MOORE-PENROSEOVA INVERZE MATICE A JEJÍ APLIKACE

LADISLAV SKULA

Abstrakt. V článku je uvedena definice pseudoinverzní matice, ukázána její existence a jednoznačnost a zmíněny dvě metody jejího výpočtu. Pomocí pseudoinverze je definováno „řešení“ systému lineárních rovnic s komplexními koeficienty vzhledem k nejmenším čtvercům s minimální normou a ukázán její význam pro aplikace. Toto řešení je demonstrováno na konkrétním příkladě.

1. Úvod

Tento příspěvek má sloužit jako doplněk k látce z lineární algebry přednášené v oboru matematického inženýrství na VUT. Motivací je následující problematika často potřebná v mnoha aplikacích matematiky ([1],[2],[4],[8],[9]):

Jestliže máme nějaký systém lineárních rovnic (koeficienty jsou reálná čísla), který máme řešit, pak pro množinu $\mathcal{R}$ všech řešení tohoto systému rovnic platí jedna z následujících možností:

(a) množina $\mathcal{R}$ je jednoprvková (systém rovnic má právě jedno řešení),
(b) množina $\mathcal{R}$ je nekonečná (systém rovnic má nekonečně mnoho řešení),
(c) množina $\mathcal{R}$ je prázdná množina, $\mathcal{R} = \emptyset$ (systém rovnic nemá žádné řešení).

Velmi často v aplikacích matematiky je ale potřeba mít nějaké řešení systému lineárních rovnic. Proto je nutno v případě a) z množiny řešení vybrat jedno v jakémsi smyslu *význačné řešení* a to používat. V případě, že systém nemá řešení (případ c), musí se nějaká n-tice reálných čísel (n je počet neznámých) určit a brát jako *významné řešení* soustavy.

Tento výběr řešení však nemůže být libovolný, musí nějakým způsobům odpovídat potřebám aplikační oblasti. V praxi takových možností se vyskytuje celá řada, ale nejvýznamnější a nejčastěji používaná metoda je metoda tzv. *metoda nejmenších čtverců*, která je založena na pojmu *pseudoinverzní matice.* Tuto metodu se pokusíme v tomto příspěvku vysvětlit.

Budeme předpokládat znalosti lineární algebry v rozsahu přednášky v prvním semestru z této oblasti ve studiu matematického inženýrství prezentované ve skriptech [6]. Tyto znalosti jenom rozšíříme o skutečnost, že uvedené výsledky platí nejenom pro reálná čísla, ale též pro čísla komplexní tvořící těleso $\mathbb{C}$ (dokonce pro lineární algebru nad libovolným komutativním tělesem).

2010 *MSC.* Primární 15A09, 93E24.

Klíčová slova. Pseudoinverzní matice, Moore-Penroseova inverze, Penroseovy podmínky, skeletní rozklad matice, metoda nejmenších čtverců, nejlepší přibližné řešení .

Práce byla podporována projektem A-Math-Net – Síť pro transfer znalostí v aplikované matematice (CZ.1.07/2.4.00/17.0100).

Tudíž prvky matice A budou komplexní čísla, transponovanou matici matice A budeme značit symbolem A^T a symbolem $\overline{A}$ budeme označovat matici vzniklou z matice A nahrazením prvků matice A, což jsou komplexní čísla, čísly komplexně sdruženými. V teorii matic s komplexními čísly se zavádí tzv. *hermitovský operátor* značený symbolem $*$ definovaný pro matici A typu $m \times n$ vztahem:

$$\boxed{A^* = (\overline{A})^T.}$$

Zřejmě matice A^* je typu $n \times m$ a pro matice A, B typů vhodných pro násobení máme:

$$\boxed{(A \cdot B)^* = B^* \cdot A^*,\ (A^*)^* = A.}$$

Hodnost matice A budeme označovat symbolem $\mathrm{r}(A)$ (z angličtiny „rank").

Poznamenejme, že partie, která pojednává o psedoinverzní matici a metodě nejmenších čtverců, je velmi dobře vysvětlena v knize [7] v kapitole 4 a 7. V této knize každá kapitola je doplněna odstavcem MATLAB Moment, ve kterém je popsáno použití systému MATLAB na výpočty uvedených pojmů.

2. Pseudoinverzní matice

V tomto odstavci zavedeme pojem pseudoinverzní matice, kterou má každá matice a tato pseudoinverzní matice je jednoznačně určena.

Definice 2.1. Nechť A je matice (nad tělesem komplexních čísel $\mathbb{C}$) typu $m \times n$. Matice X typu $n \times m$ se nazývá *pseudoinverzní matice* matice A, jestliže platí:

(1) $AXA = A$,
(2) $XAX = X$,
(3) $(AX)^* = AX$,
(4) $(XA)^* = XA$.

Podmínky (1)–(4) se nazývají *Penroseovy podmínky* a pseudoinverzní matice se často nazývá *Moore-Penroseova inverze matice A* nebo stručně *M-P inverze* ([3]). Mluvíme také jen o *pseudoinverzi matice A*.

Věta 2.2 (o jednoznačnosti pseudoinverzní matice)**.** *Jestliže matice má pseudoinverzi, pak tato pseudoinverze je určena jednoznačně.*

Důkaz. Důkaz je veden tím způsobem, že předpokládáme, že matice A má dvě pseudoinverze B a C. Pak se vhodným použitím Penroseových podmínek dokáží rovnosti:

$$B = CAB \qquad \text{a} \qquad C = CAB,$$

odkud plyne $B = C$. □

Označení 2.3. Jelikož je pseudoinverze jednoznačně určena, můžeme pro ni zavést označení. Tuto pseudoinverzi budeme označovat symbolem A^+.

Při zjišťování, zdali matice X je pseudoinverze matice A, se obvykle postupuje tak, že se ověřují Penroseovy podmínky (1)–(4). V případě jejich platnosti je pak X jednoznačně definovaná pseudoinverzní matice A^+ matice A.

Příklad 2.4. a) Jestliže A je regulární matice, pak matice $X = A^{-1}$ vyhovuje Penroseovým podmínkám, tudíž

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

b) Je-li $A = O_{m,n}$ nulová matice typu $m \times n$, pak podobně ověříme Penroseovy podmínky pro $X = O_{n,m}$, odkud plyne

$$O^+_{m,n} = O_{n,m}.$$

c) Jestliže matice A má pseudoinverzi, pak má pseudoinverzi též matice A^+ a platí:

$$\boxed{(A^+)^+ = A.}$$

d) Jestliže D je diagonální matice řádu n

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_n \end{pmatrix},$$

pak pro Moore-Penroseovu inverzi matice D máme

$$D^+ = \begin{pmatrix} d_1^+ & 0 & \dots & 0 \\ 0 & d_2^+ & \dots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_n^+ \end{pmatrix},$$

kde pro komplexní číslo c symbol c^+ značí 0, jestliže $c = 0$. V případě, že c je nenulové komplexní číslo, položíme $c^+ = c^{-1}$.

V závěrečném odstavci 4 prezentujeme větu o existenci pseudoinverzní matice pro každou matici nad tělesem komplexních čísel $\mathbb{C}$ a dvě metody výpočtu této pseudoinverze.

3. Moore-Penroseova inverze a systém lineárních rovnic

Pro řešení soustavy lineárních rovnic se používá aparát teorie matic a vektorů. V tomto příspěvku budeme pro přirozené číslo n rozumět n-rozměrným vektorovým prostorem $\mathbb{C}^n$ množinu všech matic nad komplexními čísly typu $n \times 1$. Tyto matice budeme považovat za vektory, přičemž komplexní čísla budou skaláry. Operace sčítání vektorů a násobení vektoru skalárem se definují jako tyto operace s maticemi.

Definice 3.1. Pro n-rozměrné vektory $X, Y \in \mathbb{C}^n$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

položíme

$$(X,Y) = Y^* \cdot X = \sum_{i=1}^{n} x_i \cdot \bar{y_i} \text{ a } ||X|| = \sqrt{(X,X)} = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Komplexní číslo (X,Y) se nazývá *skalární součin vektorů* X,Y a číslo $\|X\|$ *norma vektoru* X. Jestliže $(X,Y) = 0$, pak se vektory X,Y nazývají *ortogonální* ([5]).

Tvrzení 3.2. *Pro n-rozměrné vektory X,Y a komplexní číslo λ máme:*

(a) *Norma $\|X\|$ je reálné nezáporné číslo, přičemž*

$$\|X\| = 0 \iff X = 0_{n,1},$$

(b) $\|\lambda X\| = |\lambda| \cdot \|X\|$ *a platí trojúhelníková nerovnost:*

$$\boxed{\|X+Y\| \leq \|X\| + \|Y\|.}$$

Pro ortogonální vektory stejného rozměru se dá dokázat *Pythagorova věta.*

Věta 3.3 (Pythagorova věta)**.** *Pro ortogonální n-rozměrné vektory X,Y máme*

$$\boxed{\|X+Y\|^2 = \|X\|^2 + \|Y\|^2.}$$

V dalším budeme uvažovat systém $(*)$ lineárních rovnic (nad tělesem komplexních čísel $\mathbb{C}$) s maticí soustavy A typu $m \times n$, s vektorem absolutních členů B rozměru m a n-rozměrným vektorem X neznámých $x_1, \ldots, x_n$, tedy

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Tento systém $(*)$ lineárních rovnic budeme zapisovat v maticovém zápisu:

$$\boxed{A \cdot X = B.} \tag{$*$}$$

Definice 3.4. Pro systém lineárních rovnic $(*)$ položíme:

$$\boxed{X_0 = A^+ \cdot B.}$$

Pro n-rozměrný vektor X_0 platí následující věta:

Věta 3.5 (Hlavní věta)**.** *Pro každý n-rozměrný vektor $X \neq X_0$ máme:*

$$\|A \cdot X_0 - B\| \leq \|A \cdot X - B\|$$

a v případě $\|AX_0 - B\| = \|AX - B\|$ platí

$$\|X_0\| < \|X\|.$$

Tudíž vektor X_0 minimalizuje normu $\|AX - B\|$ a ze všech n rozměrných vektorů X, které minimalizují tuto normu, má nejmenší normu. Vektor X_0 (viz [3],[5]) se nazývá *řešení systému $(*)$ vzhledem k nejmenším čtvercům s minimální normou nebo nejlepší přibližné řešení systému $(*)$ (the least squares solution of minimum norm).*

Příklad 3.6. Nalezněte řešení následujícího systému vzhledem k nejmenším čtvercům s minimální normou:

$$\begin{aligned} x + 4y + 3z &= 2 \\ -x + y + 2z &= -2 \\ -2x - 2y &= 1. \end{aligned}$$

Matice této soustavy je matice

$$A = \begin{pmatrix} 1 & 4 & 3 \\ -1 & 1 & 2 \\ -2 & -2 & 0 \end{pmatrix}.$$

Vektory

$$B = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

jsou vektory absolutních členů a neznámých. Hodnost matice A soustavy je rovna 2 a hodnost rozšířené matice soustavy (AB) je rovna 3, tudíž soustava nemá řešení, ale existuje „řešení“ X_0 této soustavy vzhledem k nejmenším čtvercům s minimální normou. K získání tohoto „řešení“ použijeme pseudoinverzi matice A:

$$A^+ = \frac{1}{231} \begin{pmatrix} 3 & -43 & -54 \\ 27 & -2 & -24 \\ 24 & 41 & 30 \end{pmatrix},$$

odkud dostaneme

$$X_0 = A^+ \cdot B = \frac{1}{231} \begin{pmatrix} 38 \\ 34 \\ -4 \end{pmatrix}$$

Poznámka 3.7. Pro nejlepší přibližné řešení X_0 systému $(*)$ je sice norma $||A \cdot X - B||$ minimální, ale vektor X_0 nemusí být jediný vektor X, pro který je tato norma minimální:

Příklad 3.8. Uvažujme následující systém lineárních rovnic o dvou neznámých:

$$\begin{aligned} x + y &= 1 \\ x + y &= 0. \end{aligned}$$

Matice tohoto systému A a matice B absolutních členů jsou dány identitami:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Pro psudoinverzi A^+ máme $A^+ = \frac{1}{4} \cdot A$, odkud plyne pro nejlepší přibližné řešení

$$X_0 = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Pak

$$A \cdot X_0 - B = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Pak pro minimální normu máme

$$||A \cdot X_0 - B||^2 = \frac{1}{2}.$$

Např. pro vektor X:

$$X = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}$$

dostaneme

$$||A \cdot X - B||^2 = \frac{1}{2} \quad \text{a} \quad ||X_0|| = \frac{1}{8} \quad < \quad \frac{1}{4} = ||X||^2.$$

4. Metody výpočtu pseudoinverzní matice

V současné době existuje celá řada metod výpočtu pseudoinverzní matice, z nichž většina může sloužit jako důkaz pro existenci této Moore-Penroseovy inverze. V tomto odstavci uvedeme dvě nejčastěji používané metody ([3]).

Metoda skeletního rozkladu matice.

Definice 4.1. Matice M typu $m \times n$ se nazývá matice *úplné řádkové (sloupcové) hodnosti*, jestliže hodnost matice M je rovna počtu řádků (sloupců), tedy

$$\mathrm{r}(M) = m \quad (\mathrm{r}(M) = n).$$

Pro existenci a vyjádření pseudoinverze matice s úplnou hodností potřebujeme lemma:

Lemma 4.2. *Jestliže matice M má úplnou řádkovou (sloupcovou) hodnost, pak matice $M \cdot M^*$ $(M^* \cdot M)$ je regulární, tedy má inverzní matici.*

Následující tvrzení udává formule pro pseudoinverzi matic s úplnou hodností:

Tvrzení 4.3. *Jestliže matice M má úplnou řádkovou sloupcovou hodnost, pak má pseudoinverzi a platí:*

$$M^+ = M^* \cdot (M \cdot M^*)^{-1} \quad (M^+ = (M^* \cdot M)^{-1} \cdot M^*).$$

Pro konstrukci pseudoinverze obecné matice se většinou používá následující věty:

Věta 4.4 (O skeletním rozkladu matice – Theorem on "rank factorization" of a matrix)**.** *Nechť A je nenulová matice typu $m \times n$ s hodností r. Pak existují matice B, C typu $m \times r$, $r \times n$ takové, že platí:*

$$A = B \cdot C, \quad r(B) = r(C) = r.$$

Rozklad $A = B \cdot C$ se nazývá skeletní rozklad matice A.

Analogem pro tvrzení $(M \cdot N)^{-1} = N^{-1} \cdot M^{-1}$ pro regulární matice M, N stejného řádu je věta:

Věta 4.5. *Jestliže $A = B \cdot C$ je skeletní rozklad matice A, pak*

$$\boxed{A^+ = (B \cdot C)^+ = C^+ \cdot B^+.}$$

Uvedená formule slouží k výpočtu pseudoinverze metodou skeletního rozkladu.

Příklad 4.6. Vypočítejte pseudoinverzi matice A soustavy linárních rovnic 3 neznámých z odstavce 3. Připomeňme

$$A = \begin{pmatrix} 1 & 4 & 3 \\ -1 & 1 & 2 \\ -2 & -2 & 0 \end{pmatrix}.$$

Matice A má skeletní rozklad $A = B \cdot C$, kde

$$B = \begin{pmatrix} 1 & 3 \\ -1 & 2 \\ -2 & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Z tvrzení 4.3 dostaneme

$$B^+ = (B^*B)^{-1}B^* = \begin{pmatrix} 6 & 1 \\ 1 & 13 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 & -2 \\ 3 & 2 & 0 \end{pmatrix}$$
$$= \frac{1}{77} \begin{pmatrix} 10 & -15 & -26 \\ 17 & 13 & 2 \end{pmatrix}.$$

Podobně dostaneme

$$C^+ = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

Odtud podle vzorce $A^+ = C^+ \cdot B^+$ obdržíme

$$A^+ = \frac{1}{231} \begin{pmatrix} 3 & -43 & -54 \\ 27 & -2 & -24 \\ 24 & 41 & 30 \end{pmatrix}.$$

Poznámka 4.7. Jestliže matice B, C vystupující ve skeletním rozkladu matice A mají za prvky celá čísla, pak matice

$$\det(B^* \cdot B) \cdot \det(C \cdot C^*) \cdot A^+$$

má za prvky také celá čísla. To vysvětluje koeficient $\frac{1}{231}$ z předešlého příkladu, protože $231 = 77 \cdot 3$ a det $(B^* \cdot B) = 77$, det $(C \cdot C^*) = 3$.

Grevilleův algoritmus

Grevilleův algoritmus je založen na rekurenci vzhledem k počtu sloupců matice. Pro matici, která má jen jeden sloupec, je její pseudoinverze daná následující formulí, která se snadno dokáže ([3]).

Tvrzení 4.8. *Pro matici A typu $m \times 1$ (A má jeden sloupec) máme*

$$A^+ = 0_{1,m}, \text{ jestliže } A = 0_{m,1},$$

a v případě $A \neq 0$, máme

$$A^+ = (A^* \cdot A)^{-1} \cdot A^*.$$

Tvrzení 4.6 je první krok při použití rekurence. Pro obecnou matici pak můžeme použít hlavní větu Grevilleova algoritmu.

Věta 4.9. *Předpokládejme, že umíme sestrojit pseudoinverzi matice, která má n sloupců (n je přirozené číslo) a nechť $M = (Ns)$ je matice typu $m \times (n+1)$,* *přičemž* s *je poslední sloupec matice* M *a* N *je podmatice matice* M *typu* $m \times n$. *Položme*

$$d = N^{+} \cdot s \qquad a \qquad c = s - N \cdot d.$$

Dále položme

$$b = \begin{cases} c^{+}, & \textit{jestliže} \quad c \neq 0 \\ (1 + d^{*}d)^{-1} d^{*} N^{+} & \textit{pro} \quad c = 0. \end{cases}$$

Pak

$$M^{+} = \begin{pmatrix} N^{+} - db \\ b \end{pmatrix}.$$

Reference

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

[2] A. Ben-Israel, T. N. E. Greville: *Generalized Inverses, Theory and Applications*, 2nd ed., Springer-Verlag, New York, 2003.

[3] P. J. Davis: *Circulant Matrices*, 2nd ed., New York, 1994.

[4] K. L. Doty, C. Melchiorri, C. Bonivento: *A Theory of Generalized Inverses Applied to Robotics*, The International Journal of Robotics, 1992, 36pp.

[5] F. R. Gantmacher: *The Theory of Matrices*, Chelsea, New York, 1959, anglický překlad z ruštiny.

[6] J. Karásek, L. Skula: *Lineární algebra*, Teoretická část, 3. vyd., FSI VUT, Brno, 2012.

[7] R. Piziak, P. L. Odell: *Matrix Theory, From Generalized Inverses to Jordan Form*, Chapman & Hall/CRC, 2007.

[8] S. S. Searle: *Matrix Algebra Useful for Statistics*, New York, 1982.

[9] F. Šik: *Lineární algebra zaměřená na numerickou analýzu*, MU, Brno, 1998.

Ladislav Skula, Ústav matematiky, Fakulta strojního inženýrství, Vysoké učení technické v Brně, Technická 2, 616 69 Brno, Česká republika,
e-mail: `skula@fme.vutbr.cz`