakteristliku funktsiooni saab esitada momentide kaudu, kui m-ndat järku moment eksisteerib

$$\varphi_{\mathbf{x}}(\mathbf{t}) = 1 + \sum_{j=1}^{m}\frac{i^j}{j!}[\mathbf{t}^{\otimes(j-1)}]'m_j(\mathbf{x})\mathbf{t} + \cdots .$$

Järgmises paragrahvis vaatame, kuidas momente avaldada keskväärtustena.

§20. Momentide ja tsentraalsete momentide avaldised.

Eelmises paragrahvis toodud momentide definitsioon ei anna meile eeskirja momentide leidmiseks. Selleks on vaja leida karakteristliku funktsiooni diferentseerimisel saadavad avaldised. Definitsiooni kohaselt on esimest järku moment tuletis

$$m_1(\mathbf{x}) = \frac{1}{i}\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}}\Big|_{\mathbf{t}=\mathbf{0}}.$$

Karakteristliku funktsiooni definitsiooni kohaselt on $\varphi_{\mathbf{x}}(\mathbf{t})$ keskväärtus, seega on esimese momendi avaldises tuletis keskväärtusest

$$\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}} = \frac{dEe^{i\mathbf{t}'\mathbf{x}}}{d\mathbf{t}}.$$

Keskväärtus on integraal. Tuletise võtmisega oleks vaja minna keskväärtuse, seega integraali alla. Millal on see lubatud? Olgu $\mathbf{x}$ on pidev juhuslik vektor tõenäosustihedusega $f(\mathbf{x})$ ja olgu $h(\cdot)$ integreeruv funktsioon. Siis keskväärtuse leidmise reegli kohaselt

$$E[h(\mathbf{x})] = \int_{\mathbb{R}^p} h(\mathbf{x}) f(\mathbf{x}) d\mathbf{x},$$

mis karakteristliku funktsiooni korral saab kuju

$$Ee^{i\mathbf{t}'\mathbf{x}} = \int_{\mathbb{R}^p} e^{i\mathbf{t}'\mathbf{x}} f(\mathbf{x}) d\mathbf{x}.$$

Paremal pool on parameetrist $\mathbf{t}$ sõltuva funktsiooni integraal. Matemaatilisest analüüsist on teada, millal tuletisega parameetri järgi on lubatud minna parameetrist sõltuva funktsiooni integraali alla. Vaatame ühemõõtmelisel juhul parameetrist t sõltuva funktsiooni $f(x,t)$ integraali $\int f(x,t)dx$ ja tuletist

$$\frac{d}{dt}\Big(\int f(x,t)dx\Big).$$

Juhul kui

$$\Big|\int \frac{\partial f(x,t)}{\partial t} dx\Big| < \infty$$

ühtlaselt üle t, on tuletise võtmisega integraali alla minek lubatud. Mitmemõõtmelisel juhul on olukord sama – tuletise võtmisega võib minna integraali alla, kui integraal tuletisest on absoluutväärtuselt lõplik.

Meil on vaja näidata k-nda momendi leidmisel, et moodul tuletise keskväärtusest on lõplik (tegemist on kompleksarvulise funktsiooniga)

$$\Big|E\frac{d^k e^{i\mathbf{t}'\mathbf{x}}}{d\mathbf{t}^k}\Big| < \infty$$

iga $\mathbf{t} \in \mathbb{R}^p$. Kuna keskväärtus on vaja võtta tuletismaatriksi kõigist elementidest, siis peavad kõigi elementide tuletiste keskväärtused olema moodulilt lõplikud. Seda näitame hiljem, et keskväärtuse alla minek on lubatud. Esmalt leiame järgmise teoreemi tõestuses tuletise avaldise.

Teoreem 20.1. *Kui juhusliku vektori* $\mathbf{x}$ *kõik* k*-ndat järku absoluutsed segamomendid on lõplikud,* $E|X_1^{i_1} \times \cdots \times X_p^{i_p}| < \infty$, $\sum_{j=1}^{p} i^j = k$, $i^j \geq 0$, *siis juhusliku vektori* $\mathbf{x}$ k*-ndat järku moment avaldub kujul*

$$m_k(\mathbf{x}) = E[\mathbf{x}^{\otimes(k-1)}\mathbf{x}'].$$

Tõestus. Definitsiooni kohaselt

$$m_k(\mathbf{x}) = \frac{1}{i^k} \frac{d^k \varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^k}\bigg|_{\mathbf{t}=\mathbf{0}}.$$

Leiame tuletised ja näitame, et saame esitada tulemuse vastavalt teoreemi väitele. Tuletise $\frac{d^k}{d\mathbf{t}^k}\left(Ee^{i\mathbf{t}'\mathbf{x}}\right)$ leidmiseks näitame esmalt, et

$$\frac{d^k}{d\mathbf{t}^k}(e^{i\mathbf{t}'\mathbf{x}}) = i^k(\mathbf{x}^{\otimes(k-1)}\mathbf{x}')e^{i\mathbf{t}'\mathbf{x}}.$$

Tõestame selle väite matemaatilise induktsiooni meetodil.
Kui $k = 1$, siis

$$\frac{d(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}} = \frac{d(e^{i\mathbf{t}'\mathbf{x}})}{d(i\mathbf{t}'\mathbf{x})} \frac{d((i\mathbf{t}'\mathbf{x}))}{d\mathbf{t}} = ie^{i\mathbf{t}'\mathbf{x}}\mathbf{x}'.$$

Viimase võrduse saamisel kasutasime tuletise omadust (vii) ja seda, et $\mathbf{K}_{p,1} = \mathbf{I}_p$. Seaduspära läbinägemiseks vaatame ka juhtu $k = 2$

$$\frac{d^2(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^2} = i\frac{d(e^{i\mathbf{t}'\mathbf{x}}\mathbf{x}')}{d\mathbf{t}} = i\mathbf{x}\frac{d(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}} = i^2\mathbf{x}e^{i\mathbf{t}'\mathbf{x}}\mathbf{x}'.$$

Tulemus on väitega kooskõlas.
Oletame, et väide kehtib $k = n-1$ korral. Näitame, et see kehtib siis ka $k = n$ korral. Seega teame, et

$$\frac{d^{(n-1)}(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^{(n-1)}} = i^{n-1}\mathbf{x}^{\otimes(n-2)}\mathbf{x}'e^{i\mathbf{t}'\mathbf{x}}.$$

Siis

$$\frac{d^n(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^n} = i^{n-1}\frac{d(\mathbf{x}^{\otimes(n-2)}e^{i\mathbf{t}'\mathbf{x}}\mathbf{x}')}{d\mathbf{t}} = i^n\mathbf{x}^{\otimes(n-1)}\mathbf{x}'e^{i\mathbf{t}'\mathbf{x}}.$$

□

Kui diferentseerimine keskväärtuse all oleks lubatav, saaksime võrduse

$$E\Big(\frac{d^k(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^k}\Big) = i^k E(\mathbf{x}^{\otimes(k-1)}\mathbf{x}'e^{i\mathbf{t}'\mathbf{x}}).$$

Et tuletisega keskväärtuse alla minek oleks korrektne, pea me tõestama, et

$$\Big|E\Big(\frac{d^k(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^k}\Big)\Big| < \infty \text{ iga } \mathbf{t}\in\mathbb{R}^p \text{ korral.}$$

Kuna k-ndat järku tuletis funktsioonist $e^{i\mathbf{t}'\mathbf{x}}$ on maatriks, siis piisab näidata, et

$$\max_{l,m}\Big|E\Big(\frac{d^k(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^k}\Big)\Big|_{lm} < \infty,$$

kus $l = 1,2,\dots,p^{(k-1)}$ ja $m = 1,2,\dots,p$.
Teeme seda

$$\max_{l,m}\Big|E\Big(\frac{d^k(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^k}\Big)\Big|_{lm} = \max_{l,m}\Big|i^k E(\mathbf{x}^{\otimes(k-1)}\mathbf{x}'e^{i\mathbf{t}'\mathbf{x}})\Big|_{lm} \leq$$

$$\leq \max_{l,m} E\Big[|\mathbf{x}^{\otimes(k-1)}\mathbf{x}'||e^{i\mathbf{t}'\mathbf{x}}|\Big]_{lm}.$$

Viimase võrratuse saime, kuna imaginaarühiku i moodul on üks ja keskväärtuse moodul on väiksem-võrdne moodulite korrutise keskväärtusest. Arvestades seda, et $e^{i\mathbf{t}'\mathbf{x}}$ väärtused on ühikringjoone punktid mooduliga üks, saame võrratuse

$$\max_{l,m}\Big|E\Big(\frac{d^k(e^{i\mathbf{t}'\mathbf{x}})}{d\mathbf{t}^k}\Big)\Big|_{lm} \leq \max_{l,m} E\Big|\mathbf{x}^{\otimes(k-1)}\mathbf{x}'\Big|_{lm} < \infty,$$

kuna maatriksi $\mathbf{x}^{\otimes(k-1)}\mathbf{x}'$ suvaline element on kujul $X_1^{i_1} \times \cdots \times X_p^{i_p}$, $\sum_{j=1}^{p} i^j = k,\ i^j \geq 0$ ja tänu eeldusele on kõik k-ndat järku absoluutsed segamomendid $E|X_1^{i_1} \times \ldots \times X_p^{i_p}|$ lõplikud.
Sellega oleme näidanud, et tingimus tuletise võtmise ja integreerimise järjekorra vahetamiseks on täidetud ning seega

$$\frac{d^k \varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^k} = \frac{d^k E e^{i\mathbf{t}'\mathbf{x}}}{d\mathbf{t}^k} = E\Big(\frac{d^k e^{i\mathbf{t}'\mathbf{x}}}{d\mathbf{t}^k}\Big) = i^k E[\mathbf{x}^{\otimes(k-1)}\mathbf{x}' e^{i\mathbf{t}'\mathbf{x}}].$$

Kohal $\mathbf{t} = \mathbf{0}$ saame siit k-nda momendi avaldise

$$m_k(\mathbf{x}) = \frac{1}{i^k} \frac{d^k \varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^k}\Big|_{\mathbf{t}=\mathbf{0}} = E[\mathbf{x}^{\otimes(k-1)}\mathbf{x}'].$$

$\square$

Kuidas avalduvad esimesed momendid?

$$m_1(\mathbf{x}) = E\mathbf{x}' = (EX_1, \ldots, EX_p), \quad 1 \times p - \text{maatriks};$$

$$m_2(\mathbf{x}) = E(\mathbf{x}\mathbf{x}'), \quad p \times p - \text{maatriks};$$

$$m_3(\mathbf{x}) = E(\mathbf{x}^{\otimes 2}\mathbf{x}'), \quad p^2 \times p - \text{maatriks}.$$

Definitsiooni kohaselt saame juhusliku vektori $\mathbf{x}$ tsentraalsed momendid juhusliku vektori $\mathbf{y} = \mathbf{x} - E\mathbf{x}$ karakteristliku funktsiooni diferentseerimisel

$$\overline{m}_k(\mathbf{x}) = \frac{1}{i^k} \frac{d^k \varphi_{\mathbf{y}}(\mathbf{t})}{d\mathbf{t}^k}\Big|_{\mathbf{t}=\mathbf{0}} = m_k(\mathbf{y}).$$

Kui $k = 2$, siis teist järku tsentraalne moment on kovariatsioonimaatriks

$$\overline{m}_2(\mathbf{x}) = E[(\mathbf{x} - E\mathbf{x})(\mathbf{x} - E\mathbf{x})'] = D\mathbf{x}.$$

§21. Maatrikstuletis momentide ja hinnangute leidmisel.

Näide 1. Normaaljaotuse kaks esimest momenti.
p-mõõtmelise normaaljaotuse $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ karakteristlik funktsioon on kujul

$$\varphi_{\mathbf{x}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}.$$

Momentide leidmiseks tuleb diferentseerida karakteristlikku funktsiooni

Leiame esimese tuletise.

$$\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}} = \frac{de^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}}{d\mathbf{t}} = \frac{de^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}}{d(i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})}\frac{d(i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})}{d\mathbf{t}}$$

$$= \varphi_{\mathbf{x}}(\mathbf{t})\Big[i\boldsymbol{\mu}' - \frac{1}{2}\Big(\mathbf{t}'\boldsymbol{\Sigma} + \mathbf{t}'\frac{d(\boldsymbol{\Sigma}\mathbf{t})}{d\mathbf{t}}\Big)\Big] = \varphi_{\mathbf{x}}(\mathbf{t})\Big[i\boldsymbol{\mu}' - \frac{1}{2}(\mathbf{t}'\boldsymbol{\Sigma} + \mathbf{t}'\boldsymbol{\Sigma})\Big]$$
$$= \varphi_{\mathbf{x}}(\mathbf{t})(i\boldsymbol{\mu}' - \mathbf{t}'\boldsymbol{\Sigma}).$$

Esimene moment on

$$m_1(\mathbf{x}) = \frac{1}{i}\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}}\Big|_{\mathbf{t}=\mathbf{0}} = \boldsymbol{\mu}'.$$

Teist järku momendi saamiseks tuleb leida teist järku tuletis karakteristlikust funktsioonist kohal $\mathbf{t} = \mathbf{0}$.

$$m_2(\mathbf{x}) = \frac{1}{i^2}\frac{d^2\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^2}\Big|_{\mathbf{t}=\mathbf{0}} = \frac{1}{i^2}\frac{d\Big(\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}}\Big)}{d\mathbf{t}}\Big|_{\mathbf{t}=\mathbf{0}}.$$

Leiame tuletise.

$$\frac{d\Big(\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}}\Big)}{d\mathbf{t}} = \frac{d}{d\mathbf{t}}[\varphi_{\mathbf{x}}(\mathbf{t})(i\boldsymbol{\mu}' - \mathbf{t}'\boldsymbol{\Sigma})] = i\boldsymbol{\mu}\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}} - \frac{d(\varphi_{\mathbf{x}}(\mathbf{t})\mathbf{t}'\boldsymbol{\Sigma})}{d\mathbf{t}}.$$

Asendame avaldisse juba leitud karakteristliku funktsiooni esimese tuletise ja teisest liidetavast tuletise leidmisel kasutame korrutise tuletise valemit (omadus (xi).

$$\frac{d\Big(\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}}\Big)}{d\mathbf{t}} = i\boldsymbol{\mu}(i\boldsymbol{\mu}' - \mathbf{t}'\boldsymbol{\Sigma})\varphi_{\mathbf{x}}(\mathbf{t}) - \boldsymbol{\Sigma}\mathbf{t}\frac{d\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}} - \varphi_{\mathbf{x}}(\mathbf{t})\boldsymbol{\Sigma}$$
$$= [i^2\boldsymbol{\mu}\boldsymbol{\mu}' - i\boldsymbol{\mu}\mathbf{t}'\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{t}(i\boldsymbol{\mu}' - \mathbf{t}'\boldsymbol{\Sigma}) - \boldsymbol{\Sigma}]\varphi_{\mathbf{x}}(\mathbf{t}).$$

Kohal $\mathbf{t} = \mathbf{0}$ saame

$$\left.\frac{d^2\varphi_{\mathbf{x}}(\mathbf{t})}{d\mathbf{t}^2}\right|_{\mathbf{t}=\mathbf{0}} = i^2\boldsymbol{\mu}\boldsymbol{\mu}' - \boldsymbol{\Sigma}.$$

Siit

$$m_2(\mathbf{x}) = \frac{1}{i^2}(i^2\boldsymbol{\mu}\boldsymbol{\mu}' + i^2\boldsymbol{\Sigma}) = \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{\Sigma}.$$

Ühtlasi saame siit leida ka kovariatsioonimaatriksi.

$$D\mathbf{x} = E[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})'] = E(\mathbf{x}\mathbf{x}') - \boldsymbol{\mu}E\mathbf{x}' - E\mathbf{x}\boldsymbol{\mu}' + \boldsymbol{\mu}\boldsymbol{\mu}'$$

$$= E(\mathbf{x}\mathbf{x}') - \boldsymbol{\mu}\boldsymbol{\mu}' = m_2(\mathbf{x}) - \boldsymbol{\mu}\boldsymbol{\mu}' = \boldsymbol{\Sigma}.$$

Näide 2. Normaaljaotuse $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ parameetrite suurima tõepära hinnangud.
Normaaljaotuse $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tõenäosustihedus

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}.$$

Olgu meil valim $(\mathbf{x}_1 \ldots, \mathbf{x}_n)$ mahuga n normaaljaotusest $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
Siis tõepärafunktsioon

$$L(\mathbf{x}_1 \ldots, \mathbf{x}_n; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\boldsymbol{\mu})}$$

ja logaritmiline tõepärafunktsioon

$$l(\mathbf{x}_1 \ldots, \mathbf{x}_n; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^{n} \Big(-\frac{p}{2}\ln(2\pi) - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}(\mathbf{x}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\boldsymbol{\mu}) \Big)$$

$$= -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\boldsymbol{\mu}).$$

Suurima tõepära hinnangute leidmiseks võtame tuletised logaritmilisest tõepärafunktsioonist, võrdsustame need nulliga ja lahendame võrrandisüsteemi.

$$\frac{dl}{d\boldsymbol{\mu}} = \mathbf{0} = -\frac{1}{2}\sum_{i=1}^{n} \frac{d[(\mathbf{x}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\boldsymbol{\mu})]}{d\boldsymbol{\mu}}$$

$$= \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}.$$

Viimase võrduse saamisel rakendasime korrutise tuletise omadust (xi) ja kasutasime võrdust $\mathbf{K}_{p,1} = \mathbf{I}_p$. Seega

$$\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}) = \mathbf{0},$$

kust

$$\sum_{i=1}^{n}\mathbf{x}_i = n\boldsymbol{\mu}, \quad \widehat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i = \overline{\mathbf{x}}.$$

Keskväärtuse $\boldsymbol{\mu}$ suurima tõepära hinnangu saime lihtsalt. Kovariatsioonimaatriksi $\boldsymbol{\Sigma}$ hinnangu leidmisel saame vajadusel $\boldsymbol{\mu}$ hinnangu sisse asendada.

$$\frac{dl}{d\boldsymbol{\Sigma}} = -\frac{n}{2}\frac{d(\ln|\boldsymbol{\Sigma}|)}{d\boldsymbol{\Sigma}} - \frac{1}{2}\sum_{i=1}^{n}\frac{d((\mathbf{x}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}))}{d\boldsymbol{\Sigma}} = \mathbf{0}$$

$$= -\frac{n}{2}\mathrm{vec}'(\boldsymbol{\Sigma}^{-1}) - \frac{1}{2}\sum_{i=1}^{n}[(\mathbf{x}_i - \boldsymbol{\mu})' \otimes (\mathbf{x}_i - \boldsymbol{\mu})']\frac{d\boldsymbol{\Sigma}^{-1}}{\boldsymbol{\Sigma}}.$$

Viimase võrduse saamisel kasutasime determinandi tuletise avaldist ja tuletise omadust (vii). Pöördmaatriksi tuletise omadust (xiii) rakendades saame

$$-\frac{n}{2}\mathrm{vec}'(\boldsymbol{\Sigma}^{-1}) - \frac{1}{2}\sum_{i=1}^{n}[(\mathbf{x}_i - \boldsymbol{\mu})' \otimes (\mathbf{x}_i - \boldsymbol{\mu})'](-\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) = \mathbf{0}.$$

Korrutame paremalt saadud võrdust maatriksiga ($\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}$)ja rakendame vec-operaatori omadust (v), $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}$,

$$n\mathrm{vec}\boldsymbol{\Sigma} = \sum_{i=1}^{n}[(\mathbf{x}_i - \boldsymbol{\mu})' \otimes (\mathbf{x}_i - \boldsymbol{\mu})'].$$

Avaldame $\mathrm{vec}\boldsymbol{\Sigma}$ ja lisame paremal otsekorrutisele teguri vec1 ning rakendame vec-operaatori omadust (v)

$$\mathrm{vec}\boldsymbol{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}[(\mathbf{x}_i-\boldsymbol{\mu})'\otimes(\mathbf{x}_i-\boldsymbol{\mu})']\mathrm{vec}1 = \frac{1}{n}\mathrm{vec}\Big[\sum_{i=1}^{n}(\mathbf{x}_i-\boldsymbol{\mu})(\mathbf{x}_i-\boldsymbol{\mu})'\Big].$$

Saime vektoriseeritud maatriksite võrduse, seega

$$\boldsymbol{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i-\boldsymbol{\mu})(\mathbf{x}_i-\boldsymbol{\mu})'.$$

Asendame $\boldsymbol{\mu}$ tema hinnanguga $\widehat{\boldsymbol{\mu}} = \overline{\mathbf{x}}$

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})'.$$

Kirjandust

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis.* Third edition (2003). Wiley, New York.

Harville, D.A. (1997). *Matrix Algebra From a Statistician's Perspective.* Springer, New York.

Kilp, M. (2005). *Algebra I.* Eesti Matemaatika Selts, Tartu.

Kollo, T. (1991). *Maatrikstuletis mitmemõõtmelise statistika jaoks.* Tartu Ülikooli Kirjastus, Tartu (vene k.).

Kollo, T. & von Rosen, D. (2005, 2010). *Advanced Multivariate Statistics with Matrices.* Springer, Dordrecht.

Magnus, J.R. & Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Second edition. Wiley, Chichester.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

Neudecker, H. (1969). Some theorems on matrix differentiations with special reference to Kronecker matrix products. *J. Amer. Statist. Assoc.* **64** 953–963.

Rao, C.R. (1965). *Linear Statistical Inference and Its Applications.* Second edition (1973). Wiley, New York.

Rao, C.R. (1968). *Lineaarsed statistilised meetodid ja nende rakendused.* Nauka, Moskva (vene k.).

C.R. Rao, S.M. Mitra (1971). *Generalized Inverse of Matrices and Its Applications.* Wiley, New York.

Rao, C.R. & Rao, M.B. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics.* World Scientific, Singapore.

Schott, J.R. (1997b). *Matrix Analysis for Statistics.* Wiley, New York.

Searle, S.R. (1982). *Matrix Algebra Useful for Statistics.* Wiley, New York.

Siotani, M., Hayakawa, T. & Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* American Science Press, Columbus, Ohio.

Srivastava, M.S. & Khatri, C.G. (1979). *An Introduction to Multivariate Statistics.* North Holland, New York.