

NUMERIKUS MÓDSZEREK
(Oktatási segédlet levelező hallgatóknak)

Galántai–Jeney könyve alapján összeállította: Jeney András

2004

Tartalomjegyzék

1. A klasszikus hibaszámítás elemei	5
1.1. Az aritmetikai műveletek abszolút hibái	5
1.2. Függvényértékek hibája	6
1.3. Az aritmetikai műveletek relatív hibái	6
2. A mátrixszámítás elemei	7
2.1. Mátrixok és mátrixműveletek	7
2.2. Mátrixok inverze és determinánsa	10
2.3. Vektorok és mátrixok normája	11
3. Lineáris egyenletrendszerek megoldása	13
3.1. Lineáris egyenletrendszerek	13
3.2. Háromszögmátrixú egyenletrendszerek	14
3.3. A Gauss-módszer	15
3.4. Az LU-felbontás és alkalmazása lineáris egyenletrendszerekre	17
3.5. Lineáris egyenletrendszerek közelítő megoldása iterációval	19
4. A sajátérték-feladat	21
4.1. A hatvány módszer	22
5. Nemlineáris egyenletek közelítő megoldási módszerei	23
5.1. Egyváltozós egyenletek megoldása	23
5.1.1. Az intervallumfelező eljárás	23
5.1.2. Egyszerű iteráció	24
5.1.3. A Newton-módszer	24
5.2. Nemlineáris egyenletrendszerek megoldása Newton-módszerrel	25
6. Függvények közelítése interpolációval	27
6.1. A Lagrange-féle interpolációs feladat	27
7. Numerikus deriválás és integrálás	31
7.1. Numerikus deriválás	31
7.2. Numerikus integrálás	32
7.2.1. A trapézformula	32
7.2.2. A Simpson formula	33

8. Függvényközelítés legkisebb négyzetek módszerével	35
8.1. Diszkrét eset	35
8.2. Folytonos eset	36
9. Differenciálegyenletek numerikus megoldása	39
9.1. Az explicit Euler-módszer	39
9.2. A negyedrendű Runge-Kutta módszer	40

1. fejezet

A klasszikus hibaszámítás elemei

A klasszikus hibaszámítás alapmodellje a következő. A pontos értékeket nem ismerjük, csak adott hibakorlátú közelítéseiket. A közelítő értékekkel pontosan végzett műveletek eredményét az ismeretlen elméleti eredmény közelítésének tekintjük és azt vizsgáljuk, hogy mekkora a közelítés hibája. Például a $\sqrt{2} \approx 1.41$ közelítés hibája legfeljebb 0.01.

A következő jelöléseket és elnevezéseket használjuk: x pontos érték, a az x közelítése ($a \approx x$), $\Delta a = x - a$ a közelítés hibája, δa az a közelítő érték abszolút hibakorlátja, ha fennáll $|x - a| = |\Delta a| \leq \delta a$.

1.1. Az aritmetikai műveletek abszolút hibái

Legyen x és y két pontos érték, a az x , b pedig az y közelítése. Tegyük fel, hogy az a és b közelítések abszolút hibakorlátai δa , ill. δb . Ezekre fennáll, hogy

$$|x - a| = |\Delta a| \leq \delta a, \quad |y - b| = |\Delta b| \leq \delta b.$$

Tétel. Igazak a következő becslések:

$$\delta(a + b) \leq \delta a + \delta b, \quad (1.1)$$

$$\delta(a - b) \leq \delta a + \delta b. \quad (1.2)$$

Tétel. Fennállnak az alábbi közelítő egyenlőségek:

$$\delta(ab) \approx |a| \delta b + |b| \delta a \quad (|a| \gg \delta a, |b| \gg \delta b), \quad (1.3)$$

$$\delta(a/b) \approx \frac{|a| \delta b + |b| \delta a}{|b|^2} \quad (b \neq 0, |b| \gg \delta b). \quad (1.4)$$

Figyeljük meg, hogy osztás abszolút hibakorlátja 0-hoz közeli b esetén rendkívül nagy lehet!

1.2. Függvényértékek hibája

Külön foglalkozunk az egy- és a többváltozós esetekkel. Legyen $f : R \rightarrow R$ legalább kétszer folytonosan differenciálható függvény, $x \approx a$. Az $f(x)$ helyett $f(a)$ -t számoljuk. Az $f(a)$ közelítés $\delta(f(a))$ hibakorlátjára a másodrendű Taylor-formulából bizonyos elhanyagolással kapjuk, hogy

$$\delta(f(a)) \approx |f'(a)| \delta a. \quad (1.5)$$

Legyen $F : R^n \rightarrow R$ legalább kétszer folytonosan differenciálható függvény, $x, a \in R^n$ és $x \approx a$. Legyen $\Delta a = x - a$, $|x_i - a_i| = |\Delta a_i| \leq \delta a_i$ ($i = 1, \dots, n$) és $\delta a = [\delta a_1, \dots, \delta a_n]^T$. Az $F(x)$ helyett az $F(a)$ függvényértéket számoljuk, amelynek $\delta(F(a))$ hibakorlátja (azaz $|F(x) - F(a)|$ egy felső korlátja hasonlóan a (többváltozós) másodrendű Taylor-formulából bizonyos elhanyagolással hasonlóan adódik:

$$\delta(F(a)) \approx \sum_{i=1}^n \left| \frac{\partial F(a)}{\partial x_i} \right| \delta a_i \quad (1.6)$$

1.3. Az aritmetikai műveletek relatív hibái

Az abszolút hiba sok esetben semmitmondó. Például egy 0.001 nagyságrendű elméleti mennyiség esetén a 0.05 abszolút hibakorlát nem sokat ér. A $\pi \approx 22/7$ közelítés sok esetben jó lehet, de például a csillagászatban már bizonyosan nem.

Definíció. Az x szám a közelítő értékének relatív hibája a $\frac{\delta a}{|x|}$ mennyiség.

Az x pontos érték általában nem ismeretes. Helyette a $\frac{\delta a}{|a|}$ közelítést használjuk.

$$\frac{\delta(a+b)}{|a+b|} = \max \left\{ \frac{\delta a}{|a|}, \frac{\delta b}{|b|} \right\} \quad (ab > 0), \quad (1.7)$$

$$\frac{\delta(a-b)}{|a-b|} = \frac{\delta a + \delta b}{|a-b|} \quad (ab > 0), \quad (1.8)$$

$$\frac{\delta(ab)}{|ab|} \approx \frac{\delta a}{|a|} + \frac{\delta b}{|b|}, \quad (1.9)$$

$$\frac{\delta(\frac{a}{b})}{|\frac{a}{b}|} \approx \frac{\delta a}{|a|} + \frac{\delta b}{|b|}. \quad (1.10)$$

Figyeljük meg, hogy egymáshoz közeli a és b esetén a kivonás relatív hibája rendkívül nagy lehet!

2. fejezet

A mátrixszámítás elemei

2.1. Mátrixok és mátrixműveletek

Röviden összefoglaljuk azokat a mátrixokkal és vektorokkal kapcsolatos ismereteket, amelyekre szükségünk van.

Definíció. Egy A $m \times n$ típusú (valós) mátrixon valós a_{ij} számok alábbi táblázatát értjük:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}.$$

Az a_{ij} az A mátrix i -edik sorában és j -edik oszlopában álló mátrixelemet jelöli. Mátrixok szokásos jelölése még a következő:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}.$$

Néhány tömörebb mátrixmegadási mód:

$$A = [a_{ij}]_{i,j=1}^{m,n}, \quad A = (a_{ij})_{i,j=1}^{m,n}.$$

Az $n \times m$ típusú valós mátrixok halmazát $R^{n \times m}$ jelöli. Az A mátrixot négyzetesnek nevezzük, ha $m = n$. Ekkor a tömör megadási módok a következőképpen egyszerűsödnek:

$$A = [a_{ij}]_{i,j=1}^n, \quad A = (a_{ij})_{i,j=1}^n.$$

A mátrixok közti fontosabb műveleteket az alábbiak szerint definiáljuk.

1. Összeadás: $A, B \in R^{m \times n}$,

$$C = A + B \in R^{m \times n} \Leftrightarrow c_{ij} = a_{ij} + b_{ij} \quad (i = 1, \dots, m, j = 1, \dots, n).$$

Az összeadásra fennáll, hogy

$$A + B = B + A, \quad (A + B) + C = A + (B + C).$$

2. Számmal való szorzás: $A \in R^{m \times n}$, λ valós szám,

$$C = \lambda A \in R^{m \times n} \Leftrightarrow c_{ij} = \lambda a_{ij} \quad (i = 1, \dots, m, j = 1, \dots, n).$$

A számmal való szorzásra fennáll, hogy

$$\lambda(\mu A) = (\lambda\mu)A, \quad (\lambda + \mu)A = \lambda A + \mu A.$$

Jegyezzük meg, hogy megállapodás szerint $\lambda A = A\lambda$.

3. Transzponálás (tükrözés): $A \in R^{m \times n}$,

$$C = A^T \in R^{n \times m} \Leftrightarrow c_{ij} = a_{ji} \quad (i = 1, \dots, n, j = 1, \dots, m).$$

A transzponálásra fennáll, hogy

$$(A^T)^T = A, \quad (A + B)^T = A^T + B^T.$$

Az A mátrixot szimmetrikusnak nevezzük, ha $A^T = A$.

4. Szorzás: $A \in R^{m \times k}$, $B \in R^{k \times n}$,

$$C = AB \in R^{m \times n} \Leftrightarrow c_{ij} = \sum_{t=1}^k a_{it}b_{tj} \quad (i = 1, \dots, m, j = 1, \dots, n).$$

Vegyük észre, hogy a szorzatmátrix (i, j) indexű elemét úgy kapjuk, hogy az i -edik sort szorozzuk a j -edik oszloppal, azaz

$$c_{ij} = [a_{i1}, \dots, a_{ik}] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{kj} \end{bmatrix}.$$

A mátrixszorzás fontos tulajdonságai a következők:

$$\begin{aligned} (AB)C &= A(BC), \\ A(B + C) &= AB + AC, \\ (A + B)C &= AC + BC, \\ (AB)^T &= B^T A^T. \end{aligned}$$

Fontos megjegyezni, hogy a szorzás nem kommutatív, tehát

$$AB \neq BA \tag{2.1}$$

általában! A továbbiakban a mátrix és mátrix-vektor műveletek felírásánál feltesszük, hogy a szereplő mátrixok, ill. vektorok méretei olyanok, hogy lehetővé teszik az adott műveletet.

Definíció. Az egyetlen sorból, vagy egyetlen oszlopból álló mátrixot vektornak nevezzük.

A sorvektor szokásos megadási módjai a következők:

$$x = [x_1, \dots, x_n], \quad x = [x_1 \ x_2 \ \dots \ x_n].$$

Az oszlopvektorokat az

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in R^n$$

formában szoktuk megadni, ahol R^n az n komponensű oszlopvektorok halmaza (tkp. $R^n \equiv R^{n \times 1}$). Az R^n halmazt n -dimenziós euklideszi térnek is nevezzük. Az oszlopvektorokat meg lehet még adni az $x = [x_1, \dots, x_n]^T$, a sorvektorokat pedig az

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^T \in R^n$$

formában is.

Az i -edik egységvektor i -edik komponense 1, a többi 0. Oszlopvektornak tekintve tehát:

$$e_i = [0, \dots, 0, 1, 0, \dots, 0]^T \in R^n$$

Definíció. $x, y \in R^n$ skaláris szorzata

$$x^T y = \sum_{i=1}^n x_i y_i.$$

A következőkben speciális tulajdonságú mátrixokat vezetünk be.

Definíció. Az $I \in R^{n \times n}$ mátrix egységmátrix, ha

$$I = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}.$$

Az egységmátrixra fennáll, hogy minden $A \in R^{n \times n}$ esetén

$$AI = IA = A.$$

Definíció. A $D \in R^{n \times n}$ diagonálmátrix, ha

$$D = \begin{bmatrix} \delta_1 & 0 & \dots & \dots & 0 \\ 0 & \delta_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \delta_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \delta_n \end{bmatrix}.$$

A diagonálmátrixra fennáll, hogy minden $A \in R^{n \times m}$ és $B \in R^{m \times n}$ esetén

$$DA = D \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} = \begin{bmatrix} \delta_1 a_1^T \\ \vdots \\ \delta_n a_n^T \end{bmatrix}, \quad BD = [b_1, \dots, b_n] D = [b_1 \delta_1, \dots, b_n \delta_n].$$

A D diagonálmátrixot $diag(\delta_1, \dots, \delta_n)$, vagy $diag(\delta_i)$ ($i = 1, \dots, n$) is je-
lölheti.

Definíció. Az $0 \in R^{m \times n}$ mátrix zérusmátrix, ha 0 minden eleme, azaz

$$0 = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}.$$

A zérusmátrixra fennáll, hogy minden A mátrix esetén

$$A + 0 = A, \quad A0 = 0.$$

2.2. Mátrixok inverze és determinánsa

Definíció. Az $X \in R^{n \times n}$ mátrixot az $A \in R^{n \times n}$ mátrix inverzének nevezzük, ha $AX = XA = I$.

Ha az inverz mátrix létezik, akkor egyértelmű. Az inverz mátrix jelölése $A^{-1} = X$. Az inverz mátrixra fennállnak az alábbi tulajdonságok:

$$(A^{-1})^{-1} = A, \quad (AB)^{-1} = B^{-1}A^{-1}, \quad (A^T)^{-1} = (A^{-1})^T := A^{-T}.$$

Jelölje $A(i)$ azt az $(n-1) \times (n-1)$ -es mátrixot, amelyet az

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

mátrixból az első oszlop és az i -edik sor elhagyásával kapunk.

Definíció. Az $A \in R^{n \times n}$ négyzetes mátrix determinánsát a

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}, \quad n = 2$$

$$\det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A(i)), \quad n \geq 3.$$

előírások definiálják.

Tétel. Az $A \in R^{n \times n}$ mátrixnak akkor és csak akkor létezik inverze, ha $\det(A) \neq 0$.

2.3. Vektorok és mátrixok normája

Definíció. Az $f : R^n \rightarrow R$ függvényt vektornormának nevezzük, ha

$$f(x) \geq 0 \quad (\forall x \in R^n), \quad f(x) = 0 \Leftrightarrow x = 0, \quad (2.2)$$

$$f(\lambda x) = |\lambda| f(x) \quad (\forall x \in R^n, \forall \lambda \in R), \quad (2.3)$$

$$f(x + y) \leq f(x) + f(y) \quad (\forall x, y \in R^n). \quad (2.4)$$

A vektornorma szokásos jelölése: $\|x\| = f(x)$. A fontosabb vektornormák a következők:

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad (2.5)$$

$$\|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad (\text{euklideszi norma}), \quad (2.6)$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad (\text{maximum norma}). \quad (2.7)$$

Definíció. Az $f : R^{m \times n} \rightarrow R$ függvényt mátrixnormának nevezzük, ha

$$f(A) \geq 0 \quad (\forall A \in R^{m \times n}), \quad f(A) = 0 \Leftrightarrow A = 0, \quad (2.8)$$

$$f(\lambda A) = |\lambda| f(A) \quad (\forall A \in R^{m \times n}, \forall \lambda \in R), \quad (2.9)$$

$$f(A + B) \leq f(A) + f(B) \quad (\forall A, B \in R^{m \times n}). \quad (2.10)$$

A mátrixnorma szokásos jelölése: $\|A\| = f(A)$. A leggyakrabban használt mátrixnormák a következők:

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \quad (\text{oszlopösszeg norma}), \quad (2.11)$$

$$\|A\|_2 = \{A^T A \text{ legnagyobb sajátértéke}\}^{\frac{1}{2}} \quad (\text{spektrálnorma}), \quad (2.12)$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| \quad (\text{sorösszeg norma}), \quad (2.13)$$

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{\frac{1}{2}} \quad (\text{Frobenius norma}). \quad (2.14)$$

3. fejezet

Lineáris egyenletrendszerek megoldása

3.1. Lineáris egyenletrendszerek

Lineáris egyenletrendszerek általános alakja m egyenlet és n ismeretlen esetén:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1j}x_j + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n &= b_i \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mj}x_j + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (3.1)$$

Az egyenletrendszert megadhatjuk a tömörebb

$$Ax = b \quad (3.2)$$

formában, ahol

$$A = [a_{ij}]_{i,j=1}^{m,n} \in R^{m \times n}, \quad x \in R^n, \quad b \in R^m.$$

A megoldhatóság szempontjából három eset lehetséges:

- (i) az egyenletrendszernek nincs megoldása,
- (ii) az egyenletrendszernek pontosan egy megoldása van,
- (iii) az egyenletrendszernek végtelen sok megoldása van.

Definíció. Az $\{a_i\}_{i=1}^k \subseteq R^m$ vektorok lineárisan összefüggők, ha létezik $x \in R^k$ ($x \neq 0$), hogy

$$\sum_{i=1}^k x_i a_i = 0. \quad (3.3)$$

Ha nincs ilyen $x \neq 0$ vektor, akkor az $\{a_i\}_{i=1}^k$ vektorokat lineárisan függetlennek nevezzük.

A megoldhatóság egyik "jellemzését" megadhatjuk a rang fogalmával:

$$r(A) = \text{lineárisan független oszlop- vagy sorvektorok maximális száma} \quad (3.4)$$

Legyen $[A, b]$ az a mátrix, melyet A -ból úgy kapunk, hogy hozzáadjuk $(n + 1)$ -edik oszlopként a jobboldali b vektort. A mátrix rangjával megfogalmazva az $Ax = b$ egyenletrendszer akkor és csak akkor oldható meg, ha $r(A) = r([A, b])$. Ha $r(A) = r([A, b]) = n$, akkor az $Ax = b$ egyenletrendszernek pontosan egy megoldása van.

A továbbiakban csak négyzetes egyenletrendszerekkel foglalkozunk. Felteesszük tehát, hogy $m = n$. Ismert a következő

Tétel. Az $Ax = b$ ($A \in R^{n \times n}$, $b \in R^n$) egyenletrendszernek akkor és csak akkor van pontosan egy megoldása, ha létezik A^{-1} . Ekkor a megoldás $x = A^{-1}b$.

Tétel. Az $Ax = 0$ ($A \in R^{n \times n}$) homogén lineáris egyenletrendszernek akkor és csak akkor van $x \neq 0$ nemtriviális megoldása, ha $\det(A) = 0$.

3.2. Háromszögmátrixú egyenletrendszerek

Definíció. Az $A = [a_{ij}]_{i,j=1}^n$ mátrix alsó háromszög alakú, ha $a_{ij} = 0$ minden $i < j$ esetén.

Az alsó háromszögmátrixok alakja sematikusán a következő

$$\begin{bmatrix} * & 0 & \dots & \dots & 0 \\ * & * & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & * & 0 \\ * & \dots & \dots & * & * \end{bmatrix}.$$

Definíció. Az $A = [a_{ij}]_{i,j=1}^n$ mátrix felső háromszög alakú, ha $a_{ij} = 0$ minden $i > j$ esetén.

A felső háromszögmátrixok alakja:

$$\begin{bmatrix} * & * & \dots & \dots & * \\ 0 & * & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * & * \\ 0 & \dots & \dots & 0 & * \end{bmatrix}.$$

Könnyen igazolható, hogy alsó, vagy felső háromszögmátrixok esetén $\det(A) = a_{11}a_{22} \dots a_{nn}$. A háromszögmátrixú egyenletrendszerek megoldása igen egyszerű.

Tekintsük az

$$\begin{array}{cccccc} a_{11}x_1 + & \dots & + a_{1i}x_i + & \dots & + a_{1n}x_n & = & b_1 \\ & \ddots & \vdots & & \vdots & & \vdots \\ & & a_{ii}x_i + & \dots & + a_{in}x_n & = & b_i \\ & & & \ddots & \vdots & & \vdots \\ & & & & a_{nn}x_n & = & b_n \end{array}$$

felső háromszögmátrixú egyenletrendszert! Az egyenletrendszer akkor és csak akkor oldható meg, ha $a_{11} \neq 0, \dots, a_{nn} \neq 0$. A felső háromszögmátrixú egyenletrendszer megoldását a következő ún. visszahelyettesítő algoritmus adja:

$$\begin{aligned} x_n &= b_n/a_{nn} \\ x_i &= (b_i - \sum_{j=i+1}^n a_{ij}x_j)/a_{ii} \quad (i = n-1, n-2, \dots, 1) \end{aligned} \quad (3.5)$$

Alsóháromszögmátrixú egyenletrendszer megoldása hasonló.

3.3. A Gauss-módszer

A Gauss-féle eliminációs módszer két fázisból áll:

I. Azonos átalakításokkal az $Ax = b$ egyenletrendszert felső háromszög alakúra hozzuk:



II. A kapott felső háromszögmátrixú egyenletrendszert a (3.5) algoritmussal megoldjuk.

A felső háromszög alakra hozás a következőképpen történik.

Ha $a_{11} \neq 0$, akkor az a_{11} alatti x_1 együtthatókat nullává tesszük (kinullázzuk) úgy, hogy az i -edik sorból kivonjuk ($i = 2, \dots, n$) az első sor γ -szorosát:

$$(a_{i1} - \gamma a_{11})x_1 + (a_{i2} - \gamma a_{12})x_2 + \dots + (a_{in} - \gamma a_{1n})x_n = b_i - \gamma b_1. \quad (3.6)$$

Az $a_{i1} - \gamma a_{11} = 0$ feltételből kapjuk, hogy $\gamma = \frac{a_{i1}}{a_{11}}$. Így az első oszlop a_{11} alatti kinullázását az

$$\left. \begin{aligned} \gamma &= a_{i1}/a_{11} \\ a_{ij} &= a_{ij} - \gamma a_{1j} \quad (j = 2, \dots, n) \\ b_i &= b_i - \gamma b_1 \end{aligned} \right\} \quad (i = 2, \dots, n) \quad (3.7)$$

algoritmussal végezhetjük el. Vegyük észre, hogy az algoritmus felülírja az A mátrix $2 \leq i, j \leq n$ indexű és a b vektor $2 \leq i \leq n$ indexű elemeit (a 0-kat viszont feleslegesen nem írja be az alsó háromszög részbe). A kapott ekvivalens egyenletrendszer alakja:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \quad (3.8)$$

Ezt szétbonthatjuk az n ismeretlent tartalmazó első egyenletre és az $n-1$ ismeretlent tartalmazó kisebb $(n-1) \times (n-1)$ -es egyenletrendszerre. Ha $a_{22} \neq 0$, akkor a kisebb egyenletrendszeren megismételjük az előző lépést és így

tovább. Tegyük fel, hogy a $(k-1)$ -edik oszlopban a kinullázást már elvégeztük és az

$$\begin{array}{ccccccc} a_{11}x_1 + & \dots & \dots & + a_{1k}x_k & + & \dots & + a_{1n}x_n = b_1 \\ & & \ddots & \vdots & & & \vdots \\ & & & \vdots & & & \vdots \\ & & & a_{kk}x_k & + & \dots & + a_{kn}x_n = b_k \\ & & & \vdots & & & \vdots \\ & & & a_{ik}x_k & + & \dots & + a_{in}x_n = b_i \\ & & & \vdots & & & \vdots \\ & & & a_{nk}x_k & + & \dots & + a_{nn}x_n = b_n \end{array}$$

egyenletrendszert kaptuk. Ha $a_{kk} \neq 0$, akkor kinullázzuk az a_{kk} alatti x_k együtthatókat. Az i -edik sorból a k -edik sort m -szeresét kivonva az

$$(a_{ik} - \gamma a_{kk})x_k + (a_{i,k+1} - \gamma a_{k,k+1})x_{k+1} + \dots + (a_{in} - \gamma a_{kn})x_n = b_i - \gamma b_k \quad (3.9)$$

egyenlet adódik. Az $a_{ik} - \gamma a_{kk} = 0$ feltételből kapjuk, hogy $\gamma = \frac{a_{ik}}{a_{kk}}$. A k -edik oszlop a_{kk} alattikinullázását tehát a

$$\left. \begin{array}{l} \gamma = a_{ik}/a_{kk} \\ a_{ij} = a_{ij} - \gamma a_{kj} \quad (j = k+1, \dots, n) \\ b_i = b_i - \gamma b_k \end{array} \right\} \quad (i = k+1, \dots, n)$$

algoritmussal végezhetjük el.

A kinullázást mindaddig folytathatjuk, amíg az $a_{kk} \neq 0$ és $k \leq n-1$ feltételek fennállnak. Ha sikerül az A mátrixot felső háromszög alakra hozni, akkor a Gauss-módszer II. fázisát, a (3.5) algoritmust alkalmazzuk (visszahelyettesítünk).

A Gauss-módszer I. fázisában előfordulhat, mondjuk a k -edik lépésben, hogy az a_{kk} elem zérus. Például a

$$\begin{array}{rcl} & 4x_2 & + x_3 = 9 \\ x_1 & + x_2 & + 3x_3 = 6 \\ 2x_1 & - 2x_2 & + x_3 = -1 \end{array}$$

rendszernek ez $a_{11} = 0$. Ilyen esetekben az eliminációt csak akkor lehet folytatni, ha a sorok felcserélésével el tudjuk érni, hogy az a_{kk} helyére zérustól különböző elem kerüljön. Ha ilyen elem nincs, azaz az a_{kk} alatti elemek is mind zérusok, akkor az $[a_{ij}]_{i,j=1}^{n,k}$ részmatrix oszlopai lineárisan összefüggők, tehát A szinguláris. A fenti példában az első és harmadik sor felcserélésével kapjuk, hogy

$$\begin{array}{rcl} 2x_1 & - 2x_2 & + x_3 = -1 \\ x_1 & + x_2 & + 3x_3 = 6 \\ & 4x_2 & + x_3 = 9 \end{array}$$

Az a_{kk} elemet k -edik pivot, vagy főelemnek nevezzük. A sorok felcserélését úgy, hogy az új pivot elem zérustól különböző legyen, pivotálási, vagy főelemkiválasztási eljárásnak nevezzük. A pivot elem megválasztása nagymértékben befolyásolja az eredmények megbízhatóságát.

Részleges főelemkiválasztás: A k -adik lépésben a k -adik oszlop a_{jk} ($k \leq j \leq n$) elemei közül kiválasztjuk a maximális abszolút értékűt. Ha ennek indexe i , akkor a k -adik és az i -edik sort felcseréljük. A pivotálás után teljesül, hogy

$$|a_{kk}| = \max_{k \leq i \leq n} |a_{ik}|.$$

Teljes főelemkiválasztás: A k -adik lépésben az a_{ij} ($k \leq i, j \leq n$) mátrix-elemek közül kiválasztjuk a maximális abszolút értékűt. Ha ennek indexe (i, j) , akkor a k -adik és az i -edik sort, valamint a k -adik oszlopot és j -edik oszlopot felcseréljük. A pivotálás után teljesül, hogy

$$|a_{kk}| = \max_{k \leq i, j \leq n} |a_{ij}|.$$

Megjegyezzük, hogy oszlopcseré esetén változócsere is történik.

A főelemkiválasztásos Gauss-módszer esetén az I. fázis minden lépésében pivotálást hajtunk végre. A teljes főelemkiválasztást biztonságos stratégiának tekintjük. A részleges főelemkiválasztás egyéb technikákkal kiegészítve ugyancsak biztonságosnak tekinthető. A részleges főelemkiválasztás lényegesen kevesebb extra aritmetikai műveletet igényel mint a teljes főelemkiválasztás. Ezért a gyakorlatban inkább ezt használjuk.

3.4. Az LU-felbontás és alkalmazása lineáris egyenletrendszerekre

Definíció. Az $A \in R^{n \times n}$ mátrix LU-felbontásán a mátrix $A = LU$ szorzatalakban történő felbontását értjük, ahol $L \in R^{n \times n}$ alsó, $U \in R^{n \times n}$ pedig felső háromszögmátrix.

Ha L egység alsó háromszögmátrix (fődiagonálisában minden elem 1), akkor az LU felbontás egyértelmű.

Legyen

$$A^{(r)} = \begin{bmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{bmatrix} \quad (r = 1, \dots, n-1).$$

Az $A^{(r)}$ mátrix az A mátrix r -edik főminor mátrixa.

Tétel. Egy $A \in R^{n \times n}$ nemszinguláris mátrixnak akkor és csak akkor létezik LU-felbontása, ha

$$\det(A^{(i)}) \neq 0 \quad (i = 1, \dots, n-1). \quad (3.10)$$

Legyen $A^{(0)} = A$, $b^{(0)} = b$. A pivotálás nélküli Gauss-elimináció első fázisában (amennyiben az végrehajtható) egy

$$A^{(0)} = b^{(0)} \rightarrow A^{(1)} = b^{(1)} \rightarrow \dots \rightarrow A^{(n-1)} = b^{(n-1)}$$

ekvivalens egyenletsortozatot képezünk, ahol

$$A^{(k)} = \left[a_{ij}^{(k)} \right]_{i,j=1}^n.$$

Fennáll, hogy

$$A^{(n-1)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix}$$

Ha az A mátrix nonszinguláris és létezik LU -felbontása, úgy $a_{kk}^{(k-1)} \neq 0$, ($k = 1, 2, \dots, n$). Amennyiben a zérusokat nem írjuk be, hanem a otthagyjuk a már ottlévő értéket, akkor utoljára az

$$\tilde{A}^{(n-1)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1,n-1}^{(0)} & a_{1n}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(1)} & \cdots & a_{2,n-1}^{(1)} & a_{2n}^{(1)} \\ \vdots & a_{32}^{(1)} & & & \vdots \\ \vdots & \vdots & \ddots & & \vdots \\ & & & a_{n-1,n-1}^{(n-2)} & \\ a_{n1}^{(0)} & a_{n2}^{(1)} & \cdots & a_{n,n-1}^{(n-2)} & a_{nn}^{(n-1)} \end{bmatrix},$$

táblázatot kapjuk. Kimutatható, hogy az A mátrix LU -felbontásában

$$U = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix},$$

valamint

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21}^{(0)}/a_{11}^{(0)} & \ddots & & \\ a_{31}^{(0)}/a_{11}^{(0)} & 1 & & \vdots \\ \vdots & a_{k+1,k}^{(1)}/a_{kk}^{(1)} & \ddots & \\ \vdots & \vdots & & 1 & 0 \\ a_{n1}^{(0)}/a_{11}^{(0)} & \cdots & a_{nk}^{(1)}/a_{kk}^{(1)} & \cdots & a_{n,n-1}^{(n-2)}/a_{n-1,n-1}^{(n-2)} & 1 \end{bmatrix}$$

Összegezve: a Gauss-módszer az I. fázisban előállítja az A mátrix LU -felbontását, illetve az ekvivalens

$$Ux = L^{-1}b \quad (3.11)$$

egyenletrendszert. A felbontás U tényezője közvetlenül megjelenik az utolsó táblázat felső háromszög részében, az L tényezőt pedig az alsó háromszög részből úgy kapjuk, hogy a főátlóbeli elemekkel végigosztjuk az oszlopokat. Tehát a Gauss-módszer az $A = LU$ speciális szorzatfelbontáson (faktorizáción) alapul.

Megjegyzés: Ha az eljárás során főelemkiválasztást kell végrehajtani, akkor a Gauss-módszer a PA mátrix LU -felbontását adja meg, ahol P permutáció mátrix (azaz PA az A -ból a sorok permutációjával áll elő).

AZ LU-MÓDSZER ALGORITMUSA:

1. Határozzuk meg az $A = LU$ felbontást!
2. Oldjuk meg az $Ly = b$ egyenletrendszert!
3. Oldjuk meg az $Ux = y$ egyenletrendszert!

Az eljárás az eredeti Gauss-módszerhez képest egy extra alsó háromszög-mátrixú egyenletrendszer megoldását is megköveteli. Az eljárás akkor előnyös, ha egyénél több

$$Ax = b_1, Ax = b_2, \dots$$

alakú egyenletrendszert kell megoldani közös A együtthatómátrixszal. Ekkor elég az A mátrix LU -felbontását egyszer meghatározni.

3.5. Lineáris egyenletrendszerek közelítő megoldása iterációval

Alakítsuk át az $Ax = b$ egyenletrendszert a vele ekvivalens $x = Bx + c$ alakúra. Ez sokféleképpen lehetséges, pl. úgy, hogy a i -edik egyenletből kifejezzük az x_i ismeretlent ($i = 1, 2, \dots, n$):

$$\begin{aligned} x_1 &= (b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n)/a_{11} \\ x_2 &= (b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n)/a_{22} \\ &\vdots \\ x_n &= (b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n,n-1}x_{n-1})/a_{nn} \end{aligned} \quad , \quad (3.12)$$

azaz

$$\begin{aligned} c_1 &= b_1/a_{11} & 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \dots & -a_{1n}/a_{11} \\ c_2 &= b_2/a_{22} & -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & \dots & -a_{2n}/a_{22} \\ &\vdots & \vdots & \vdots & \vdots & & \vdots \\ c_n &= b_n/a_{nn} & -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \dots & & 0 \end{aligned} \quad , \quad B = \begin{pmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \dots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & \dots & -a_{2n}/a_{22} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \dots & & 0 \end{pmatrix} \quad (3.13)$$

Legyen $x^{(0)} \in R^n$ valamilyen kezdeti vektor. Tekintsük az

$$\begin{aligned} x^{(1)} &= Bx^{(0)} + c \\ x^{(2)} &= Bx^{(1)} + c \\ &\vdots \\ x^{(k)} &= Bx^{(k-1)} + c \\ &\vdots \end{aligned}$$

sorozatot.

Tétel. Ha $\|B\|_\infty < 1$ akkor az $x^{(0)}, x^{(1)}, \dots, x^{(k)} \dots$ sorozat bármely $x^{(0)} \in R^n$ esetén az egyenletrendszer x megoldásához konvergál; a k -adik iterált hibájára érvényes az alábbi becslés:

$$\|x - x^{(k)}\|_\infty \leq \frac{\|B\|_\infty}{1 - \|B\|_\infty} \|x^{(k)} - x^{(k-1)}\|_\infty .$$

Megjegyzés: Az említett átalakítással a $\|B\|_\infty < 1$ akkor és csak akkor teljesül, ha

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| , \quad (i = 1, 2, \dots, n)$$

Ekkor azt mondjuk, hogy a mátrix diagonálisan domináns.

4. fejezet

A sajátérték-feladat

A mátrixok sajátérték-feladatának megfogalmazásához szükségünk van a komplex elemű mátrixok és vektorok bevezetésére. A komplex elemű n -dimenziós oszlopvektorokat C^n -el jelöljük. Hasonlóképpen az $m \times n$ típusú komplex elemű mátrixok halmazát $C^{m \times n}$ jelöli. Nyilvánvalóan fennáll, hogy $R^n \subset C^n$ és $R^{m \times n} \subset C^{m \times n}$.

Definíció. Legyen A $n \times n$ valós, vagy komplex mátrix. A $\lambda \in C$ számot az A mátrix sajátértékének, az $x \in C^n$ ($x \neq 0$) vektort pedig a λ sajátértékhez tartozó sajátvektornak nevezzük, ha

$$Ax = \lambda x. \quad (4.1)$$

A sajátvektor egy olyan vektor, amelyet az $x \rightarrow Ax$ leképezés a saját hatásvonalán hagy (irányítás, nagyság változhat). A sajátérték-feladat megoldása a sajátértékek és a hozzájuk tartozó sajátvektorok meghatározását jelenti.

Egy x sajátvektor t -szerese is sajátvektor $t \neq 0$ esetén, ui. $A(tx) = tAx = t\lambda x = \lambda(tx)$. Az $Ax = \lambda x$ egyenletrendszer átrendezéssel az ekvivalens $(A - \lambda I)x = 0$ alakra hozható. Ennek a homogén egyenletrendszernek akkor és csak akkor van zérustól különböző megoldása, ha $\det(A - \lambda I) = 0$. A

$$\phi(\lambda) = \det(A - \lambda I) = \det \left(\begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix} \right) = 0 \quad (4.2)$$

egyenletet az A mátrix karakterisztikus egyenletének nevezzük. A determinánst kifejtve a λ változó n -ed fokú polinomját kapjuk, az algebra alaptétele miatt tehát az A $n \times n$ mátrixnak a multiplicitásukat is beleszámítva pontosan n sajátértéke van. Egy λ_k sajátértékhez tartozó lineárisan független sajátvektorok száma legfeljebb annyi, mint λ_k multiplicitása a $\phi(\lambda) = 0$ karakterisztikus egyenletben. Különböző sajátértékekhez tartozó sajátvektorok lineárisan függetlenek.

Előfordulhat, hogy egy valós együtthatós mátrixnak komplex sajátértékei és sajátvektorai vannak. Valós aritmetika esetén ezeket csak speciális technikákkal kaphatjuk meg. Az összes sajátérték és sajátvektor meghatározása elvileg sem könnyű feladat.

4.1. A hatvány módszer

Nevezzük a legnagyobb abszolút értékű sajátértéket domináns sajátértéknek. Indexeljük a sajátértékeket úgy, hogy $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ teljesüljön. A von Mises-től származó módszer alap gondolata a következő. Tegyük fel, hogy az A $n \times n$ típusú valós mátrixnak csak valós sajátértéke van. Legyen $v^{(0)} \in R^n$ adott! Képezzük a $v^{(k)} = Av^{(k-1)} = A^k v^{(0)}$ ($k = 1, 2, \dots$) sorozatot! (Innen ered a hatványmódszer elnevezés.) Megmutatható, hogy bizonyos feltételek esetén a $v^{(k)}$ és a $v^{(k-1)}$ vektorok (azaz két szomszédos közelítés) megfelelő komponenseinek hányadosa a λ_1 (azaz a domináns) sajátértékhez tart, a $v^{(k)}$ sorozat pedig a λ_1 -hez tartozó sajátvektorhoz. Ha $|\lambda_1| > 1$, akkor a $v^{(k)}$ komponensei gyorsan nőnek abszolút értékben, $|\lambda_1| < 1$ esetben pedig gyorsan zérushoz tartanak. Ezért időnként (legjobb minden lépésben) célszerű normálni.

A hatványmódszer algoritmus ($v^{(0)} \in R^n, y \in R^n$):

Hajtsuk végre az alábbiakat a $k = 1, 2, \dots$ értékekre:

$$z^{(k)} = Av^{(k-1)}$$

$$v^{(k)} = z^{(k)} / \|z^{(k)}\|_\infty$$

Megmutatható, hogy alkalmas feltételek mellett $\|z^{(k)}\|_\infty \rightarrow |\lambda_1|$ és $v^{(k)} \rightarrow$ a λ_1 -hez tartozó sajátvektorhoz.

Az eljárás a $|\lambda_2/\lambda_1|$ nagyságrendtől függő konvergencia sebességgel rendelkezik. Erősen érzékeny a $v^{(0)}$ kezdővektor megválasztására is. Komplex sajátértékek, illetve többszörös λ_1 esetén az eljárást módosítani kell.

A hatványmódszert, amely igen előnyös lehet nagyméretű ritka mátrixok esetén, leginkább a legnagyobb, ill. a legkisebb abszolút értékű sajátértékek meghatározására használjuk. Ez utóbbi a következőképpen történhet. Az A^{-1} sajátértékei: $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$. Ezek közül a legnagyobb abszolút értékű sajátérték $\frac{1}{\lambda_n}$ lesz, ezért alkalmazva a hatványmódszert A^{-1} -re megkapjuk a λ_n reciprokát.

Végül megjegyezzük, hogy rangszámcsökkentő eljárásokkal és egyéb módosításokkal, a Mises eljárás alkalmassá tehető az összes sajátérték-sajátvektor meghatározására is.

5. fejezet

Nemlineáris egyenletek közelítő megoldási módszerei

5.1. Egyváltozós egyenletek megoldása

Az $f(x) = 0$ ($f : R \rightarrow R$) alakú valós egyenletek megoldását vizsgáljuk. Három módszert ismertetünk.

5.1.1. Az intervallumfelező eljárás

Tegyük fel, hogy $f : R \rightarrow R$ folytonos az $[a, b]$ intervallumon és fennáll, hogy

$$f(a)f(b) < 0. \quad (5.1)$$

Ekkor a Bolzano-tétel miatt az $f(x) = 0$ egyenletnek van legalább egy $x^* \in (a, b)$ gyöke. Ezt a Bolzano-tétel bizonyításából ismert eljárással kaphatjuk meg. Legyen $c = (a + b)/2$ és vizsgáljuk az $f(c)$ értékét. Ha $f(a)f(c) < 0$, akkor az $[a, c]$ intervallumban van gyök. Egyébként a $[c, b]$ intervallum tartalmaz gyököt. Az új intervallumot újra megfelezzük és így tovább. Az egymásba skatulyázott zárt intervallumok ráhúzódnak az egyenlet egy gyökére. Algoritmikus formában az eljárás a következő:

Az intervallum felező algoritmus:

Input $[a_1, b_1] = [a, b]$ Hajtsuk végre (I), (II), (III)-at $i = 1, 2, \dots$ -re a Stop feltételig:

(I.) $x_i = \frac{a_i + b_i}{2}$

(II.) ha $b_i - a_i < 2\varepsilon$, akkor Stop

(III.) ha $f(a_i)f(x_i) < 0$, akkor $a_{i+1} = a_i$, $b_{i+1} = x_i$,
egyébként $a_{i+1} = x_i$, $b_{i+1} = b_i$

Az előírt pontosság ε , ugyanis nyilvánvaló az alábbi becslés:

$$|x^* - x_i| \leq \frac{b_i - a_i}{2} = \frac{b - a}{2^i} \quad (i = 1, 2, \dots). \quad (5.2)$$

Itt $x_i \approx x^*$ és az algoritmust a Stop feltétele akkor állítja le, ha az x_i közelítés hibája kisebb mint egy előre megadott $\varepsilon > 0$ hibakorlát. Megjegyezzük, hogy a módszer konvergenciája csak folytonos $f(x)$ esetén áll fenn.

5.1.2. Egyszerű iteráció

A módszert az $f(x) = x - g(x) = 0$ alakú, vagy ilyen alakra hozható egyenletek esetén alkalmazzuk. Az $f(x) = 0$ egyenlet ekvivalens az

$$x = g(x) \quad (5.3)$$

egyenlettel. Legyen $x_0 \in R$ egy kezdeti közelítés és tekintsük az alábbi sorozatot:

$$\begin{aligned} x_1 &= g(x_0) \\ x_2 &= g(x_1) \\ &\vdots \\ x_i &= g(x_{i-1}) \\ &\vdots \end{aligned}$$

Tétel. Ha $g(x)$ folytonosan differenciálható $[a, b]$ -n, mégpedig $g'(x) \leq q < 1$ korláttal, továbbá $a \leq g(x) \leq b$ minden $x \in [a, b]$ esetén, akkor az $x = g(x)$ egyenletnek pontosan egy x^* gyöke van $[a, b]$ -n és az x_i sorozat határértéke éppen ez az x^* . Igaz továbbá az alábbi hibabecslés:

$$|x_i - x^*| \leq \frac{q}{1 - q} |x_i - x_{i-1}| \quad (i = 1, 2, \dots). \quad (5.4)$$

A tételben szereplő hibabecslést használjuk fel a kilépési feltételhez.

5.1.3. A Newton-módszer

Tegyük fel, hogy $f : R \rightarrow R$ folytonosan differenciálható. A módszer lényege, hogy az x_{i-1} pontban a függvényhez érintőt húzunk és ennek az érintőnek a zérushelye adja meg a keresett gyök i -edik közelítését, azaz x_i -t. Az érintő irántangense $f'(x_{i-1})$ és egyenlete

$$y - f(x_i) = f'(x_i)(x - x_i). \quad (5.5)$$

Az $y = 0$ egyenlet megoldása:

$$x_i = x_{i-1} - \frac{f(x_{i-1})}{f'(x_{i-1})}, \quad (5.6)$$

feltéve, hogy $f'(x_{i-1}) \neq 0$.

A Newton módszer tehát a következő: Adott egy $x_0 \in R$ kezdeti közelítés és képezzük az

$$x_i = x_{i-1} - \frac{f(x_{i-1})}{f'(x_{i-1})} \quad (i = 1, 2, \dots) \quad (5.7)$$

sorozatot. Vegyük észre, hogy a Newton-módszer tkp. az $x = x - f(x)/f'(x) = g(x)$ egyenletre alkalmazott iterációs eljárás. A Newton-módszer konvergenciájára vonatkozik az alábbi

Tétel. Legyen $f : [a, b] \rightarrow R$ kétszer folytonosan differenciálható, $|f''(x)| \leq M$ és $|f'(x)| \geq m > 0$ ($x \in [a, b]$). Ha az $f(a)f(b) < 0$, valamint $f'(x), f''(x) \neq 0$ az $[a, b]$ intervallumban, akkor pontosan egy x^* van az intervallumban, továbbá $x_0 = a$, vagy $x_0 = b$ választással $x_i \rightarrow x^*$ ($i \rightarrow +\infty$) és

$$|x_i - x^*| \leq \frac{M}{2m} |x_i - x_{i-1}|^2 \quad (i = 1, 2, \dots). \quad (5.8)$$

x_0 -ra az $f(x_0)f''(x_0) > 0$ feltételnek kell teljesülni.

Azt mondjuk, hogy a Newton-módszer konvergenciája lokális, mert az x_1 kezdeti közelítésnek az x^* gyök "közelében" kell lennie.

5.2. Nemlineáris egyenletrendszerek megoldása Newton-módszerrel

Az $f(x) = 0$ ($x \in R^n$) egyenletrendszer alakja koordinátás alakban

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, \dots, x_n) &= 0. \end{aligned}$$

Az $x_i = [x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}]^T \in R^n$ pontban linearizáljuk az $f_k(x)$ koordináta függvényt ($k = 1, \dots, n$):

$$\begin{aligned} f_1(x) &\approx f_1(x_i) + \nabla f_1(x_i)^T(x - x_i) \\ &\vdots \\ f_n(x) &\approx f_n(x_i) + \nabla f_n(x_i)^T(x - x_i). \end{aligned}$$

Az $f(x) = 0$ egyenletrendszer megoldása helyett keressük a linearizált egyenletrendszer

$$\begin{aligned} f_1(x_i) + \nabla f_1(x_i)^T(x - x_i) &= 0 \\ &\vdots \\ f_n(x_i) + \nabla f_n(x_i)^T(x - x_i) &= 0 \end{aligned} \quad (5.9)$$

közös megoldását, amely az x_{i+1} új közelítést definiálja. Vegyük észre, hogy az $y = f_k(x_i) + \nabla f_k(x_i)^T(x - x_i)$ egyenlet tkp. az $y = f_k(x)$ függvény érintősíkjá az x_i pontban. Ha bevezetjük az

$$J(x) = \left[\frac{\partial f_i(x)}{\partial x_j} \right]_{i,j=1}^n = \begin{bmatrix} \nabla f_1(x_i)^T \\ \vdots \\ \nabla f_n(x_i)^T \end{bmatrix} \quad (5.10)$$

jelölést ($J(x)$ az $f(x)$ függvény Jacobi-mátrixa), akkor a linearizált egyenletrendszer átmege a tömörebb

$$f(x_i) + J(x_i)(x - x_i) = 0 \quad (5.11)$$

alakba, amelynek megoldása adja a Newton-módszert:

$$x_{i+1} = x_i - [J(x_i)]^{-1} f(x_i) \quad (i = 1, \dots). \quad (5.12)$$

A gyakorlatban soha nem invertáljuk a $J(x_i)$ Jacobi-mátrixot. Helyette a lineáris egyenletrendszer alakot oldjuk meg.

Az itt használható kilépési feltételek:

(A) $\|f(x_i)\| \leq \varepsilon_1$.

(B) $\|x_{i+1} - x_i\| \leq \varepsilon_2$.

(C) $i = i_{\max}$.

feltételek és ezek kombinációi.

Az eljárás konvergenciája alkalmas feltételek esetén lokális és másodrendű.

6. fejezet

Függvények közelítése interpolációval

Az interpoláció feladatát a következőképpen fogalmazhatjuk meg. Ismerjük egy $y = f(x)$ ($f : R \rightarrow R$) függvény

$$a \leq x_1 < x_2 < \dots < x_n \leq b \quad (6.1)$$

pontokban felvett értékeit, az

$$y_i = f(x_i) \quad (i = 1, \dots, n) \quad (6.2)$$

függvényértékeket. Az $f(x)$ függvényt, amely lehet ismert, vagy akár ismeretlen is, egy olyan, általában könnyen számítható $h(x)$ függvénnyel közelítjük (tkp. helyettesítjük), amelyre fennáll, az $y_i = h(x_i)$ ($i = 1, \dots, n$) interpolációs feltétel.

Az $\{x_i\}_{i=1}^n$ pontokat interpolációs alappontoknak nevezzük. Az interpolációs feltétel teljesülése esetén azt reméljük, hogy a $h(x)$ interpoláló függvény az (x_i, x_{i+1}) intervallumokban jól közelíti az $f(x)$ függvényt. Ha $f(x)$ -et a $h(x)$ függvénnyel az (x_1, x_n) intervallumon kívül közelítjük, akkor extrapolációról beszélünk.

A $h(x)$ függvény megválasztásától függően beszélünk különböző típusú interpolációkról. Itt csak az egyik legfontosabbat tárgyaljuk.

6.1. A Lagrange-féle interpolációs feladat

A feladat szokásos megfogalmazása a következő. Adottak az $x_1 < x_2 < \dots < x_n$ alappontok és az $y_i = f(x_i)$ ($i = 1, \dots, n$) függvényértékek. Határozzuk meg azt a legfeljebb $(n - 1)$ -edfokú

$$p(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \quad (6.3)$$

polinomot, amelyre teljesül a

$$y_i = p(x_i) \quad (i = 1, \dots, n) \quad (6.4)$$

interpolációs feltétel.

A Lagrange-féle interpolációs polinom létezése és egyértelmősége könnyen belátható. A polinom többféle ekvivalens alakban is felírható. Különösen fontos azonban a Lagrange-féle előállítás. Legyen

$$l_i(x) = \prod_{k=1, k \neq i}^n \frac{x - x_k}{x_i - x_k} \quad (i = 1, \dots, n) \quad (6.5)$$

az i -edik Lagrange-féle alappolinom. Ekkor az interpolációs polinom előáll

$$p(x) = \sum_{i=1}^n y_i l_i(x) \quad (6.6)$$

alakban. Ennek igazolására vegyük észre, hogy

$$l_i(x_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (6.7)$$

és

$$p(x_j) = \sum_{i=1}^n y_i l_i(x_j) = y_j l_j(x_j) = y_j \quad (j = 1, \dots, n). \quad (6.8)$$

A Lagrange-féle interpolációs polinom hibájára vonatkozik a következő **Tétel** (Cauchy). Ha $f \in C^n[a, b]$, $[x_1, x_n] \subset [a, b]$ és $x \in [a, b]$, akkor

$$f(x) - p(x) = \frac{f^{(n)}(\xi_x)}{n!} (x - x_1)(x - x_2) \dots (x - x_n),$$

ahol $\xi_x = \xi(x)$ az x és az x_1, x_n pontok által kifeszített intervallumban van.

Következmény. Ha $|f^{(n)}(x)| \leq M_n$ ($x \in [a, b]$), akkor

$$|f(x) - p(x)| \leq \frac{M_n}{n!} (b - a)^n. \quad (6.9)$$

Az interpolációs eljárásoktól elvárjuk, hogy a pontok számának növelése esetén a közelítés hibája csökken. Ez azonban nem minden esetben van így.

A Lagrange-féle interpolációs polinom másképpen is származtatható. Tekintsük a következő függvényrendszert:

$$\phi_1(x) = 1, \quad \phi_2(x) = x, \quad \dots, \quad \phi_n(x) = x^{n-1} \quad (6.10)$$

A Lagrange-féle interpolációs függvény alakja

$$h(x) = a_1 \phi_1(x) + a_2 \phi_2(x) + \dots + a_n \phi_n(x) = \sum_{i=1}^n a_i \phi_i(x). \quad (6.11)$$

Az ismeretlen a_1, \dots, a_n együtthatókat az interpolációs feltételből határozhatjuk meg. Ekkor teljesülnie kell az alábbi n feltételnek

$$\begin{aligned} a_1 \phi_1(x_1) + a_2 \phi_2(x_1) + \dots + a_n \phi_n(x_1) &= f(x_1), \\ &\vdots \\ a_1 \phi_1(x_n) + a_2 \phi_2(x_n) + \dots + a_n \phi_n(x_n) &= f(x_n), \end{aligned} \quad (6.12)$$

amely lineáris egyenletrendszer az ismeretlen a_1, \dots, a_n együtthatókra nézve. Legyen

$$B = [\phi_j(x_i)]_{i,j=1}^n = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix} \quad (6.13)$$

és

$$a = [a_1, \dots, a_n]^T, \quad c = [f(x_1), \dots, f(x_n)]^T. \quad (6.14)$$

A fenti feltétel tömör alakban

$$Ba = c. \quad (6.15)$$

Ha $\det(B) \neq 0$, akkor az egyenletrendszernek pontosan egy megoldása van: $a = B^{-1}c$.

A gyakorlatban másféle $\{\phi_i(x)\}_{i=1}^n$ alapfüggvényeket is alkalmaznak, de nem minden $\{\phi_i(x)\}_{i=1}^n$ függvényrendszer és $x_1 < x_2 < \dots < x_n$ alappontrendszer esetén van megoldása a lineáris interpolációs feladatnak.

7. fejezet

Numerikus deriválás és integrálás

7.1. Numerikus deriválás

A numerikus deriválás alapproblémája az $f : R \rightarrow R$ függvény deriváltjának kiszámítása egy, vagy több adott pontban. A probléma kézenfekvő megoldása az, hogy az $f(x)$ függvényt egy $h(x)$ függvénnyel (pl. interpolációval) közelítjük és az $f'(x)$ közelítésének a közelítő függvény $h'(x)$ deriváltját tekintjük. Sematikusan: ha $f(x) \approx h(x)$, akkor $f'(x) \approx h'(x)$ és általában $f^{(j)}(x) \approx h^{(j)}(x)$.

A Lagrange-interpoláció alapján az $f(x)$ függvény x -pontbeli j -edik deriváltjának közelítését az

$$f^{(j)}(x) \approx p^{(j)}(x) = \sum_{i=1}^n y_i l_i^{(j)}(x) \quad (7.1)$$

összefüggés adja meg.

Legyen $n = 2k + 1$, az alappontok pedig legyenek

$$x - kh, \dots, x - h, x, x + h, \dots, x + kh. \quad (7.2)$$

Ha $p(x)$ az $f(x)$ függvényt ezen pontokban interpoláló Lagrange-polinom, akkor igaz, hogy

$$|f'(x) - p'(x)| < \frac{Kh^{2k}}{\binom{2k}{k} (2k+1)}, \quad (7.3)$$

ahol K az $|f^{(2k+1)}(x)|$ korlátja az $[x - kh, x + kh]$ intervallumon.

Az $n = 3$ esetben az

$$f'(x) \approx \frac{1}{2h} [f(x+h) - f(x-h)], \quad (7.4)$$

A differencia hányadosok alkalmazása a közelítő deriválás legelterjedtebb módszere. Legcélszerűbben a Taylor-sorfejtés felhasználásával vezethetünk le közelítő formulákat.

Tegyük fel, hogy $f \in C^2$ és írjuk fel $f(x)$ másodfokú Taylor-polinomját:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(\xi) \quad (x < \xi < x+h).$$

Egyszerű számolással adódik, hogy

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2}f''(\xi),$$

ahonnan a

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \quad (7.5)$$

közelítést kapjuk.

Ennél pontosabb közelítést kaphatunk, ha $f \in C^3$:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \quad (7.6)$$

A második derivált ismert közelítései az

$$f''(x) \approx \frac{1}{h^2} (f(x) - 2f(x+h) + f(x+2h)) \quad (7.7)$$

és az

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}, \quad (7.8)$$

képletek.

Közelítő parciális differencia hányadosokat hasonló módon vezethetünk le.

7.2. Numerikus integrálás

Numerikus integrálást általában akkor végzünk, ha:

- a primitív függvény nem ismert, vagy nem állítható elő könnyen,
- az $f(x)$ függvénynek csak véges sok értéke ismert.

A numerikus eljárások alapötlete a következő:

$$f(x) \approx h(x) \Rightarrow \int_a^b f(x)dx \approx \int_a^b h(x)dx. \quad (7.9)$$

7.2.1. A trapézformula

Legyen $x_1 = a$ és $x_2 = b$. A két pontra támaszkodó elsőfokú Lagrange-féle interpolációs polinom

$$p(x) = f(x_1) \frac{x - x_2}{x_1 - x_2} + f(x_2) \frac{x - x_1}{x_2 - x_1}, \quad (7.10)$$

amelynek határozott integrálja

$$\begin{aligned} \int_a^b p(x)dx &= \int_{x_1}^{x_2} p(x)dx = \left[f(x_1) \frac{(x-x_2)^2}{2(x_1-x_2)} + f(x_2) \frac{(x-x_1)^2}{2(x_2-x_1)} \right]_{x_1}^{x_2} \\ &= \frac{x_2-x_1}{2} [f(x_1) + f(x_2)]. \end{aligned} \quad (7.11)$$

Ez tkp. egy trapéz területe. A közelítés hibájára fennáll, hogy

$$\left| \int_a^b f(x)dx - \frac{b-a}{2} [f(a) + f(b)] \right| \leq \frac{M_2}{12} (b-a)^3. \quad (7.12)$$

Ha az $[a, b]$ intervallumot felbontjuk az

$$a = x_1 < x_2 < \dots < x_{n+1} = b \quad (7.13)$$

pontokkal n részintervallumra, akkor az összetett trapézformula a következő:

$$\int_a^b f(x)dx \approx T_n(f) = \sum_{i=1}^n \frac{x_{i+1} - x_i}{2} [f(x_i) + f(x_{i+1})]. \quad (7.14)$$

A képlet hibájára $f \in C^2[a, b]$ esetén fennáll, hogy

$$\left| \int_a^b f(x)dx - \sum_{i=1}^n \frac{x_{i+1} - x_i}{2} [f(x_i) + f(x_{i+1})] \right| \leq \frac{M_2}{12} \sum_{i=1}^n (x_{i+1} - x_i)^3. \quad (7.15)$$

Ha az alappontok ekvidisztansak, azaz $x_i = x_1 + (i-1)h$ ($h = \frac{b-a}{n}$, $i = 1, \dots, n+1$), akkor a képlet alakja egyszerűsödik:

$$\int_a^b f(x)dx \approx T_n(f) = \frac{h}{2} \left[f(x_1) + 2 \sum_{i=2}^n f(x_i) + f(x_{n+1}) \right]. \quad (7.16)$$

A képlet hibájára pedig $nh = b - a$ miatt fennáll, hogy

$$\left| \int_a^b f(x)dx - T_n(f) \right| \leq \frac{M_2(b-a)h^2}{12} = \frac{M_2(b-a)^3}{12n^2}. \quad (7.17)$$

Ha az integrálközelítését az intervallum $2n$ részre osztásával is (azaz a $T_{2n}(f)$ -et is) meg tudjuk határozni, akkor a gyakorlatban rendszerint alkalmazhatjuk a következő becslést:

$$\left| \int_a^b f(x)dx - T_{2n}(f) \right| \leq |T_{2n} - T_n|. \quad (7.18)$$

7.2.2. A Simpson formula

Legyen $x_1 = a$, $x_2 = \frac{a+b}{2}$ és $x_3 = b$. Tekintsük a három pontra támaszkodó másodfokú Lagrange-féle interpolációs polinomot:

$$p(x) = f(x_1) \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + f(x_2) \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + f(x_3) \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} \quad (7.19)$$

Ennek az $[a, b]$ intervallumon vett integrálja adja a következő közelítő formulát

$$\int_a^b f(x)dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right], \quad (7.20)$$

amelynek hibájára $f \in C^4[a, b]$ esetén fennáll, hogy

$$\left| \int_a^b f(x) dx - \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \right| \leq M_4 \frac{(b-a)^5}{2880}. \quad (7.21)$$

Ha az $[a, b]$ intervallumot itt is felbontjuk az

$$a = x_1 < x_2 < \dots < x_{n+1} = b$$

pontokkal n részintervallumra, akkor az összetett Simpson formula a következő:

$$\int_a^b f(x) dx \approx S_n(f) = \sum_{i=1}^n \frac{x_{i+1} - x_i}{6} \left[f(x_i) + 4f\left(\frac{x_i + x_{i+1}}{2}\right) + f(x_{i+1}) \right].$$

Ennek hibájára fennáll, hogy

$$\left| \int_a^b f(x) dx - S_n(f) \right| \leq \frac{M_4}{2880} \sum_{i=1}^n (x_{i+1} - x_i)^5.$$

Ha az alappontok ekvidisztansak, azaz $x_i = x_1 + (i-1)h$ ($h = \frac{b-a}{n}$, $i = 1, \dots, n+1$), akkor a képlet alakja

$$S_n(f) = \frac{h}{6} \left[f(x_1) + 2 \sum_{i=2}^n f(x_i) + 4 \sum_{i=1}^n f\left(x_i + \frac{h}{2}\right) + f(x_{n+1}) \right] \quad (7.22)$$

amelynek képlethibájára fennáll, hogy

$$\left| \int_a^b f(x) dx - S_n(f) \right| \leq \frac{M_4(b-a)}{2880} h^4 = \frac{M_4(b-a)^5}{2880n^4}. \quad (7.23)$$

Ha az integrálközelítést az intervallum $2n$ részre osztásával is (azaz az $S_{2n}(f)$ -et is) meg tudjuk határozni, akkor a gyakorlatban rendszerint alkalmazhatjuk a következő becslést:

$$\left| \int_a^b f(x) dx - S_{2n}(f) \right| \leq |S_{2n} - S_n|. \quad (7.24)$$

8. fejezet

Függvényközelítés legkisebb négyzetek módszerével

Tekintjük az interpolációnál már bevezetett

$$h(x) = a_1\phi_1(x) + a_2\phi_2(x) + \dots + a_m\phi_m(x) = \sum_{i=1}^m a_i\phi_i(x). \quad (8.1)$$

alakú közelítő függvényt, ahol a $\{\phi_i(x)\}_{i=1}^n$ ismert függvények. Az $f(x) \approx h(x)$ legkisebb négyzetes közelítést külön tárgyaljuk diszkrét és folytonos esetben.

8.1. Diszkrét eset

Ismert az $[a, b]$ intervallum

$$a = x_1 < x_2 < \dots < x_n = b$$

felosztása. A legkisebb négyzetes közelítés alapelve:

$$d(a_1, a_2, \dots, a_m) = \sum_{i=1}^n [f(x_i) - h(x_i)]^2 = \sum_{i=1}^n [f(x_i) - \sum_{j=1}^m a_j\phi_j(x_i)]^2 = \min \quad (8.2)$$

Tehát az m -változós valós $d(a_1, a_2, \dots, a_m)$ függvény minimumhelyét kell meghatároznunk. Az $[\hat{a}_1, \dots, \hat{a}_m]^T$ akkor lehet minimumhely, ha kielégíti a

$$\frac{\partial d(a_1, \dots, a_m)}{\partial a_j} = 0 \quad (j = 1, \dots, m) \quad (8.3)$$

egyenletrendszert. Vezessük be a

$$g_j = [\phi_j(x_1), \phi_j(x_2), \dots, \phi_j(x_n)]^T, \quad (j = 1, \dots, m) \quad (8.4)$$

és az $f = [f(x_1), f(x_2), \dots, f(x_n)]^T$ vektorokat. A deriválások elvégzése és rendezés után a feltételi egyenletrendszer

$$\begin{aligned}
a_1 g_1^T g_1 + a_2 g_1^T g_2 + \dots + a_m g_1^T g_m &= g_1^T f \\
a_1 g_2^T g_1 + a_2 g_2^T g_2 + \dots + a_m g_2^T g_m &= g_2^T f \\
&\vdots \\
a_1 g_m^T g_1 + a_2 g_m^T g_2 + \dots + a_m g_m^T g_m &= g_m^T f
\end{aligned} \tag{8.5}$$

alakú. Ezt az egyenletrendszert normál egyenletrendszernek nevezzük.

Tétel. Ha a g_j vektorok ($j = 1, \dots, m$) lineárisan függetlenek, akkor a normál egyenletrendszer egyértelműen megoldható és az $[\hat{a}_1, \dots, \hat{a}_n]^T$ megoldás egyben a $d(a_1, a_2, \dots, a_m)$ függvénynek minimumhelye, ami ez esetben szintén egyértelműen létezik.

Megmutatható, hogy polinomközelítés, azaz

$$\phi_1(x) = 1, \phi_2(x) = x, \dots, \phi_m(x) = x^{m-1} \tag{8.6}$$

esetén a tétel feltételei fennállnak, tehát a feladatnak egyértelmű megoldása van.

Megjegyzés: Általában $m < n$. Az $m = n$ esetben a legkisebb négyzetes közelítés és az interpolációs közelítés megegyezik, ha mindkettő létezik.

8.2. Folytonos eset

Ebben az esetben a legkisebb négyzetes közelítés alapelve:

$$r(a_1, \dots, a_n) = \int_a^b (f(x) - h(x))^2 dx = \int_a^b \left(f(x) - \sum_{j=1}^m a_j \phi_j(x) \right)^2 dx = \min \tag{8.7}$$

Vezessük be a

$$g_{ij} = \int_a^b \phi_i(x) \phi_j(x) dx, \quad (i, j = 1, \dots, m) \tag{8.8}$$

és a

$$c_i = \int_a^b \phi_i(x) f(x) dx, \quad (i = 1, \dots, m) \tag{8.9}$$

jelölést. Ezekkel hasonló normál egyenletrendszer vezethető le, mint diszkrét esetben:

$$\begin{aligned}
a_1 g_{11} + a_2 g_{12} + \dots + a_m g_{1m} &= c_1 \\
a_1 g_{21} + a_2 g_{22} + \dots + a_m g_{2m} &= c_2 \\
&\vdots \\
a_1 g_{m1} + a_2 g_{m2} + \dots + a_m g_{mm} &= c_m
\end{aligned} \tag{8.10}$$

Tétel. Ha a $\phi_j(x)$ függvények ($j = 1, \dots, m$) lineárisan függetlenek, akkor a normál egyenletrendszer egyértelműen megoldható és az $[\hat{a}_1, \dots, \hat{a}_n]^T$ megoldás

egyben az $r(a_1, a_2, \dots, a_m)$ függvénynek minimumhelye, ami ez esetben szintén egyértelműen létezik.

Itt is igaz, hogy polinomközelítés, azaz

$$\phi_1(x) = 1, \phi_2(x) = x, \dots, \phi_m(x) = x^{m-1} \quad (8.11)$$

esetén a tétel feltételei fennállnak, tehát a feladatnak egyértelmű megoldása van.

9. fejezet

Differenciálegyenletek numerikus megoldása

Csak közönséges differenciálegyenletek kezdetiérték feladata megoldására szolgáló Runge-Kutta típusú módszerekkel foglalkozunk. Az

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (f : \mathbb{R}^2 \rightarrow \mathbb{R}) \quad (9.1)$$

alakú kezdetiérték feladatokat vizsgáljuk, ahol $f(x, y)$ folytonos a

$$D = \{(x, y) \mid |x - x_0| < k_x, |y - y_0| < k_y\} \subseteq \mathbb{R}^2 \quad (9.2)$$

nyílt tartományon, k_x és k_y pozitív konstansok és létezik olyan $L > 0$ konstans, hogy

$$|f(x, y) - f(x, z)| \leq L|y - z| \quad ((x, y), (x, z) \in D). \quad (9.3)$$

Ekkor minden $(x_0, y_0) \in D$ esetén a kezdetiérték feladatnak létezik pontosan egy megoldása valamely $[x_0, B]$ intervallumon, azaz

$$y'(x) = f(x, y(x)), \quad x \in [x_0, B]. \quad (9.4)$$

A feladat numerikus megoldásán a következőt értjük. A megoldást egy $[x_0, b]$ intervallum ($b \leq B$) diszkrét pontjaiban keressük. Ezek a pontok legyenek

$$x_0 = t_0 < t_1 < \dots < t_j < \dots < t_N = b. \quad (9.5)$$

A $\{t_i\}_{i=0}^N$ alappont halmazt az $[x_0, b]$ intervallum felosztásának nevezzük. A felosztás ekvidisztans (egyenávolságú), ha

$$t_i = t_0 + ih \quad (i = 0, 1, \dots, N), \quad h = \frac{b - t_0}{N}. \quad (9.6)$$

Az $y(t_i)$ elméleti megoldás t_i pontbeli közelítését jelölje y_i . Értelmszerűen $y(t_0) = y_0$. A $h_i = t_{i+1} - t_i > 0$ mennyiséget i -edik lépéshossznak nevezzük.

9.1. Az explicit Euler-módszer

A vizsgált feladatokban, ha egy x pontban ismert $y = y(x)$, akkor ismert az $y'(x) = f(x, y(x))$ derivált érték is. Ennek következtében az x pont egy

környezetében az $y(x)$ pontbeli érintőközelítést azonnal megadhatjuk:

$$y(x+h) \approx y(x) + hy'(x) = y(x) + hf(x, y(x)).$$

Ha az x pontban az $y(x) \approx \hat{y}$ közelítő érték ismert, akkor a fenti képlet átmegy az

$$y(x+h) \approx \hat{y} + hf(x, \hat{y})$$

közelítésbe.

Az Euler-módszer alapgondolata ezek után a következő: A $t_1 = t_0 + h_0$ pontban közelítsük az $y(t_1)$ elméleti megoldást a görbe (x_0, y_0) pontbeli érintőjével, azaz legyen

$$y(t_0 + h_0) \approx y_0 + hy'(t_0) = y_1 = y_0 + hf(t_0, y_0). \quad (9.7)$$

A t_1 pontbeli y_1 közelítést felhasználva kapjuk, hogy

$$y(t_2) \approx y(t_1) + h_1 f(t_1, y(t_1)) \approx y_2 = y_1 + h_1 f(t_1, y_1). \quad (9.8)$$

Az eljárást folytatva kapjuk hogy

$$y_{i+1} = y_i + h_i f(t_i, y_i) \quad (i = 0, 1, \dots, N-1), \quad (9.9)$$

ahol $y_i \approx y(t_i)$. Ezt a képletet nevezzük explicit Euler-módszernek.

9.2. A negyedrendű Runge-Kutta módszer

Az Euler-módszernek számtalan továbbfejlesztése ismeretes. Ezek közül az egyik legfontosabb az explicit egylépéses módszerek osztálya, ezekből is csak a negyedrendű Runge-Kutta módszert említjük ekvidisztans h lépésközzel (t_i helyett most x_i jelölést használva):

$$y_{i+1} = y_i + h\left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4\right) \quad (i = 0, 1, \dots, N-1), \quad (9.10)$$

ahol

$$\begin{aligned} k_1 &= f(x_i, y_i) \\ k_2 &= f(x_i + h/2, y_i + hk_1/2) \\ k_3 &= f(x_i + h/2, y_i + hk_2/2) \\ k_4 &= f(x_i + h, y_i + hk_3) \end{aligned} \quad (9.11)$$