MAATRIKSID STATISTIKAS. LISA 1

L1.1. Maatriksalgebra mõisteid ja tulemusi

Maatriksalgebra annab vahendid mitmemõõtmelise statistika esitamiseks kompaktsel ja ülevaatlikul kujul. üleminek maatrikskeelele cá 60 aastat tagasi 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). 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

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

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

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

Maatriksi transponeerimine on seotud liitmise ja korrutamisega järgmiselt:

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

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

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

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

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

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

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

ja *idempotentne* , kui

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

Ruutmaatriks $\mathbf{A} : n \times n$ on *ülemine kolmnurkmaatriks*, kui tema allpool peadiagonaali asetsevad elemendid võrduvad nulliga:

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

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

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 $\vec{a}$ moodustatud diagonaalmaatriks on sarnase tähistusega:

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

Determinant ja pöördmaatriks. Ruutmaatriksi $\mathbf{A} : n \times n$ korral on tähtsaks arvuliseks karakteristikuks tema *determinant* $|\mathbf{A}|$:

$$|\mathbf{A}| = \sum_{j_1,\ldots,j_n} (-1)^{N(j_1,\ldots,j_n)} \prod_{i=1}^{n} a_{ij_i}, \tag{L1.1}$$

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 nii, et suurem on pärast väiksemat. Determinandi arvutamiseks on valemit (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 **A** determinandi esitada arendusena i-nda rea või j-nda veeru kaudu:

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}(-1^{i+j}|\mathbf{A}_{(ij)}| \quad \text{iga } i \text{ korral}, \tag{L1.2}$$

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}(-1^{i+j}|\mathbf{A}_{(ij)}| \quad \text{iga } j \text{ korral}. \tag{L1.3}$$

Maatriksi **A** esimesest r reast ja veerust moodustatud $r \times r$-alammaatriksi determinanti nimetatakse r-järku *nurgamiinoriks*. Esitame järgnevalt mõned determinandi olulisemad omadused:

$$|\mathbf{A}'| = |\mathbf{A}|; \tag{L1.4}$$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|; \tag{L1.5}$$

$$|\mathbf{I}_m + \mathbf{AB}| = |\mathbf{I}_n + \mathbf{BA}|, \tag{L1.6}$$

kus **A** on $m \times n$-maatriks ja **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 **A** pöördmaatriks $\mathbf{A}^{-1}$, mis on määratud võrdusega

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

Pöördmaatriksil on järgmised omadused.

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

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}; \tag{L1.7}$$

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}; \tag{L1.8}$$

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

Pöördmaatriksi üldelement avaldub valemiga

$$(\mathbf{A}^{-1})_{ij} = \frac{(-1^{i+j}|\mathbf{A}_{(ji)}|}{|\mathbf{A}|}. \tag{L1.9}$$

Kahe maatriksi summa pöördmaatriksi avaldis on keerulisem.
Teoreem L1.1 (binomiaalne pöördteoreem). *Olgu järgnevas võrduses maatriksid* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *ja* $\mathbf{D}$ *sobivate dimensioonidega ja eksisteerigu seal esinevad pöördmaatriksid. Siis*

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

Astak ja jälg Vektorid $\vec{x}_i$, $i = 1,\ldots,r$ on *lineaarselt sõltumatud*, kui

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

parajasti siis, kui $c_i = 0$, $i = 1,\ldots,r$
Definitsioon L1.1. *Maatriksi* $\mathbf{A} : m\times n$ *astakuks* $r(\mathbf{A})$ *nimetatakse tema lineaarselt sõltumatute veergude maksimaalset arvu.*
Loetleme alljärgnevalt astaku olulisemad omadused:

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

(ii) $r(\mathbf{A}) \le \min(m,n)$;

(iii) $r(\mathbf{AB}) \le \min(r(\mathbf{A}), r(\mathbf{B}))$;

(iv) $r(\mathbf{A}+\mathbf{B}) \le r(\mathbf{A}) + r(\mathbf{B})$;

(v) kui $\mathbf{A}$ ja $\mathbf{C}$ on pööratavad, siis $r(\mathbf{ABC}) = r(\mathbf{B})$;

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

Definitsioon L1.2. *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) $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{B}^{-1}\mathbf{A}\mathbf{B})$, kui $\mathbf{B}$ on pööratav;

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

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

(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) $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}') = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2$;

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

(x) $r(\mathbf{A}) \geq \frac{(\mathrm{tr}\mathbf{A})^2}{\mathrm{tr}(\mathbf{A}^2)}$, kui $\mathbf{A}$ on sümmeetriline.

Omaväärtused, omavektorid.
Olgu $\mathbf{A} : n \times n$ reaalarvuline maatriks. Vaatame seda kui lineaarteisendust ruumis $\mathbb{R}^n$, mis teisendab suvalise vektori $\vec{x}$ teiseks vektoriks $\vec{y}$,

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

Vektorid, mille siht säilib selle teisendusega, on määratud võrrandiga

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

Sel võrrandil on mittetriviaalne lahend, kui

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0. \tag{L1.10}$$

Võrrandit (10) nimetatakse maatriksi $\mathbf{A}$ *karakteristlikuks võrrandiks* ja selle lahendeid λ_i maatriksi $\mathbf{A}$ *omaväärtusteks*. Vektorit $\vec{x}_i$, mis rahuldab võrdust

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

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 komplekssed, kusjuures nende hulgas võib olla kordseid. Seega on reaalarvulisel maatriksil n omaväärtust, mis võivad olla nii reaalarvulised kui kompleksarvud. Omaväärtustele vastavad omavektorid ei ole üheselt määratud, varieeruda võivad nende pikkus ja suund. 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 –). Esitame järgnevalt olulisemad omaväärtuste ja omavektorite omadused.

(i) Reaalarvulise sümmeetrilise maatriksi kõik omaväärtused λ_i on reaalarvulised ja neile vastavad omavektorid $\vec{x}_i$ saab valida reaalarvuliste koordinaatidega.

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

(iii) Maatriksid $\mathbf{A}$ ja $\mathbf{CAC}^{-1}$ on samade omaväärtustega suvalise pööratava maatriksi $\mathbf{C}$ korral.

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

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

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

(vii) 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 $\mathbf{A}\mathbf{A}'$ 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) Sümmeetrilise maatriksi $\mathbf{A} : n \times n$ korral $\sum_{i=1}^{n} \lambda_i = \mathrm{tr}\mathbf{A}$.

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

Mitmemõõtmelise statistika jaoks ülioluline on järgmine tulemus.

Teoreem L1.2. *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.*

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

$$\mathbf{A} = \sum_{i_1}^{n} \lambda_i \vec{x}_i \vec{x}_i',$$

kus $\vec{x}_i$ on omaväärtusele λ_i vastav normeeritud omavektor. Kasulikuks osutub ka järgmine tulemus, mida tuntakse *Rayleigh suhte* nime all.

Teoreem L1.3 (Rayleigh suhe). *Olgu* $\mathbf{A} : n \times n$ *sümmeetriline maatriks omaväärtustega* λ_i. *Siis iga* $\vec{x} \in \mathbb{R}^n$ *korral*

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

Positiivne määratus.

Definitsioon L1.3. *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on positiivselt (negatiivselt) määratud, kui* $\vec{x}'\mathbf{A}\vec{x} > 0$ (< 0) *iga vektori* $\vec{x} \neq \vec{0} \in \mathbb{R}$ *korral.*

Kui $\mathbf{A}$ on positiivselt (negatiivselt) määratud, tähistame seda $\mathbf{A} > 0$, $(\mathbf{A} < 0)$.

Definitsioon L1.4. *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on positiivselt poolmääratud (negatiivselt poolmääratud), kui* $\vec{x}'\mathbf{A}\vec{x} \geq 0$ (≤ 0) *iga vektori* $\vec{x} \in \mathbb{R}$ *korral ja leidub vähemalt üks vektor* $\vec{x}_0 \neq \vec{0}$ *nii, et* $\vec{x}_0'\mathbf{A}\vec{x}_0 = 0$.

Kui $\mathbf{A}$ on positiivselt poolmääratud (negatiivselt poolmääratud), tähistame seda $\mathbf{A} \geq 0$, $(\mathbf{A} \leq 0)$.

Definitsioon L1.5. *Sümmeetriline maatriks* $\mathbf{A} : n \times n$ *on mittenegatiivselt (mittepositiivselt) määratud, kui* $\mathbf{A} > 0$ *või* $\mathbf{A} \geq 0$ *(*$\mathbf{A} < 0$ *või* $\mathbf{A} \leq 0$*).*

Järgnevalt esitame mõned positiivselt määratud maatriksite 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 $\mathbf{A} : n \times n$, $\mathbf{A} > 0$ ja $\mathbf{B} : n \times m$, $m \leq n$ ja $r(\mathbf{B}) = m$. Siis

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

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

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

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

(viii) 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 $r < m$.

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

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

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 $\vec{x} \in \mathbb{R}^n$ avaldub summana

$$\vec{x} = \vec{x}_1 + \vec{x}_2,$$

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

operatsioon. Tähistame P lineaarse operaatori, mis projekteerib ruumi $\mathbb{R}^n$ vektorid alamruumi $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$,

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

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 $\mathbb{M}$. Projektorit iseloomustav omadus on idempotentsus.

Teoreem L1.4. *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.*

Vaatame maatriksit $\mathbf{A} : n \times m$ kui teisendust ruumist $\mathbb{R}^m$ ruumi $\mathbb{R}^n$. Kujutised $\mathbf{A}\vec{x}$ moodustavad ruumis $\mathbb{R}^n$ alamruumi, mis erijuhul võib kokku langeda ka kogu ruumiga. Lineaarteisenduse $\mathbf{A}$ *kujutisruumiks* $\mathcal{C}(\mathbf{A})$ nimetatakse hulka

$$\mathcal{C}(\mathbf{A}) = \{\vec{x} : \vec{x} = \mathbf{A}\vec{y},\ \vec{y} \in \mathbb{R}^m\}.$$

Teisenduse $\mathbf{A}$ *nullruum* $\mathcal{N}(\mathbf{A})$ on määratud võrdusega

$$\mathcal{N}(\mathbf{A}) = \{\vec{x} : \mathbf{A}\vec{y} = \vec{0},\ \vec{y} \in \mathbb{R}^m\}.$$

Teisiti öeldes on lineaarteisenduse $\mathbf{A}$ kujutisruum maatriksi $\mathbf{A}$ veeruvektorite lineaarne kate. Kuidas leida projektor, mis projekteerib etteantud kujutisruumile? Olgu $\mathbf{A} : n \times m,\ m \leq n$ täisastakuga $r(\mathbf{A}) = m$ maatriks. Projektor, mis projekteerib suvalise vektori ruumist $\mathbb{R}^m$ maatriksi $\mathbf{A}$ veeruvektorite $\vec{a}_i,\ i = 1, \ldots, m$ poolt määratud kujutisruumi paralleelselt nullruumiga $\mathcal{N}(\mathbf{A})$ on antud järgmises teoreemis.

Teoreem L1.5. *Olgu* $\mathbf{A} : n \times m,\ m \leq n$ *täisastakuga* $r(\mathbf{A}) = m$ *maatriks veeruvektoritega* $\vec{a}_i,\ i = 1, \ldots, m$. *Siis projektor kujutisruumile* $\mathcal{C}(\mathbf{A})$ *on kujul*

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

Üldistatud pöördmaatriks.

Üldistatud pöördmaatriksi mõiste on vajalik lineaarsete võrrandisüsteemide lahendamisel, kui võrrandeid on vähem kui tundmatuid või tundmatutele seatud tingimused on lineaarses sõltuvuses.

Definitsioon L1.6. *Maatriksi* $\mathbf{A} : m \times n$ *üldistatud pöördmaatriksiks nimetatakse* $n \times m$*-maatriksit* $\mathbf{A}^-$, *kui*

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}. \tag{L1.11}$$

Üldistatud pöördmaatriks ei ole üheselt määratud, selliseid maatrikseid, mis rahuldavad võrdust (L1.11), võib olla lõpmata palju. Kahtlemata rahuldab võrdust (L1.11) ruutmaatriksi $\mathbf{A}$ pöördmaatriks, kui viimane eksisteerib ja on siis ka ühene. Selleks, et üldistatud pöördmaatriksite hulgast välja eraldada üks kindlate omadustega esindaja, on vaja lisatingimusi.

Definitsioon L1.7. *Maatriksi* $\mathbf{A} : m \times n$ *Moore-Penrose üldistatud pöördmaatriksiks nimetatakse* $n \times m$*-maatriksit* $\mathbf{A}^+$, *kui ta rahuldab järgmist nelja tingimust:*

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

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

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

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

Moore-Penrose üldistatud pöördmaatriks on ühene, leidub üks ja ainult üks maatriks, mis rahuldab tingimusi (i) – (iv). Moore-Penrose'i üldistatud pöördmaatriksi seos lineaarsete võrrandite süsteemi lahendamisega on ära toodud järgmises teoreemis.

Teoreem L1.6. *Olgu* $\mathbf{A}$ $m \times n$*-maatriks,* $\vec{x}$ n*-vektor ja* $\vec{y}$ m*-vektor. Võrrand* $\mathbf{A}\vec{x} = \vec{y}$ *on lahenduv parajasti siis, kui*

$$\mathbf{A}\mathbf{A}^{+}\vec{y} = \vec{y}$$

ning selle üldlahend on kujul

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

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

L1.2. Plokkmaatriksid

Tehted ja tähistused.
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" nimetust (vt. näiteks Magnus & Neudecker (1999)). Selle all peetakse silmas eelkõige mõisteid otsekorrutis, kommutatsioonimaatriks, vec-operaator ja maatrikstuletis ning nende omadusi ja vastastikuseid seoseid.
Definitsioon L1.8. *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*

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

Plokkmaatriksi jaoks kasutame tähistust

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

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

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

Plokkmaatriks $\mathbf{A} = [\mathbf{A}_{ij}], \quad i = 1, \ldots, u;\ j = 1, \ldots, v$ on plokkdiagonaalne, kui $u = v$ ja $\mathbf{A}_{ij} = 0$, kui $i \neq j$. Eespool vaadeldud tehted maatriksitega esitatakse arvestades plokkstruktuuri:

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

(ii) $\mathbf{A} + \mathbf{B} = [\mathbf{A}_{ij} + \mathbf{B}_{ij}], \quad i = 1,\dots,u;\ j = 1,\dots,v,$ kui plokid on samade mõõtmetega;

(iii) $\mathbf{A} = [\mathbf{A}_{ij}] : m \times n, \quad i = 1,\dots,u;\ j = 1,\dots,v$ ja $\mathbf{B} = [\mathbf{B}_{jl}] : n \times r, \quad j = 1,\dots,v;\ l = 1,\dots,w$ 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:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}. \tag{L1.12}$$

Sel juhul on võimalik esitada determinant ja pöördmaatriks plokkide kaudu.

Teoreem L1.7. *Olgu pööratav plokkmaatriks* $\mathbf{A}$ *antud kujul (L1.12). Kui* $\mathbf{A}_{22}$ *on pööratav, siis*

$$|\mathbf{A}| = |\mathbf{A}_{22}| \cdot |\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|; \tag{L1.13}$$

kui $\mathbf{A}_{11}$ *on pööratav, siis*

$$|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|. \tag{L1.14}$$

Valemites (L1.13) ja (L1.14) esinevad plokkide vahed kannavad *Schuri täiendite* nime. Ploki $\mathbf{A}_{ii}$ Schuri täiendiks $\mathbf{A}_{jj\cdot i}$ nimetatakse maatriksit

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

.

Teoreem L1.8. *Olgu pööratav plokkmaatriks* $\mathbf{A}$ *antud kujul (L1.12). Siis*

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

Plokkdiagonaalse maatriksi pöördmaatriks avaldub lihtsamalt:

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

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

Definitsioon L1.9. *Plokkmaatriksit* $\mathbf{K}_{m,n}$: $mn \times mn$, *mis koosneb* $n \times m$*-plokkidest, nimetatakse kommuteerimismaatriksiks, kui*

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

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

$$\mathbf{K}_{2,3} = \left(\begin{array}{ccccccccc} 1 & \cdots & 0 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & \cdots & 0 & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & \cdots & 0 & \vdots & 0 & 0 & \vdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & \cdots & 0 & \vdots & 0 & 1 & \vdots & 0 & 0 \\ 0 & \cdots & 0 & \vdots & 0 & 0 & \vdots & 0 & 1 \end{array} \right).$$

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: *kommuteerimismaatriks* $\mathbf{K}_{m,n}$ *on* $mn \times mn$*-plokkmaatriks, mis koosneb* $n \times m$*-plokkidest, kusjuures* ij*-ndas plokis on* ji*-s element 1 ja ülejäänud elemendid on nullid,* $i = 1 \ldots, m$; $j = 1, \ldots, n$.
Toome ära mõned $\mathbf{K}_{m,n}$ omadused:

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

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

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

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

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

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 L1.10. *Maatriksite* $\mathbf{A} : m \times n$ *ja* $\mathbf{B} : r \times s$ *otsekorrutis* $\mathbf{A} \otimes \mathbf{B}$ *on* $mr \times ns$*-plokkmaatriks*

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

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

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

Otsekorrutise põhiomadused on

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

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

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

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

(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

$$(\mathbf{A}\otimes\mathbf{C})(\mathbf{B}\otimes\mathbf{D}) = (\mathbf{AB})\otimes(\mathbf{CD}); \tag{L1.15}$$

(vi) olgu $\mathbf{A}$ ja $\mathbf{B}$ pööratavad, siis

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

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

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

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

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

(viii) vektorite $\vec{a}$ ja $\vec{b}$ korral

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

vec-operaator.
Definitsioon L1.11. *Olgu* $\mathbf{A} = (\vec{a}_1,\dots,\vec{a}_n)$ $m\times n$*-maatriks. Vektoriseerimisoperaator* $\mathrm{vec}(\cdot)$ *on operaator maatriksite ruumist* $\mathbb{R}^{m\times n}$ *ruumi* $\mathbb{R}^{mn}$ *kujul*

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

Järgnevad omadused seovad vec-operaatori teiste maatriksoperatsioonidega:

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

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

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

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

(iii) kui $\mathbf{A}$ on $n \times n$ sümmeetriline maatriks, siis

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

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

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\mathrm{vec}\mathbf{B}; \tag{L1.16}$$

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

$$\mathrm{tr}(\mathbf{A}'\mathbf{B}) = (\mathrm{vec}\mathbf{A})'\mathrm{vec}\mathbf{B}; \tag{L1.17}$$

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

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

(vi) olgu $\mathbf{A} : m \times n$ ja $\mathbf{B} : r \times s$, maatriksid, siis

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

Maatrikstuletis.
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 ülalkirjeldatud statistikaprobleemide lahendamiseks ja tema esitus maatriksite kaudu kannab maatrikstuletise nimetust. Kasutame osatuletiste järjestamiseks definitsiooni, mille võttis kasutusele Heinz Neudecker (Neudecker, 1969).

Definitsioon L1.12. *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 ja*

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\mathrm{vec}'\mathbf{X}} \otimes \mathrm{vec}\mathbf{Y}, \tag{L1.18}$$

kus

$$\frac{d}{d\mathrm{vec}'\mathbf{X}} = \Big(\frac{\partial}{\partial x_{11}}, \ldots, \frac{\partial}{\partial x_{p1}}, \frac{\partial}{\partial x_{12}}, \ldots, \frac{\partial}{\partial x_{p2}}, \ldots, \frac{\partial}{\partial x_{1q}}, \ldots, \frac{\partial}{\partial x_{pq}}\Big)$$

ja osatuletise $\frac{\partial y_{kl}}{\partial x_{gh}}$ paiknemine maatriksis on määratud otsekorrutisega (L1.18).

Kirjutame plokkmaatriksi (L1.18) välja elementide kaudu:

$$
\mathbf{A} = \begin{pmatrix}
\frac{\partial y_{11}}{\partial x_{11}} & \cdots & \frac{\partial y_{11}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{11}}{\partial x_{1q}} & \cdots & \frac{\partial y_{11}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{r1}}{\partial x_{11}} & \cdots & \frac{\partial y_{r1}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{r1}}{\partial x_{1q}} & \cdots & \frac{\partial y_{r1}}{\partial x_{pq}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial y_{12}}{\partial x_{11}} & \cdots & \frac{\partial y_{12}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{12}}{\partial x_{1q}} & \cdots & \frac{\partial y_{12}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{r2}}{\partial x_{11}} & \cdots & \frac{\partial y_{r2}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{r2}}{\partial x_{1q}} & \cdots & \frac{\partial y_{r2}}{\partial x_{pq}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial y_{1s}}{\partial x_{11}} & \cdots & \frac{\partial y_{1s}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{1s}}{\partial x_{1q}} & \cdots & \frac{\partial y_{1s}}{\partial x_{pq}} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_{rs}}{\partial x_{11}} & \cdots & \frac{\partial y_{rs}}{\partial x_{p1}} & \vdots & \cdots & \vdots & \frac{\partial y_{rs}}{\partial x_{1q}} & \cdots & \frac{\partial y_{rs}}{\partial x_{pq}}
\end{pmatrix},
$$

Maatrikstuletise omadused esitame loeteluna, maatriksite mõõtmed toome ära, kui need erinevad definitsioonis esinevatest:

(i)

$$\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{I}_{pq};$$

(ii) kui $\mathbf{Y} = c\mathbf{X}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = c\mathbf{I}_{pq};$$

(iii) kui $\mathrm{vec}\mathbf{Y} = \mathbf{A}\mathrm{vec}\mathbf{X}$, kus $\mathbf{A}$ on konstantne maatriks, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \mathbf{A};$$

(iv) kui $\mathbf{Y} = \mathbf{U} + \mathbf{V}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathbf{U}}{d\mathbf{X}} + \frac{d\mathbf{V}}{d\mathbf{X}};$$

(v) olgu $\mathbf{Z} : m \times n$, $\mathbf{Y} : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}} \frac{d\mathbf{Y}}{d\mathbf{X}};$$

(vi) olgu $\mathbf{Y} = \mathbf{AXB}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A});$$

(vii) olgu $\mathbf{Z} = \mathbf{AYB}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid ja $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A}) \frac{d\mathbf{Y}}{d\mathbf{X}};$$

(viii)

$$\frac{d\mathbf{X}'}{d\mathbf{X}} = \mathbf{K}_{p,q}, \qquad \frac{d\mathbf{X}}{d\mathbf{X}'} = \mathbf{K}_{q,p};$$

(ix) kui $\mathbf{X}$ on $p \times p$-maatriks, siis

$$\frac{d\mathbf{X}_d}{d\mathbf{X}} = (\mathbf{K}_{p,p})_d;$$

(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 $\mathbf{X}$ funktsioonid, siis

$$\frac{d\mathbf{W}}{d\mathbf{X}} = \left.\frac{d\mathbf{W}}{d\mathbf{Y}}\right|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}} + \left.\frac{d\mathbf{W}}{d\mathbf{Z}}\right|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}};$$

(xi) kui $\mathbf{Y}$ ja $\mathbf{Z}$ on $\mathbf{X}$ funktsioonid, siis

$$\frac{d(\mathbf{YZ})}{d\mathbf{X}} = \left.\frac{d(\mathbf{YZ})}{d\mathbf{Y}}\right|_{\mathbf{Z}=const} \frac{d\mathbf{Y}}{d\mathbf{X}} + \left.\frac{d(\mathbf{YZ})}{d\mathbf{Z}}\right|_{\mathbf{Y}=const} \frac{d\mathbf{Z}}{d\mathbf{X}};$$

(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 $\mathbf{Y}^0 = \mathbf{I}_r,\ n \geq 1$;

(xiii) olgu $p \times p$-maatriks $\mathbf{X}$ pööratav, siis

$$\frac{d\mathbf{X}^{-1}}{d\mathbf{X}} = -(\mathbf{X}')^{-1} \otimes \mathbf{X}^{-1};$$

(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 $\mathbf{Y}^0 = \mathbf{I}_r,\ n \geq 1$;

(xv) olgu $\mathbf{Z} : m \times n$, $\mathbf{Y} : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = (\mathbf{I}_s \otimes \mathbf{K}_{n,r} \otimes \mathbf{I}_m) \Big[(\mathbf{I}_{rs} \otimes \mathrm{vec}\mathbf{Z}) \frac{d\mathbf{Y}}{d\mathbf{X}} + (\mathrm{vec}\mathbf{Y} \otimes \mathbf{I}_{mn}) \frac{d\mathbf{Z}}{d\mathbf{X}} \Big];$$

(xvi) kui $\mathbf{X}$ on pööratav $p \times p$-maatriks, siis

$$\frac{|d\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}| \mathrm{vec}'(\mathbf{X}')^{-1};$$

(xvii) kui $\mathbf{X}$ on $p \times p$-maatriks, siis

$$\frac{d(\mathrm{tr}\mathbf{X})}{d\mathbf{X}} = \mathrm{vec}'\mathbf{I}_p.$$

Viidatud kirjandus

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). *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.).

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.