Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

František Navrátil

Maticová algebra ve statistice

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Doc. Mgr. Michal Kulich, Ph.D.
Studijní program: Matematika
Studijní obor: Finanční matematika

Praha 2011

Na tomto místě bych rád poděkoval svému vedoucímu, panu doc. Mgr. Michalu Kulichovi, Ph.D., za poskytnuté odborné materiály, ochotně předané cenné rady a velmi efektivní komunikaci.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne .... Podpis autora

Název práce: Maticová algebra ve statistice

Autor: František Navrátil

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Doc. Mgr. Michal Kulich, Ph.D.

Abstrakt: Práce se zabývá teorií maticové algebry, kterou lze uplatnit v pravděpodobnosti a statistice. Cílem práce je tuto látku srozumitelně a přehledně shrnout, aby student seznámený se základy teorie matic mohl rozšířit své znalosti a využít je při dalším studiu. Proto práce obsahuje množství definic a dokazovaných vět, příklady pro usnadnění pochopení látky, zmiňuje aplikace a uvádí odkazy na další literaturu. Práce začíná uvedením základních poznatků maticové algebry, které jsou součástí běžných kurzů lineární algebry. Následující kapitoly jsou již specifické (mimo jiné) pro pravděpodobnost a statistiku - zaměřují se zejména na speciální typy matic a jejich vlastnosti, důležité rozklady matic, funkce matic a maticové derivování.

Klíčová slova: maticová algebra, statistika, idempotentní matice, spektrální rozklad, Kroneckerův součin

Title: Matrix Algebra in Statistics

Author: František Navrátil

Department: Department of Probability and Mathematical Statistics

Supervisor: Doc. Mgr. Michal Kulich Ph.D.

Abstract: The thesis deals with the theory of matrix algrebra, which is applicable in probability and statistics. The aim of the thesis is to summarize it in a clear and understandable way, so that the student familiar with the basics of matrix theory can expand his knowledge and use it in further studies. Therefore, the thesis contains many definitions and proved theorems, and examples to help understanding the theory. Applications are mentioned. It also provides references for further reading. The thesis begins with a brief summary of basic definitions and results in matrix algebra, which are covered in the usual courses on linear algebra. Subsequent chapters are specific, inter alia, for probability and statistics - in particular, they focus on special types of matrices and their properties, important matrix decompositions, functions of matrices and matrix differentiation.

Keywords: matrix algebra, statistics, idempotent matrix, spectral decomposition, Kronecker product

Obsah

Úvod

Maticová algebra se uplatňuje v mnoha oblastech matematiky. Tato práce se zaměřuje na tu část teorie matic, která se běžně používá v pravděpodobnosti, statistice a souvisejících oborech. Student se s ní tedy setká na různých přednáškách, často v podobě pouhých poznámek, bez důkazů a souvislostí. Přednášky lineární algebry sice osvětlují základy maticového počtu, jsou však naplněny především jinými tématy, a tak prostor pro další partie teorie matic nezbývá.

Smyslem práce je tyto užitečné informace vyhledat v literatuře a přehledně je shrnout. Proto obsahuje definice používaných pojmů, související tvrzení včetně důkazů, pro ověření porozumění spočítané ilustrační příklady a v poznámkách naznačené souvislosti s užitím ve statistice. Student by po přečtení této práce měl být schopen hlouběji proniknout do látky probírané v dalších předmětech. Absolvování základních přednášek lineární algebry a matematické analýzy je nutnou a zároveň postačující podmínkou ke čtení a pochopení tohoto materiálu. Pro případné zájemce jsou však zmíněny i knihy k dalšímu možnému studiu.

Úvodní kapitola shrnuje základní poznatky teorie matic, se kterými je student obeznámen již po prvním ročníku studia. Začíná zavedením pojmu matice a uvádí typy matic. Dále obsahuje výčet základních poznatků o stopě, hodnosti, determinantu, vlastních číslech a vektorech matice a o inverzní a transponované matici. Velká část uvedeného je užita v následujících kapitolách.

V druhé kapitole se nachází množství definic speciálních typu matic a jejich vlastností. Mimo jiné se zaměřuje na idempotentní matici. Část je věnována také ortogonální matici, jejíž vlastnosti uplatníme i v následujících kapitolách.

Třetí kapitola se zabývá rozkladem matice na dvojici (či trojici) jiných matic, které mají užitečné vlastnosti (jsou ortogonální, diagonální apod.). Hlavní částí je spektrální rozklad matice. Všechna uvedená tvrzení jsou dokázána.

Čtvrtá kapitola obsahuje několik částí, z nichž v každé je popsáno jedno užitečné a zároveň velmi zajímavé téma - funkce matic, maticové derivování, Kroneckerův součin a vektorizace matic.

1. Základní pojmy

V této kapitole přehledně shrneme základní definice a důležité věty z teorie matic. Utvoříme tak základní aparát pro vyjadřování v dalších, hlubších částech práce. Pro studenta, který absolvoval kurz lineární algebry, by měla být první část (definice matice a výčet terminologie) snadným opakováním. Ve druhé části, kde rozebereme pojmy stopa matice, determinant, hodnost, vlastní čísla a vlastní vektory, je shrnuto množství vlastností uvedených pojmů (bez výjimky jsou to užitečná fakta - často z pohledu výpočetního i teoretického). Důkazy lze zpravidla dohledat v základní literatuře [2], [3].

1.1 Matice, maticové operace

Definice. Obdélníkové schéma čísel

$$\mathbf{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

nazýváme *matice typu* $m \times n$ (matice o m řádcích a n sloupcích). Jednotlivá čísla nazýváme *prvky matice*. Prvek v i-tém řádku v j-tém sloupci značíme a_{ij} (v práci předpokládáme $a_{ij} \in \mathbb{R}$, kde $\mathbb{R}$ značí množinu reálných čísel).

Některé matice mají pro svou zvláštní důležitost pojmenování, v textu se budou často vyskytovat. Proto uvádíme přehled běžně užívané terminologie.

- Matici o jednom sloupci nazýváme *sloupcový vektor*.
- Matici o jednom řádku nazýváme *řádkový vektor*.
- Sloupcový vektor, jehož všechny prvky jsou 1 nazýváme *jednotkový vektor*, značíme $\mathbf{1}$.
- Matici, jejíž každý prvek je 0 nazýváme *nulová matice*, značíme $\mathbf{0}$.
- Matici typu $n \times n$ nazýváme *čtvercová matice* (*řádu* n). Prvky a_{ii} nazýváme *diagonální*.
- Čtvercovou matici splňující $a_{ij} = a_{ji}$ pro každé i, j nazýváme *symetrická matice*.
- Čtvercovou matici splňující $a_{ij} = 0$ pro každé $i > j$ nazýváme *horní trojúhelníková matice* (analogicky definujeme *dolní trojúhelníkovou matici*).
- Čtvercovou matici, jejíž všechny nenulové prvky jsou diagonální, nazýváme *diagonální matice*.
- Diagonální matici splňující $a_{ii} = 1$ pro každé i nazýváme *jednotková matice*, značíme $\mathbf{I}$.

V textu se budeme držet standardního značení, pro matice budeme používat velká tučná písmena ($\mathbf{A}$, $\mathbf{B}$, $\ldots$), pro vektory malá tučná písmena ($\mathbf{a}$, $\mathbf{b}$, $\ldots$). Nebude-li uvedeno jinak, máme na mysli vektorem vektor sloupcový. Bude-li třeba zdůraznit rozměry matice resp. vektoru, budeme značit $\mathbf{A}_{mn}$ (m počet řádků, n počet sloupců) resp. $\mathbf{a}_m$.

Shrneme si rovněž základní maticové operace.

- Buď $k \in \mathbb{R}$. k-*násobkem* matice $\mathbf{A}$ rozumíme matici, jejíž prvky jsou k-násobky prvků matice $\mathbf{A}$ (tj. $k\mathbf{A} = (ka_{ij})$).
- Buďte $\mathbf{A}$ a $\mathbf{B}$ matice stejného typu. *Součtem* matic $\mathbf{A}$ a $\mathbf{B}$ rozumíme matici $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$.
- Buď $\mathbf{A}$ matice typu $m \times n$ a $\mathbf{B}$ matice typu $n \times p$. *Součinem* matic $\mathbf{A}$ a $\mathbf{B}$ rozumíme matici $\mathbf{AB} = (\sum_{l=1}^{n} a_{il} b_{lj})$ (typu $m \times p$).
- *Transponovanou maticí* k matici $\mathbf{A}$ typu $m \times n$ rozumíme matici $\mathbf{A}^T = (a_{ji})$ typu $n \times m$.
- Buď $\mathbf{A}$ čtvercová matice. Existuje-li matice $\mathbf{A}^{-1}$ splňující $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, nazveme ji *inverzní matice* k matici $\mathbf{A}$, o matici $\mathbf{A}$ pak řekneme, že je *invertibilní*. Inverzní matice je určena jednoznačně.

Věta 1. *(Vlastnosti transponované matice.) Buďte* $\mathbf{A}$ *a* $\mathbf{B}$ *matice takového typu, aby následující výrazy byly definovány. Pak platí:*

(i) $(\mathbf{A}^T)^T = \mathbf{A}$.

(ii) $(k\mathbf{A})^T = k\mathbf{A}^T$, $k \in \mathbb{R}$.

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

(iv) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

(v) Je-li $\mathbf{A}$ *invertibilní matice, pak* $\mathbf{A}^T$ *je rovněž invertibilní a platí* $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

Věta 2. *(Vlastnosti inverzní matice.) Buďte* $\mathbf{A}$ *a* $\mathbf{B}$ *invertibilní matice. Pak platí:*

(i) $(k\mathbf{A})^{-1} = k^{-1}\mathbf{A}^{-1}$, $k \in \mathbb{R}\backslash\{0\}$.

(ii) Matice $\mathbf{AB}$ *je invertibilní a platí* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

(iii) Matice $\mathbf{A}^{-1}$ *je invertibilní a platí* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

Definice. Matici $\mathbf{A}$ nazýváme *bloková matice*, pokud platí

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

$\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ a $\mathbf{E}$ jsou matice takové, že $\mathbf{B}$ a $\mathbf{C}$ mají stejný počet řádků, $\mathbf{B}$ a $\mathbf{D}$ mají stejný počet sloupců. Nejčastěji jsou $\mathbf{B}$ a $\mathbf{E}$ matice čtvercové, není to však podmínkou.

Poznámka. Předchozí definici používáme často, užitečný je například zápis matice jako sloupcového vektoru řádků (nebo řádkového vektoru sloupců).

1.2 Základní vlastnosti matic

Definice. Buď $\mathbf{A}$ čtvercová matice řádu n. Výraz $a_{11} + ... + a_{nn} = \sum a_{ii}$ (součet diagonálních prvků $\mathbf{A}$) nazýváme *stopa matice* $\mathbf{A}$, označujeme $\mathrm{tr}(\mathbf{A})$.

Věta 3. *(Vlastnosti stopy matice.) Buďte $\mathbf{A}$ a $\mathbf{B}$ čtvercové matice stejného řádu. Pak platí:*

(i) $\mathrm{tr}(k\mathbf{A}) = k\mathrm{tr}(\mathbf{A}),\ k \in \mathbb{R}.$

(ii) $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A}).$

(iii) $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B}).$

(iv) Buďte $\mathbf{A}$ matice typu $m \times n$ a $\mathbf{B}$ matice typu $n \times m$. Pak platí $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} b_{ji}.$

Definice. Buď $\mathbf{A}$ čtvercová matice řádu n. Číslo

$$\det(\mathbf{A}) = \sum_{P \in \mathbb{S}_n} \mathrm{zn}(P) \prod_{i=1}^{n} a_{P(i)i},$$

kde $\mathbb{S}_n$ je množina všech permutací n-prvkové množiny a $\mathrm{zn}(P)$ je znaménko permutace P, nazýváme *determinant matice* $\mathbf{A}$. Podrobnosti viz [2, kap. 14].

Definice. Řekneme, že matice $\mathbf{A}$ je *regulární*, pokud $\det(\mathbf{A}) \neq 0$. Řekneme, že matice je *singulární*, pokud $\det(\mathbf{A}) = 0$.

Věta 4. *(Vlastnosti determinantu matice.) Mimo [2] a [3] také [12, dod. A].*

(i) $\det(k\mathbf{A}) = k^n \det(\mathbf{A}).$

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

(iii) $\det(\mathbf{A}^{-1}) = (\det(\mathbf{A}))^{-1}.$

(iv) $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}).$

(v) Je-li $\mathbf{A}$ horní trojúhlelníková matice, dolní trojúhelníková matice nebo diagonální matice, pak $\det(\mathbf{A}) = \prod a_{ii}.$

(vi) Buďte $\mathbf{A}$ a $\mathbf{B}$ čtvercové matice libovolného řádu, pak platí:

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} = \det(\mathbf{A}) \det(\mathbf{B}).$$

(vii) $$\det \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{pmatrix} = \det(\mathbf{B}) \det(\mathbf{E} - \mathbf{D}\mathbf{B}^{-1}\mathbf{C}) = \det(\mathbf{E}) \det(\mathbf{B} - \mathbf{C}\mathbf{E}^{-1}\mathbf{D}).$$

(viii) Buďte $\mathbf{A}$ regulární matice řádu m, $\mathbf{B}$ matice typu $m \times n$ a $\mathbf{C}$ matice typu $n \times m$, pak platí:

$$\det(\mathbf{A} + \mathbf{BC}) = \det(\mathbf{A}) \det(\mathbf{I}_{mm} + \mathbf{A}^{-1}\mathbf{BC}) = \det(\mathbf{A}) \det(\mathbf{I}_{nn} + \mathbf{C}\mathbf{A}^{-1}\mathbf{B}).$$

Definice. Řekneme, že vektory $\mathbf{x}_1, ..., \mathbf{x}_n$ jsou *lineárně nezávislé*, pokud neexistují konstanty $k_1, ..., k_n$, z nichž je aspoň jedna nenulová, aby platilo:

$$k_1\mathbf{x}_1 + ... + k_n\mathbf{x}_n = \mathbf{0}.$$

Definice. *Lineární kombinací* vektorů $\mathbf{x}_1, ..., \mathbf{x}_n$ rozumíme libovolný vektor tvaru $k_1\mathbf{x}_1 + ... + k_n\mathbf{x}_n$, $k_i \in \mathbb{R}$. Jsou-li všechna k_i rovna nule, řekneme, že lineární kombinace je *triviální*.

Definice. *Hodnost matice* definujeme jako nejvyšší počet lineárně nezávislých sloupců matice $\mathbf{A}$, značíme $\mathrm{h}(\mathbf{A})$.

Věta 5. *(Vlastnosti hodnosti matice.) [12, dod. A].*

(i) $0 \leq \mathrm{h}(\mathbf{A}_{mn}) \leq \min(m, n)$.

(ii) $\mathrm{h}(\mathbf{A}) = \mathrm{h}(\mathbf{A}^T)$.

(iii) $\mathrm{h}(\mathbf{A} + \mathbf{B}) \leq \mathrm{h}(\mathbf{A}) + \mathrm{h}(\mathbf{B})$.

(iv) $\mathrm{h}(\mathbf{AB}) \leq \min(\mathrm{h}(\mathbf{A}), \mathrm{h}(\mathbf{B}))$.

(v) $\mathrm{h}(\mathbf{A}) = \mathrm{h}(\mathbf{A}^T\mathbf{A}) = \mathrm{h}(\mathbf{A}\mathbf{A}^T)$.

(vi) *Nechť* $\mathrm{h}(\mathbf{B}_{nn}) = n$ *a* $\mathrm{h}(\mathbf{C}_{pp}) = p$, *potom* $\mathrm{h}(\mathbf{BAC}) = \mathrm{h}(\mathbf{A})$.

Definice. Buď $\mathbf{A}$ čtvercová matice řádu n. Polynom $\det(\mathbf{A} - \lambda\mathbf{I})$ nazýváme *charakteristický polynom matice* $\mathbf{A}$. Jeho (obecně komplexní) kořeny $\lambda_1, ..., \lambda_n$ nazýváme *vlastní hodnoty (čísla) matice* $\mathbf{A}$ (může nastat $\lambda_i = \lambda_j$, ačkoli $i \neq j$), n-tici $\lambda_1, ..., \lambda_n$ nazýváme *spektrum* matice $\mathbf{A}$. Nenulový vektor $\mathbf{u}$ splňující $\mathbf{Au} = \lambda_i\mathbf{u}$ nazýváme *vlastní vektor matice* $\mathbf{A}$ (*příslušný* λ_i).

Věta 6. *(Vlastnosti vlastních čísel a vektorů.) Buď* $\mathbf{A}$ *čtvercová matice řádu* n, $\lambda_1, ..., \lambda_n$ *její vlastní hodnoty a* $\mathbf{u}_i$ *vlastní vektor příslušný* $\lambda_i, i \in \{1, ..., n\}$. *Pak platí:*

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

(ii) $\prod_{i=1}^{n} \lambda_i = \det(\mathbf{A})$.

(iii) $k\mathbf{u}_i$, $k \in \mathbb{R}\backslash\{0\}$ *je rovněž vlastním vektorem příslušným* λ_i.

(iv) *Matice* $k\mathbf{A}$ *(*$k \in \mathbb{R}\backslash\{0\}$*) má vlastní čísla* $k\lambda_1, ..., k\lambda_n$ *a vlastní vektory* $\mathbf{u}_i$.

(v) *Matice* $\mathbf{A}^k$ *má vlastní čísla* $\lambda_1^k, ..., \lambda_n^k$ *a vlastní vektory* $\mathbf{u}_i$.

(vi) *Pokud je* $\mathbf{A}$ *invertibilní, pak* $\mathbf{A}^{-1}$ *má vlastní čísla* $\lambda_1^{-1}, ..., \lambda_n^{-1}$ *a vlastní vektory* $\mathbf{u}_i$.

(vii) *Matice* $\mathbf{A}^T$ *má vlastní čísla* $\lambda_1, ..., \lambda_n$.

(viii) *Vlastní čísla horní trojúhelníkové matice, dolní trojúhelníkové matice a diagonální matice jsou diagonální prvky matice, tj* $\lambda_1 = a_{11}, ..., \lambda_n = a_{nn}$.

(ix) *Buď* $\mathbf{B}$ *čtvercová matice řádu* n. *Pak vlastní čísla matic* $\mathbf{AB}$ *a* $\mathbf{BA}$ *jsou stejná.*

(x) *Jsou-li* $\mathbf{u}$ *a* $\mathbf{v}$ *vlastní vektory matice* $\mathbf{A}$ *příslušné* λ_i*, pak i vektor* $\mathbf{u}+\mathbf{v}$ *je vlastní vektor matice* $\mathbf{A}$ *příslušný* λ_i*.*

Věta 7. *(Souvislost mezi uvedenými pojmy.) Buď* $\mathbf{A}$ *čtvercová matice řádu* n*. Pak následující tvrzení jsou ekvivalentní:*

(i) $\mathbf{A}$ *je regulární* ($\det(\mathbf{A}) \neq 0$)*.*

(ii) $\mathbf{A}$ *má hodnost* n*.*

(iii) $\mathbf{A}$ *je invertibilní.*

(iv) 0 *není vlastní číslo matice* $\mathbf{A}$*.*

(v) *Sloupce (či řádky uvažované jako sloupcové vektory) matice* $\mathbf{A}$ *jsou lineárně nezávislé.*

Definice. Symbol *Kroneckerovo delta* (δ_{ij}) definujeme $\delta_{ij}=0$ pro $i \neq j$, $\delta_{ij}=1$ pro $i=j$ (tedy $\mathbf{I}=(\delta_{ij})$).

Definice. *Skalární součin* vektorů $\mathbf{x}_n$ a $\mathbf{y}_n$ definujeme jako $\mathbf{x}^T\mathbf{y}=\sum_{i=1}^{n} x_i y_i$.

Poznámka. Je zásadní rozdíl mezi výrazy $\mathbf{x}^T\mathbf{y}$ (skalární součin - reálné číslo) a $\mathbf{x}\mathbf{y}^T$ (matice typu $n \times n$).

Poznámka. V práci nedefinujeme základní pojmy teorie vektorových prostorů, jako je dimenze, báze, podprostor, generátory atd. Užijeme je totiž pouze okrajově (v několika důkazech). V základní literatuře [2], [3] jsou samozřejmě tyto definice velmi podrobně probrány.

2. Matice užívané ve statistice

2.1 Idempotentní matice

Definice. Čtvercovou matici $\mathbf{A}$, pro niž platí $\mathbf{A}^2 = \mathbf{A}$, nazýváme *idempotentní matice.*

Příklad. Diagonální matice, jejíž diagonální prvky jsou čísla 1 nebo 0 v libovolném množství je idempotentní (tedy i $\mathbf{0}$, $\mathbf{I}$). Příkladem nediagonální matice, která je idempotentní, je $\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$, je totiž $\frac{1}{n}\mathbf{1}\mathbf{1}^T\frac{1}{n}\mathbf{1}\mathbf{1}^T = \frac{1}{n}\mathbf{1}(\frac{1}{n}\mathbf{1}^T\mathbf{1})\mathbf{1}^T = \frac{1}{n}\mathbf{1}\mathbf{1}^T$.

Věta 8. *(Vlastnosti idempotentní matice.) Buď $\mathbf{A}_{nn}$ idempotentní matice. Pak platí:*

(i) $\mathbf{I} - \mathbf{A}$, $\mathbf{A}^T$ *jsou idempotentní matice.*

(ii) $\mathrm{tr}(\mathbf{A}) = \mathrm{h}(\mathbf{A})$ *[1, kap. 4].*

(iii) $\mathrm{h}(\mathbf{A}) + \mathrm{h}(\mathbf{I} - \mathbf{A}) = n$.

(iv) Je-li $\mathbf{A}$ *symetrická, je pozitivně semidefinitní (viz část 2.5)[1, kap. 4].*

(v) Vlastní čísla matice $\mathbf{A}$ *mohou být pouze 1 nebo 0. Násobnost 1 je rovna* $\mathrm{h}(\mathbf{A})$*, násobnost 0 je rovna* $n - \mathrm{h}(\mathbf{A})$ *[6, věta 21.8.2].*

Důkaz.

(i) $(\mathbf{I} - \mathbf{A})^2 = \mathbf{I} - 2\mathbf{A} + \mathbf{A}^2 = \mathbf{I} - 2\mathbf{A} + \mathbf{A} = \mathbf{I} - \mathbf{A}$. S využitím vlastnosti (iv) transponované matice platí $(\mathbf{A}^T)^2 = \mathbf{A}^T\mathbf{A}^T = (\mathbf{A}\mathbf{A})^T = (\mathbf{A})^T$.

(ii) Nechť $\mathrm{h}(\mathbf{A}) = r \neq 0$ (pokud $r = 0$, je uvažovaná rovnost triviální). Nejdříve využijeme toho, že matici $\mathbf{A}$ lze zapsat jako součin matic $\mathbf{B}_{nr}\mathbf{C}_{rn}$, jejichž hodnost je r (toto tvrzení je dokázané v kapitole 3, věta o skeletním rozkladu). Z vlastnosti (v) hodnosti matice víme, že $\mathrm{h}(\mathbf{B}^T\mathbf{B}) = r$, tedy matice je invertibilní, proto můžeme označit $\mathbf{L} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$. Platí $\mathbf{LB} = \mathbf{I}$. Obdobně můžeme najít matici $\mathbf{P}$, pro kterou platí $\mathbf{CP} = \mathbf{I}$. Protože matice $\mathbf{A}$ je idempotentní, platí $\mathbf{A} = \mathbf{A}^2 = \mathbf{BCBC} = \mathbf{BC}$. Uvažujme následující dvě uzávorkování výrazu $\mathbf{LBCBCP}$ a upravme:

$$\mathbf{L}(\mathbf{BCBC})\mathbf{P} = \mathbf{L}(\mathbf{BC})\mathbf{P} = (\mathbf{LB})(\mathbf{CP}) = \mathbf{II} = \mathbf{I}$$

$$(\mathbf{LB})\mathbf{CB}(\mathbf{CP}) = \mathbf{ICBI} = \mathbf{CB}.$$

Tedy $\mathbf{CB} = \mathbf{I}_{rr}$. Využitím vlastnosti (iv) stopy matice dostáváme: $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(BC) = \mathrm{tr}(\mathbf{CB}) = \mathrm{tr}(\mathbf{I}) = r = \mathrm{h}(\mathbf{A})$ [1, kap. 4].

(iii) S využitím předchozích dvou bodů a vlastností stopy matice platí $\mathrm{h}(\mathbf{A}) + \mathrm{h}(\mathbf{I} - \mathbf{A}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{I} - \mathbf{A}) = \mathrm{tr}(\mathbf{A} + \mathbf{I} - \mathbf{A}) = \mathrm{tr}\mathbf{I} = n$.

(iv) Buď $\mathbf{x}_n$ libovolně zvolený vektor. S postupným využitím předpokladu idempotentnosti, symetrie a vlastnosti (iv) transponované matice dostáváme: $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}\mathbf{A}^2\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) \geq 0$ [1, kap. 4].

(v) Předpokládejme, že λ je vlastní číslo matice $\mathbf{A}$, $\mathbf{u}$ vlastní vektor náležící λ. Pak platí: $\lambda\mathbf{u} = \mathbf{Au} = \mathbf{A}^2\mathbf{u} = \mathbf{A}(\mathbf{Au}) = \mathbf{A}(\lambda\mathbf{u}) = \lambda\mathbf{Au} = \lambda^2\mathbf{u}$. Protože $\mathbf{u}$ je nenulový, platí $\lambda = \lambda^2$, tedy $\lambda = 1$ nebo $\lambda = 0$. Násobnost plyne z (ii) a obecně platné rovnosti součtu vlastních čísel a stopy matice [6, věta 21.8.2].

□

Poznámka. Buď $\mathbf{A}$ regulární idempotentní matice. Pak $\mathbf{A}$ je nutně jednotková. Můžeme totiž psát $\mathbf{A} = \mathbf{AI} = \mathbf{AAA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$.

Příklad. Podle vlastnosti (i) a předchozího příkladu je matice $\mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}^T$ idempotentní, zřejmě je také symetrická. Její hodnost (a stopa) je $n-1$. Navíc má následující vlastnosti: pro libovolný vektor $\mathbf{x}_n$ platí:

$$\mathbf{x}^T(\mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}^T)\mathbf{x} = \mathbf{x}^T\mathbf{x} - \frac{1}{n}\mathbf{x}^T\mathbf{1}\mathbf{1}^T\mathbf{x} = \sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2 = \sum_{i=1}^{n} x_i^2 - 2n\bar{x}_n^2 + n\bar{x}_n^2 =$$

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

kde $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$ (aritmetický průměr x_i). Také platí:

$$(\mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}^T)\mathbf{x} = \mathbf{x} - \frac{1}{n}\sum_{i=1}^{n} x_i\mathbf{1} = \mathbf{x} - \bar{x}_n\mathbf{1},$$

výsledkem je tedy vektor odchylek od průměru [12, dod. A].

2.2 Ortogonální matice

S pojmem ortogonální matice se student do hloubky seznámil na úvodní přednášce lineární algebry. Přesto si uvedeme několik nových vlastností. Ortogonální matice je důležitým pojmem v kapitole 3.

Definice. Řekneme, že čtvercová matice $\mathbf{A}$ je *ortogonální*, pokud platí $\mathbf{AA}^T = \mathbf{I}$ (tedy $\mathbf{A}^T = \mathbf{A}^{-1}$).

Definice. Řekneme, že vektory $\mathbf{x}$ a $\mathbf{y}$ jsou navzájem *ortogonální*, je-li $\mathbf{x}^T\mathbf{y} = 0$. Pokud je navíc $\mathbf{x}^T\mathbf{x} = 1$ a $\mathbf{y}^T\mathbf{y} = 1$ řekneme, že jsou *ortonormální*. Řekneme, že množina vektorů je ortogonální (ortonormální), jsou-li všechny navzájem různé dvojice vektorů v ní obsažené ortogonální (ortonormální).

Poznámka. Množina ortogonálních vektorů je lineárně nezávislá [2, věta 26.13].

Poznámka. Pokud je libovolná z matic $\mathbf{A}, \mathbf{A}^T, \mathbf{A}^{-1}$ ortogonální jsou ortogonální i zbývající dvě [2, věta 27.8]. Ortogonální matice je regulární, platí $|\det(\mathbf{A})| = 1$ [2, věta 27.7]. Označme sloupce (či řádky) ortogonální matice $\mathbf{A}$ uvažované jako sloupcové vektory $\mathbf{a}_1, ..., \mathbf{a}_n$, pak $\mathbf{a}_i^T\mathbf{a}_j = \delta_{ij}$ (tedy je to množina ortonormálních vektorů).

Věta 9. *Buďte* $\mathbf{A}, \mathbf{B}$ *ortogonální matice. Pak platí:*

(i) Matice $\mathbf{AB}$ *je ortogonální.*

(ii) **A** *nemůže mít jiné reálné vlastní číslo než* 1 *nebo* -1 *[6, věta 21.8.1].*

Důkaz.

(i) S využitím vlastnosti (iv) transponované matice máme: $(\mathbf{AB})(\mathbf{AB})^T = \mathbf{ABB}^T\mathbf{A}^T = \mathbf{AA}^T = \mathbf{I}$ (poslední dvě rovnosti plynou z předpokladu ortogonality matic).

(ii) Buď λ reálné vlastní číslo **A**. Pak existuje nenulový vektor **u** takový, že platí $\mathbf{Au} = \lambda\mathbf{u}$, můžeme psát: $\mathbf{u}^T\mathbf{u} = \mathbf{u}^T\mathbf{Iu} = \mathbf{u}^T\mathbf{A}^T\mathbf{Au} = (\mathbf{Au})^T(\mathbf{Au}) = (\lambda\mathbf{u})^T(\lambda\mathbf{u}) = \lambda^2\mathbf{u}^T\mathbf{u}$. $\mathbf{u}^T\mathbf{u}$ je nenulové, tedy $1 = \lambda^2$ [6, věta 21.8.1].

□

Poznámka. (O permutačních maticích.) Příkladem ortogonální matice je libovolná čtvercová matice, která má v každém řádku a v každém sloupci právě jeden prvek roven 1 a všechny ostatní nulové. Takové matici říkáme *permutační matice*. Označme $\mathbf{P}_{nn}$ permutační matici a uvažujme libovolnou matici $\mathbf{A}_{np}$. Nechť je v řádku r permutační matice číslo 1 ve sloupci s. Pak r-tým řádkem **PA** je s-tý řádek matice **A** (řádky matice **PA** jsou permutované řádky matice **A**). Obdobně lze ukázat, že násobíme-li vhodnou matici **B** permutační maticí zprava, součinem je matice jejíž sloupce jsou permutované sloupce **B**.

Poznámka. (O rozložitelnosti matic.) Uvažujme čtvercovou matici $\mathbf{C}_{nn}$. O matici **C** řekneme, že je *reducibilní* (*rozložitelná*), existuje-li permutační matice **P** taková, že $\mathbf{PCP}^T$ je ve tvaru $\mathbf{C} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{E} & \mathbf{F} \end{pmatrix}$, kde matice $\mathbf{D}, \mathbf{F}$ jsou čtvercové. Není-li matice reducibilní, nazýváme ji *ireducibilní* (*nerozložitelná*) [14, kap 1c.3].

2.3 Stochastická matice

Základním pojmem teorie Markovových řetězců je stochastická matice. Z pohledu maticové algebry má několik zajímavých vlastností, které si ukážeme.

Definice. *Stochastická matice* je čtvercová matice splňující $a_{ij} \geq 0$, $\sum_{j=1}^{n} a_{ij} = 1$ pro všechna $i \in \{1, ..., n\}$. *Stochastický vektor* definujeme jako vektor, jehož složky jsou nezáporné a jejich součet je 1.

Poznámka. Někdy se rozlišuje *pravá* a *levá* stochastická matice. Pravá stochastická matice je uvedena v definici, pro levou stochastickou maticí je podmínka $\sum_{j=1}^{n} a_{ij} = 1$ nahrazena podmínkou $\sum_{i=1}^{n} a_{ij} = 1$. Také existuje *dvojitě* stochastická matice (splňující obě podmínky).

Věta 10. *(Vlastnosti stochastické matice.) Buďte* **A** *a* **B** *stochastické matice, pak platí:*

(i) Matice **AB** *je stochastická. Tedy speciálně* $\mathbf{A}^2, \mathbf{A}^3, ...$ *jsou stochastické matice.*

(ii) Každé vlastní číslo λ *matice* **A** *splňuje* $|\lambda| \leq 1$.

(iii) 1 je vlastním číslem matice **A**. *[10, kap. 6]*

Důkaz.

(i) Nezápornost prvků matice **AB** je zřejmá. Protože matice **A** a **B** jsou stochastické, platí $\sum_{i=1}^{n} b_{ji} = 1, \forall j \in \{1, ..., n\}$ a $\sum_{j=1}^{n} a_{kj} = 1, \forall k \in \{1, ..., n\}$. Ověřme pro libovolné k, že součet prvků v k-tém řádku matice **AB** je 1. Tuto podmínku můžeme zapsat jako $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{kj} b_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{kj} b_{ji} = \sum_{j=1}^{n} a_{kj} \sum_{i=1}^{n} b_{ji} = \sum_{j=1}^{n} a_{kj} = 1$.

(ii) Uvažujme vlastní číslo λ a jemu příslušný vlastní vektor **u**. Označme l index takový, že platí $|x_l| \geq |x_j|, \forall j \in \{1, ..., n\}$. Ukážeme, že platí $|\lambda - a_{ll}| \leq \sum_{j \neq l} |a_{lj}|$ (toto tvrzení je známo pod názvem Geršgorinova věta). Máme $\mathbf{Au} = \lambda \mathbf{u}$, psáno po složkách: $\sum_{j=1}^{n} a_{ij} u_j = \lambda u_i, \forall i \in \{1, ..., n\}$. Speciálně tedy $\sum_{j=1}^{n} a_{lj} u_j = \lambda u_l$, odečtením členu $a_{ll} u_l$ dostáváme $\sum_{j \neq l}^{n} a_{lj} u_j = \lambda u_l - a_{ll} u_l$. Aplikujme na rovnici absolutní hodnotu a vydělme ji u_l, dostáváme $|\lambda - a_{ll}| = |u_l^{-1} \sum_{j \neq l} a_{lj} u_j| \leq |\sum_{j \neq l} a_{lj}|$, protože $|u_j u_l^{-1}| \leq 1$. Nyní již podle trojúhelníkové nerovnosti, Geršgorinovy věty a z předpokladu, že **A** je stochastická, platí:

$$|\lambda| \leq |\lambda - a_{ll}| + |a_{ll}| \leq \sum_{j \neq l} |a_{lj}| + |a_{ll}| = \sum_{j=1}^{n} |a_{lj}| = 1.$$

(iii) Tvrdíme, že platí $\mathbf{Au} = \lambda \mathbf{u}$ pro $\lambda = 1$ a jistý vektor **u**. Uvažujme $\mathbf{u} = \mathbf{1}$, uvedená rovnost platí, protože $\sum_{j=1}^{n} a_{ij} 1 = 1, \forall i \in \{1, ..., n\}$. Tedy **1** je vlastní vektor matice příslušný vlastnímu číslu 1. [10, kap. 6]

□

Poznámka. (Perronova a Frobeniova věta.) Každá *kladná* (tj. $a_{ij} > 0 \, \forall i, j$) čtvercová matice **A** má kladné vlastní číslo λ_M splňující: λ_M je jednoduchým kořenem charakteristické rovnice, λ_M je větší než absolutní hodnoty všech ostatních vlastních čísel, vlastní vektor příslušný λ_M má všechny složky kladné. S ohledem na předchozí větu je pro stochastické matice $\lambda_M = 1$. Je-li matice **A** ireducibilní nezáporná (oslabení předpokladu), pak věta pořád platí, jen λ_M je větší nebo rovno absolutní hodnotě všech ostatních vlastních čísel. Ireducibilita matice není vždy patrná, pokud ale najdeme nějakou mocninu m matice **A** takovou, že $\mathbf{A}^m$ je kladná, splníme silnější předpoklad a s využitím vlastnosti (vii) vlastních čísel matice dostáváme informaci o vlastních číslech a vektorech matice **A**. Podrobněji viz [14, kap. 1c.3], náznak důkazu je v [13, kap. 14].

Poznámka. (Souvislost s Markovovy řetězci.) Stacionární rozdělení je *levým* vlastním vektorem příslušným vlastnímu číslu 1 (kde levý vlastní vektor je řádkový vektor **v** splňující pro nějaké λ rovnost $\mathbf{vA} = \lambda \mathbf{v}$). Markovův řetězec je rozložitelný, je-li rozložitelná jeho matice.

2.4 Symetrická matice

V úvodní kapitole jsme již zmínili pojem symetrické matice ($\mathbf{A} = \mathbf{A}^T$). Požadavek symetrie je nutný při spektrálním rozkladu matice (viz kapitola 3). Zde zmíníme další užitečné vlastnosti symetrické matice.

Věta 11. *(Vlastnosti symetrické matice.) Buď* $\mathbf{A}$ *čtvercová matice řádu* n*, pak platí:*

(i) Pro libovolné přirozené p *a pro každou matici* $\mathbf{X}_{pn}$ *je matice* $\mathbf{XAX}^T$ *symetrická. Speciálně pro každou matici* $\mathbf{B}$ *platí, že* $\mathbf{BB}^T$ *je symetrická matice.*

(ii) Všechna vlastní čísla $\mathbf{A}$ *jsou reálná [3, věta 15.10].*

(iii) Vlastní vektory odpovídající navzájem různým vlastním číslům matice $\mathbf{A}$ *jsou navzájem ortogonální [6, věta 21.4.5].*

Důkaz.

(i) $(\mathbf{XAX}^T)^T = (\mathbf{X}^T)^T\mathbf{A}^T\mathbf{X}^T = \mathbf{XA}^T\mathbf{X}^T = \mathbf{XAX}^T$. Kde první rovnost vyplývá z opakovaného užití vlastnosti (iv) transponované matice a poslední ze symetrie matice $\mathbf{A}$. Ve speciálním případě stačí uvažovat $\mathbf{BIB}^T = \mathbf{BB}^T$, kde $\mathbf{I}$ je jednotková matice patřičného řádu.

(ii) Viz [3, věta 15.10]

(iii) Buďte $\lambda_1 \neq \lambda_2$ vlastní čísla matice $\mathbf{A}$ a $\mathbf{u}_1, \mathbf{u}_2$ jim odpovídající vlastní vektory. Můžeme psát $\lambda_1\mathbf{u}_1^T\mathbf{u}_2 = (\lambda\mathbf{u}_1)^T\mathbf{u}_2 = (\mathbf{Au}_1)^T\mathbf{u}_2 = \mathbf{u}_1^T\mathbf{A}^T\mathbf{u}_2 = \mathbf{u}_1^T\mathbf{Au}_2 = \mathbf{u}_1^T(\lambda_2\mathbf{u}_2) = \lambda_2\mathbf{u}_1^T\mathbf{u}_2$. Vzhledem k tomu, že $\lambda_1 \neq \lambda_2$, nutně $\mathbf{u}_1^T\mathbf{u}_2 = 0$ [6, věta 21.4.5].

□

2.5 Pozitivně (semi)definitní matice

Pojem pozitivně semidefinitní matice je ve statistice nezbytný - každá rozptylová matice je pozitivně semidefinitní. Více se o této matici dozvíme také v kapitole 3. Čerpali jsme z [6, kap. 14 a 21].

Definice. Buď $\mathbf{A}_{nn}$ symetrická čtvercová matice. Pokud pro každý nenulový vektor $\mathbf{x}_n$ platí $\mathbf{x}^T\mathbf{Ax} = \sum_{ij} a_{ij}x_ix_j > 0$ řekneme, že matice $\mathbf{A}$ je *pozitivně definitní* (značíme $\mathbf{A} > 0$). Pokud pro každý nenulový vektor $\mathbf{x}$ o n složkách platí $\mathbf{x}^T\mathbf{Ax} \geq 0$ řekneme, že matice $\mathbf{A}$ je *pozitivně semidefinitní* (značíme $\mathbf{A} \geq 0$). Obdobně lze definovat *negativně definitní* a *negativně semidefinitní* matice.

Věta 12. *(Vlastnosti pozitivně definitní a pozitivně semidefinitní matice.)*

(i) Pro libovolnou matici $\mathbf{A}$ *platí* $(\mathbf{AA}^T) \geq 0$, $(\mathbf{A}^T\mathbf{A}) \geq 0$.

(ii) Vlastní čísla pozitivně definitní matice jsou kladná.

(iii) Vlastní čísla pozitivně semidefinitní matice jsou nezáporná.

(iv) Pozitivně definitní matice je invertibilní a její inverzní matice je rovněž pozitivně definitní.

(v) Nechť $\mathbf{A}_{nn} \geq 0$*, pak pro každou matici* $\mathbf{B}_{nm}$ *platí* $\mathbf{B}^T\mathbf{AB} \geq 0$*. Pokud navíc* $\mathbf{A} > 0$ *a matice* $\mathbf{B}$ *je čtvercová a regulární, pak* $\mathbf{B}^T\mathbf{AB} > 0$ *[12, dodatek A].*

Důkaz.

(i) S využitím vlastnosti (iv) transponované matice můžeme psát $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T\mathbf{A}\mathbf{x} \geq 0$, kde $\mathbf{x}$ je vektor o odpovídajícím počtu složek. Obdobně lze dokázat pro $\mathbf{A}\mathbf{A}^T$.

(ii) Pozitivně definitní matice je symetrická, tedy víme z vlastnosti (ii) symetrické matice, že její čísla jsou reálná. Buď tedy λ vlastní číslo matice $\mathbf{A}$ a $\mathbf{u}$ jemu příslušný vlastní vektor. Víme $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$, tedy $\mathbf{u}^T\mathbf{A}\mathbf{u} = \lambda\mathbf{u}^T\mathbf{u}$. Proto $\lambda = (\mathbf{u}^T\mathbf{A}\mathbf{u})/(\mathbf{u}^T\mathbf{u})$. Z definice je vlastní vektor nenulový a tudíž je jmenovatel kladný. Matice je pozitivně definitní a tedy $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pro každý nenulový $\mathbf{x}$, tedy speciálně i pro $\mathbf{u}$, tedy čitatel je kladný.

(iii) Obdobně jako (ii).

(iv) Stačí využít vlastnosti (vi) vlastních čísel. Invertibilita matice plyne z nenulovosti vlastních čísel.

(v) Zřejmě pro každý nenulový vektor $\mathbf{x}_m$ platí $\mathbf{x}^T\mathbf{B}^T\mathbf{A}\mathbf{B}\mathbf{x} = (\mathbf{B}\mathbf{x})^T\mathbf{A}(\mathbf{B}\mathbf{x}) \geq 0$, tedy $\mathbf{B}^T\mathbf{A}\mathbf{B} \geq 0$. Druhá část tvrzení plyne z faktu $\mathbf{B}\mathbf{x} \neq 0$ (neboť $\mathbf{B}$ je regulární).

□

2.6 Projekční matice

V této části si zavedeme pojem projekční matice a popíšeme její základní vlastnosti. Rovněž stručně vysvětlíme její geometrickou interpretaci. V kapitole 4 osvětlíme souvislost s metodou nejmenších čtverců. Čerpali jsme z [6, kap. 12].

Definice. Buď $\mathbf{A}_{mn}$ matice hodnosti n (sloupce jsou lineárně nezávislé vektory). Matici $\mathbf{P_A} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ nazýváme (*ortogonální*) *projekční matice* (na prostor generovaný sloupci matice $\mathbf{A}$).

Poznámka. Inverzní matice v předchozí definici existuje podle vlastnosti (v) hodnosti matice. Požadavek lineární nezávislosti sloupců matice $\mathbf{A}$ lze vypustit, pak sice inverzní matice neexistuje, ale existuje tzv. *pseudoinverzní* matice, s níž si definice zachová dobrý smysl i vlastnosti uvedené v následující větě. Více o pseudoinverzní matici lze nalézt např. v [2, kap. 31].

Věta 13. *Vlastnosti projekční matice. Buď* $\mathbf{A}_{mn}$ *matice hodnosti* n*, pak pro matici* $\mathbf{P_A} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ *platí:*

(i) $\mathbf{P_A}\mathbf{A} = \mathbf{A}$,

(ii) $\mathbf{P_A}$ *je idempotentní,*

(iii) $\mathbf{P_A}$ *je symetrická,*

(iv) $\mathrm{h}(\mathbf{P_A}) = \mathrm{h}(\mathbf{A})$.

Důkaz.

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

(ii) $\mathbf{P}_\mathbf{A}^2 = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{P_A}$.

(iii) $\mathbf{P}_\mathbf{A}^T = (\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)^T = (\mathbf{A}^T)^T((\mathbf{A}^T\mathbf{A})^{-1})^T\mathbf{A}^T = \mathbf{A}((\mathbf{A}^T\mathbf{A})^T)^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{P_A}$.

(iv) $\mathrm{h}(\mathbf{A}) = \mathrm{h}(\mathbf{P_A A}) \leq \mathrm{h}(\mathbf{P_A})$ (z vlastnosti (iv) hodnosti matice) obdobně $\mathrm{h}(\mathbf{P_A}) = \mathrm{h}(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T) \leq \mathrm{h}(\mathbf{A})$.

□

Poznámka. Uvažujme bod $\mathbf{a}$ v n-dimenzionálním prostoru a jeho podprostor dimenze m, $m < n$, který je generován sloupci jisté matice $\mathbf{A}$. Hledáme-li bod tohoto podprostoru, který je nejblíže bodu $\mathbf{a}$, pak tímto bodem je $\hat{\mathbf{a}} = \mathbf{P_A a}$ (uvažujeme vzdálenost bodů $\mathbf{x}$ a $\mathbf{y}$ definovanou jako $\sqrt{\sum(x_i - y_i)^2}$). Pro $n = 3$, $m = 2$ si jako $\mathbf{a}$ můžeme představit libovolný bod prostoru a jako $\hat{\mathbf{a}}$ patu kolmice, která byla spuštěna z bodu $\mathbf{a}$ do jisté roviny. Bodu $\hat{\mathbf{a}}$ pak říkáme *ortogonální projekce* bodu $\mathbf{a}$ na danou rovinu. Příklady lze nalézt v [7, kap. 12].

3. Rozklady matice

V této kapitole se budeme zabývat problémem tzv. rozkladu matice. Rozložení matice znamená její zapsání jako součinu několika jiných matic, které mají jisté vlastnosti. Uplatnění mnohých rozkladů je nejen teoretické (např. při důkazech), ale i praktické (při výpočtu determinantů, inverzních matic apod.) [12, dodatek A].

3.1 Skeletní rozklad

Věta 14. *(Skeletní rozklad.) Buď* $\mathbf{A}_{mn}$ *libovolná matice,* $\mathrm{h}(\mathbf{A}) = r \geq 1$*. Pak existují matice* $\mathbf{B}_{mr}$, $\mathbf{C}_{rn}$ *hodnosti* r *takové, že platí*

$$\mathbf{A} = \mathbf{B}\mathbf{C}.$$

Důkaz. Vytvořme matici **B** jako matici, jejíž sloupce jsou lineárně nezávislé sloupce matice **A** (r takových podle definice hodnosti existuje). Každý sloupec matice **A** tedy můžeme napsat jako lineární kombinaci sloupců matice **B**. Koeficienty lineární kombinace vytvářející j-tý sloupec matice **A** označme $c_{1j}, ..., c_{nj}$, vezmeme-li je za j-tý sloupec matice **C**, bude splněna požadovaná rovnost. $\mathrm{h}(\mathbf{B}) \leq r$ podle vlastnosti (i) hodnosti matice, $\mathrm{h}(\mathbf{B}) \geq r$ podle vlastnosti (iv) hodnosti matice. Stejné argumenty platí i pro matici **C**. [1, kap. 4] □

3.2 Spektrální rozklad

Poznámka. Z věty o vlastnostech vlastních čísel a vektorů (konkrétně bodů (iii) a (x)) lze snadno dovodit, že přidáme-li ke všem (reálným) vlastním vektorům příslušným jisté (reálné) vlastní hodnotě λ nulový vektor, vznikne aritmetický vektorový prostor nad $\mathbb{R}$ (viz [3, kap. 1]). Takový prostor nazýváme *vlastní prostor* matice **A** příslušný λ. *Geometrickou násobností* vlastního čísla λ rozumíme dimenzi příslušného vlastního prostoru. *Algebraickou násobností* vlastního čísla λ rozumíme násobnost λ jako kořene charakteristického polynomu matice **A**. Zřejmě součet algebraických ani geometrických násobností všech vlastních hodnot matice nemůže překročit řád matice n. Obecně platí, že algebraická násobnost vlastního čísla je větší než jeho geometrická násobnost nebo jsou si obě násobnosti rovny (důkaz viz [6, věta 21.3.4]). V následující významné větě ukážeme, že pro symetrické matice vždy nastává rovnost.

Věta 15. *(Spektrální rozklad.) Buď* **A** *symetrická čtvercová matice. Pak existují ortogonální matice* **U** *a diagonální matice* $\mathbf{\Lambda}$ *takové, že platí*

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T.$$

Navíc diagonální prvky matice $\mathbf{\Lambda}$ *jsou vlastní čísla matice* **A** *a sloupce matice* **U** *jim odpovídající vlastní vektory.*

Důkaz. Nejprve předpokládejme, že všechny vlastní čísla matice **A** mají algebraickou násobnost 1. Podle vlastnosti (ii) symetrické matice víme, že tedy existuje

n-tice navzájem různých reálných čísel $\lambda_1, ..., \lambda_n$ a jim odpovídající vlastní vektory $\mathbf{u}_1, ..., \mathbf{u}_n$, které jsou podle (i) téže věty navzájem ortogonální (a tedy lineárně nezávislé). Můžeme pro $i, j = 1, ..., n$ psát $\mathbf{A}\mathbf{u}_j = \lambda_j \mathbf{u}_j$, $\mathbf{u}_i^T \mathbf{A} \mathbf{u}_j = \lambda_j \mathbf{u}_i^T \mathbf{u}_j = \lambda_j \delta_{ij}$, v maticové formě $\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{\Lambda}$. Tuto rovnost stačí upravit následovně: $\mathbf{U}\mathbf{U}^T \mathbf{A} \mathbf{U} \mathbf{U}^T = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$, tedy $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$.

Nechť nyní nejsou všechna vlastní čísla navzájem různá. Řekněme, že existuje k ($k < n$) navzájem různých vlastních čísel s algebraickými násobnostmi $m_1, ..., m_k$. Těmto vlastním číslům odpovídají vlastní prostory $V_1, ..., V_k$, které mají dimenze $d_1, ..., d_k$. Označme $d = \sum_{i=1}^k d_i$. Můžeme tedy najít ortonormální množinu vektorů $\mathbf{u}_1, ..., \mathbf{u}_d$ (díky ortogonalitě $V_1, ..., V_k$), že $\mathbf{u}_1, ..., \mathbf{u}_{m_1}$ tvoří bázi V_1, dále $\mathbf{u}_{m_1+1}, ..., \mathbf{u}_{m_1+m_2}$ tvoří bázi V_2 atd. Po případném přečíslení můžeme dále psát $\mathbf{A}\mathbf{u}_i = \lambda \mathbf{u}_i, i = 1, ..., d$. Nyní sporem ukážeme, že není možné, aby $d < n$ (pokud $d = n$, jsme hotovi). Bez újmy na obecnosti můžeme předpokládat, že všechna vlastní čísla $\mathbf{A}$ jsou kladná. Pokud ne, můžeme totiž vzít v úvahu matici $\mathbf{A} + (|\lambda_{min}| + 1)\mathbf{I}$, kde λ_{min} je nejmenší vlastní číslo matice $\mathbf{A}$. Zřejmě tato nová matice má stejné vlastní vektory jako matice původní. Označme nyní $\mathbf{B} = \mathbf{A} - \sum_{i=1}^d \lambda_i \mathbf{u}_i \mathbf{u}_i^T$. Z vlastností stopy matice a vlastnosti (i) vlastních čísel matice můžeme psát $\mathrm{tr}(\mathbf{B}) = \mathrm{tr}(\mathbf{A}) - \sum_{i=1}^d \lambda_i (\mathbf{u}_i^T \mathbf{u}_i) = \sum_{i=d+1}^n \lambda_i \geq 0$ (stačí si uvědomit, že obecně platí $\mathbf{x}^T \mathbf{x} = \mathrm{tr}(\mathbf{x}\mathbf{x}^T)$). Tedy $\mathbf{B}$ má nějaké nenulové vlastní číslo β a jemu příslušný vlastní vektor $\mathbf{x}$. Potom pro $1 \leq j \leq d$ platí: $\beta \mathbf{u}_j^T \mathbf{x} = \mathbf{u}_j^T \mathbf{B} \mathbf{x} = (\lambda_j \mathbf{u}_j^T - \sum_{i=1}^d \lambda_i (\mathbf{u}_j^T \mathbf{u}_i) \mathbf{u}_i^T) \mathbf{x} = 0$. Což znamená, že $\mathbf{x}$ je ortogonální k $\mathbf{u}_1, ..., \mathbf{u}_d$. Zároveň ale můžeme snadno ukázat, že $\mathbf{x}$ je i vlastní vektor $\mathbf{A}$: $\beta \mathbf{x} = \mathbf{B}\mathbf{x} = (A - \sum_{i=1}^d \lambda_i \mathbf{u}_i \mathbf{u}_i^T)\mathbf{x} = \mathbf{A}\mathbf{x} - \sum_{i=1}^d \lambda_i (\mathbf{u}_i^T \mathbf{x}) \mathbf{u}_i = \mathbf{A}\mathbf{x}$. Takže β je vlastním číslem $\mathbf{A}$ a $\mathbf{x}$ jemu příslušný vlastní vektor. To ale znamená, že je obsažen v nějakém z prostorů $V_1, ..., V_k$ a nemůže být ortogonální na jeho bázi (kterou tvoří neprázdná podmnožina vektorů $\mathbf{u}_1, ..., \mathbf{u}_n$) [12, dodatek A]. $\square$

Poznámka. (O mocninách matice.) Pro symetrickou matici $\mathbf{A}$ podle předchozí věty můžeme psát $\mathbf{A}^2 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^T$, kde $\mathbf{\Lambda}^2$ lze snadno spočítat, protože $\mathbf{\Lambda}$ je diagonální matice. Podobně můžeme počítat $\mathbf{A}^n$ pro $n \in \mathbb{N}$. Dále zřejmě $\mathbf{A}^{-1} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T$ (je-li $\mathbf{\Lambda}^{-1}$ definována, tedy jsou-li diagonální prvky $\mathbf{\Lambda}$ nenulové), protože $\mathbf{A}\mathbf{A}^{-1} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T = \mathbf{I}$. Obecně můžeme definovat $\mathbf{A}^{\frac{p}{q}}$ (kde p, q jsou celá čísla), pokud je tato mocnina definována pro všechny diagonální prvky matice $\mathbf{\Lambda}$. Aplikací vlastnosti (i) symetrické matice na $\mathbf{\Lambda}^{\frac{p}{q}}$ vidíme, že $\mathbf{A}^{\frac{p}{q}}$ je opět symetrická matice.

Poznámka. (O rozptylových maticích.) Buď $\mathbf{A}$ libovolná pozitivně semidefinitní matice. Pak má smysl $\mathbf{A}^{\frac{1}{2}}$. Buď $\mathbf{X}$ náhodný vektor s jednotkovou rozptylovou maticí. Pak rozptylovou matici náhodného vektoru $\mathbf{A}^{\frac{1}{2}}\mathbf{X}$ spočítáme: $\mathrm{var}(\mathbf{A}^{\frac{1}{2}}\mathbf{X}) = \mathbf{A}^{\frac{1}{2}}\mathrm{var}(\mathbf{X})(\mathbf{A}^{\frac{1}{2}})^T = \mathbf{A}$. Což znamená, že každá pozitivně semidefinitní matice je i rozptylovou maticí nějakého náhodného vektoru.

Důsledek. *Buď* $\mathbf{A}_{nn} \geq 0$, $\mathrm{h}(\mathbf{A}) = r \geq 1$. *Pak existuje matice:*

(i) $\mathbf{B}_{nr}$ *hodnosti* r *taková, že platí*

$$\mathbf{A} = \mathbf{B}\mathbf{B}^T.$$

(ii) Symetrická $\mathbf{C}_{nn}$ *taková, že platí*

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

Důkaz.

(i) Protože $\mathbf{A}$ je pozitivně semidefinitní, je podle definice symetrická, tedy má nějaký spektrální rozklad $\mathbf{U\Lambda U}^T$. Podle vlastnosti (vi) hodnosti matice vidíme, že $\mathrm{h}(\mathbf{A}) = \mathrm{h}(\mathbf{\Lambda})$, tedy $\mathbf{\Lambda}$ má r nenulových prvků na diagonále, označme je $\lambda_1, ..., \lambda_r$. Navíc podle vlastnosti (iii) pozitivně semidefinitní matice jsou všechna tato čísla kladná. Můžeme tedy označit:

$$\tilde{\mathbf{\Lambda}}^{\frac{1}{2}}_{mr} = \begin{pmatrix} \lambda_1^{\frac{1}{2}} & 0 & \dots & 0 \\ 0 & \lambda_2^{\frac{1}{2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_r^{\frac{1}{2}} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Platí: $\mathbf{A} = \mathbf{U\Lambda U}^T = \mathbf{U}\tilde{\mathbf{\Lambda}}^{\frac{1}{2}}(\tilde{\mathbf{\Lambda}}^{\frac{1}{2}})^T\mathbf{U}^T = \mathbf{BB}^T$, kde jsme označili $\mathbf{B} = \mathbf{U}\tilde{\mathbf{\Lambda}}^{\frac{1}{2}}$. $\mathbf{U}$ je ortogonální, tedy regulární, takže $\mathbf{B}$ má opravdu hodnost r. [1, kap. 4]

(ii) Stačí zvolit $\mathbf{C} = \mathbf{U\Lambda}^{\frac{1}{2}}\mathbf{U}^T$. Symetrii jsme zdůvodnili v poznámce o mocninách matice. Rovnost $\mathbf{A} = \mathbf{C}^2$ je zřejmá. [12, dodatek A].

□

3.3 Rozklad podle singulárních hodnot

Věta 16. *(Rozklad podle singulárních hodnot.) Buď $\mathbf{A}_{mn}$ libovolná matice nenulové hodnosti r. Pak existují matice $\mathbf{L}_{rr}$, $\mathbf{U}_{mr}$, $\mathbf{V}_{nr}$ takové, že platí*

$$\mathbf{A} = \mathbf{ULV}^T.$$

Navíc $\mathbf{L}$ je diagonální matice s kladnými prvky na diagonále a $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$.

Důkaz. Pro jiný důkaz viz [6, kap. 21], tento je založený spíše na [14, kap. 1c.3] a [12, dodatek A]. Podle vlastnosti (i) pozitivně semidefinitní matice víme, že $\mathbf{A}^T\mathbf{A}$ je pozitivně semidefinitní. Tedy podle důkazu (i) předchozího důsledku a vlastnosti (v) hodnosti matice víme, že má r kladných vlastních čísel. Označme je $\lambda_1^2, ..., \lambda_r^n$. Větou o spektrálním rozkladu máme zaručenu existenci r ortogonálních vektorů $\mathbf{v}_1, ..., \mathbf{v}_r$ jim příslušných (tedy podle definice při tomto značení máme $(\mathbf{A}^T\mathbf{A})\mathbf{v}_i = \lambda_i^2\mathbf{v}_i)$. Označme $\mathbf{V}$ matici, jejíž sloupce jsou $\mathbf{v}_1, ..., \mathbf{v}_r$. Díky ortogonalitě platí $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. Dále pro $i = 1, ..., r$ označme $\mathbf{u}_i = \lambda_i^{-1}\mathbf{A}\mathbf{v}_i$ a matici, jejíž sloupce jsou tyto vektory označme $\mathbf{U}$. Ukážeme ortogonalitu množiny $\mathbf{u}_1, ..., \mathbf{u}_r$: $\mathbf{u}_i^T\mathbf{u}_j = \lambda_i^{-1}\mathbf{v}_i^T\mathbf{A}^T\lambda_j^{-1}\mathbf{A}\mathbf{v}_j = \lambda_i^{-1}\lambda_j^{-1}\mathbf{v}_i^T(\mathbf{A}^T\mathbf{A})\mathbf{v}_j = \lambda_i^{-1}\lambda_j^{-1}\mathbf{v}_i^T\lambda_j^2\mathbf{v}_j = \frac{\lambda_j}{\lambda_i}\mathbf{v}_i^T\mathbf{v}_j = \delta_{ij}$. Tedy platí $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. Označme $\mathbf{L}$ diagonální matici, jejíž diagonální prvky jsou λ_i. Chceme ukázat $\mathbf{A} = \mathbf{ULV}^T$, tedy $\mathbf{U}^T\mathbf{AV} = \mathbf{U}^T\mathbf{ULV}^T\mathbf{V} = \mathbf{L}$.

$$\mathbf{U}^T\mathbf{AV} = \mathbf{U}^T \begin{pmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \dots & \mathbf{A}\mathbf{v}_r \end{pmatrix} = \mathbf{U}^T \begin{pmatrix} \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \dots & \lambda_r\mathbf{u}_r \end{pmatrix} = \mathbf{L}.$$

□

3.4 Choleského rozklad

Věta 17. *(Choleského rozklad.) Buď $\mathbf{A}_{nn} > 0$. Pak existuje jednoznačně určená dolní trojúhelníková matice $\mathbf{L}_{nn}$ s kladnými diagonálními prvky, že platí:*

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T.$$

Důkaz. Dokážeme matematickou indukcí podle n. Pro $n = 1$ definujme $\mathbf{L} = \sqrt{a_{11}}$. Pro $n > 1$ pišme $\mathbf{A}$ jako: $\begin{pmatrix} \mathbf{A}_{n-1} & \tilde{\mathbf{a}} \\ \tilde{\mathbf{a}}^T & a_{nn} \end{pmatrix}$, kde $\mathbf{A}_{n-1}$ vznikne z $\mathbf{A}$ vypuštěním posledního sloupce a řádku, $\tilde{\mathbf{a}}$ je poslední sloupec $\mathbf{A}$ bez posledního prvku. Z definice víme, že platí $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0, \forall \mathbf{x}_n \neq \mathbf{0}$, speciálně pro všechny $\mathbf{x}_n$, jejichž poslední složka je 0, tedy $\mathbf{A}_{n-1}$ je PD. Dále platí $(0, ..., 0, 1)^T\mathbf{A}(0, ..., 0, 1) = a_{nn} > 0$. Z indukčního předpokladu víme, že existuje jednoznačně určená matice $\mathbf{L}_{n-1}$ taková, že $\mathbf{A}_{n-1} = \mathbf{L}_{n-1}\mathbf{L}_{n-1}^T$ a diagonální prvky $\mathbf{L}_{n-1}$ jsou kladné. Hledejme tedy matici $\mathbf{L}$ splňující:

$$\begin{pmatrix} \mathbf{A}_{n-1} & \tilde{\mathbf{a}} \\ \tilde{\mathbf{a}}^T & a_{nn} \end{pmatrix} = \begin{pmatrix} \mathbf{L}_{n-1} & \mathbf{0} \\ \tilde{\boldsymbol{\pi}}^T & l_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{n-1}^T & \tilde{\boldsymbol{\pi}} \\ \mathbf{0}^T & l_{nn} \end{pmatrix} = \mathbf{L}\mathbf{L}^T.$$

Neznámé pro nás jsou $\tilde{\boldsymbol{\pi}}$ a l_{nn}. Máme $\mathbf{L}_{n-1}\tilde{\boldsymbol{\pi}} = \tilde{\mathbf{a}}$, $\mathbf{L}_{n-1}$ je dolní trojúhelníková matice a na diagonále má kladné prvky, tedy má kladný determinant (je regulární). Existuje tedy jednoznačné určená matice $\mathbf{L}_{n-1}^{-1}$ taková, že $\tilde{\boldsymbol{\pi}} = \mathbf{L}_{n-1}^{-1}\tilde{\mathbf{a}}$. Dále z vlastností (iv) a (vi) determinantu máme $\det(\mathbf{A}) = (\det(\mathbf{L}_{n-1}))^2 l_{nn}^2$. Protože $\det(\mathbf{A}) > 0$ a $(\det(\mathbf{L}_{n-1}))^2 > 0$, existuje jednoznačně určené kladné číslo

$$l_{nn} = \sqrt{\frac{\det(\mathbf{A})}{(\det(\mathbf{L}_{n-1}))^2}}.$$

[5, kap. 2].

□

Příklad. Choleského rozklad lze snadno spočítat i bez využití inverzních matic. Rovnost napíšeme pro jednotlivé prvky matice $\mathbf{A}$ a rovnice řešíme ve správném pořadí, abychom znali potřebné proměnné. Jednou možností je postupovat po sloupcích matice $\mathbf{L}$ shora dolů. Pro matici

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 20 & 26 \\ 3 & 26 & 70 \end{pmatrix} = \mathbf{L}\mathbf{L}^T = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \cdot \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix}$$

máme

$$\begin{aligned} 1 &= l_{11}^2, \\ 2 &= l_{11}l_{21}, \\ 3 &= l_{11}l_{31}, \\ 20 &= l_{21}^2 + l_{22}^2, \\ 26 &= l_{21}l_{31} + l_{22}l_{32}, \\ 70 &= l_{31}^2 + l_{32}^2 + l_{33}^2. \end{aligned}$$

Postupně dostaneme: $l_{11} = 1, \quad l_{21} = 2, \quad l_{31} = 3, \quad l_{22} = 4, \quad l_{32} = 5, \quad l_{33} = 6.$

3.5 QR rozklad

Věta 18. *(QR rozklad.) Buď* $\mathbf{A}_{mn}$ *matice,* $\mathrm{h}(\mathbf{A}) = n \geq 1$. *Pak existuje matice* $\mathbf{Q}_{mn}$*, jejíž sloupce jsou ortonormální vektory, a horní trojúhelníková matice s kladnými diagonálními prvky* $\mathbf{R}_{nn}$ *hodnosti* n*, že platí*

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

Důkaz. Označme $\mathbf{a}_1, ..., \mathbf{a}_n$ sloupce matice $\mathbf{A}$ (lineárně nezávislé vektory). Pak existují nenulové ortogonální vektory $\mathbf{b}_1, ..., \mathbf{b}_n$, které získáme následujícím postupem:

$$\begin{aligned}
\mathbf{b}_1 &= \mathbf{a}_1 \\
\mathbf{b}_2 &= \mathbf{a}_2 - x_{12}\mathbf{b}_1 \\
&\vdots \\
\mathbf{b}_n &= \mathbf{a}_n - x_{n-1n}\mathbf{b}_{n-1} - ... - x_{1n}\mathbf{b}_1,
\end{aligned}$$

kde $x_{ij} = (\mathbf{a}_j^T \mathbf{b}_i)/(\mathbf{b}_i^T \mathbf{b}_i)$. $\mathbf{b}_1, ..., \mathbf{b}_n$ jsou nenulové, protože je lze vyjádřit jako netriviální lineární kombinaci vektorů $\mathbf{a}_1, ..., \mathbf{a}_n$. Ortogonalitu dokážeme matematickou indukcí. Pro jeden vektor máme $\mathbf{b}_1 = \mathbf{a}_1$, tedy tvrzení platí. Nechť nyní tvrzení platí pro $k-1$, podle indukčního předpokladu máme pro $j = 1, ..., k-1$: $\mathbf{b}_k^T \mathbf{b}_j = \mathbf{a}_k^T \mathbf{b}_j - x_{k-1k}\mathbf{b}_{k-1}^T \mathbf{b}_j - ... - x_{1k}\mathbf{b}_1^T \mathbf{b}_j = \mathbf{a}_k^T \mathbf{b}_j - x_{jk}\mathbf{b}_j^T \mathbf{b}_j = 0 \Leftrightarrow x_{jk} = (\mathbf{a}_k^T \mathbf{b}_j)/(\mathbf{b}_j^T \mathbf{b}_j)$. Označme $\mathbf{B}$ matici, jejíž sloupce jsou $\mathbf{b}_1, ..., \mathbf{b}_n$ a $\mathbf{X}$ matici, jejíž prvky x_{ij} jsou výše uvedené pro $i < j$ a dále 0 pro $i > j$ a 1 pro $i = j$. Pak platí $\mathbf{A} = \mathbf{BX}$ (jak lze vidět, když v soustavě definující $\mathbf{b}_1, ..., \mathbf{b}_n$ vyjádříme $\mathbf{a}_1, ..., \mathbf{a}_n$). Dále definujme diagonální matici $\mathbf{D}$, jejíž diagonální prvky jsou $d_{ii} = (\mathbf{b}_i^T \mathbf{b}_i)^{-1}$. Pak můžeme psát $\mathbf{A} = \mathbf{BDD}^{-1}\mathbf{X}$. Stačí označit $\mathbf{Q} = \mathbf{BD}$ a $\mathbf{R} = \mathbf{D}^{-1}\mathbf{X}$ a matice splňují požadavky uvedené ve znění věty. [2, kap. 26], [6, kap. 6]

□

4. Další kapitoly teorie matic

4.1 Funkce matic

V této části se seznámíme s velmi zajímavým tématem, které popisuje, jak zobecnit pojem funkce (řekněme reálné funkce reálné proměnné) tak, aby jejím argumentem i hodnotou byla matice (stejného rozměru). Jinými slovy pro známou funkci $f : \mathbb{R} \to \mathbb{R}$ chceme zadefinovat $f : \mathbb{A}_{nn} \to \mathbb{A}_{nn}$ (kde $\mathbb{A}_{nn}$ je množina všech matic řádu n). Samozřejmě cílem je, aby při tomto zobecnění byly vlastnosti funkcí ve velké míře zachovány. Příkladem funkcí, které již dokážeme v tomto smyslu uvažovat jsou např. polynomiální funkce (protože přirozená mocnina matice a násobení matice skalárem jsou základní maticové operace), nebo v kapitole 3 zmíněná odmocninová matice $\mathbf{A}^{\frac{1}{2}}$. Vysvětlíme, jak obecně tento problém vyřešit a více se zaměříme na elementární funkce (exponenciála, logaritmus, goniometrické funkce). K tomuto bude potřeba zopakovat některé pojmy týkající se Jordanova tvaru matice (podrobněji viz [2, kap. 18], [3, kap. 17]). Kapitolu o exponenciále matice můžeme nalézt v ([13, kap. 11]). Stručně problém funkcí matic nalezneme v [9, kap. 3.13] uceleně v [4, kap. 5]. Velmi detailně se teorií i praktickými aspekty maticových funkcí zaobírá kniha [8]. Maticové funkce mají rozsáhlé uplatnění, např. v Markovových řetězcích se spojitým časem hraje důležitou roli exponenciála matice.

Definice. Řekneme, že čtvercové matice $\mathbf{A}_{nn}$ a $\mathbf{B}_{nn}$ jsou *podobné*, existuje-li regulární matice $\mathbf{C}_{nn}$ taková, že platí $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$.

Poznámka. Podobné matice mají stejný charakteristický polynom (nahlédneme z $\det(\mathbf{C}^{-1}\mathbf{A}\mathbf{C} - \lambda\mathbf{I}) = \det(\mathbf{C}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{C}) = \det(\mathbf{C}^{-1})\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{C}) = \det(\mathbf{A} - \lambda\mathbf{I})$), vlastní čísla, stopu, hodnost a determinant [2, kap. 18].

Definice. Řekneme, že matice je *diagonalizovatelná*, pokud je podobná nějaké diagonální matici.

Definice. Čtvercovou blokově diagonální matici $\mathbf{J}$ ve tvaru

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_k \end{pmatrix}, \text{ kde } \mathbf{J}_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}$$

nazveme *Jordanova matice*, matice $\mathbf{J}_i$ nazýváme *Jordanovy buňky*. Je-li jistá matice $\mathbf{A}$ podobná $\mathbf{J}$, řekneme, že $\mathbf{A}$ má *Jordanův kanonický tvar* $\mathbf{J}$.

Věta 19. *Každá čtvercová matice má Jordanův kanonický tvar (prvky Jordanovy matice mohou být komplexní čísla) [2, věta 18.16].*

Poznámka. Matice je diagonalizovatelná, je-li její Jordanův kanonický tvar $\mathbf{J}$ složen pouze z buněk řádu 1. Pro diagonalizovatelné matice se problémy maticových funkcí výrazně zjednodušují.

Definice. Nechť má matice $\mathbf{A}$ Jordanův kanonický tvar složený z Jordanových buněk $\mathbf{J}_1, ..., \mathbf{J}_k$ řádu $j_1, ..., j_k$. Řekneme, že funkce f je *definovaná na spektru* $\mathbf{A}$, pokud existují všechny

$$f^{(d)}(\lambda_i), \quad d = 0, ..., j_i - 1, \quad i = 1, ..., k.$$

Definice. Nechť je f definovaná na spektru $\mathbf{A}$ a nechť $\mathbf{A} = \mathbf{CJC}^{-1}$. Pak definujeme maticovou funkci $f(\mathbf{A})$ takto:

$$f(\mathbf{A}) = \mathbf{C}f(\mathbf{J})\mathbf{C}^{-1} = \mathbf{C}\mathrm{diag}(f(\mathbf{J}_k))\mathbf{C}^{-1},$$

kde

$$f(\mathbf{J}_i) = \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \dots & \frac{f^{(j_i-1)}(\lambda_i)}{(j_i-1)!} \\ 0 & f(\lambda_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & f'(\lambda_i) \\ 0 & \dots & 0 & f(\lambda_i) \end{pmatrix}.$$

Poznámka. Tato na první pohled složitá definice má svůj smysl, jak později uvidíme. Definování maticové funkce intuitivně např. ve smyslu aplikování funkce f na prvky matice a_{ij} obecně nepřinese očekávané výsledky. Snadno vidíme např. že označíme-li $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 9 & 16 \end{pmatrix}$ a $\mathbf{A}^{\frac{1}{2}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, pak $\mathbf{A} \neq (\mathbf{A}^{\frac{1}{2}})^2$.

Poznámka. Předchozí definice se výrazně zjednodušuje pro diagonalizovatelné matice, f se pouze aplikuje na diagonální prvky Jordanova kanonického tvaru. Užitečné pozorování je, že na zápis $f(\mathbf{A}) = \mathbf{C}f(\mathbf{J})\mathbf{C}^{-1}$ se můžeme dívat jako na spektrální rozklad. Vlastní vektory matic $\mathbf{A}$ a $f(\mathbf{A})$ se shodují, vlastní čísla matice $f(\mathbf{A})$ získáme aplikací f na vlastní čísla $\mathbf{A}$. [8, kap. 1]

Věta 20. *Nechť má f Taylorův rozvoj*

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!}(z - \alpha)^k$$

s poloměrem konvergence r. *Pak pro čtvercovou matici $\mathbf{A}$ je $f(\mathbf{A})$ definováno a dáno vzorcem*

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!}(\mathbf{A} - \alpha\mathbf{I})^k,$$

právě když každá vlastní hodnota λ_i matice $\mathbf{A}$ splňuje jednu z podmínek:

(i) $|\lambda_i - \alpha| < r$,

(ii) $|\lambda_i - \alpha| = r$ *a Taylorova řada $f^{(n_i-1)}(\lambda)$ (kde n_i je řád Jordanovy buňky, kterému λ_i odpovídá) je konvergentní v bodě $\lambda = \lambda_i$.*

Důkaz.
Viz [8, věta 4.7]. □

Poznámka. Předchozí věta umožňuje zapsat elementární funkce matic takto:

$$\begin{aligned} e^{\mathbf{A}} &= \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + ... \\ \cos \mathbf{A} &= \mathbf{I} - \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^4}{4!} - ... \\ \sin \mathbf{A} &= \mathbf{A} - \frac{\mathbf{A}^3}{3!} + \frac{\mathbf{A}^5}{5!} - ... \\ \ln(\mathbf{I} + \mathbf{A}) &= \mathbf{A} - \frac{\mathbf{A}^2}{2} + \frac{\mathbf{A}^3}{3} - ..., \quad |\lambda_i| < 1. \end{aligned}$$

Dá se ukázat, že takto definované funkce mají velkou část očekávaných vlastností zachovánu. Např. exponenciála za jistých okolností splňuje rovnost $e^{\mathbf{AB}} = e^{\mathbf{A}} e^{\mathbf{B}}$ (postačující podmínkou je komutativita matic, tj. $\mathbf{AB} = \mathbf{BA}$). Platí i Eulerův vzorec ($e^{i\mathbf{A}} = \cos \mathbf{A} + i \sin \mathbf{A}$) nebo součtové vzorce funkcí sinus a cosinus. Vlastnosti logaritmu, např. $\ln(\mathbf{AB}) = \ln \mathbf{A} + \ln \mathbf{B}$ jsou zachovány (avšak za dalších předpokladů, viz [8, kap. 11]).

Příklad. Ilustrujme platnost vztahu $\sin^2 \mathbf{A} + \cos^2 \mathbf{A} = \mathbf{I}$ pro $\mathbf{A} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$. Tato matice je obecná Jordanova buňka řádu 2, je tedy sama sobě Jordanovým kanonickým tvarem. Podle definice máme: $\sin^2 \mathbf{A} + \cos^2 \mathbf{A} = \begin{pmatrix} \sin a & \cos a \\ 0 & \sin a \end{pmatrix}^2 +$

$$\begin{pmatrix} \cos a & -\sin a \\ 0 & \cos a \end{pmatrix}^2 = \begin{pmatrix} \sin^2 a & 2 \sin a \cos a \\ 0 & \sin^2 a \end{pmatrix} + \begin{pmatrix} \cos^2 a & -2 \sin a \cos a \\ 0 & \cos^2 a \end{pmatrix} =$$

$$\begin{pmatrix} \sin^2 a + \cos^2 a & 0 \\ 0 & \sin^2 a + \cos^2 a \end{pmatrix} = \mathbf{I}.$$

4.2 Maticové derivování

V této části ukážeme, jak elegantně využít zápisu pomocí vektorů a matic při derivování. Dokážeme některé základní vzorce. Budeme uvažovat funkce $f : \mathbb{R}^{m \times n} \to \mathbb{R}$, později matice takových funkcí. Poznatky aplikujeme na odvození řešení problému nejmenších čtverců. Předpokládáme základní znalosti (parciálního) derivování z matematické analýzy. Ucelenou teorii maticového derivování najdeme např. v [6, kap. 15]. Věnuje se mu celá kniha [11], a to včetně aplikací ve statistice a ekonometrii.

Definice. Buď $\mathbf{X}_{mn}$ matice, jejíž prvky x_{ij} jsou reálné proměnné a f reálná funkce těchto mn proměnných ($f : \mathbb{R}^{m \times n} \to \mathbb{R}$). Pak *derivací f podle* $\mathbf{X}$ rozumíme matici parciálních derivací $\left(\frac{\partial f(\mathbf{X})}{\partial x_{ij}} \right)$, značíme $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}}$. Ve speciálním případě $n = 1$ píšeme $\frac{\partial f(\mathbf{X})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{X})}{\partial x_j} \right)$.

Příklad. Mějme $f : \mathbb{R}^{2 \times 2} \to \mathbb{R}$ definovanou předpisem $f(\mathbf{X}) = \sin x_{11} + e^{x_{21} + x_{22}} + x_{21}$, $\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$. Pak $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} = \begin{pmatrix} \cos x_{11} & 0 \\ e^{x_{21}+x_{22}} + 1 & e^{x_{21}+x_{22}} \end{pmatrix}$.

Věta 21. *(O derivaci forem.) Buďte* $\mathbf{a}$, $\mathbf{y}$ *a* $\mathbf{A}$ *vektory a matice konstant takových rozměrů, aby následující součiny byly definovány. Pak platí (vybráno z [12, dod. A]):*

(i) $\dfrac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}, \quad \dfrac{\partial \mathbf{x}^T\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}, \quad \dfrac{\partial \mathbf{x}^T\mathbf{A}\mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}\mathbf{y},$

(ii) $\dfrac{\partial \mathbf{x}^T\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^T)\mathbf{x}$ *(*$\mathbf{A}$ *je čtvercová matice, je-li navíc symetrická, je* $(\mathbf{A} + \mathbf{A}^T)\mathbf{x} = 2\mathbf{A}\mathbf{x}$*).*

Důkaz.

(i) Ve všech případech výrazy stačí rozepsat, zderivovat a opět zapsat pomocí uvedené definice. V první části máme $\frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \left(\frac{\partial(a_1x_1 + \ldots + a_nx_n)}{\partial x_j}\right) = (a_j) = \mathbf{a}$. Ostatní případy jsou obdobné.

(ii) Opět je klíčové uvědomit si, že $\mathbf{x}^T\mathbf{A}\mathbf{x}$ je součet $\sum_{ik} a_{ik}x_ix_k$. Derivaci tohoto výrazu podle x_j můžeme názorně zapsat jako:

$$\begin{array}{rcl} 0 + \ldots + 0 + & a_{1j}x_1 & +0 + \ldots + 0 + \\ & \vdots & \\ +a_{j1}x_1 + \ldots + a_{jj-1}x_{j-1} + & 2a_{jj}x_j & +a_{jj+1}x_{j+1} + \ldots + a_{jn}x_n + \\ & \vdots & \\ +0 + \ldots + 0 + & a_{nj}x_n & +0 + \ldots + 0. \end{array}$$

Tedy $\sum_i a_{ji}x_i + \sum_k a_{kj}x_k$ (člen $2a_{jj}x_j$ je rozdělen do obou sum jako $a_{jj}x_j$). Nyní si stačí uvědomit, že tyto sumy jsou j-tým prvkem vektoru $\mathbf{A}\mathbf{x}$ resp. $\mathbf{A}^T\mathbf{x}$.

□

Věta 22. *(Nejmenší čtverce.) Funkci* $f(\mathbf{x}) = (\mathbf{y} - \mathbf{A}\mathbf{x})^T(\mathbf{y} - \mathbf{A}\mathbf{x})$, *kde* $\mathbf{A}$ *je matice plné sloupcové hodnosti, minimalizuje jednoznačně určený vektor* $\mathbf{x}$ *daný vztahem*

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

Důkaz. Upravme nejdříve: $f(\mathbf{x}) = (\mathbf{y}-\mathbf{A}\mathbf{x})^T(\mathbf{y}-\mathbf{A}\mathbf{x}) = (\mathbf{y}^T - \mathbf{x}^T\mathbf{A}^T)(\mathbf{y}-\mathbf{A}\mathbf{x}) = \mathbf{y}^T\mathbf{y} - \mathbf{x}^T\mathbf{A}^T\mathbf{y} - \mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{y} - 2\mathbf{x}^T\mathbf{A}^T\mathbf{y} + \mathbf{x}^T(\mathbf{A}^T\mathbf{A})\mathbf{x}$. Zderivujme s využitím předchozí věty tento výraz podle $\mathbf{x}$: $-2\mathbf{A}^T\mathbf{y} + 2\mathbf{A}^T\mathbf{A}\mathbf{x}$ (podle vlastnosti (i) symetrické matice víme, že $\mathbf{A}^T\mathbf{A}$ je symetrická). Z matematické analýzy víme, že v bodě, kde položíme první derivaci rovnu $\mathbf{0}$, se může nacházet minimum. Je ale $-2\mathbf{A}^T\mathbf{y} + 2\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{y} \Leftrightarrow \mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$. Tedy $\mathbf{x}$ je jednoznačně určený bod. Vzhledem k tomu, že f je shora neomezená spojitá funkce, nabývá f v tomto bodě minima. Založeno na [12, dod. A]. □

Poznámka. Výsledek předchozí věty není vůbec překvapivý, na konci části o projekčních maticích jsme popsali (bez zdůvodnění) řešení ekvivalentního problému (s odmocninou navíc). Je totiž $f(\mathbf{x}) = \sum_j (y_j - \sum_i a_{ji}x_i)^2$ a $\hat{\mathbf{y}} = \mathbf{P}_{\mathbf{A}}\mathbf{y} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y} = \mathbf{A}\mathbf{x}$.

Poznámka. Připomeňme, že *algebraický doplněk prvku* a_{ij} matice $\mathbf{A}$ je číslo $c_{ij} = (-1)^{i+j}\det(\mathbf{M}_{ij})$, kde $\mathbf{M}_{ij}$ je matice, která vznikla z matice $\mathbf{A}$ vynecháním i-tého řádku a j-tého sloupce. *Adjungovaná matice* je transponovaná matice algebraických doplňků (značíme $\mathrm{adj}(\mathbf{A}) = (c_{ij})^T$). *Laplaceův rozvoj determinantu podle* i*-tého řádku* můžeme při tomto značení zapsat jako $\det(\mathbf{A}) = \sum_j a_{ij}c_{ij}$.

Poznámka. Dosud jsme derivovali podle vektoru. V následující větě popíšeme několik užitečných výsledků derivování podle matice. Rozlišíme zvlášť derivování podle matice mn různých proměnných a symetrické matice (tedy matice, kde $x_{ij} = x_{ji}$). Podrobněji se tomuto tématu věnuje [6, kap. 15], kde je odvozeno i velké množství dalších výsledků, vybrané uvádíme v následující větě.

Věta 23. *Buď* $\mathbf{A}$ *matice konstant a* $\mathbf{X}$ *čtvercová matice.*

(i) $$\frac{\partial \mathrm{tr}(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{I},$$

$$\frac{\partial \mathrm{tr}(\mathbf{XA})}{\partial \mathbf{X}} = \begin{cases} \mathbf{A}^T, & \text{není-li } \mathbf{X} \text{ symetrická} \\ \mathbf{A} + \mathbf{A}^T - \mathrm{diag}(\mathbf{A}), & \text{je-li } \mathbf{X} \text{ symetrická,} \end{cases}$$

(ii) $$\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} = \begin{cases} (\mathrm{adj}(\mathbf{X}))^T, & \text{není-li } \mathbf{X} \text{ symetrická} \\ 2\mathrm{adj}(\mathbf{X}) - \mathrm{diag}(\mathrm{adj}(\mathbf{X})), & \text{je-li } \mathbf{X} \text{ symetrická,} \end{cases}$$

$$\frac{\partial \det(\mathbf{X})}{\partial x_{ij}} = \mathrm{tr}\left(\mathrm{adj}(\mathbf{X})\frac{\partial \mathbf{X}}{\partial x_{ij}}\right)$$

Důkaz.

(i) V obou případech stačí rozepsat stopu. V první části máme $\mathrm{tr}(\mathbf{X}) = \sum_i x_{ii}$. Ve druhé $\mathrm{tr}(\mathbf{XA}) = \sum_j \sum_i x_{ji}a_{ij}$, tedy pro $\mathbf{X}$ nesymetrickou je výsledek zřejmý. Je-li $\mathbf{X}$ symetrická, máme pro $i \neq j$ v součtu členy $x_{ji}a_{ij}$ a $x_{ij}a_{ji}$, jejichž derivace podle $x_{ji} = x_{ij}$ je a_{ij} a a_{ji}, avšak člen $x_{ii}a_{ii}$ se v součtu vyskytuje jen jednou, výsledek lze tedy šikovně zapsat v uvedeném tvaru.

(ii) Ukážeme, že $\frac{\partial \det(\mathbf{X})}{\partial x_{ij}} = c_{ij}$, kde c_{ij} je algebraickým doplňkem prvku x_{ij}. Rozepíšeme $\det(\mathbf{X}) = \sum_k x_{ik}c_{ik}$. Stačí si všimnout, že všechny algebraické doplňky v sumě jsou vzhledem k x_{ij} konstantní a $\frac{\partial x_{ik}}{\partial x_{ij}} = \delta_{jk}$. Důkaz pro symetrickou matici $\mathbf{X}$ a další vztah viz [6, kap. 15].

$\square$

Definice. Buď $\mathbf{X}_{pq}$ matice, jejíž prvky x_{ij} jsou reálné proměnné a $\mathbf{Y}_{mn}$ matice, jejíž prvky y_{ij} jsou funkce těchto proměnných. Pak *derivací* $\mathbf{X}$ *podle* $\mathbf{Y}$ (značíme $\frac{\partial \mathbf{Y}}{\partial \mathbf{X}}$ rozumíme blokovou $mp \times nq$ matici:

$$\frac{\partial \mathbf{Y}}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial \mathbf{Y}}{\partial x_{11}} & \frac{\partial \mathbf{Y}}{\partial x_{12}} & \cdots & \frac{\partial \mathbf{Y}}{\partial x_{1q}} \\ \frac{\partial \mathbf{Y}}{\partial x_{21}} & \frac{\partial \mathbf{Y}}{\partial x_{22}} & \cdots & \frac{\partial \mathbf{Y}}{\partial x_{2q}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{Y}}{\partial x_{p1}} & \frac{\partial \mathbf{Y}}{\partial x_{p2}} & \cdots & \frac{\partial \mathbf{Y}}{\partial x_{pq}} \end{pmatrix}, \quad \text{kde } \frac{\partial \mathbf{Y}}{\partial x_{ij}} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{ij}} & \frac{\partial y_{12}}{\partial x_{ij}} & \cdots & \frac{\partial y_{1n}}{\partial x_{ij}} \\ \frac{\partial y_{21}}{\partial x_{ij}} & \frac{\partial y_{22}}{\partial x_{ij}} & \cdots & \frac{\partial y_{2n}}{\partial x_{ij}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x_{ij}} & \frac{\partial y_{m2}}{\partial x_{ij}} & \cdots & \frac{\partial y_{mn}}{\partial x_{ij}} \end{pmatrix}.$$

Příklad. Spočítejme $\frac{\partial \mathbf{Y}}{\partial \mathbf{X}}$ pro

$$\mathbf{Y} = \begin{pmatrix} x_{11} & \sin x_{22} & x_{11}x_{21}^2 \\ x_{12} + x_{21} & x_{22}x_{11} & x_{11}x_{12} + x_{21}x_{22} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}.$$

Máme

$$\frac{\partial \mathbf{Y}}{\partial \mathbf{X}} = \begin{pmatrix} \frac{\partial \mathbf{Y}}{\partial x_{11}} & \frac{\partial \mathbf{Y}}{\partial x_{12}} \\ \frac{\partial \mathbf{Y}}{\partial x_{21}} & \frac{\partial \mathbf{Y}}{\partial x_{22}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & x_{21}^2 & 0 & 0 & 0 \\ 0 & x_{22} & x_{12} & 1 & 0 & x_{11} \\ 0 & 0 & 2x_{11}x_{21} & 0 & \cos x_{22} & 0 \\ 1 & 0 & x_{22} & 0 & x_{11} & x_{21} \end{pmatrix}.$$

Věta 24. *Buď* $\mathbf{A}$ *matice konstant.*

(i) $\dfrac{\partial \mathbf{x}^T \mathbf{A}}{\partial \mathbf{x}} = \mathbf{A}$,

(ii) $\dfrac{\partial \mathbf{Y}^{-1}}{\partial x} = -\mathbf{Y}^{-1}\dfrac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}$, *kde* $\mathbf{Y}$ *je regulární matice, jejíž prvky jsou funkce proměnné* x.

Důkaz.

(i) Zobecnění případu $\frac{\partial \mathbf{a}^T \mathbf{x}}{\partial \mathbf{x}}$, stačí rozepsat: $\mathbf{x}^T \mathbf{A} = \begin{pmatrix} \sum_i x_i a_{i1} & \dots & \sum_i x_i a_{in} \end{pmatrix}$.

(ii) Jak lze dohledat např. v [6, kap. 15], maticové derivování zachovává mnohé známé vlastnosti. Zde využijeme obdobu derivace součinu $(uv)' = u'v + uv'$ v maticové podobě. Protože $\mathbf{Y}$ je regulární, platí $\mathbf{I} = \mathbf{Y}\mathbf{Y}^{-1}$, $\frac{\partial \mathbf{I}}{\partial x} = \frac{\partial \mathbf{Y}\mathbf{Y}^{-1}}{\partial x}$, $\mathbf{0} = \mathbf{Y}\frac{\partial \mathbf{Y}^{-1}}{\partial x} + \frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}$. Nyní stačí rovnici vynásobit zleva maticí $\mathbf{Y}^{-1}$ a odečíst od rovnosti druhého sčítance.[6, kap. 15]

□

4.3 Kroneckerův součin, vektorizace matic

Kromě tradičního maticového násobení, který celou dobu používáme, existuje několik dalších způsobů, jak definovat součin matic a dosáhnout užitečných výsledků. Jedním z nich je např. *Hadamardův součin* (součin po složkách). V této kapitole zavedeme Kroneckerův součin a úzce související pojmy vektorizace matic, shrneme základní vlastnosti (výběr z [12, dod. A], [6, kap. 16]).

Definice. Buďte $\mathbf{A}_{mn}$ a $\mathbf{B}_{pq}$ libovolné matice. *Kroneckerovým součinem* matic $\mathbf{A}$ a $\mathbf{B}$ (značíme $\mathbf{A} \otimes \mathbf{B}$) rozumíme $mp \times nq$ matici

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix}.$$

Poznámka. $\mathbf{A} \otimes \mathbf{B}$ můžeme vnímat jako blokovou matici, jejíž bloky jsou $a_{ij}\mathbf{B}$, ale také jako matici součinů prvku matice $\mathbf{A}$ s prvkem matice $\mathbf{B}$. Kroneckerův součin je sice (na rozdíl od obyčejného maticového součinu) definován pro matice libovolných rozměrů, ale obecně neplatí $\mathbf{A} \otimes \mathbf{B} = \mathbf{B} \otimes \mathbf{A}$.

Definice. Buď $\mathbf{A}_{mn}$ libovolná matice, jejíž sloupce jsou vektory $\mathbf{a}_1, ..., \mathbf{a}_n$. *Vektorizaci* matice $\mathbf{A}$ (značíme vec($\mathbf{A}$)) definujeme jako vektor

$$\mathrm{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

Pro $\mathbf{A}$ symetrickou označme $\mathbf{a}_i^* = \begin{pmatrix} a_{ii} & a_{i+1i} & ... & a_{ni} \end{pmatrix}^T$ (sloupce bez naddiagonálních prvků) a definujme vech($\mathbf{A}$) jako

$$\mathrm{vech}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \\ \vdots \\ \mathbf{a}_n^* \end{pmatrix}.$$

Příklad. Spočítejme $\mathbf{A} \otimes \mathbf{B}$, vec($\mathbf{A}$), vech($\mathbf{B}$) pro

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Máme

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} & a_{13}b_{11} & a_{13}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} & a_{13}b_{21} & a_{13}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} & a_{23}b_{11} & a_{23}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} & a_{23}b_{21} & a_{23}b_{22} \end{pmatrix},$$

$$\mathrm{vec}(\mathbf{A}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \\ a_{13} \\ a_{23} \end{pmatrix}, \mathrm{vech}(\mathbf{B}) = \begin{pmatrix} b_{11} \\ b_{21} \\ b_{22} \end{pmatrix}.$$

Věta 25. *(Vlastnosti Kroneckerova součinu.) Buďte* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ *libovolné matice,* $\mathbf{a}$ *a* $\mathbf{b}$ *libovolné vektory a* k *libovolná konstanta. Pak platí:*

(i) $k \otimes \mathbf{A} = k\mathbf{A}$,

(ii) $\mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1\mathbf{b} \\ a_2\mathbf{b} \\ \vdots \\ a_m\mathbf{b} \end{pmatrix}$, $\mathbf{a} \otimes \mathbf{b}^T = \mathbf{a}\mathbf{b}^T$, $\mathbf{a}^T \otimes \mathbf{b}^T = \begin{pmatrix} a_1\mathbf{b}^T & a_2\mathbf{b}^T & \dots & a_m\mathbf{b}^T \end{pmatrix}$,

(iii) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$, $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$,
$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$, $\mathbf{I} \otimes \mathbf{A} = \mathrm{diag}(\mathbf{A}, ..., \mathbf{A})$,

(iv) $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$,

(v) $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$ *(pro* $\mathbf{A}$ *a* $\mathbf{B}$ *regulární),*

(vi) $\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})$,

(vii) $\mathrm{h}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{h}(\mathbf{A})\mathrm{h}(\mathbf{B})$,

(viii) $\det(\mathbf{A}_{mm} \otimes \mathbf{B}_{nn}) = (\det(\mathbf{A}))^m(\det(\mathbf{B}))^n$,

(ix) *buď λ (resp. μ) vlastní číslo matice $\mathbf{A}$ (resp. $\mathbf{B}$) a $\mathbf{u}$ (resp. $\mathbf{v}$) jemu odpovídající vlastní vektor, pak $\lambda\mu$ je vlastní číslo matice $\mathbf{A} \otimes \mathbf{B}$ a $\mathbf{u} \otimes \mathbf{v}$ jemu odpovídající vlastní vektor.*

Důkaz.
(i) - (iv) zřejmé. (v) s pomocí (iii). (vi) zřejmé. (vii) viz [6, kap. 16]. (viii) plyne z (ix), k důkazu (ix) se využije (iii). □

Věta 26. *(Vlastnosti vec, vech.)*

(i) $\mathrm{vec}(c\mathbf{A}) = c\,\mathrm{vec}(\mathbf{A})$,

(ii) $\mathrm{vec}(\mathbf{a}^T) = \mathrm{vec}(\mathbf{a}) = \mathbf{a}$,

(iii) $\mathrm{vec}(\mathbf{b}\mathbf{a}^T) = \mathbf{a} \otimes \mathbf{b}$,

(iv) $\mathrm{vec}(\mathbf{A}_{mn}\mathbf{B}_{np}) = \mathrm{diag}(\mathbf{A}, ..., \mathbf{A})\mathrm{vec}(\mathbf{B}) = (\mathbf{I}_{pp} \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$,

(v) $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$,

(vi) *Buďte $\mathbf{A}_{mn}$ a $\mathbf{B}_{mp}$ matice konstant a $\mathbf{X}_{np}$ matice neznámých, pak matice $\mathbf{X}$ řeší $\mathbf{AX} = \mathbf{B}$ právě tehdy, když řeší $(\mathbf{I}_{pp} \otimes \mathbf{A})\mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\mathbf{B})$.*

Důkaz.
(i) - (iv) zřejmé. (v) a (vi) viz [6, kap. 16]. □

Závěr

Cílem práce bylo sestavit přehled výsledků z maticové algebry, které se používají ve statistice (lineárních modelech, mnohorozměrné analýze, Markovových procesech) v didakticko-pedagogickém pojetí.

Práce začíná přehledem výsledků maticové algebry, které student získal při studiu lineární algebry. Tyto základy jsme v dalších kapitolách značně rozšířili. Uvedené poznatky lze využít pro hlubší pochopení dalších předmětů. Práce tak může být použita jako doplňkový studijní materiál pro ty, kteří se na přednáškách mnohých předmětů nespokojí s pouhými náznaky důkazů či s odkazy na další literaturu na místě, kde by mohl být úplný důkaz (postavený na větší znalosti maticové algebry).

Práce čerpá mimo jiné z objemných knih ([4], [6], [8] aj.), které běžný student nevlastní a informace zdaleka nenalezne ani v běžných učebnicích lineární algebry ([2], [3], [13] atd.). Student tak díky této práci na jednom místě najde množství poznatků, které nejsou snadno dohledatelné v literatuře nebo na internetu. Vybraná tvrzení obsahují důkazy, které byly důsledně a podrobně zpracovány (případně modifikovány, abychom si vystačili s námi uvedenou teorií), aby si čtenář nemusel nic domýšlet. Zároveň jsme tak zamezili nebezpečí neporozumění. Příklady usnadní pochopení látky a poznámky uvádí možné užití či motivaci k dalšímu studiu. Toto je zcela na místě, protože zejména jednotlivé části poslední kapitoly by bez obtíží vydaly na neméně zajímavou samostatnou práci.

Seznam použité literatury

[1] Anděl, J. *Matematická statistika.* 2. vydání. Praha: SNTL, 1985.

[2] Bečvář, J. *Lineární algebra.* 3. vydání. Praha: MatfyzPress, 2005. ISBN 80-86732-57-6.

[3] Bican, L. *Lineární algebra a geometrie.* Dotisk 1. vydání. Praha: Academia, 2002. ISBN 80-200-0843-8.

[4] Gantmacher, F. R. *The Theory of Matrices, Volume 1.* 1. vydání. New York: Chelsea, 1959. ISBN 0-8284-0131-4.

[5] Hämmerlin, G., Hoffmann, K. H. *Numerical Mathematics.* 1. vydání. New York: Springer-Verlag, 1991. ISBN 0-387-97494-6.

[6] Harville, D. A. *Matrix Algebra From a Statistician's Perspective.* 3. vydání. New York: Springer-Verlag, 1997. ISBN 0-387-94978-X.

[7] Harville, D. A. *Matrix Algebra: Exercises and Solutions.* 1. vydání. New York: Springer-Verlag, 2001. ISBN 0-387-95318-3.

[8] Higham, N. J. *Functions of Matrices: Theory and Computation.* 1. vydání. Philadelphia: SIAM, 2008. ISBN 978-0-89871-646-7.

[9] Khuri, A. I. *Linear Model Methodology.* 1. vydání. Boca Raton: CRC Press, 2009. ISBN 978-1-58488-481-1.

[10] Kwak, J. H., Hong, S. *Linear Algebra.* 2. vydání. Boston: Birkhäuser, 2004. ISBN 0-8176-4294-3.

[11] Magnus, J. R., Neudecker, H. *Matrix Differential Calculus with Applications in Statistics and Econometrics.* 2. revidované vydání. Chichester: John Wiley & Sons, 1999. ISBN 0-471-98633-X.

[12] Mardia, K. V., Kent, J. T., Bibby, J. M. *Multivariete Analysis.* 1. vydání. London: Academic Press, 1979. ISBN 0-12-471252-5.

[13] Motl, L., Zahradník, M.. *Pěstujeme lineární algebru.* První dotisk 3. vydání. Praha: Karolinum, 2003. ISBN 80-246-0421-3.

[14] Rao, R. C. *Lineární metody statistické indukce a jejich aplikace.* 1. vydání překladu. Praha: Academia, 1978.