

KÖZELÍTŐ ÉS SZIMBOLIKUS SZÁMÍTÁSOK

5. GYAKORLAT

Iterációs módszerek: Jacobi és Gauss-Seidel iteráció

Készítette:

Gelle Kitti

Csendes Tibor

Somogyi Viktor

London András

Deák Gábor

jegyzetei alapján

1. Iterációs módszerek

Korábban már vettünk lineáris egyenletrendszerek megoldására szolgáló algoritmusokat (LU, Cholesky és QR felbontások). Az eddig tanult módszereket *direkt módszerek*nek nevezzük. Az ilyen típusú eljárások véges sok meghatározott számú lépésben megadják a megoldást, úgy, hogy az egyenletrendszert olyan alakra hozzák elemi átalakításokkal, hogy a megoldások könnyen leolvashatóak legyenek. Ezek a kisebb vagy közepes méretű lineáris egyenletrendszerek megoldására (esetleg olyanokra, amelyek nagyobb sávmátrix alakúak) ajánlottak, mivel a műveletigényük ott még elviselhető.

Viszont a nagy méretű, nem feltétlenül ritka együttható-mátrixú egyenletrendszerek megoldásakor ezek igen nagy műveletigényűek lehetnek. Ezekre alkalmasabbak az *iterációs módszerek* (továbbá olyan esetekben is használják, amikor az eliminációs módszerek kerekítési hibáival terhelt eredményeket pontosítani kell). Az iterációs módszerek olyan eljárások, amelyek adott kezdeti értékből kiindulva minden iterációval jobb közelítést adják a megoldásnak, de azt általában véges sok lépésben nem érik el.

Az iterációs módszerek lényege, hogy egy olyan $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k)}, \dots$ vektorsorozatot generálunk, mely az adott egyenletrendszer megoldásához konvergál. Mivel számításaink minden esetben véges pontosságúak, így a kívánt pontosságot gyakran igen kevés lépésben elérhetjük.

Az iterációs módszerek alapja az ún. fixpontkeresés, melyet a következőképpen csinálunk: a Banach-féle fixponttétel szerint, ha f egy „jó tulajdonságú” függvény, akkor egyetlen olyan x^* pont létezik, hogy $x^* = f(x^*)$ (ezt hívjuk az f függvény fixpontjának), és ez megkapható úgy, hogy egy tetszőleges x_0 pontból kiindulva képezzük az $x^{(0)}, x^{(1)} = f(x^{(0)}), x^{(2)} = f(x^{(1)}), \dots, x^{(k+1)} = f(x^{(k)}), \dots$ sorozatot és vesszük a határértékét. Ekkor $x^* = \lim_{(k \rightarrow \infty)} x^{(k)}$. Tehát veszünk egy nekünk szimpatikus kezdővektort, és az új közelítő értéket úgy kapjuk, hogy az előzőn végrehajtjuk minden iterációban ugyanazt a műveletet. Az $x^* = \lim_{(k \rightarrow \infty)} x^{(k)}$ pedig azt jelenti, hogy ha „elég sok” iterációt csinálunk, akkor meg fogjuk kapni a megoldást (vagy legalábbis annak egy tetszőleges pontosságú közelítését).

Egy iterációs módszernek két fontos tulajdonsága van. Egyrészt, hogy konvergál-e, azaz hogy tetszőleges kezdővektorból indulva tényleg a (jelen esetben egyenletrendszer) megoldásának közelítését kapjuk. Másrészt, hogy mit választunk *leállási kritérium*nak. Általában akkor fejezünk be egy iterációt, ha annak két egymás utáni közelítése elég közel van egymáshoz, azaz $\|x^{(m)} - x^{(m+1)}\| \leq \epsilon$ teljesül, adott ϵ -ra. E mellé szokás még megadni egy maximális iterációs számot is.

2. Jacobi iteráció

A feladat tehát egy $A\mathbf{x} = \mathbf{b}$ egyenletrendszer megoldása. Itt az egyenletrendszert átrendezzük olyan alakra, hogy a bal oldalon a változók álljanak (x_1, \dots, x_n) , jobb oldalon pedig az egyenletrendszer többi része és az eredeti jobb oldal. Ezt minden egyenletnél úgy tudjuk megoldani, hogy az i -edik egyenletben az i -edik változó együtthatójával osztunk, majd az i -edik tagon kívül mindegyiket kivonjuk az egyenletből.

Formálisan a következő alakban írhatjuk fel:

$$\mathbf{x}^{(k+1)} = D^{-1}\mathbf{b} - D^{-1}(A - D)\mathbf{x}^{(k)}.$$

Itt a D egy diagonális mátrix, amely az A főátlóbeli elemeit tartalmazza. Ennek az inverze pont az a mátrix, ahol D diagonális elemeit felváltják azok reciprokai. Vegyük észre, hogy a D^{-1} -zel balról való szorzás pont azt eredményezi, mintha az i -edik egyenletet elosztottuk volna az i -edik változó együtthatójával. A képlet kifejezi továbbá a kivonást is, ugyanis az új egyenletünkben az eredeti A főátlóbeli elemein kívül minden tagját átvittük a jobb oldalra kivonással (ezt fejezi ki az $A - D$ tag).

A továbbiakban legyen $\mathbf{c} = D^{-1}\mathbf{b}$ és $B = -D^{-1}(A - D)$. Ekkor az iterációs egyenletünk

$$\mathbf{x}^{(k+1)} = \mathbf{c} + B\mathbf{x}^{(k)}.$$

Hogyan is néz ki ez egy konkrét sorra az egyenletrendszerből? Nézzük az i -edik egyenletet:

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad 1 \leq i \leq n.$$

Tegyük fel, hogy $a_{ii} \neq 0$ ($i = 1, \dots, n$). Az i -edik egyenletből fejezzük ki az i -edik változót, s ezzel megkapjuk a fenti alakú iterációs formát:

$$x_i^{(k+1)} = \frac{b_i}{a_{ii}} - \sum_{(\ell=1, \ell \neq i)}^n \frac{a_{i\ell}}{a_{ii}} x_\ell^{(k)}$$

ahol $i = 1, 2, \dots, n$; $k = 1, 2, 3 \dots$ az iterációs index és $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)}$ kezdőértékek.

Példa. Hajtsunk végre néhány Jacobi iterációt a következő $A\mathbf{x} = \mathbf{b}$ alakú egyenletrendszerünkön:

$$\begin{pmatrix} 5 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 2 \end{pmatrix}$$

Megoldás. Írjuk fel egyenletrendszer formájában a fenti mátrixos alakot:

$$5x_1 + x_2 = 7$$

$$x_1 + 2x_2 = 5$$

$$x_2 + 3x_3 = 2$$

Rendezzük át az egyenletrendszerünket $\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$ alakúra:

$$x_1^{(k+1)} = -\frac{1}{5}x_2^{(k)} + \frac{7}{5}$$

$$x_2^{(k+1)} = -\frac{1}{2}x_1^{(k)} + \frac{5}{2}$$

$$x_3^{(k+1)} = -\frac{1}{3}x_2^{(k)} + \frac{2}{3}$$

Írjuk vissza mátrix alakra:

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -1/5 & 0 \\ -1/2 & 0 & 0 \\ 0 & -1/3 & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} + \begin{pmatrix} 7/5 \\ 5/2 \\ 2/3 \end{pmatrix}$$

Induljunk el az $x^{(0)} = (1 \ 1 \ 1)^T$ kezdővektorból.

1. ITERÁCIÓ.

$$x^{(1)} = \begin{pmatrix} 0 & -1/5 & 0 \\ -1/2 & 0 & 0 \\ 0 & -1/3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 7/5 \\ 5/2 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 2 \\ 1/3 \end{pmatrix}$$

2. ITERÁCIÓ.

$$x^{(2)} = \begin{pmatrix} 0 & -1/5 & 0 \\ -1/2 & 0 & 0 \\ 0 & -1/3 & 0 \end{pmatrix} \begin{pmatrix} 6/5 \\ 2 \\ 1/3 \end{pmatrix} + \begin{pmatrix} 7/5 \\ 5/2 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1.9 \\ 0 \end{pmatrix}$$

3. ITERÁCIÓ.

$$x^{(3)} = \begin{pmatrix} 0 & -1/5 & 0 \\ -1/2 & 0 & 0 \\ 0 & -1/3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1.9 \\ 0 \end{pmatrix} + \begin{pmatrix} 7/5 \\ 5/2 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 1.02 \\ 2 \\ 0.03 \end{pmatrix}$$

A pontos gyökök az $x = (1 \ 2 \ 0)^T$ vektor elemei.

3. Gauss-Seidel iteráció

A módszer a Jacobi iteráció egy olyan módosítása, ahol a $(k + 1)$ -edik közelítő vektor i -edik koordinátájának kiszámításakor figyelembe vesszük, hogy az $1., 2., \dots, (i - 1)$. koordinátákat már kiszámítottuk, így azokat használjuk fel $\mathbf{x}^{(k)}$ megfelelő koordinátái helyett. Tehát az $\mathbf{x}^{(k+1)}$ i . komponensének kiszámításához az ebben az iterációs lépésben kapott értékeket is felhasználjuk és csak i -nél nagyobb komponensek esetén használjuk az $\mathbf{x}^{(k)}$ i -nél nagyobb komponenseinek értékét. Az $\mathbf{x}^{(k+1)}$ első komponensét a Jacobi eljárással határozzuk meg. (Ez egyébként egy bevett módszer az algoritmusok javítására, bár nem biztos, hogy minden esetben tényleg jobb lesz.) Képlete:

$$x_i^{(k+1)} = \frac{b_i}{a_{ii}} - \sum_{\ell=1}^{(i-1)} \frac{a_{i\ell}}{a_{ii}} x_\ell^{(k+1)} - \sum_{(\ell=i+1)}^n \frac{a_{i\ell}}{a_{ii}} x_\ell^{(k)}$$

ahol $i = 1, 2, \dots, n$; $k = 1, 2, 3 \dots$ az iterációs index és $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)}$ kezdőértékek.

Példa. Az előző feladat így nézne ki Gauss-Seidel iterációval:

$$\begin{aligned} x_1^{(k+1)} &= -\frac{1}{5}x_2^{(k)} + \frac{7}{5} \\ x_2^{(k+1)} &= -\frac{1}{2}x_1^{(k+1)} + \frac{5}{2} \\ x_3^{(k+1)} &= -\frac{1}{3}x_2^{(k+1)} + \frac{2}{3} \end{aligned}$$

Induljunk el az $x^{(0)} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$ kezdővektorból.

1. ITERÁCIÓ.

$$\begin{aligned} x_1^{(1)} &= -\frac{1}{5}1 + \frac{7}{5} \implies x_1^{(1)} = 1.2 \\ x_2^{(1)} &= -\frac{1}{2}1.2 + \frac{5}{2} \implies x_2^{(1)} = 1.9 \\ x_3^{(1)} &= -\frac{1}{3}1.9 + \frac{2}{3} \implies x_3^{(1)} = 0.03 \end{aligned}$$

2. ITERÁCIÓ.

$$\begin{aligned} x_1^{(2)} &= -\frac{1}{5}1.9 + \frac{7}{5} \implies x_1^{(2)} = 1.02 \\ x_2^{(2)} &= -\frac{1}{2}1.02 + \frac{5}{2} \implies x_2^{(2)} = 1.99 \\ x_3^{(2)} &= -\frac{1}{3}1.99 + \frac{2}{3} \implies x_3^{(2)} = 0.003 \end{aligned}$$

4. Az iterációk konvergenciája

Mint az elején említettük, az iterációs módszerek egy fontos tulajdonsága, hogy konvergensek-e, azaz „elég sok” iterációt végrehajtva közelítjük-e a keresett megoldást. Ahhoz, hogy tudjunk beszélni a fenti iterációs módszerek konvergenciájáról, szükségünk van a következő fogalomra:

Definíció 4.1. *Szigorúan diagonális domináns mátrix: az $n \times n$ -es A mátrix soronként (szigorúan) domináns főátlójú, ha a főátló minden eleme abszolút értékben nagyobb a sorában lévő többi elem abszolút értékeinek összegénél, azaz képletben:*

$$|a_{ii}| > \sum_{(j=1, j \neq i)}^n |a_{ij}|$$

minden $i = 1, 2, \dots, n$ -re.

Definíció 4.2. *Akkor mondjuk, hogy egy iterációs sorozat globálisan konvergens, ha tetszőleges kezdővektorból indulva a módszer a megoldáshoz konvergál.*

Elegendő feltétel az iterációk konvergenciájára: Ha az n egyenletből álló n -ismeretlenes egyenletrendszer együtthatómátrixa szigorúan diagonális domináns, akkor bármely indulóvektor esetén a Jacobi- és a Gauss-Seidel iteráció is konvergens (tehát globálisan konvergens).

Ez a feltétel csak elegendő, nem pedig szükséges a konvergenciához, tehát előfordulhat olyan, hogy valamelyik módszer konvergál olyan egyenletrendszer esetén, melynek nem szigorúan diagonális domináns az együttható mátrixa.

Továbbá általánosan elmondható, hogy a Jacobi eljárás tetszőleges kezdeti vektor esetén konvergál, ha:

$$\max_k \sum_{(i=1, i \neq k)}^n \left| \frac{a_{ik}}{a_{ii}} \right| < 1$$

oszlopösszeg feltétel, vagy a

$$\max_i \sum_{(i=1, i \neq k)}^n \left| \frac{a_{ik}}{a_{ii}} \right| < 1$$

sorösszeg feltétel vagy a

$$\left(\sum_{(i=1, i \neq k)}^n \left| \frac{a_{ik}}{a_{ii}} \right|^2 \right)^{1/2} < 1$$

feltétel teljesül.

Tétel(Sassenfeld kritérium¹): Tegyük fel hogy a fent definiált B mátrix eleget tesz a következő feltételnek:

$$p = \max_{j=1, \dots, n} p_j < 1$$

, ahol a p_j számokat a:

$$p_1 = \sum_{(k=2)}^n \left| \frac{a_{1k}}{a_{11}} \right|, \quad p_j = \sum_{(k=1)}^{(j-1)} \left| \frac{a_{jk}}{a_{jj}} \right| p_k + \sum_{(k=j+1)}^n \left| \frac{a_{jk}}{a_{jj}} \right|$$

($j = 2, \dots, n$) rekurzióval határozzuk meg. Akkor a Gauss-Seidel iteráció minden kezdővektorból indítva konvergál a megoldáshoz annál gyorsabban, minél kisebb ez a p érték.

Szükséges és elegendő feltétel az iterációk konvergenciájára: Mindkét iterációs eljárás pontosan akkor konvergens, ha a $\rho(B)$ spektrálsugár (azaz a legnagyobb abszolútértékű sajátérték abszolútértékére) teljesül, hogy $\rho(B) < 1$. (Erről még beszélünk egy későbbi gyakorlaton, az előadáson ennek a bizonyítása szerepel.)

¹R. Kress: Numerical Analysis, Springer Verlag, 1988