Közelítő és szimbolikus számítások

4. gyakorlat

Mátrix invertálás Gauss-eliminációval, Cholesky felbontás, QR felbontás

Készítette:

Gelle Kitti

Csendes Tibor

Somogyi Viktor

London András

Deák Gábor

jegyzetei alapján

1. Mátrix invertálás Gauss-eliminációval

Definíció 1.1. *Egy $n \times n$-es A mátrix invertálható, ha létezik olyan $n \times n$-es B mátrix, melyre igaz, hogy:*

$$AB = BA = I_n,$$

ahol I_n az $n \times n$-es egységmátrixot jelöli. Ekkor B-t az A mátrix inverzének nevezzük. Jelölése: A^{-1}.

1. Megjegyzés. Vegyük észre, hogy ez $n \times n$-es mátrixokra igaz, viszont $n \times m$-es mátrixok esetében nem ilyen egyszerű a helyzet. Azt mondjuk, hogy a nem négyzetes mátrixok nem invertálhatóak, de létezhet *bal-* vagy *jobbinverzük.* Ha az A $n \times m$-es mátrix rangja m, akkor létezik egy B mátrix, hogy $BA = I_m$. Ez a B mátrix A balinverze. Hasonlóan, ha az A $n \times m$-es mátrix rangja n, akkor létezik egy B mátrix, hogy $AB = I_n$. Ez a B mátrix A jobbinverze.

A Gauss-eliminációt a lineáris egyenletrendszerek megoldásán túl mátrixok invertálására is használhatjuk. Legyen A egy $n \times n$-es mátrix. Tekintsük az inverz mátrix definícióját, ahol most B-t, tehát az inverzet keressük, ez az ismeretlen. Legyen a könnyebb érthetőség miatt $B = X$. Ilyenkor tehát így néz ki az inverz egyenlete:

$$AX = I \tag{1}$$

Ezt szétbonthatjuk úgy, hogy az X és I mátrix oszlopait vesszük:

$$\begin{aligned} A\boldsymbol{x_1} &= \boldsymbol{e_1} \\ &\vdots \\ A\boldsymbol{x_n} &= \boldsymbol{e_n}, \end{aligned} \tag{2}$$

ahol $\boldsymbol{x_i}$ az X mátrix i-edik oszlopa, $\boldsymbol{e_i}$ pedig az I mátrix i-edik oszlopa (tehát az a vektor, aminek i-edik koordinátája 1 és a többi 0).

Ha a (2) egyenletben az A mátrixot háromszögmátrix alakra hozzuk, akkor az $MA = U$ alakra támaszkodva n darab felső háromszögmátrix alakú egyenletrendszert kell megoldani ahhoz, hogy megkapjuk az inverz mátrixot. Tehát balról szorzunk az M mátrixszal (ahol M az $M_{n-1} \dots M_1$ Gauss-transzformációs mátrixok szorzata ebben a sorrendben), így megkapjuk a kívánt alakot:

$$\begin{gathered} MA\boldsymbol{x_i} = M\boldsymbol{e_i} \\ \Downarrow \\ U\boldsymbol{x_i} = M\boldsymbol{e_i}. \end{gathered} \tag{3}$$

Példa. Határozzuk meg a lenti mátrix inverzét!

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix}$$

Megoldás. Először tehát írjuk fel, hogy milyen (1) alakú egyenletet kell megoldanunk:

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ezután bontsuk szét (2) alakú egyenletekre. Ezek (3) alakú formáit kell megoldanunk, amit a fent látható módon állítunk elő.

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Számoljuk ki U-t és M-et:

$$M_1 A = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -5/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 10/3 & 8/3 \\ 0 & -1/3 & -2/3 \end{pmatrix} = A_1$$

$$M_2 A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/10 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 0 & 10/3 & 8/3 \\ 0 & -1/3 & -2/3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 10/3 & 8/3 \\ 0 & 0 & -2/5 \end{pmatrix} = U$$

$$M = M_2 M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/10 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -5/3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -17/10 & 1/10 & 1 \end{pmatrix}$$

Oldjuk meg egyenként a fenti szétbontott egyenleteket. A megoldást itt csak röviden fogjuk vázolni.
Az első egyenlet megoldása:

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{11}\\ x_{21}\\ x_{31}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ -1/3 & 1 & 0\\ -17/10 & 1/10 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}$$

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{11}\\ x_{21}\\ x_{31}\end{pmatrix} = \begin{pmatrix}1\\ -1/3\\ -17/10\end{pmatrix}$$

$$x_{11} = 5/4,\ x_{21} = -7/2,\ x_{31} = 17/4$$

A második egyenlet megoldása:

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{12}\\ x_{22}\\ x_{32}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ -1/3 & 1 & 0\\ -17/10 & 1/10 & 1\end{pmatrix}\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}$$

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{12}\\ x_{22}\\ x_{32}\end{pmatrix} = \begin{pmatrix}0\\ 1\\ 1/10\end{pmatrix}$$

$$x_{12} = -1/4,\ x_{22} = 1/2,\ x_{32} = -1/4$$

A harmadik egyenlet megoldása: (érdemes észrevenni, hogy $M \cdot e_i$ az pont az M mátrix i. oszlopa lesz)

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{13}\\ x_{23}\\ x_{33}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ -1/3 & 1 & 0\\ -17/10 & 1/10 & 1\end{pmatrix}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

$$\begin{pmatrix}3 & 2 & 1\\ 0 & 10/3 & 8/3\\ 0 & 0 & -2/5\end{pmatrix}\begin{pmatrix}x_{13}\\ x_{23}\\ x_{33}\end{pmatrix} = \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

$$x_{13} = -1/2,\ x_{23} = 2,\ x_{33} = -5/2$$

Állítsuk össze az inverz mátrixot.

$$\begin{pmatrix}5/4 & -1/4 & -1/2\\ -7/2 & 1/2 & 2\\ 17/4 & -1/4 & -5/2\end{pmatrix}$$

Ellenőrzés. Végezzük el az $A^{-1}A$ műveletet.

$$\begin{pmatrix} 5/4 & -1/4 & -1/2 \\ -7/2 & 1/2 & 2 \\ 17/4 & -1/4 & -5/2 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. Cholesky felbontás

A lineáris egyenletrendszerek megoldása során megkülönböztetünk néhány speciális szerkezetű mátrixot, ezek könnyebbé teszik az életünket, a számolásokat így speciális esetekhez is köthetjük, ami gyakran egyszerűsít és gyorsít az eljáráson.

A Cholseky felbontást akkor használjuk, ha az adott A mátrix *szimmetrikus* és *pozitív definit.*

Definíció 2.1. *Egy A mátrix akkor szimmetrikus, ha $A = A^T$, azaz megegyezik a transzponáltjával.*

Definíció 2.2. *Egy A mátrix akkor pozitív definit, ha szimmetrikus, valamint $\boldsymbol{x}^T A\boldsymbol{x} > 0$ minden nem nulla $\boldsymbol{x}$ vektorra.*

Ez gyakorlatban úgy néz ki, hogy megvizsgáljuk a mátrix diagonális (főátlóban lévő) elemeit. Egy mátrix pontosan akkor pozitív definit, ha az összes főminorja pozitív.

Definíció 2.3. *Az A mátrix i-edik főminorja az A első i sorából és első i oszlopából képzett A_i mátrix determinánsa.*

Példa. Az $\begin{pmatrix} 5 & 2 & 6 \\ 4 & 9 & 1 \\ 2 & 1 & 7 \end{pmatrix}$ 2. főminorja a $det \begin{pmatrix} 5 & 2 \\ 4 & 9 \end{pmatrix} = 37.$

És például a pozitív definit mátrixokra az is igaz, hogy a főátlóban minden elem pozitív.

Ha az $A \in \mathbb{R}^{n\times n}$ mátrix szimmetrikus és pozitív definit, akkor létezik $U = L^T$ alakban felbontás, tehát

$$A = LU = LL^T$$

ahol L alsó trianguláris mátrix, melynek diagonális elemei pozitív számok. (Emlékezzünk, hogy az LU-nál csak egyesek lehettek.)

Definíció 2.4. *Az $A = LL^T$ felbontást Cholesky felbontásnak hívjuk. (Egyes szakirodalmakban* $\mathrm{R^TR}$ *felbontásként olvashatunk róla.)*

Példa. Vizsgáljuk meg a lenti mátrixot, és ha megfelel a feltételeknek, akkor határozzuk meg a Cholesky felbontását.

$$\begin{pmatrix} 9 & -6 & 3 \\ -6 & 20 & 18 \\ 3 & 18 & 62 \end{pmatrix}$$

Megoldás. Először azt kell megvizsgálnunk, hogy szimmetrikus-e (ez könnyen eldönthető, akár ránézésre is).

Ezután azt vizsgáljuk, hogy pozitív definit-e. Ehhez meg kell nézni, hogy a főminorjaink pozitívak-e:

$$\det(A_1) = 9$$

$$\det(A_2) = 144$$

$$\det(A_3) = 5184$$

tehát ez a feltétel is teljesül.

A determináns számítása részletezve:

$$\det \begin{pmatrix} 9 & -6 \\ -6 & 20 \end{pmatrix} = 9 \cdot 20 - ((-6) \cdot (-6)) = 144$$

$$\det \begin{pmatrix} 9 & -6 & 3 \\ -6 & 20 & 18 \\ 3 & 18 & 62 \end{pmatrix} =$$

$$9 \cdot 20 \cdot 62 + ((-6) \cdot 18 \cdot 3) + (3 \cdot (-6) \cdot 18) - (3 \cdot 20 \cdot 3) - ((-6) \cdot (-6) \cdot 62) - (9 \cdot 18 \cdot 18) = 5184$$

(Sarrus-szabállyal)

Ezután felírjuk az egyenletet, amit meg kell oldanunk:

$$\begin{pmatrix} 9 & -6 & 3 \\ -6 & 20 & 18 \\ 3 & 18 & 62 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix}$$

Végezzük el a jobb oldalon a szorzást.

$$\begin{pmatrix} 9 & -6 & 3 \\ -6 & 20 & 18 \\ 3 & 18 & 62 \end{pmatrix} = \begin{pmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} \\ l_{11}l_{21} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} \\ l_{11}l_{31} & l_{21}l_{31} + l_{22}l_{32} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{pmatrix}$$

Tehát az A mátrixunk egy, a jobb oldalon szereplő alakú szimmetrikus mátrix, azaz a megoldáshoz úgy jutunk, hogy kifejezzük a bal oldal segítségével a jobb oldali mátrix elemeit (először l_{11}-et, aztán l_{21}, $l_{31}, \ldots, l_{n1}$, aztán l_{22}, $l_{32}, \ldots, l_{n2}, \ldots$ - tehát oszloponként mindig egy-egy új ismeretlent fejezünk ki - sorrendben tudjuk megkapni ezeket):

$$l_{11} = \sqrt{l_{11}^2} = \sqrt{9} = 3$$

$$l_{11}l_{21} = -6 \Rightarrow 3l_{21} = -6 \Rightarrow l_{21} = -2$$

$$l_{11}l_{31} = 3 \Rightarrow 3l_{31} = 3 \Rightarrow l_{31} = 1$$

$$l_{21}^2 + l_{22}^2 = 20 \Rightarrow 4 + l_{22}^2 = 20 \Rightarrow l_{22}^2 = 16 \Rightarrow l_{22} = 4$$

$$l_{21}l_{31} + l_{22}l_{32} = 18 \Rightarrow -2 + 4l_{32} = 18 \Rightarrow 4l_{32} = 20 \Rightarrow l_{32} = 5$$

$$l_{31}^2 + l_{32}^2 + l_{33}^2 = 62 \Rightarrow 1^2 + 5^2 + l_{33}^2 = 62 \Rightarrow 26 + l_{33}^2 = 62 \Rightarrow l_{33}^2 = 36 \Rightarrow l_{33} = 6$$

A végeredmény tehát:

$$L = \begin{pmatrix} 3 & 0 & 0 \\ -2 & 4 & 0 \\ 1 & 5 & 6 \end{pmatrix}$$

Ellenőrzés.

$$LL^T = \begin{pmatrix} 3 & 0 & 0 \\ -2 & 4 & 0 \\ 1 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 & -2 & 1 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 9 & -6 & 3 \\ -6 & 20 & 18 \\ 3 & 18 & 62 \end{pmatrix} = A$$

2.1 Lineáris egyenletrendszer megoldása

Tegyük fel, hogy adott egy $A\boldsymbol{x} = \boldsymbol{b}$ alakú lineáris egyenletrendszerünk. Ekkor a következő lépésekkel oldhatjuk meg a Cholesky felbontás segítségével:

(1.) $LL^T\boldsymbol{x} = \boldsymbol{b}$ — az A mátrixot felbontjuk Cholesky alakba
(2.) $L\boldsymbol{y} = \boldsymbol{b}$ — az $L^T\boldsymbol{x} = \boldsymbol{y}$ helyettesítéssel megoldjuk $\boldsymbol{y}$-ra
(3.) $L^T\boldsymbol{x} = \boldsymbol{y}$ — végül az $\boldsymbol{y}$ segítségével kifejezzük $\boldsymbol{x}$-et.

Miért szeretjük? Ha alkalmazható a Cholesky felbontás, akkor a lineáris egyenletrendszerek megoldásánál közel kétszer olyan gyors, mint az LU felbontás. További előnye az LU-val szemben, hogy numerikusan stabilis (ez kb azt jelenti, hogy ,,kicsit" változtatva az inputot a két eredménymátrix is csak ,,kicsit" változik).

2.2 Cholesky felbontás Matlabbal

```
>> A = [1 2 3; 2 8 12; 3 12 27];
>> b = [14 54 108]';
>> L = chol(A, 'lower'); %Cholesky felbontas, also haromszogmatrix
>> y = L\b;
>> x = L'\y

x =

     1
     2
     3
```

3. QR felbontás

Definíció 3.1. *Egy Q négyzetes mátrixot ortogonálisnak nevezünk, ha $QQ^T = Q^TQ = I$, tehát a mátrix inverze megegyezik a transzponáltjával.*

Definíció 3.2. *Tetszőleges $A \in \mathbb{R}^{n \times n}$ reguláris mátrixnak létezik az $A = QR$ felbontása, ahol Q egy ortogonális mátrix, R pedig egy felső háromszögmátrix.*

3.1 Lineáris egyenletrendszer megoldása

Tegyük fel, hogy adott egy $A\boldsymbol{x} = \boldsymbol{b}$ alakú lineáris egyenletrendszerünk. Ekkor ha rendelkezésünkre áll az $A = QR$ felbontás, a Q^T-tal balról való beszorzás után a következő egyenletrendszert kell megoldanunk: $Q^TA\boldsymbol{x} = Q^TQR\boldsymbol{x} = R\boldsymbol{x} = Q^T\boldsymbol{b}$. Írjuk ki lépésenként a lineáris egyenletrendszerek megoldását a QR felbontás segítségével:

(1.) $QR\boldsymbol{x} = \boldsymbol{b}$ — az $A = QR$ alak megadása
(2.) $Q\boldsymbol{y} = \boldsymbol{b}$ — az $R\boldsymbol{x} = \boldsymbol{y}$ helyettesítéssel megoldjuk $\boldsymbol{y}$-ra
(3.) $R\boldsymbol{x} = \boldsymbol{y}$ — végül az $\boldsymbol{y}$ segítségével kifejezzük $\boldsymbol{x}$-et.

Ez azért lesz jó, mert az ortogonális Q^T-tal való szorzás nem növeli az esetlegesen A-ban vagy b-ben lévő hibákat. Viszont a QR felbontást valamikor elő kell állítanunk, ennek a költsége pedig nagyjából háromszorosa az LU felbontás előállításának.

Akkor mégis miért használjuk? Az LU felbontás sok esetben (főleg, ha főelemkiválasztás nélkül alkalmazzuk) nem fog létezni, és ha a mátrixunk nem szimmetrikus és pozitív definit, Cholesky felbontást sem tudunk csinálni, ilyenkor hasznos a QR felbontás. Viszont mivel lassabb és nehezebben programozható, egyenletrendszerek megoldására általában nem használják (kivéve persze, ha szinguláris az egyenletrendszer), viszont sok más dologra igen (látunk majd használatára példát néhány gyakorlat múlva).

3.2 QR felbontás Matlabban

```
A=[2 1 5; 2 1 -1; 3 3 -2];
b=[2,1,-1]';
[Q,R] = qr(A);
x = R\(Q'*b)

x =
     2.2778
    -2.7222
    -0.1667
```