ELTE IK - Programtervező Informatikus BSc

Záróvizsga tételek

3.1. Lineáris egyenletrendszerek iterációs módszerei

0.1 Lineáris egyenletrendszerek iterációs módszerei

A lineáris egyenletrendszert (LER) vektorsorozatokkal közelítjük, törekedve a minél gyorsabb konvergenciára. Az iterációs módszereknek a lényege az $Ax = b \iff x = Bx + c$ átalakítás. Ilyen alak létezik, sőt nem egyértelmű, hanem sokféle lehet, és a különböző átalakítások szolgáltatják a különféle iterációs módszereket.

Definíció (kontrakció): Az $F : \mathbb{R}^n \to \mathbb{R}^n$ függvény kontrakció, ha $\exists 0 \leq q < 1 : \forall x, y \in \mathbb{R}^n$:

$$\|F(x) - F(y)\| \leq q\|x - y\|$$

.

A q értéket kontrakciós együtthatónak nevezzük.

Banach-féle fixponttétel: Legyen $F : \mathbb{R}^n \to \mathbb{R}^n$ kontrakció a q kontrakciós együtthatóval. Ekkor a következő állítások igazak:

1. $\exists! x^* \in \mathbb{R}^n : x^* = f(x^*)$. Azt mondjuk, hogy x^* az f függvény fixpontja.
2. $\forall x^{(0)} \in \mathbb{R}^n$ kezdőérték esetén az $x^{(k+1)} = f(x^{(k)})$ sorozat konvergens, és $\lim\limits_{k\to\infty} x^{(k)} = x^*$.
3. $\|x^{(k)} - x^*\| \leq \frac{q^k}{1-q}\|x^{(1)} - x^0\|$.
4. $\|x^{(k)} - x^*\| \leq q^k\|x^{(0)} - x^*\|$.

Vegyük észre, hogy az $Ax = b \iff x = Bx + c$ átírással megteremtettük a kapcsolatot a Banach-féle fixponttétellel, hisz most az $F(x) = Bx + c$ függvény fixpontját keressük. A fenti felírásban B-t átmenetmátrixnak nevezzük.

Tétel (elégséges feltétel a konvergenciára): Ha a LER B átmenetmátrixára $\|B\| < 1$, akkor tetszőleges $x^{(0)}$-ból indított $x^{(k+1)} := Bx^{(k)} + c$ iteráció konvergál az $Ax = b$ LER megoldásához.

Tétel (Szükséges és elégséges feltétel a konvergenciára): Tetszőleges $x^{(0)}$-ból indított $x^{(k+1)} := Bx^{(k)} + c$ iteráció konvergál az $Ax = b$ LER megoldásához $\iff \varrho(B) < 1$, ahol $\varrho(B) = \max_{1\leq i\leq n} |\lambda_i(B)|$ a B mátrix spektrálsugara.

0.1.1 Jacobi-iteráció

Tekintsük az $A \in \mathbb{R}^{n\times n}$ mátrix $L + D + U$ felbontását, ahol L a mátrix szigorú alsó része, U a szigorú felső része, D pedig a diagonális része. Ennek segítségével konstruáljuk meg a következő átírást:

$$Ax = b \iff (L + D + U)x = b \iff Dx = -(L + U)x + b \iff x = -D^{-1}(L + U)x + D^{-1}b$$

A Jacobi-iteráció átmenetmátrixa tehát $B_J = -D^{-1}(L+U)$, maga az iteráció pedig:

$$x^{(k+1)} := -D^{-1}(L+U)x^{(k)} + D^{-1}b$$

Koordinátás alakban felírva:

$$x_i^{(k+1)} := -\frac{1}{a_{ii}}\left[\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}x_j^{(k)} - b_i\right] (i=1,...,n)$$

Tétel: Ha az A mátrix szigorúan diagonálisan domináns a soraira, akkor $\|B_J\|_\infty < 1$ (azaz konvergens a módszer).

Tétel: Ha az A mátrix szigorúan diagonálisan domináns az oszlopaira, akkor $\|B_J\|_1 < 1$ (azaz konvergens a módszer).

0.1.2 Csillapított Jacobi-iteráció

Továbbra is a Jacobi-iterációval foglalkozunk, csak egy plusz ω paraméter bevezetésével próbáljuk finomítani a módszert. Tekintsük a $Dx = -(L+U)x + b$ egyenletet, valamint a triviális $Dx = Dx$ egyenletet. Ezeket rendre szorozzuk meg ω, illetve $1-\omega$ értékekkel, majd adjuk össze a két egyenletet:

$$Dx = (1-\omega)Dx - \omega(L+U)x + \omega b$$

Szorozzunk D^{-1}-zel:

$$x = (1-\omega)Ix - \omega D^{-1}(L+U)x + \omega D^{-1}b \Longleftrightarrow x = ((1-\omega)I - \omega D^{-1}(L+U))x + \omega D^{-1}b$$

Ez alapján $B_{J(\omega)} = (1-\omega)I - \omega D^{-1}(L+U)$ és $c_J(\omega) = \omega D^{-1}b$.

Észrevehető, hogy $\omega = 1$ esetén pont a Jacobi-iterációt kapjuk vissza.

Koordinátás alakban felírva:

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} - \frac{\omega}{a_{ii}}\left[\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}x_j^{(k)} - b_i\right] (i=1,...,n)$$

Tétel: Ha $J(1)$ konvergens, akkor $\omega \in (0,1)$-re $J(\omega)$ is az.

0.1.3 Gauss-Seidel iteráció

Egy másik lehetséges iteráció konstruálásának az ötlete a következő:

$$Ax = b \Longleftrightarrow (L+D+U)x = b \Longleftrightarrow (L+D)x = -Ux + b \Longleftrightarrow x = -(L+D)^{-1}Ux + (L+D)^{-1}b$$

Ez az ötlet szüli a Gauss-Seidel iterációt, vagyis:

$$x^{(k+1)} := -(L+D)^{-1}Ux^{(k)} + (L+D)^{-1}b$$

Az iteráció átmenetmátrixa tehát $B_S = -(L+D)^{-1}U$. A koordinátás alak felírásához kicsit átírjuk az iterációt:

$$(L+D)x^{(k+1)} = -Ux^{(k)} + b$$

$$Dx^{(k+1)} = -Lx^{(k+1)} - Ux^{(k)} + b$$

$$x^{(k+1)} = -D^{-1}\big[Lx^{(k+1)} + Ux^{(k)} - b\big]$$

$$x_i^{(k+1)} = -\frac{1}{a_{ii}}\left[\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} - b_i\right] \; (i = 1, ..., n)$$

Megjegyzés: Az implementáció során elég egyetlen x vektort eltárolni, és annak a komponenseit sorban felülírni, ugyanis láthatjuk, hogy az első $i-1$ komponenst már az "új", $x^{(k+1)}$ vektorból vesszük.

Tétel: Ha A szigorúan diagonálisan domináns

1. a soraira, akkor $\|B_S\|_\infty \leq \|B_J\|_\infty < 1$.
2. az oszlopaira, akkor $\|B_S\|_1 \leq \|B_J\|_1 < 1$.

Azaz a Gauss-Seidel is konvergens, és legalább olyan gyors, mint a Jacobi.

0.1.4 Relaxációs módszer

A relaxációs módszer lényegében a csillapított Gauss-Seidel iterációt jelenti. Ennek megkonstruálásához tekintsük az $(L + D)x = -Ux + b$ és $Dx = Dx$ egyenleteket. Ezeket rendre szorozzuk meg ω, illetve $1 - \omega$ értékekkel, majd adjuk össze a két egyenletet:

$$(D + \omega L)x = (1 - \omega)Dx - \omega Ux + \omega b$$

$$x = (D + \omega L)^{-1}\big[(1 - \omega)D - \omega U\big]x + \omega(D + \omega L)^{-1}b$$

Az iteráció tehát: $x^{(k+1)} = (D+\omega L)^{-1}\big[(1-\omega)D-\omega U\big]x^{(k)} + \omega(D+\omega L)^{-1}b$, ahol az átmenetmátrix: $B_{S(\omega)} = (D + \omega L)^{-1}\big[(1 - \omega)D - \omega U\big]$. A koordinátás alak felírásához itt is átírjuk kicsit az iterációt:

$$(D + \omega L)x^{(k+1)} = (1 - \omega)Dx^{(k)} - \omega Ux^{(k)} + \omega b$$

$$Dx^{(k+1)} = -\omega Lx^{(k+1)} - \omega Ux^{(k)} + \omega b + (1 - \omega)Dx^{(k)}$$

$$x^{(k+1)} = -\omega D^{-1}\big[Lx^{(k+1)} + Ux^{(k)} - b\big] + (1 - \omega)x^{(k)}$$

$$x_i^{(k+1)} = -\frac{\omega}{a_{ii}}\left[\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} - b_i\right] + (1 - \omega)x_i^{(k)} \; (i = 1, ..., n)$$

Vegyük észre, hogy $\omega = 1$ esetén a Gauss-Seidel iterációt kapjuk.

Tétel: Ha a relaxációs módszer konvergens minden kezdővektorból indítva, akkor $\omega \in (0, 2)$.

Megjegyzés: Ha $\omega \notin (0, 2)$, akkor általában nem konvergens a módszer (bár adott feladat esetén előfordulhat, hogy találunk olyan kezdővektort, amelyből indítva konvergál a módszer).

Tétel: Ha A szimmetrikus és pozitív definit és $\omega \in (0, 2)$, akkor a relaxációs módszer konvergens. Ennek következménye a Gauss-Seidel iteráció konvergenciája ($\omega = 1$ eset).

Tétel: Ha A tridiagonális, akkor $\varrho(B_S) = \varrho(B_J)^2$, azaz a Jacobi és Gauss-Seidel iteráció egyszerre konvergens, illetve divergens.

Tétel: Ha A szimmetrikus, pozitív definit és tridiagonális, akkor a $J(1)$, $S(1)$ és $S(\omega)$ $\omega \in (0,2)$- re konvergens, és $S(\omega)$-ra az optimális paraméter értéke:

$$\omega_0 = \frac{2}{1+\sqrt{1-\varrho(B_J)^2}}$$

0.1.5 Richardson-iteráció

Legyen $p \in \mathbb{R}$. Így

$$Ax = b \Longleftrightarrow 0 = -Ax + b \Longleftrightarrow 0 = -pAx + pb \Longleftrightarrow x = (I - pA)x + pb$$

Az iteráció tehát $x^{(k+1)} := (I-pA)x^{(k)} + pb$. Az átmenetmátrix: $B_{R(p)} = I - pA)$. Az $r^{(k)} := b - Ax^{(k)}$ vektort maradékvektornak (reziduumvektornak) nevezzük, hiszen

$$x^{(k+1)} = x^{(k)} - pAx^{(k)} + pb = x^{(k)} + pr^{(k)}$$

$$r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + pr^{(k)}) = r^{(k)} - pAr^{(k)}$$

Tekintsük az előállítás algoritmusát: $r^{(0)} := b - Ax^{(0)}$, továbbá a fentiek miatt:

$$x^{(k+1)} := x^{(k)} + pr^{(k)}$$

$$r^{(k+1)} := r^{(k)} - pAr^{(k)}$$

Tétel: Ha A szimmetrikus, pozitív definit, a sajátértékei pedig a következők:

$$0 < m := \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n =: M$$

akkor $p \in (0, \frac{2}{M})$ esetén $R(p)$ konvergens, és az optimális paraméter: $p_0 = \frac{2}{m+M}$. Továbbá igaz, hogy: $\varrho(B_{R(p_0)}) = \frac{M-m}{M+m}$.

0.2 Spline-interpoláció

Az eddig említett interpolációs módszerekben polinomokkal dolgoztunk. Lehetőség van arra is, hogy a megadott pontrendszerre más típusú függvényt próbáljunk illeszteni. Igen előnyös tulajdonságokkal rendelkeznek a bizonyos folytonossági előírásoknak is megfelelő, szakaszonként polinom függvények, a spline-ok.

l-edfokú spline: Legyen adott $\Omega_n = \{x_0, x_1, ..., x_n\}$ az $[a,b]$ intervallum egy felosztása, ahol $x_0 = a, x_n = b$ és $l \in \mathbb{N}$. Az $s : [a,b] \to \mathbb{R}$ függvény egy l-edfokú spline az Ω_n-re vonatkozóan, ha:

1. $s_{|[x_{k-1},x_k]}$ egy l-edfokú polinom $\forall k = 1,..,n$
2. $s \in C^{l-1}[a,b]$, tehát a teljes intervallumon $(l-1)$-szer folytonosan derviálható

Jelölés: $s_l(\Omega_n)$ az Ω_n-hez tartozó l-edfokú spline-ok halmaza.

Spline-interpoláció: Legyenek adottak $x_k, f(x_k)$ értékek $k = 0,1,..,n$-re és $l \in \mathbb{N}$. Keressük azt az $s \in s_l(\Omega_n)$ spline-t, amelyre $s(x_k) = f(x_k)$. Ehhez elő kell állítanunk minden intervallumra egy l-edfokú polinomot. Ha a polinomokat az együtthatóikkal reprezentáljuk, akkor ez $n(l+1)$ ismeretlen. Az előírt feltételek száma: $2n$ interpolációs és $(l-1)(n-1)$ folytonossági feltétel, hiszen csak a belső pontokban kell előírni az illető deriváltakra vonatkozó megfelelő folytonossági feltételt. Az így kapott összes feltétel darabszáma $(l+1)n - (l-1)$, tehát az egyértelműséghez $l-1$ feltétel hiányzik még. Ezeket úgynevezett peremfeltételekkel adjuk meg. Pl. a harmadfokú spline-interpolációhoz 2 peremfeltétel szükséges. Ezek a következők (ezek közül elég egyet választani, mert mindegyik 2 feltételt tartalmaz):

1. Természetes peremfeltétel: $s''(a) = s''(b) = 0$.
2. Hermite-féle peremfeltétel: $s'(a) = s_a, s'(b) = s_b$, ahol s_a, s_b előre megadott számok.
3. Periodikus peremfeltétel (ekkor feltételezzük, hogy $s(a) = s(b)$ is teljesül): $s'(a) = s'(b)$ és $s''(a) = s''(b)$

Elsőfokú spline előállítása: Az elsőfokú spline előállítása triviális szakaszonkénti lineáris Lagrange-interpolációval.

Másodfokú spline előállítása: Egyetlen peremfeltétel szükséges, legyen a következő: $s'(a) = s_a$ valamilyen s_a számra. Az $[x_0, x_1]$ szakaszon Hermite-interpolációval előállítjuk azt a H_2 másodfokú polinomot, amely megfelel az interpolációs feltételeknek és a peremfeltételnek. Az így kapott polinom x_1-beli deriváltja meghatározott, tehát a folytonos deriválhatóság miatt az $[x_1, x_2]$ szakaszon a bal végpontban adott a derivált értéke. Ismét Hermite-interpolációt alkalmazva megkapjuk az $[x_1, x_2]$ szakaszhoz tartozó polinomot. Ezt az eljárást ismételve állíthatjuk elő a másodfokú interpolációs spline-t.

Függvény tartója: A $supp(f) := \overline{\{x \in \mathbb{R} : f(x) \neq 0\}}$ halmazt az f függvény tartójának nevezzük.

Számegyenes felosztása: $\Omega_\infty := \{..., x_{-2}, x_{-1}, x_0, x_1, ...x_n, x_{n+1}, ...\}$

B-spline: A $B_{l,k}, k \in \mathbb{Z}$ l-edfokú spline függvények rendszerét B-spline függvényeknek nevezzük, ha az alábbi feltételek teljesülnek:

1. $supp(B_{l,k}) = [x_k, x_{k+l+1}]$, azaz a tartója minimális
2. $B_{l,k}(x) \geq 0$
3. $\sum\limits_{k \in \mathbb{Z}} B_{l,k}(x) = 1$