LINEÁRIS EGYENLETRENDSZEREK MEGOLDÁSA INTERVALLUMARITMETIKAI MÓDSZEREKKEL

Diplomamunka

Írta: Lerchner Szilvia Zsuzsanna

Alkalmazott matematikus szak

Témavezető:

Gergó Lajos, docens

Numerikus Analízis Tanszék

Eötvös Loránd Tudományegyetem, Informatikai Kar

Eötvös Loránd Tudományegyetem

Természettudományi Kar

2011

Tartalomjegyzék

1. fejezet

Bevezetés

1.1. Célkitűzés, motiváció

Az intervallumaritmetikai kutatások témaköre az utóbbi évtizedekben ismét a kutatók figyelmébe került. Számos cikk foglalkozik a numerikus matematika intervallum alapú kiterjesztésével, a különböző intervallumfeladatok megoldhatóságával, megoldási algoritmusok megadásával. Néhány programozási nyelv esetében intervallumaritmetikai kiterjesztés is született, ami azt jelenti, hogy könnyen lehet intervallumaritmetikai számításokat elvégezni (Fortran, Algol, C). Az említett terület megbízható numerikus számítások címen került be a szakirodalomba és jelenleg is sokan foglalkoznak a témával, sok publikáció jelenik meg évről évre. Célom a diplomamunkával az, hogy az intervallumaritmetika alapvető bevezetése után az intervallumos lineáris egyenletrendszerek témakörében áttekintést adjak a mai kutatási eredményekről.

Dolgozatom első fejezetében összefoglalom az alapvető intervallumaritmetikai fogalmakat és állításokat, és bevezetem az intervallummátrixokat illetve intervallumvektorokat. Dolgozatom további részeiben az

$$\mathbf{A}x = \mathbf{b} \tag{1.1}$$

úgynevezett intervallum-együtthatós lineáris egyenletrendszerekkel foglalkozom. A második fejezetben egy elméleti kérdésre térek ki, mégpedig az intervallum-együtthatós lineáris egyenletrendszerek megoldhatóságának kérdésére. Itt a legfontosabb eredmény az, hogy ez egy NP-nehéz feladat. Ha (1.1)-ben az $\mathbf{A}$ mátrix reguláris, akkor létezik megoldása (1.1)-nek, ami egy kompakt és összefüggő halmaz, ha viszont $\mathbf{A}$ szinguláris, akkor a megoldáshalmaz vagy üres halmaz, vagy minden komponense nem korlátos. Mivel a megoldáshalmaz még reguláris esetben is általában egy bonyolult nemkonvex struktúra, ezért innentől kezdve nem ezt keressük, hanem az őt tartalmazó legszűkebb intervallumvektort. A harmadik fejezetben leírom, hogy milyen feltételek mellett lehet alkalmazni reguláris baloldal esetén a jól ismert Gauss-eliminációt az intervallumos feladatokra. Itt az derül ki, hogy igazán csak akkor érdemes ezt használni, ha az $\mathbf{A}$ intervallummátrix elemeinek szélessége elég kicsi. Ezután a negyedik fejezetben egy olyan eredményt írok le, amelyet akkor érdemes alkalmazni, ha az $\mathbf{A}$ intervallummátrix elemeinek szélessége viszonylag nagy, és ekkor ugyan több számolás árán, mint a Gauss-eliminációval, de jobb eredményre jutunk. Végül az ötödik fejezetben azokat az eredményeket írom le, melyek arról szólnak, hogy mit lehet tenni, ha nem tudjuk, hogy az $\mathbf{A}$ mátrix reguláris-e. Itt a végső eredmény egy konkrét algoritmus, amely tetszőleges baloldalú intervallum-együtthatós lineáris egyenletrendszer esetén, ha $\mathbf{A}$ szinguláris, akkor ad egy szinguláris részmátrixot, ellenkező esetben pedig megadja a megoldásokat tartalmazó legszűkebb intervallumvektort.

Összefoglalva tehát dolgozatomban kitérek mind az elméleti kérdésekre (megoldás létezése), mind a megoldási módszerekre, algoritmusokra. 2009-es és 2010-es megjelenésű friss cikkek eredményeit dolgoztam fel egy lehetséges megoldási algoritmus bemutatására. Matlab programot is készítettem az algoritmusok működésének az ellenőrzésére, kipróbálására.

1.2. Alapfogalmak és jelölések

A valós számok halmazát a szokásos $\mathbb{R}$-rel fogjuk jelölni, míg a tagjait kisbetűkkel. $\mathbb{R}$-nek egy speciális részhalmaza zárt valós intervallum, vagy röviden csak intervallum, ami a következő:

$$\mathcal{A} = [a_1, a_2] = \{t \in \mathbb{R} : a_1 \leq t \leq a_2, a_1, a_2 \in \mathbb{R}\}.$$

1.2.1. Valós intervallumaritmetika

Jelöljük $I(\mathbb{R})$-rel a zárt valós intervallumok halmazát. Ekkor speciáliasan $[x, x] \in I(\mathbb{R})$ pontintervallum.

1. Definíció. *Két intervallum, $\mathcal{A} = [a_1, a_2]$ és $\mathcal{B} = [b_1, b_2]$ egyenlők, azaz $\mathcal{A} = \mathcal{B}$, ha halmazelméleti értelemben egyenlők.*

Tehát $\mathcal{A} = \mathcal{B} \Leftrightarrow a_1 = b_1$ és $a_2 = b_2$.

1. Állítás. *Az "$=$" reláció két $I(\mathbb{R})$-beli elemen reflexív, szimmetrikus és tranzitív.*

2. Definíció. *Legyen $* \in \{+, -, \cdot, :\}$ egy bináris művelet $\mathbb{R}$-en. Ha $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$, akkor $\mathcal{A} * \mathcal{B} := \{z = a * b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$. Ez bináris műveletet definiál $I(\mathbb{R})$-en.*

Megjegyezzük, hogy az osztás esetén mindig feltesszük, hogy $0 \notin \mathcal{B}$, és ezt a továbbiakban nem jegyezzük meg. Továbbá azt is megjegyezzük, hogy azonos szimbólumokat használunk az $\mathbb{R}$-beli és az $I(\mathbb{R})$-beli műveletekre.

2. Állítás. *Legyen $\mathcal{A} = [a_1, a_2]$ és $\mathcal{B} = [b_1, b_2]$. Ekkor*

1. $\mathcal{A} + \mathcal{B} = [a_1 + b_1, a_2 + b_2]$,
2. $\mathcal{A} - \mathcal{B} = [a_1 - b_2, a_2 - b_1] = \mathcal{A} + [-1, -1] \cdot \mathcal{B}$,
3. $\mathcal{A}\mathcal{B} = [\min\{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}, \max\{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}]$,
4. $\mathcal{A} : \mathcal{B} = [a_1, a_2] \cdot \left[\frac{1}{b_2}, \frac{1}{b_1}\right]$.

A fenti állításból az is látszik, hogy $I(\mathbb{R})$ zárt ezekre a műveletekre nézve. Továbbá világos, hogy a valós számok halmaza izomorf a pontintervallumok halmazával, ezért $[x, x] * \mathcal{A}$-t egyszerűbben $x * \mathcal{A}$-val jelöljük.

3. Definíció. *Legyen $\mathcal{X} \in I(\mathbb{R})$, $r(x)$ folytonos függvény $\mathbb{R}$-en. Ekkor*

$$r(\mathcal{X}) := \left[\min_{x\in\mathcal{X}} r(x), \max_{x\in\mathcal{X}} r(x)\right].$$

1. Tétel. *Legyen $\mathcal{A}, \mathcal{B}, \mathcal{C} \in I(\mathbb{R})$. Ekkor az alábbiak teljesülnek.*

1. *$\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$ és $\mathcal{AB} = \mathcal{BA}$, azaz az összeadás és a szorzás kommutatív.*

2. *$(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$ és $(\mathcal{AB})\mathcal{C} = \mathcal{A}(\mathcal{BC})$, azaz ezek a műveletek asszociatívak is.*

3. *$\mathcal{X} = [0, 0]$ és $\mathcal{Y} = [1, 1]$ az egyértelmű neutrális elemek az összeadásra és a szorzásra nézve, azaz*

$$\mathcal{X} + \mathcal{A} = \mathcal{A} + \mathcal{X} = \mathcal{A} \quad \forall\mathcal{A} \in I(\mathbb{R}) \Leftrightarrow \mathcal{X} = [0, 0],$$

$$\mathcal{YA} = \mathcal{AY} = \mathcal{A} \quad \forall\mathcal{A} \in I(\mathbb{R}) \Leftrightarrow \mathcal{Y} = [1, 1].$$

4. *$\mathcal{AB} = 0 \Leftrightarrow \mathcal{A} = [0, 0]$ vagy $\mathcal{B} = [0, 0]$, azaz $I(\mathbb{R})$-nek nincs nullosztója.*

5. *Nem minden $\mathcal{A} = [a_1, a_2] \in I(\mathbb{R})$, $a_1 \neq a_2$ esetén létezik $\mathcal{A}$-nak inverze.*

6. *Minden $\mathcal{A} \in I(\mathbb{R})$ esetén $0 \in \mathcal{A} - \mathcal{A}$ és $1 \in \mathcal{A} : \mathcal{A}$.*

7. *$\mathcal{A}(\mathcal{B} + \mathcal{C}) \subseteq \mathcal{AB} + \mathcal{AC}$ (szubdisztributivitás),*
   $a(\mathcal{B} + \mathcal{C}) = a\mathcal{B} + a\mathcal{C}$, ha $a \in \mathbb{R}$ és
   $\mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{AB} + \mathcal{AC}$, ha $bc \geq 0 \quad \forall b \in \mathcal{B}$ és $c \in \mathcal{C}$ esetén.

A következő állítás az $\mathcal{AX} = \mathcal{B}$ intervallumegyenlet megoldhatóságáról szól, ahol $\mathcal{A}$ és $\mathcal{B}$ adottak, és $\mathcal{X}$ az ismeretlen. Természetesen feltesszük, hogy $\mathcal{A} \neq [0, 0]$. Ehhez először bevezetünk egy segédfüggvényt.

4. Definíció. *Legyen $\mathcal{A} = [a_1, a_2] \in I(\mathbb{R})$. Ekkor*

$$\Phi\mathcal{A} := \begin{cases} \frac{a_1}{a_2} & \text{, ha } |a_1| \leq |a_2|, \\ \frac{a_2}{a_1} & \text{különben.} \end{cases}$$

3. Állítás. *Akkor és csak akkor $\exists\mathcal{X} \in I(\mathbb{R})$, amely kielégíti az $\mathcal{AX} = \mathcal{B}$ egyenletet, ha*

$$\Phi\mathcal{A} \geq \Phi\mathcal{B}.$$

A megoldás akkor és csak akkor nem egyértelmű, ha

$$\Phi\mathcal{A} = \Phi\mathcal{B} \leq 0.$$

Ha tekintjük az $ax = b$, $a \in \mathcal{A}$, $b \in \mathcal{B}$ egyenlet megoldáshalmazát:

$$\widetilde{\mathcal{X}} = \left\{ x = \frac{b}{a} : a \in \mathcal{A},\ b \in \mathcal{B} \right\} = \mathcal{B} : \mathcal{A},$$

akkor ez lényegesen különbözik az $\mathcal{X}$ intervallumtól, ami kilégíti az $\mathcal{A}\mathcal{X} = \mathcal{B}$ egyenletet. Ezért $\mathcal{X}$-et nem az $\mathcal{A}\mathcal{X} = \mathcal{B}$ megoldásának, hanem algebrai megoldásának szokták nevezni. A következő állítás ennek a két megoldásnak a kapcsolatáról szól.

4. Állítás. *Legyen $\mathcal{A}\mathcal{X} = \mathcal{B}$, $0 \notin \mathcal{A}$ valamely $\mathcal{X} \in I(\mathbb{R})$-re. Ekkor $\mathcal{X} \subseteq \mathcal{B} : \mathcal{A}$.*

1. Megjegyzés. *Az $\mathcal{A}\mathcal{X} = \mathcal{B}$-nek akkor is létezhet algebrai megoldása, ha $\mathcal{B} : \mathcal{A}$ nem értelmes.*

A következő tétel egy alapvető intervallumaritmetikai tulajdonságról, a tartalmazási monotonitásról szól.

2. Tétel. *(Tartalmazási monotonitás) Legyen $\mathcal{A}^1, \mathcal{A}^2, \mathcal{B}^1, \mathcal{B}^2 \in I(\mathbb{R})$, és tegyük fel, hogy $\mathcal{A}^1 \subseteq \mathcal{B}^1$ és $\mathcal{A}^2 \subseteq \mathcal{B}^2$. Ekkor $* \in \{+, -, \cdot, :\}$ esetén $\mathcal{A}^1 * \mathcal{A}^2 \subseteq \mathcal{B}^1 * \mathcal{B}^2$.*

Speciálisan, ha $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$, $a \in \mathcal{A}$, $b \in \mathcal{B}$, akkor $a * b \in \mathcal{A} * \mathcal{B}$.

2. Megjegyzés. *Az egyértékű operátorokra is igaz, hogy ha $\mathcal{X} \subseteq \mathcal{Y}$, akkor $r(\mathcal{X}) \subseteq r(\mathcal{Y})$, és ha $x \in \mathcal{X}$, akkor $r(x) \in r(\mathcal{X})$.*

Ahhoz, hogy folytonosságról, illetve konvergenciáról beszélhessünk, először a távolság fogalmát kell definiálnunk.

5. Definíció. *Legyen $\mathcal{A} = [a_1, a_2], \mathcal{B} = [b_1, b_2] \in I(\mathbb{R})$. Ekkor $\mathcal{A}$ és $\mathcal{B}$ távolsága:*

$$q(\mathcal{A}, \mathcal{B}) := \max\{|a_1 - b_1|, |a_2 - b_2|\}.$$

5. Állítás. *A fent definiált távolság metrika $I(\mathbb{R})$-en.*

3. Megjegyzés. *Ha ezt a távolságot a pontintervallumokra alkalmazzuk, akkor a valós számokon értelmezett távolságfogalmat kapjuk.*

6. Állítás. *$I(\mathbb{R})$ topologikus tér, ahol a konvezgenciára a következő teljesül:*

$$\lim_{k \to \infty} \mathcal{A}^{(k)} = \mathcal{A} \Leftrightarrow \lim_{k \to \infty} a_1^{(k)} = a_1, \quad \lim_{k \to \infty} a_2^{(k)} = a_2.$$

3. Tétel. *$(I(\mathbb{R}), q)$ teljes metrikus tér.*

4. Tétel. *Tegyük fel, hogy $(\mathcal{A}^{(k)})$ olyan intervallumsorozat, melyre teljesül, hogy $\mathcal{A}^{(0)} \supseteq \mathcal{A}^{(1)} \supseteq \mathcal{A}^{(2)} \supseteq ....$ Ekkor $(\mathcal{A}^{(k)})$ kovergens, és*

$$\lim_{k\to\infty} \mathcal{A}^{(k)} = \mathcal{A} = \bigcap_{k=0}^{\infty} \mathcal{A}^{(k)}.$$

1. Következmény. *Tegyük fel, hogy $(\mathcal{A}^{(k)})$ olyan intervallumsorozat, melyre teljesül, hogy $\mathcal{A}^{(0)} \supseteq \mathcal{A}^{(1)} \supseteq \mathcal{A}^{(2)} \supseteq ... \supseteq \mathcal{B}$. Ekkor $(\mathcal{A}^{(k)})$ kovergens, és*

$$\lim_{k\to\infty} \mathcal{A}^{(k)} = \mathcal{A} \supseteq \mathcal{B}.$$

5. Tétel. *Az $+, -, \cdot, :$ műveletek az intervallumokon folytonosak.*

2. Következmény. *Legyen r folytonos függvény, és legyen*

$$r(\mathcal{X}) = \left[\min_{x\in\mathcal{X}} r(x), \max_{x\in\mathcal{X}} r(x)\right].$$

Ekkor $r(\mathcal{X})$ folytonos intervallum-függvény.

A következőkben definiálni fogjuk egy intervallum abszolútértékét, de a szokásostól eltérően nem az intervallum hossza lesz ez, hanem, mint a valós számoknál, a nullától vett távolsága.

6. Definíció. *Legyen $\mathcal{A} = [a_1, a_2] \in I(\mathbb{R})$. Ekkor*

$$|\mathcal{A}| := q(\mathcal{A}, [0,0]) = \max\{|a_1|, |a_2|\} = \max_{a\in\mathcal{A}} |a|.$$

Könnyen látszik a definícióból, hogy ha $\mathcal{A}, \mathcal{B} \in I(\mathbb{R}), \quad \mathcal{A} \subseteq \mathcal{B}$, akkor $|\mathcal{A}| \leq |\mathcal{B}|$.

6. Tétel. *Legyen $\mathcal{A} = [a_1, a_2], \mathcal{B} = [b_1, b_2], \mathcal{C} = [c_1, c_2], \mathcal{D} = [d_1, d_2] \in I(\mathbb{R})$. Ekkor a következők teljesülnek.*

1. $q(\mathcal{A}+\mathcal{B}, \mathcal{A}+\mathcal{C}) = q(\mathcal{B}, \mathcal{C})$,

2. $q(\mathcal{A}+\mathcal{B}, \mathcal{C}+\mathcal{D}) \leq q(\mathcal{A}, \mathcal{C}) + q(\mathcal{B}, \mathcal{D})$,

3. $q(a\mathcal{B}, a\mathcal{C}) = |a| q(\mathcal{B}, \mathcal{C})$, *ahol* $a \in \mathbb{R}$,

4. $q(\mathcal{A}\mathcal{B}, \mathcal{A}\mathcal{C}) \leq |\mathcal{A}| q(\mathcal{B}, \mathcal{C})$.

7. Állítás. *Legyen $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$. Ekkor a következők igazak.*

1. $|\mathcal{A}| \geq 0$, *és* $|\mathcal{A}| = 0 \Leftrightarrow \mathcal{A} = [0, 0]$,

2. $|\mathcal{A} + \mathcal{B}| \leq |\mathcal{A}| + |\mathcal{B}|$,

3. $\forall x \in \mathbb{R}$ *esetén* $|x\mathcal{A}| = |x||\mathcal{A}|$,

4. $|\mathcal{A}\mathcal{B}| = |\mathcal{A}||\mathcal{B}|$.

7. Definíció. *Az $\mathcal{A} = [a_1, a_2]$ intervallum szélessége a következő:*

$$d(\mathcal{A}) := a_2 - a_1 \geq 0.$$

Tehát a pontintervallumok jellemezhetők úgy is, hogy $\{\mathcal{A} \in I(\mathbb{R}) : d(\mathcal{A}) = 0\}$. A következő állítások a definícióból triviálisan következnek.

8. Állítás. *Legyen $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$. Ekkor*

1. *ha* $\mathcal{A} \subseteq \mathcal{B}$, *akkor* $d(\mathcal{A}) \leq d(\mathcal{B})$.

2. $d(\mathcal{A} \pm \mathcal{B}) = d(\mathcal{A}) \pm d(\mathcal{B})$.

7. Tétel. *Legyen $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$. Ekkor*

1. $d(\mathcal{A}\mathcal{B}) \leq d(\mathcal{A})|\mathcal{B}| + |\mathcal{A}|d(\mathcal{B})$.

2. $d(\mathcal{A}\mathcal{B}) \geq \max\{d(\mathcal{A})|\mathcal{B}|, |\mathcal{A}|d(\mathcal{B})\}$.

3. $d(a\mathcal{B}) = |a|d(\mathcal{B})$.

4. $d(\mathcal{A}^n) \leq n|\mathcal{A}^{n-1}|a(\mathcal{A})$, $n = 1, 2, ...$, *ahol* $\mathcal{A}^n = \underbrace{\mathcal{A} \cdot \mathcal{A} \cdot ... \cdot \mathcal{A}}_{n\ db}$.

5. $d((\mathcal{A} - a)^n) \leq 2(d(\mathcal{A}))^n$, *ahol* $a \in \mathcal{A}$, $n = 1, 2, ...$

6. *Ha* $\mathcal{C} \in I(\mathbb{R})$ *és* $0 \in \mathcal{C}$, *akkor* $|\mathcal{C}| \leq d(\mathcal{C}) \leq 2|\mathcal{C}|$.

8. Tétel. *Legyen* $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$, *és tegyük fel, hogy* $\mathcal{A}$ *szimmetrikus intervallum, azaz* $\mathcal{A} = -\mathcal{A}$. *Ekkor*

1. $\mathcal{A}\mathcal{B} = |\mathcal{B}|\mathcal{A}$, *és*

2. $d(\mathcal{A}\mathcal{B}) = |\mathcal{B}|d(\mathcal{A})$.

Ez utóbbi állítás akkor is igaz, ha nem tesszük fel azt, hogy $\mathcal{A}$ *szimmetrikus, hanem csak annyit, hogy* $0 \in \mathcal{A}$ *és* $b_1 \geq 0$ *vagy* $b_2 \leq 0$.

9. Tétel. *Legyen* $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$. *Ekkor*

1. $d(\mathcal{A}) = |\mathcal{A} - \mathcal{A}|$, *és*

2. $\mathcal{A} \subseteq \mathcal{B}$ *esetén* $\frac{1}{2}(d(\mathcal{B}) - d(\mathcal{A})) \leq q(a, \mathcal{B}) \leq d(\mathcal{B}) - d(\mathcal{A})$.

8. Definíció. *Legyen* $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$. *Ekkor* $\mathcal{A}$ *és* $\mathcal{B}$ *metszete:*

$$\mathcal{A} \cap \mathcal{B} := \{c : c \in \mathcal{A} \wedge c \in \mathcal{B}\}.$$

Azaz a halmazelméleti metszete a két intervallumnak.

Megjegyezük, hogy az eredmény akkor és csak akkor intervallum, ha ez a halmazelméleti metszet nem üres, és ebben az esetben

$$\mathcal{A} \cap \mathcal{B} = [\max\{a_1, b_1\}, \min\{a_2, b_2\}].$$

9. Állítás. *Legyen* $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \in I(\mathbb{R})$, *és tegyük fel, hogy* $\mathcal{A} \subseteq \mathcal{C}$ *és* $\mathcal{B} \subseteq \mathcal{D}$. *Ekkor* $\mathcal{A} \cap \mathcal{B} \subseteq \mathcal{C} \cap \mathcal{D}$.

Ez az állítás a metszet definíciójából, míg a következő a fent írt alakból látszik jól.

10. Állítás. *A metszet művelet, ha értelmezhető, akkor folytonos* $I(\mathbb{R})$*-en.*

1.2.2. Intervallumokon értelmezett függvények értékkészlete és kiértékelése

Legyen f folytonos valós függvény. Az f-hez tartozó $f(x)$ kifejezés egy kiszámítási módot jelent, ami meghatároz egy értéket az f minden argumentumához. Ha f tartalmaz konstansokat is, akkor ezt a következőképpen jelöljük: $f(x; a^{(0)}, a^{(1)}, ..., a^{(m)})$. Az egyszerűség kedvéért tegyük fel, hogy minden konstans csak egyszer fordul elő egy kifejezésben. Ezt új konstans bevezetésével, vagy a kifejezés egyszerűsítésével érhetjük el.

9. Definíció.

$$W(f,\mathcal{X};\mathcal{A}^{(0)},\mathcal{A}^{(1)},...,\mathcal{A}^{(m)}) := \{f(x;a^{(0)},a^{(1)},...,a^{(m)}) : x\in\mathcal{X}, a^{(k)}\in\mathcal{A}^{(k)}, 0\le k\le m\} =$$

$$= \left[\min_{x\in\mathcal{X},a^{(k)}\in\mathcal{A}^{(k)},0\le k\le m} f(x;a^{(0)},a^{(1)},...,a^{(m)}), \max_{x\in\mathcal{X},a^{(k)}\in\mathcal{A}^{(k)},0\le k\le m} f(x;a^{(0)},a^{(1)},...,a^{(m)})\right],$$

ahol $x\in\mathcal{X}, a^{(k)}\in\mathcal{A}^{(k)}, 0\le k\le m$ *egymástól függetlenek.*

10. Definíció. *Legyen adott az* f *függvény egy kifejezése. Az* $f(X;\mathcal{A}^{(0)},\mathcal{A}^{(1)},...,\mathcal{A}^{(m)})$ *az a kifejezés, amit úgy kapunk, hogy az eredeti kifejezésben az összes változó és konstans helyébe intervallumokat helyettesítünk és a műveleteket intervallumműveletekkel cseréljük fel. Ha ez értelmes, akkor ezt nevezzük az* f *függvény intervallumkiértékelésének, vagy intervallumaritmetikai kiértékelésének.*

4. Megjegyzés. *Az* f *intervallumkiértékelése függ a kifejezéstől, és* $W(f,\mathcal{X};\mathcal{A}^{(0)},A^{(1)},...,\mathcal{A}^{(m)})$ *csak magától a függvénytől függ.*

Példa: Legyen a g függvény két kifejezése a következő:

$$g^{(1)}(x;a) = \frac{ax}{1-x}, x\neq 1, x\neq 0,$$

és

$$g^{(2)}(x;a) = \frac{a}{\frac{1}{x}-1}, x\neq 1, x\neq 0.$$

Ekkor

$$W(g,[2,3];[0,1]) = \left\{\frac{ax}{1-x} : 2\le x\le 3, 0\le a\le 1\right\} = [-2,0],$$

$$g^{(1)}([2,3];[0,1]) = \frac{[2,3][0,1]}{1-[2,3]} = \frac{[0,3]}{[-2,-1]} = [-3,0],$$

$$g^{(2)}([2,3];[0,1]) = \frac{[0,1]}{\frac{1}{[2,3]}-1} = \frac{[0,1]}{[-\frac{2}{3},-\frac{1}{2}]} = [-2,0].$$

A fent bevezetett jelölések több változó esetén is alkalmazhatók.

10. Tétel. *Legyen* f *folytonos függvény, és legyen*

$$f(x^{(1)},x^{(2)},...,x^{(n)};a^{(0)},a^{(1)},...,a^{(m)})$$

az f *egy kifejezése. Továbbá tegyük fel, hogy létezik az* f *intervallumkiértékelése:*

$$f(\mathcal{Y}^{(1)},\mathcal{Y}^{(2)},...,\mathcal{Y}^{(n)};\mathcal{B}^{(0)},\mathcal{B}^{(1)},...,\mathcal{B}^{(m)}),$$

az $\mathcal{Y}^{(1)},\mathcal{Y}^{(2)},...,\mathcal{Y}^{(n)},\mathcal{B}^{(0)},\mathcal{B}^{(1)},...,\mathcal{B}^{(m)}$ *intervallumokkal. Ekkor*

1. $\forall \mathcal{X}^{(k)} \subseteq \mathcal{Y}^{(k)}$, $\mathcal{A}^{(j)} \subseteq \mathcal{B}^{(j)}$, $1 \leq k \leq n$, $0 \leq j \leq m$ *esetén*

$$W(f, \mathcal{X}^{(1)}, ..., \mathcal{X}^{(n)}; \mathcal{A}^{(0)}, ..., \mathcal{A}^{(m)}) \subseteq f(\mathcal{X}^{(1)}, ..., \mathcal{X}^{(n)}; \mathcal{A}^{(0)}, ..., \mathcal{A}^{(m)}),$$

2. $\forall \mathcal{X}^{(k)} \subseteq \mathcal{Z}^{(k)} \subseteq \mathcal{Y}^{(k)}$, $\mathcal{A}^{(j)} \subseteq \mathcal{C}^{(j)} \subseteq \mathcal{B}^{(j)}$, $1 \leq k \leq n$, $0 \leq j \leq m$ *esetén*

$$f(\mathcal{X}^{(1)}, ..., \mathcal{X}^{(n)}; \mathcal{A}^{(0)}, ..., \mathcal{A}^{(m)}) \subseteq f(\mathcal{Z}^{(1)}, ..., \mathcal{Z}^{(n)}; \mathcal{C}^{(0)}, ..., \mathcal{C}^{(m)}).$$

A tétel első pontjában bizonyos esetekben egyenlőség áll, a következő tétel egy ilyen esetről szól.

11. Tétel. *Tekintsük az $x \in \mathbb{R}$ változó p polinomjának a következő kifejezését.*

$$p(x; a^{(0)}, a^{(1)}, ..., a^{(m)}) = (...((a^{(m)}x + a^{(m-1)})^{n_m-1} + a^{(m-2)})^{n_m-2} + ... + a^{(1)})^{n_1} + a^{(0)},$$

ahol $n_l \geq 2$, $1 \leq l \leq m-1$. Ha a kifejezésben megjelenő intervallumok hatványát a következőképpen értelmezzük:

$$\mathcal{X}^k = \left[\min_{x \in \mathcal{X}} x^k, \max_{x \in \mathcal{X}} x^k\right],$$

akkor

$$W(p, \mathcal{X}; \mathcal{A}^{(0)}, \mathcal{A}^{(1)}, ..., \mathcal{A}^{(m)}) = p(\mathcal{X}; \mathcal{A}^{(0)}, \mathcal{A}^{(1)}, ..., \mathcal{A}^{(m)}).$$

Speciális esetekben adhatunk becslést egy függvény intervallum kiértékelése és intervallum-értékkészlete közti eltérésére.

12. Tétel. *Legyen f folytonos függvény, és legyen $f(x; a^{(0)}, a^{(1)}, ..., a^{(m)})$ az f egy kifejezése. Az $\widetilde{f}(x^{(1)}, x^{(2)}, ..., x^{(n)}; a^{(0)}, a^{(1)}, ..., a^{(m)})$ kifejezést úgy definiáljuk, hogy az eredeti függvényben az x változó minden előfordulását egy új változóval, $x^{(k)}$-val, $1 \leq k \leq n$ jelöljük az újban. Tegyük fel, hogy létezik az f intervallumkiértékelése: $f(\mathcal{Y}; \mathcal{B}^{(0)}, \mathcal{B}^{(1)}, ..., \mathcal{B}^{(m)})$, az $\mathcal{Y}, \mathcal{B}^{(0)}, \mathcal{B}^{(1)}, ..., \mathcal{B}^{(m)}$ intervallumokkal. Továbbá tegyük fel, hogy $\widetilde{f}(x^{(1)}, x^{(2)}, ..., x^{(n)}; a^{(0)}, a^{(1)}, ..., a^{(m)})$ $\forall x^{(k)} \in \mathcal{Y}$, $1 \leq k \leq n$ esetén kielégíti a Lipschitz-feltételt az $x^{(j)} \in \mathcal{Y}$, $1 \leq j \leq n$, $j \neq k$ és $a^{(j)} \in \mathcal{A}^{(j)}$, $0 \leq j \leq m$ tetszőleges választása esetén. Ekkor $\forall X \subseteq \mathcal{Y}$ esetén*

$$q(W(f, \mathcal{X}; \mathcal{A}^{(0)}, \mathcal{A}^{(1)}, ..., \mathcal{A}^{(m)}), f(\mathcal{X}; \mathcal{A}^{(0)}, \mathcal{A}^{(1)}, ..., \mathcal{A}^{(m)})) \leq \gamma d(\mathcal{X}),$$

valamely $\gamma \geq 0$-ra.

11. Definíció. *Az f következő kifejezését:*

$$f(x) = f(z) + (x - z)h(x - z)$$

, az f függvény z pont körüli centrális alakjának nevezünk.

13. Tétel. *Legyen $f : \mathbb{R} \to \mathbb{R}$ és legyen $f(x) = f(z) + (x - z)h(x - z)$ egy kifejezés f-re a centrális alakban. A $\widetilde{h}(x^{(1)} - z, x^{(2)} - z, ..., x^{(n)} - z)$ kifejezést úgy definiáljuk, hogy az eredeti függvényben az $x - z$ változó minden előfordulását egy új változóval, $x^{(k)} - z$-vel, $1 \leq k \leq n$ jelöljük az újban. Tegyük fel, hogy az f intervallumkiértékelése létezik valamilyen $\mathcal{Y} \in I(\mathbb{R})$-re, és hogy $\widetilde{h}(x^{(1)} - z, x^{(2)} - z, ..., x^{(n)} - z)$ teljesíti a Lipschitz-feltételt minden változójában úgy, mint az előző tételben. Ekkor $\forall \mathcal{X} \subseteq \mathcal{Y}$ esetén*

$$q(W(f, \mathcal{X}), f(\mathcal{X})) \leq c(d(\mathcal{X}))^2$$

valamely $c \geq 0$-ra.

Megjegyezzük, hogy az előző tételt többváltozós függvényekre is ki lehet mondani.

14. Tétel. *Legyen f folytonos függvény, és legyen $f(x; a^{(0)}, a^{(1)}, ..., a^{(m)})$ az f egy kifejezése. Az $\widetilde{f}(x^{(1)}, x^{(2)}, ..., x^{(n)}; a^{(0)}, a^{(1)}, ..., a^{(m)})$ kifejezést úgy definiáljuk, hogy az eredeti függvényben az x változó minden előfordulását egy új változóval, $x^{(k)}$-val, $1 \leq k \leq n$ jelöljük az újban. Tegyük fel, hogy létezik az f intervallumkiértékelése: $f(\mathcal{Y}; \mathcal{B}^{(0)}, \mathcal{B}^{(1)}, ..., \mathcal{B}^{(m)})$, az $\mathcal{Y}, \mathcal{B}^{(0)}, \mathcal{B}^{(1)}, ..., \mathcal{B}^{(m)}$ intervallumokkal. Továbbá tegyük fel, hogy $\widetilde{f}(x^{(1)}, x^{(2)}, ..., x^{(n)}; a^{(0)}, a^{(1)}, ..., a^{(m)})$ $\forall x^{(k)} \in \mathcal{Y}$, $1 \leq k \leq n$ esetén kielégíti a Lipschitz-feltételt az $x^{(j)} \in \mathcal{Y}$, $1 \leq j \leq n$, $j \neq k$ és $a^{(j)} \in \mathcal{A}^{(j)}$, $0 \leq j \leq m$ tetszőleges választása esetén. Ekkor $\forall \mathcal{X} \subseteq \mathcal{Y}$ esetén*

$$d(f(\mathcal{X})) \leq cd(\mathcal{X}),$$

valamely $c \geq 0$-ra.

5. Megjegyzés. *Többváltozós függvények esetén a következő becslés adható:*

$$d(f(\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, ..., \mathcal{X}^{(n)})) \leq \sum_{k=1}^{n} c^{(k)} d(\mathcal{X}^{(k)}) \leq c \max_{1 \leq k \leq n} d(\mathcal{X}^{(k)}).$$

15. Tétel. *Legyen $f : \mathbb{R} \to \mathbb{R}$ differenciálható függvény az $\mathcal{X} = [x_1, x_2]$ intervallumon. Továbbá legyen $f'(x)$ az f' egy kifejezése, ami az $\mathcal{X}$ intervallumon kiértékelhető. Tegyük fel, hogy az előző tétel összes feltétele teljesül f'-re. Ekkor $\forall y \in \mathcal{X}$ esetén $\exists c \geq 0$, hogy*

1. $W(f,\mathcal{X}) \subseteq f(y) + f'(\mathcal{X})(\mathcal{X}-y)$,

2. $q(W(f,\mathcal{X}), f(y) + f'(\mathcal{X})(\mathcal{X}-y)) \leq c(d(\mathcal{X}))^2$.

1.3. Intervallummátrixok

Az $m \times n$-es valós mátrixok halmazát a szokásos $\mathbb{R}^{m\times n}$, az egy oszlopból álló mátrixokat, azaz az oszlopvektorokat $\mathbb{R}^n$ jelöli. Jelölje $I(\mathbb{R})^{m\times n}$ az olyan $m \times n$-es mátrixok halmazát, melyek komponensei intervallumok, az intervallumvektorokat pedig $I(\mathbb{R})^n$.

12. Definíció. *$\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{m\times n}$ és $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{m\times n}$ egyenlők, azaz $\mathbf{A} = \mathbf{B}$ pontosan akkor, ha minden komponensük egyenlő, azaz $\mathcal{A}_{ij} = \mathcal{B}_{ij}$, $1 \leq i \leq m$,*
$1 \leq j \leq n$.

Definiálunk egy részbenrendezést $I(\mathbb{R})^{m\times n}$-en.

13. Definíció. *Legyen $\mathbf{A} = (\mathcal{A}_{ij})$ és $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{m\times n}$. Ekkor azt mondjuk, hogy $\mathbf{A} \subseteq \mathbf{B}$, ha $\mathcal{A}_{ij} \subseteq \mathcal{B}_{ij}$ $1 \leq i \leq m$, $1 \leq j \leq n$.*

6. Megjegyzés. *Ha $\mathbf{A}$ pontmátrix, azaz $\mathbf{A} \in \mathbb{R}^{m\times n}$, akkor az $\mathbf{A} \in \mathbf{B}$ jelölést használjuk.*

14. Definíció. *1. Ha $\mathbf{A} = (\mathcal{A}_{ij})$ és $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{m\times n}$, akkor*

$$\mathbf{A} \pm \mathbf{B} := (\mathcal{A}_{ij} \pm \mathcal{B}_{ij}).$$

2. Ha $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{m\times r}$ és $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{r\times n}$, akkor

$$\mathbf{AB} := \left(\sum_{k=1}^{r} \mathcal{A}_{ik}\mathcal{B}_{kj} \right).$$

Speciálisan, ha $\mathbf{u} = (\mathcal{U}_i) \in I(\mathbb{R})^n$, akkor

$$\mathbf{Au} = \left(\sum_{k=1}^{r} \mathcal{A}_{ik}\mathcal{U}_{k} \right).$$

3. Ha $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{m\times n}$ és $\mathcal{X} \in I(\mathbb{R})$, akkor

$$\mathcal{X}\mathbf{A} = \mathbf{A}\mathcal{X} := (\mathcal{X}\mathcal{A}_{ij})\,.$$

11. Állítás. *Legyen* $\mathbf{A} \in I(\mathbb{R})^{m\times r}$ *és* $\mathbf{B} \in I(\mathbb{R})^{r\times n}$. *Ekkor*

$$\{AB : A \in \mathbf{A}, B \in \mathbf{B}\} \subseteq \{C : C \in \mathbf{AB}\}.$$

Egyenlőség általában nem igazolható.

12. Állítás. *Legyen* $\mathbf{A}, \mathbf{B} \in I(\mathbb{R})^{m\times n}$ *és* $c \in \mathbb{R}^n$. *Ekkor*

1. $\{A + B : A \in \mathbf{A}, B \in \mathbf{B}\} = \mathbf{A} + \mathbf{B}$, *és*

2. $\{Ac : A \in \mathbf{A}\} = \mathbf{A}c$.

Tehát az intervallummátrixok halmaza zárt az előző definícióban bevezetett műveletekre.

16. Tétel. *Legyen* $\mathbf{A}, \mathbf{B}$ *és* $\mathbf{C}$ *intervallummátrix. Ekkor*

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$,

3. $\mathbf{A} + 0 = 0 + \mathbf{A} = \mathbf{A}$, *ahol* 0 *a megfelő méretű nullmátrix.*

4. $\mathbf{A}I = I\mathbf{A} = \mathbf{A}$, *ahol* I *a megfelelő méretű egységmátrix.*

5. $(\mathbf{A} + \mathbf{B})\mathbf{C} \subseteq \mathbf{AC} + \mathbf{BC}$ *és* $\mathbf{C}(\mathbf{A} + \mathbf{B}) \subseteq \mathbf{CA} + \mathbf{CB}$.

6. $(\mathbf{A} + \mathbf{B})C = \mathbf{A}C + \mathbf{B}C$ *és* $C(\mathbf{A} + \mathbf{B}) = C\mathbf{A} + C\mathbf{B}$, *ahol* $C \in \mathbb{R}^{k\times m}$.

7. $\mathbf{A}(BC) \subseteq (\mathbf{A}B)C$, *ahol* $B \in \mathbb{R}^{n\times k}$ *és* $C \in \mathbb{R}^{k\times l}$.

8. $(A\mathbf{B})\mathbf{C} \subseteq A(\mathbf{BC})$, *ha* $\mathbf{C} = -\mathbf{C}$, *és* $A \in \mathbb{R}^{k\times m}$.

9. $A(\mathbf{B}C) = (A\mathbf{B})C$, *ahol* $C \in \mathbb{R}^{n\times k}$.

10. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$, *ha* $\mathbf{B} = -\mathbf{B}$ *és* $\mathbf{C} = -\mathbf{C}$.

17. Tétel. *Legyenek* $\mathbf{A}^{(k)}, \mathbf{B}^{(k)}$, $k = 1, 2$ *intrevallummátrixok és* $\mathcal{X}, \mathcal{Y}$ *intervallumok. Továbbá tegyük fel, hogy* $\mathbf{A}^{(k)} \subseteq \mathbf{B}^{(k)}$, $k = 1, 2$ *és* $\mathcal{X} \subseteq \mathcal{Y}$. *Ekkor*

1. $\mathbf{A}^{(1)} * \mathbf{A}^{(2)} \subseteq \mathbf{B}^{(1)} * \mathbf{B}^{(2)}$, *ahol* $* = \{+, -, \cdot\}$, *és*

2. $\mathcal{X}\mathbf{A}^{(1)} \subseteq \mathcal{Y}\mathbf{B}^{(1)}$.

7. Megjegyzés. *Ha speciálisan* $A \in \mathbf{A}$, $B \in \mathbf{B}$ *és* $x \in \mathcal{X}$, *akkor*

1. $A * B \in \mathbf{A} * \mathbf{B}$, *ahol* $* = \{+, -, \cdot\}$, *és*

2. $xA \in \mathcal{X}\mathbf{A}$.

Az intervallumokhoz hasonlóan a következőkben definiáljuk az intervallummátrixok szélességét és abszolútértékét.

15. Definíció. *Legyen* $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{m\times n}$. *Ekkor*

$$d(\mathbf{A}) := (d(\mathcal{A}_{ij}))$$

az $\mathbf{A}$ *szélességmátrixa.*

16. Definíció. *Legyen* $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{m\times n}$. *Ekkor*

$$|\mathbf{A}| := (|\mathcal{A}_{ij}|)$$

az $\mathbf{A}$ *abszolútérték-mátrixa.*

17. Definíció. *Legyen* $X = (x_{ij}), Y = (y_{ij}) \in \mathbb{R}^{m\times n}$. *Ekkor azt mondjuk, hogy* $X \leq Y$, *ha* $x_{ij} \leq y_{ij}$ $\forall 1 \leq i \leq m$ *és* $1 \leq j \leq n$ *esetén.*

13. Állítás. *Legyen* $\mathbf{A}$ *és* $\mathbf{B}$ *intervallummátrix, ekkor a következők teljesülnek.*

1. *Ha* $\mathbf{A} \subseteq \mathbf{B}$, *akkor* $d(\mathbf{A}) \leq d(\mathbf{B})$.

2. $d(\mathbf{A} \pm \mathbf{B}) = d(\mathbf{A}) \pm d(\mathbf{B})$.

3. $d(\mathbf{A}) = \sup_{A,A' \in \mathbf{A}} |A - A'|$.

4. $\mathbf{A} \subseteq \mathbf{B}$ *esetén* $|\mathbf{A}| \leq |\mathbf{B}|$.

5. $|\mathbf{A}| = \sup_{A \in \mathbf{A}} |A|$.

6.
   - $|\mathbf{A}| \geq 0$ *és* $|\mathbf{A}| = 0 \Leftrightarrow \mathbf{A} = 0$,
   - $|\mathbf{A} + \mathbf{B}| \leq |\mathbf{A}| + |\mathbf{B}|$,
   - $|x\mathbf{A}| = |\mathbf{A}x| = |x||\mathbf{A}|$ $\forall x \in \mathbb{R}$ *és*
   - $|\mathbf{AB}| \leq |\mathbf{A}||\mathbf{B}|$.

7. $d(\mathbf{AB}) \leq d(\mathbf{A})|\mathbf{B}| + |\mathbf{A}|d(\mathbf{B})$.

8. $d(\mathbf{AB}) \geq |\mathbf{A}|d(\mathbf{B})$ *és* $d(\mathbf{AB}) \geq d(\mathbf{A})|\mathbf{B}|$.

9. • $d(a\mathbf{B}) = |a|d(\mathbf{B})$ $\forall a \in \mathbb{R}$ *esetén,*

   • $d(A\mathbf{B}) = |A|d(\mathbf{B})$, *ha A megfelelő méretű valós mátrix.*

   • $d(\mathbf{B}A) = d(\mathbf{B})|A|$, *ha A megfelelő méretű valós mátrix.*

10. *Ha a* 0 *a nullmátrixot jelöli, akkor* $0 \in \mathbf{A}$ *esetén* $|\mathbf{A}| \leq d(\mathbf{A}) \leq 2|\mathbf{A}|$.

11. *Ha* $\mathbf{A} = -\mathbf{A}$, *akkor* $\mathbf{AB} = \mathbf{A}|\mathbf{B}|$.

12. *Legyen* $\mathbf{B} = (\mathcal{B}_{ij})$ *és tegyük fel, hogy* $0 \in \mathbf{A}$ *és* $0 \notin \mathcal{B}_{ij}$. *Ekkor* $d(\mathbf{AB}) = d(\mathbf{A})|\mathbf{B}|$.

18. Definíció. *Legyen* $\mathbf{A} = (\mathcal{A}_{ij})$ *és* $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{m\times n}$. *Ekkor az* $\mathbf{A}$ *és* $\mathbf{B}$ *intervallummátrixok távolsága*

$$q(\mathbf{A}, \mathbf{B}) := (q(\mathcal{A}_{ij}, \mathcal{B}_{ij})).$$

14. Állítás. *Legyenek* $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ *intervallummátrixok. Ekkor*

1. $q(\mathbf{A}, \mathbf{B}) = 0 \Leftrightarrow \mathbf{A} = \mathbf{B}$,

2. $q(\mathbf{A}, \mathbf{B}) \leq q(\mathbf{A}, \mathbf{C}) + q(\mathbf{B}, \mathbf{C})$,

3. $q(\mathbf{A} + \mathbf{C}, \mathbf{B} + \mathbf{C}) = q(\mathbf{A}, \mathbf{B})$,

4. $q(\mathbf{A} + \mathbf{B}, \mathbf{C} + \mathbf{D}) = q(\mathbf{A}, \mathbf{C}) + q(\mathbf{B}, \mathbf{D})$,

5. $q(\mathbf{AB}, \mathbf{AC}) \leq |\mathbf{A}|q(\mathbf{B}, \mathbf{C})$.

A fent definiált távolságfogalommal és egy tetszőleges monoton mátrixnormával metrikát kapunk $I(\mathbb{R})^{m\times n}$-en. Mivel $I(\mathbb{R})^{m\times n}$ felfogható úgy, hogy $I(\mathbb{R}) \times I(\mathbb{R}) \times ... \times I(\mathbb{R})$ (nm db) és $I(\mathbb{R})$ teljes metrikus tér, ezért $I(\mathbb{R})^{m\times n}$ is az. A konvergencia a pontonkénti konvergencia, azaz

$$\lim_{k\to\infty} \mathbf{A}^{(k)} = \mathbf{A} \Leftrightarrow \lim_{k\to\infty} \mathcal{A}_{ij}^{(k)} = \mathcal{A}_{ij},$$

$1 \leq i \leq m$, $1 \leq j \leq n$.

3. Következmény. *Legyen* $\{\mathbf{A}^{(k)}\}_{k=0}^{\infty}$ *olyan intervallummátrix-sorozat, melyre* $\mathbf{A}^{(0)} \supseteq \mathbf{A}^{(1)} \supseteq \dots$. *Ekkor* $\{\mathbf{A}^{(k)}\}_{k=0}^{\infty}$ *konvergens, és*

$$\lim_{k\to\infty} \mathbf{A}^{(k)} = \mathbf{A} = (\mathcal{A}_{ij}),$$

ahol

$$\mathcal{A}_{ij} = \bigcap_{k=0}^{\infty} \mathcal{A}_{ij}^{(k)}.$$

4. Következmény. *Az* $I(\mathbb{R})^{m\times n}$*-en definiált műveletek* $(+,-,\cdot)$ *folytonosak.*

15. Állítás. *Legyen* $\mathbf{X} \subseteq \mathbf{Y} \in I(\mathbb{R})^{m\times n}$. *Ekkor*

$$\frac{1}{2}(d(\mathbf{Y}) - d(\mathbf{X})) \leq q(\mathbf{X},\mathbf{Y}) \leq d(\mathbf{Y}) - d(\mathbf{X}).$$

19. Definíció. *Legyen* $\mathbf{A},\mathbf{B} \in I(\mathbb{R})^{m\times n}$. *Ekkor*

$$\mathbf{A}\cap\mathbf{B} := \{C : C \in \mathbf{A}, C \in \mathbf{B}\},$$

azaz a halmazelméleti metszete a két mátrixnak.

16. Állítás. *Legyen* $\mathbf{A} = (\mathcal{A}_{ij})$ *és* $\mathbf{B} = (\mathcal{B}_{ij}) \in I(\mathbb{R})^{m\times n}$. *Ekkor* $\mathbf{A}\cap\mathbf{B}$ *pontosan akkor* $I(\mathbb{R})^{m\times n}$*-beli, ha nem üres. Ebben az esetben*

$$\mathbf{A}\cap\mathbf{B} = (\mathcal{A}_{ij}\cap\mathcal{B}_{ij}),$$

$1 \leq i \leq m$, $1 \leq j \leq n$.

5. Következmény. *(Tartalmazási monotonitás) Legyenek* $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ *intervallummátrixok. Továbbá tegyük fel, hogy* $\mathbf{A}\subseteq\mathbf{C}$ *és* $\mathbf{B}\subseteq\mathbf{D}$. *Ekkor*

$$\mathbf{A}\cap\mathbf{B} \subseteq \mathbf{C}\cap\mathbf{D}.$$

A következőkben olyan $\mathbf{A}x = \mathbf{b}$ lineáris egyenletrendszerekkel fogunk foglalkozni, melyek $\mathbf{A}$ mátrixa intervallummátrix és a jobb oldal intervallumvektor.

2. fejezet

Intervallum-együtthatós lineáris egyenletrenszerek megoldása

Ebben a fejezetben az intervallum-együtthatós lineáris egyenletrendszerek megoldhatóságának kérdését tárgyaljuk általános esetben a [2] cikk alapján. Legyen

$$\mathbf{A} = [\underline{A}, \overline{A}] \in I(\mathbb{R})^{m\times n}, \qquad \mathbf{b} = [\underline{b}, \overline{b}] \in I(\mathbb{R})^{m}.$$

20. Definíció. *Egy*

$$\mathbf{A}x = \mathbf{b}$$

intervallum-együtthatós lineáris egyenletrendszert megoldhatónak nevezünk, ha

$$Ax = b$$

megoldható minden $A \in \mathbf{A}$ és $b \in \mathbf{b}$ esetén.

A következő jelöléseket fogjuk használni a továbbiakban. Legyen

$$A_c := \frac{1}{2}(\underline{A} + \overline{A})$$

az $\mathbf{A}$ intervallummátrix centrummátrixa,

$$\Delta := \frac{1}{2}(\overline{A} - \underline{A})$$

a sugármátrix. Ekkor

$$\mathbf{A} = [A_c - \Delta, A_c + \Delta].$$

Ugyanígy a jobb oldali **b** vektor esetén

$$b_c := \frac{1}{2}(\underline{b} + \overline{b})$$

és

$$\delta := \frac{1}{2}(\overline{b} - \underline{b}),$$

és így

$$\mathbf{b} = [b_c - \delta, b_c + \delta].$$

Továbbá legyen

$$Y_m := \{y \in \mathbb{R}^m : y_j \in \{-1, 1\} \forall j\},$$

azaz Y_m tartalmazza az összes m-dimenziós ± 1 vektort. Y_m elemszáma 2^m. Végül $\forall y \in Y_m$ vektor esetén jelölje

$$T_y = diag(y_1, ..., y_m).$$

Már most megjegyezzük, hogy $\forall y \in Y_m$ esetén

$$A_c - T_y\Delta \in \mathbf{A}, \quad A_c + T_y\Delta \in \mathbf{A}, \quad b_c + T_y\delta \in \mathbf{b}.$$

Most kimondjuk azt a két állítást, amit a megoldhatóságról szóló tétel bizonyításánál használni fogunk. Az első a jól ismert Farkas-lemma.

1. Lemma. *(Farkas) Legyen $A \in \mathbb{R}^{m\times n}$ és $b \in \mathbb{R}^m$. Ekkor az*

$$Ax = b,$$

$$x \geq 0$$

rendszernek akkor és csak akkor létezik megoldása, ha $\forall p \in \mathbb{R}^m$ esetén, melyre

$$A^T p \geq 0,$$

$$b^T p \geq 0.$$

18. Tétel. *(Oettli-Prager [9]) Legyen*

$$X = \{x : |A_c x - b_c| \leq \Delta|x| + \delta\}.$$

Ekkor minden $x \in X$ esetén létezik $A \in \mathbf{A}$ és $b \in \mathbf{b}$, melyre $Ax = b$.

Attól az esettől eltekintve, amikor $\underline{A} = \overline{A}$ és $\underline{b} = \overline{b}$ az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrenszer végtelen sok lineáris egyenletrendszert tartalmaz. A most következő tétel, ami egyébként ennek a fejezetnek a legfontosabb állítása, azt mondja ki, hogy az $\mathbf{A}x = \mathbf{b}$ megoldása karakterizálható véges sok nemnegatív megoldással. Persze ezek száma általában exponenciális a mátrix méretében.

19. Tétel. *Az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrendszernek akkor és csak akkor megoldható, ha $\forall y \in Y_m$ esetén az*

$$\begin{aligned} (A_c - T_y\Delta)x^1 - (A_c + T_y\Delta)x^2 &= b_c + T_y\delta, \\ x^1 \geq 0, \quad x^2 &\geq 0, \end{aligned} \tag{2.1}$$

rendszernek létezik x_y^1, x_y^2 megoldása. Továbbá ebben az esetben $\forall A \in \mathbf{A}$, $b \in \mathbf{b}$ esetén az $Ax = b$ egyenletrendszernek létezik megoldása a

$$Conv\{x_y^1 - x_y^2 : y \in Y_m\}$$

halmazban.

Bizonyítás: Először nézzük a szükségességet. Tegyük fel, hogy az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrendszer megoldható, és indirekt tegyük fel, hogy (2.1) rendszernek nem létezik megoldása. Ekkor a Farkas-lemma szerint $\exists p \in \mathbb{R}^m$, melyre

$$(A_c - T_y\Delta)^T p \geq 0, \tag{2.2}$$

$$(A_c + T_y\Delta)^T p \leq 0, \tag{2.3}$$

$$(b_c + T_y\delta)^T p < 0. \tag{2.4}$$

Ekkor (2.2) és (2.3) szerint

$$\Delta^T T_y p \leq A_c^T p \leq -\Delta^T T_y p,$$

így

$$|A_c^T p| \leq -\Delta^T T_y p = |-\Delta^T T_y p| \leq \Delta^T |p|.$$

Mivel

$$p \in \{x : |A_c^T x| \leq \Delta^T |x|\},$$

ezért az Oettli-Prager-tételt az

$$[A_c^T - \Delta^T, A_c^T + \Delta^T]z = [0,0]$$

intervallum-együtthatós lineáris egyenletrendszerre alkalmazva azt kapjuk, hogy $\exists A \in \mathbf{A}$, melyre

$$A^T p = 0. \tag{2.5}$$

Tehát $\exists p \in \mathbb{R}^m$, melyre (2.4) és (2.5) teljesül. Ha erre alkalmazzuk a Farkas-lemmát, akkor azt kapjuk, hogy $\nexists x \in \mathbb{R}^n$, melyre

$$Ax = b_c + T_y \delta.$$

Ez ellentmond annak a feltételnek, miszerint az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrendszer megoldható, ugyanis $A \in \mathbf{A}$ és $b_c + T_y\delta \in \mathbf{b}$.
Most nézzük az elégségesség bizonyítását. Tegyük fel, hogy $\forall y \in Y_m$ esetén (2.1) rendszernek létezik megoldása: x_y^1, x_y^2. Legyen $A \in \mathbf{A}$ és $b \in \mathbf{b}$ tetszőleges. Azt kell megmutatni, hogy ekkor az $Ax = b$ lineáris egyenletrendszernek létezik megoldása. Ehhez először azt mutatjuk meg, hogy $\forall y \in Y_m$ esetén

$$T_y A x_y \geq T_y b, \tag{2.6}$$

ahol $x_y = x_y^1 - x_y^2$. Tehát legyen $y \in Y_m$ tetszőleges. Ekkor

$$T_y(Ax_y - b) = T_y(A_c x_y - b_c) + T_y(A - A_c)x_y + T_y(b_c - b).$$

Mivel

$$|T_y(A - A_c)x_y| \leq \Delta|x_y|,$$

ezért

$$T_y(A - A_c)x_y \geq -\Delta|x_y|$$

, és ugyanígy, mivel

$$|T_y(b_c - b)| \leq \delta,$$

ezért

$$T_y(b_c - b) \geq -\delta,$$

és így

$$T_y(Ax_y - b) \geq T_y(A_c x_y - b_c) - \Delta|x_y| - \delta =$$

$$= T_y(A_c(x_y^1 - x_y^2) - b_c) - \Delta|x_y^1 - x_y^2| - \delta \geq$$

$$\geq T_y(A_c(x_y^1 - x_y^2) - b_c) - \Delta(x_y^1 - x_y^2) - \delta.$$

Ha felbontjuk a zárójeleket és kiemeljük x_y^1-t és x_y^2-t, akkor azt kapjuk, hogy

$$T_y(Ax_y - b) \geq T_y((A_c - T_y\Delta)x_y^1 - (A_c + T_y\Delta)x_y^2 - (b_c + T_y\delta)).$$

Mivel x_y^1, x_y^2 megoldása a (2.1) renszernek, ezért

$$T_y(Ax_y - b) \geq T_y((A_c - T_y\Delta)x_y^1 - (A_c + T_y\Delta)x_y^2 - (b_c + T_y\delta)) = 0,$$

ami igazolja (2.6)-ot.

Ezt felhasználva megmutatjuk, hogy ha $\lambda_y \geq 0$ és $y \in Y_m$, akkor a

$$\begin{aligned} \sum_{y \in Y_m} \lambda_y A x_y &= b, \\ \sum_{y \in Y_m} \lambda_y &= 1 \end{aligned} \tag{2.7}$$

lineáris egyenletrendszernek létezik megoldása. A Farkas-lemma szerint elég azt megmutatni, hogy $\forall p \in \mathbb{R}^m$, $p_0 \in \mathbb{R}$ esetén ha

$$p^T A x_y + p_0 \geq 0 \qquad \forall y \in Y_m, \tag{2.8}$$

akkor

$$p^T b + p_0 \geq 0. \tag{2.9}$$

Tegyük fel tehát, hogy $p \in \mathbb{R}^m$ és $p_0 \in \mathbb{R}$ kielégíti (2.8)-at. Definiáljuk $y \in Y_m$-t a következő módon

$$y_i = \begin{cases} -1, & \text{ha } p_i \geq 0, \\ 1 & \text{különben}, \end{cases}$$

$(i = 1, 2, ..., m)$. Mivel $p = -T_y|p|$ és $T_y = T_y^T$, ezért

$$p^T b + p_0 = -|p|^T T_y b + p_0.$$

(2.6) miatt

$$p^T b + p_0 \geq -|p|^T T_y A x_y + p_0 = p^T A x_y + p_0.$$

Végül (2.8) miatt

$$p^T b + p_0 \geq p^T A x_y + p_0 \geq 0,$$

ami igazolja (2.9)-et. Így ha $\lambda_y \geq 0$ és $y \in Y_m$, akkor a (2.7) egyenletrendszernek létezik megoldása.

Legyen

$$x = \sum_{y \in Y_m} \lambda_y x_y,$$

ekkor (2.7) miatt $Ax = b$ és

$$x \in Conv\{x_y : y \in Y_m\} = Conv\{x_y^1 - x_y^2 : y \in Y_m\},$$

és ezzel a tétel bizonyítása teljes. $\square$

A következőkben megnézzük, hogy mit is mond valójában az imént belátott tétel. Ha $y_i = 1$, akkor az $A_c - T_y\Delta$ és az $A_c + T_y\Delta$ i-edik sora megegyezik $\underline{A}$ és $\overline{A}$ i-edik sorával, és $(b_c + T_y\delta)_i = \overline{b}_i$. Ez azt jelenti, hogy ebben az esetben (2.1) i-edik egyenlete a következő

$$(\underline{A}x^1 - \overline{A}x^2)_i = \overline{b}_i. \tag{2.10}$$

Ugyanígy, ha $y_i = -1$, akkor

$$(\overline{A}x^1 - \underline{A}x^2)_i = \underline{b}_i. \tag{2.11}$$

Tehát $\forall y \in Y_m$-re a (2.1) rendszerek családja megegyezik az olyan rendszerek családjával, ahol az i-edik egyenlet vagy a (2.10), vagy a (2.11) alakban van, $i = 1, ..., m$. A különböző ilyen rendszerek száma pontosan 2^q, ahol a q a (Δ, δ) mátrix nemnulla sorainak számát jelöli. Így a megoldandó rendszerek száma exponenciális, ezért az előző tétel a gyakorlatban csak akkor használható, ha q viszonylag kicsi.

Most megmutatjuk, hogy hogyan lehet konstruálni tetszőleges $A \in \mathbf{A}$ és $b \in \mathbf{b}$ esetén az $Ax = b$ azon megoldását, amelyik a $Conv\{x_y^1 - x_y^2 : y \in Y_m\}$ halmazban van. Ehhez az Y_m elemeinek egy speciális sorrendjére lesz szükség, amit indukcióval definiálunk a következőképpen.

1. Az Y_1 sorrendje legyen a következő: $-1, 1$.

2. Ha az Y_j sorrendje $y^{(1)}, y^{(2)}, ..., y^{(2^j)}$, akkor az Y_{j+1} sorrendje legyen

$$((y^{(1)})^T, -1)^T, ((y^{(2)})^T, -1)^T, ..., ((y^{(2^j)})^T, -1)^T,$$

$$((y^{(1)})^T, 1)^T, ((y^{(2)})^T, 1)^T, ..., ((y^{(2^j)})^T, 1)^T.$$

Továbbá egy $z^{(1)}, z^{(2)}, ..., z^{(2h)}$ páros elemszámú sorozatban a $z^{(j)}, z^{(j+h)}$ párokat konjugált pároknak nevezzük. Legyen minden $y \in Y_m$ esetén x_y^1, x_y^2 a (2.1) rendszer megoldása. Ekkor az algoritmus a következő.

1. Válasszunk egy tetszőleges $A \in \mathbf{A}$-t és $b \in \mathbf{b}$-t.

2. Az $((x_{-y}^1 - x_{-y}^2)^T, (A(x_{-y}^1 - x_{-y}^2) - b)^T)^T$ vektorokat tegyük a nekik megfelelő y-ok Y_m-beli sorrendjébe.

3. Az aktuális sorban minden x, x' konjugált párhoz legyen
$$\lambda = \begin{cases} \frac{x'_k}{x'_k - x_k}, & \text{ha } x'_k \neq x_k, \\ 1 & \text{különben,} \end{cases}$$
ahol k az aktuális utolsó komponens indexe. Legyen
$$x := \lambda x + (1 - \lambda)x'.$$

4. Töröljük a sorozat második felét, majd a megmaradó részben töröljük a vektorok utolsó koordinátáját.

5. Ha egyetlen x vektor maradt, akkor x megoldása $Ax = b$-nek és
$$x \in Conv\{x_y^1 - x_y^2 : y \in Y_m\}.$$
Ellenkező esetben menjünk vissza a 3. lépésre.

Az algoritmus 2^m db $n + m$ hosszú vektorral indul, és minden lépésben megfelezi a vektorok számát, illetve eggyel csökkenti a dimenzióját. Így a végére egyetlen $x \in \mathbb{R}^n$ vektor marad. Az 5. pontban tett kijelentést a [10] cikk második tétele indokolja.

A megoldhatóság ellenőrzését szolgáló rendszerek (2.1) száma általában exponenciális az $\mathbf{A}$ intervallummátrix sorában. Ez az eredmény valószínűleg lényegesen nem javítható a következő tétel miatt.

20. Tétel. *Az intervallum-együtthatós lineáris egyenletrendszerek megoldhatóságának ellenőrzése NP-nehéz feladat.*

Az állítás abból a tényből következik, hogy egy intervallummátrix regularitásának ellenőrzése NP-teljes a [11] cikk alapján. Ez nyilvánvalóan polinom időben visszavezethető az intervallum-együtthatós lineáris egyenletrendszerek megoldhatóságának kérdésére, ami így NP-nehéz.

3. fejezet

Gauss-elimináció

3.1. Gauss-elimináció algoritmusa intervallummátrixokra

Legyen $\mathbf{A} = (\mathcal{A}_{ij})$ intervallummátrix, $\mathbf{b} = (\mathcal{B}_i)$ intervallumvektor. Feltesszük, hogy A^{-1} létezik minden $A \in \mathbf{A}$ esetén. Keressük a

$$\Sigma = \{x : Ax = b, A \in \mathbf{A}, b \in \mathbf{b}\}$$

halmazt. Mivel ez a halmaz általában túl bonyolúlt, ezért ehelyett egy olyan intervallumvektort keresünk, ami ezt tartalmazza. A Gauss-eliminációt fogjuk alkalmazni az intervallum-együtthatós rendszerre. A kezdőtáblázatunk a következő:

$$\begin{matrix} \mathcal{A}_{11} & \cdots & \mathcal{A}_{1n} & \mathcal{B}_1 \\ \vdots & & \vdots & \vdots \\ \mathcal{A}_{n1} & \cdots & \mathcal{A}_{nn} & \mathcal{B}_n \end{matrix} .$$

Ha feltesszük, hogy $0 \notin \mathcal{A}_{11}$, akkor az első eliminációs lépés után a következő táblázatot kapjuk:

$$\begin{matrix} \mathcal{A}'_{11} & \mathcal{A}'_{12} & \cdots & \mathcal{A}'_{1n} & \mathcal{B}'_1 \\ 0 & \mathcal{A}'_{22} & \cdots & \mathcal{A}'_{2n} & \mathcal{B}'_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \mathcal{A}'_{n2} & \cdots & \mathcal{A}'_{nn} & \mathcal{B}'_n \end{matrix} ,$$

ahol az első sor ugyanaz, mint az előző táblázat első sora, és az i-edik sort úgy kapjuk, hogy az előző tábla i-edik sorából kivonjuk az első sor $\mathcal{A}_{i1}/\mathcal{A}_{11}$-szeresét $2 \leq i \leq n$, azaz

$$\begin{aligned}
\mathcal{A}'_{1j} &= \mathcal{A}_{1j} && 1 \leq j \leq n,\\
\mathcal{B}'_1 &= \mathcal{B}_1 &&\\
\mathcal{A}'_{ij} &= \mathcal{A}_{ij} - \mathcal{A}_{1j}(\mathcal{A}_{i1}/\mathcal{A}_{11}) && 2 \leq i, j \leq n,\\
\mathcal{B}'_i &= \mathcal{B}_i - \mathcal{B}_1(\mathcal{A}_{i1}/\mathcal{A}_{11}) && 2 \leq i \leq n,\\
\mathcal{A}'_{i1} &= 0 && 2 \leq i \leq n.
\end{aligned}$$

17. Állítás. *Az eredeti rendszer megoldáshalmaza része az új rendszer megoldáshalmazának, azaz*

$$\{x : Ax = b, A \in \mathbf{A}, b \in \mathbf{b}\} \subseteq \{y : A'y = b', A' \in \mathbf{A}', b' \in \mathbf{b}'\}.$$

Bizonyítás: Legyen $A = (a_{ij}) \in \mathbf{A}$ és $b = (b_i) \in \mathbf{b}$, és tekintsük az alábbi lineáris egyenletrenszert:

$$Ax = b.$$

Legyen $A' := (a'_{ij})$ és $b' := (b'_i)$, ahol

$$\begin{aligned}
a'_{1j} &= a_{1j} && 1 \leq j \leq n,\\
b'_1 &= b_1 &&\\
a'_{ij} &= a_{ij} - a_{1j}(a_{i1}/a_{11}) && 2 \leq i, j \leq n,\\
b'_i &= b_i - b_1(a_{i1}/a_{11}) && 2 \leq i \leq n,\\
a'_{i1} &= 0 && 2 \leq i \leq n.
\end{aligned}$$

Ismert, hogy az $A'y = b'$ lineáris egyenletrendszer megoldása ugyan az, mint az $Ax = b$ rendszeré. A tartalmazási monotonitás miatt $A' \in \mathbf{A}'$ és $b' \in \mathbf{b}'$, ami bizonyítja az állítást. □

Ha ezt a lépést $n-1$-szer elvégezzük, akkor az erdeti táblából egy felső háromszög alakút kapunk:

$$\begin{array}{ccccc}
\widetilde{\mathcal{A}}_{11} & \widetilde{\mathcal{A}}_{12} & \cdots & \widetilde{\mathcal{A}}_{1n} & \widetilde{\mathcal{B}}_1\\
 & \widetilde{\mathcal{A}}_{22} & \cdots & \widetilde{\mathcal{A}}_{2n} & \widetilde{\mathcal{B}}_2\\
 & & \ddots & \vdots & \vdots\\
 & & & \widetilde{\mathcal{A}}_{nn} & \widetilde{\mathcal{B}}_n
\end{array},$$

melyre igaz, hogy

$$\{x : Ax = b, A \in \mathbf{A}, b \in \mathbf{b}\} \subseteq \{\widetilde{x} : \widetilde{A}\widetilde{x} = \widetilde{b}, \widetilde{A} \in \widetilde{\mathbf{A}}, \widetilde{b} \in \widetilde{\mathbf{b}}\}.$$

Legyen

$$\mathcal{X}_n := \frac{\widetilde{\mathcal{B}}_n}{\widetilde{\mathcal{A}}_{nn}},$$

$$\mathcal{X}_i := \frac{\widetilde{\mathcal{B}}_i - \sum_{j=i+1}^{n} \widetilde{\mathcal{A}}_{ij}\mathcal{X}_j}{\widetilde{\mathcal{A}}_{ii}}, \qquad 1 \le i \le n-1.$$

Ekkor $\mathbf{x} := (\mathcal{X}_i)$ intervallumvektor esetén

$$\Sigma = \{x : Ax = b, A \in \mathbf{A}, b \in \mathbf{b}\} \subseteq \mathbf{x}.$$

A következőkben a Gauss-eliminációval kapott intervallumvektor néhány tulajdonságával foglalkozunk, majd megnézzük, hogy milyen feltételek mellett hajtható végre. Azt már most megjegyezzük, hogy ha speciálisan $A = (a_{ij})$ nemszinguláris pontmátrix, akkor a Gauss-elimináció minden jobb oldali intervallumvektor esetén végrehajtható. Esetleg szükség lehet sorcserékre az elimináció alatt, de ez a megoldást nem befolyásolja.

Legyen

$$g : \mathbb{R}^{n\times n} \times \mathbb{R}^n \to \mathbb{R}^n$$

olyan leképezés, ami egy nemszinguláris A mátrixhoz és egy tetszőleges b vektorhoz az $Ax = b$ lineáris egyenletrendszer Gauss-eliminációval kapott megoldását rendeli, azaz

$$x = g(A, b).$$

A g leképezés egyértelmű, de több kifejezése is lehet. Például $g(A, b) = A^{-1}b$. g kifejezése függ attól is, hogy a Gauss-elimináció során hogy választjuk a pivotelemeket. A következő állításban szereplő tulajdonságok függetlenek a pivotelemek választásától.

18. Állítás. *Legyen $g(\mathbf{A}, \mathbf{b})$ a fent definiált leképezés intervallumkiértékelése. Az $\mathbf{x} = g(\mathbf{A}, \mathbf{b})$ intervallumvektor a fent leírt módon, Gauss-eliminációval kiszámítható.*

1. *Legyen $\mathbf{A}, \mathbf{B} \in I(\mathbb{R})^{n\times n}$ és $\mathbf{a}, \mathbf{b} \in I(\mathbb{R})^n$. Tobábbá tegyük fel, hogy $\mathbf{A} \subseteq \mathbf{B}$ és $\mathbf{a} \subseteq \mathbf{b}$. Ekkor*

$$g(\mathbf{A}, \mathbf{a}) \subseteq g(\mathbf{B}, \mathbf{b}).$$

2. *Legyen $A \in \mathbb{R}^{n\times n}$ és $\mathbf{b} = \mathbf{u} + \mathbf{v} \in I(\mathbb{R})^n$. Ekkor*

$$g(A, \mathbf{b}) = g(A, \mathbf{u}) + g(A, \mathbf{v}).$$

3. *Legyen $A \in \mathbb{R}^{n\times n}$ és $\mathbf{b} \in I(\mathbb{R})^n$. Ekkor*

$$A^{-1}\mathbf{b} \subseteq g(A, \mathbf{b}).$$

4. *Legyen $A \in \mathbb{R}^{n\times n}$ és $\mathbf{a}, \mathbf{b} \in I(\mathbb{R})^n$. Továbbá tegyük fel, hogy létezik $\alpha \geq 0$, hogy $d(\mathbf{a}) \leq \alpha d(\mathbf{b})$. Ekkor*

$$d(g(\mathbf{A}, \mathbf{a})) \leq \alpha d(g(A, \mathbf{b})).$$

Bizonyítás:

1. A tartalmazási monotonitás miatt triviális.

2. Mivel $\mathbf{A}, \mathbf{B} \in I(\mathbb{R})^{n\times n}$ és tudjuk, hogy $a(\mathcal{B} + \mathcal{C}) = a\mathcal{B} + a\mathcal{C}$ $\forall a \in \mathbb{R}, \mathcal{B}, \mathcal{C} \in I(\mathbb{R})$, ezért ha ezt a Gauss-elimináció képleteibe beírjuk, akkor megkapjuk az állítást.

3. Ismeretes, hogy ha f_1 és f_2 az f függvény két kifejezése, melyekre f_1-ben a változó pontosan egyszer fordul elő, míg f_2-ben m-szer, akkor $f_1(\mathcal{X}) \subseteq f_2(\mathcal{X})$. Ez igaz többváltozós függvényekre is. Tekintsük az i-edik ($1 \leq i \leq n$) komponensét $A^{-1}\mathbf{b}$-nek és $g(A, \mathbf{b})$-nek. A Gauss-elimináció képleteiben a $\mathbf{b}$ intervallumvektor komponensei többször is előfordulnak, míg $A^{-1}\mathbf{b}$ i-edik komponensének kiszámítása során csak egyszer.

4. Ismeretes, hogy $d(\mathcal{A} \pm \mathcal{B}) = d(\mathcal{A}) \pm d(\mathcal{B})$ és $d(a\mathcal{B}) = |a|d(\mathcal{B})$ minden $a \in \mathbb{R}$ és $\mathcal{A}, \mathcal{B} \in I(\mathbb{R})$ esetén. Valamint feltettük, hogy létezik $\alpha \geq 0$, amelyre $d(\mathbf{a}) \leq \alpha d(\mathbf{b})$. Ezeket a Gauss-elimináció algoritmusában használva rögtön megkapjuk az állítást. $\square$

3.2. Gauss-elimináció elvégezhetősége

Most térjünk rá a Gauss-elimináció elvégezhetőségének kérdésére. A következő tétel az 1 illetve a 2-dimenziós esetről szól.

21. Tétel. *Legyen $1 \leq n \leq 2$, és tegyük fel, hogy $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{n\times n}$ nem tartalmaz szinguláris mátrixot. Ekkor a Gauss-elimináció algoritmusa elvégezhető.*

Bizonyítás:

1. $n = 1$ eset: Ebben az esetben $\mathbf{A} = \mathcal{A}_{11}$ és a tétel feltétele ekvivalens azzal, hogy $0 \notin \mathcal{A}_{11}$, ami bizonyítja az állítást.

2. $n = 2$ eset: Az egyenletrendszerünk a következő:

$$\begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{pmatrix} \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix} = \begin{pmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{pmatrix}.$$

   Ekkor $\mathcal{A}_{11}$ és $\mathcal{A}_{21}$ közül legalább az egyik nem tartalmazza a 0-t, mert ellnkező esetben létezne $A \in \mathbf{A}$, ami szinguláris. Esetleges sorcserével elérhetjük, hogy $0 \notin \mathcal{A}_{11}$. A Gauss-elimináció szerint

$$\mathcal{A}'_{22} = \mathcal{A}_{22} - (1/\mathcal{A}_{11})\mathcal{A}_{21}\mathcal{A}_{12}.$$

   Tekinthetjük $\mathcal{A}'_{22}$-t egy f függvény intervallumaritmetikai kiértékelésének, ahol az f változói a_{11}, a_{12}, a_{21} és a_{22},

$$f(a_{11}, a_{12}, a_{21}, a_{22}) = a_{22} - (1/a_{11})a_{21}a_{12}.$$

   Mivel feltettük, hogy minden $A \in \mathbf{A}$-ra

$$\det(A) = a_{11}a_{22} - a_{21}a_{12} \neq 0,$$

   ezért

$$f(a_{11}, a_{12}, a_{21}, a_{22}) = (1/a_{11})\det(A) \neq 0.$$

   Az intervallumkiértékelés a pontos értéket adja, ha a_{11}-et $\mathcal{A}_{11}$-gyel, a_{12}-t $\mathcal{A}_{12}$-vel, a_{21}-et $\mathcal{A}_{21}$-gyel és a_{22}-t $\mathcal{A}_{22}$-vel helyettesítjük, mivel minden változó pontosan egyszer fordul elő a kifejezésben. Tehát $0 \notin \mathcal{A}'_{22}$, ami azt jelenti, hogy a Gauss-elimináció elvégezhető. $\square$

A fenti bizonyítás $n \geq 3$ esetre nem általánosítható. A fejezet további részében szeretnénk megkapni az intervallummátrixok egy olyan osztályát, amelyre a Gauss-elimináció esetleges sorcserékkel mindig elvégezhető. Mostantól az intervallumokat nem a kezdő és végpontjukkal adjuk meg, hanem a középpontjával és a sugarával, vagy más néven a félszélességével. Azaz $\mathcal{A} = [a_1, a_2]$ a következő alakban is felírható:

$$\mathcal{A} = [a - r, a + r] =: \langle a, r \rangle,$$

ahol

$$a = \frac{1}{2}(a_1 + a_2), \qquad r = \frac{1}{2}d(\mathcal{A}) = \frac{1}{2}(a_2 - a_1).$$

Könnyen igazolható, hogy ha $\mathcal{A} = \langle a, r\rangle, \mathcal{B} = \langle b, s\rangle \in I(\mathbb{R})$, akkor

$$\mathcal{A} \pm \mathcal{B} = \langle a \pm b, r + s\rangle.$$

A szorzás esetében csak a következő egyenlőségre lesz szükségünk:

$$[-r, r][-s, s] = \langle 0, r\rangle\langle 0, s\rangle = \langle 0, rs\rangle.$$

Tegyük fel, hogy $0 \notin \mathcal{A} = \langle a, r\rangle$. Mivel

$$\frac{1}{\mathcal{A}} = \left[\frac{1}{a+r}, \frac{1}{a-r}\right] = \left[\frac{a}{a^2-r^2} - \frac{r}{a^2-r^2}, \frac{a}{a^2-r^2} + \frac{r}{a^2-r^2}\right],$$

ezért

$$\frac{1}{\mathcal{A}} = \left\langle \frac{a}{a^2-r^2}, \frac{r}{a^2-r^2} \right\rangle.$$

Az $\mathcal{A}$ abszolútértékét a következőképpen számolhatjuk:

$$|\mathcal{A}| = \max\{a_1, a_2\} = |a| + r.$$

Továbbá az is igaz, hogy

$$0 \notin \mathcal{A} \Leftrightarrow |a| - r > 0.$$

És végül

$$\mathcal{A} = \langle a, r\rangle \subseteq \langle 0, |\mathcal{A}|\rangle = \langle 0, |a| + r\rangle.$$

2. Lemma. *Legyenek* $\mathcal{A} = \langle a, r_1\rangle, \mathcal{B} = \langle b, r_2\rangle, \mathcal{C} = \langle c, r_3\rangle$ *és* $\mathcal{D} = \langle d, r_4\rangle$ *valós intervallumok. Továbbá tegyük fel, hogy* $0 \notin \mathcal{D}$. *Ekkor*

$$\mathcal{Z} = \langle z, r_5\rangle = \mathcal{A} - \frac{1}{\mathcal{D}}\mathcal{B}\mathcal{C}$$

esetén

$$|a| - r_1 - \frac{|\mathcal{B}||\mathcal{C}|}{|d| - r_4} \leq |z| - r_5.$$

Bizonyítás: A tartalmazási monotonitás miatt

$$\mathcal{Z} = \langle z, r_5 \rangle = \mathcal{A} - \mathcal{B}\mathcal{C}\frac{1}{\mathcal{D}} \subseteq \mathcal{A} - \langle 0, |\mathcal{B}| \rangle \langle 0, |\mathcal{C}| \rangle \left\langle \frac{d}{d^2 - r_4^2}, \frac{r_4}{d^2 - r_4^2} \right\rangle =$$

$$= \langle a, r_1 \rangle - \left\langle 0, |\mathcal{B}||\mathcal{C}|\frac{|d|}{d^2 - r_4^2} + |\mathcal{B}||\mathcal{C}|\frac{r_4}{d^2 - r_4^2} \right\rangle =$$

$$= \langle a, r_1 \rangle - \left\langle 0, |\mathcal{B}||\mathcal{C}|\frac{1}{|d| - r_4} \right\rangle =$$

$$= \left\langle a, r_1 + |\mathcal{B}||\mathcal{C}|\frac{1}{|d| - r_4} \right\rangle =: \langle a, r_6 \rangle.$$

Mivel $\mathcal{Z} \subseteq \langle a, r_6 \rangle$, ezért

$$|a| - |z| \leq |a - z| \leq r_6 - r_5.$$

ezt átrendezve

$$|z| - r_5 \geq |a| - r_6 = |a| - r_1 - |\mathcal{B}||\mathcal{C}|\frac{1}{|d| - r_4},$$

és ez volt az állítás. □

21. Definíció. *Legyen $B = (b_{ij}) \in \mathbb{R}^{n \times n}$. Ekkor B M-mátrix, ha*

1. $b_{ij} \leq 0$, *ha* $i \neq j$ *és*

2. $B^{-1} \geq 0$.

Ismeretes, hogy a definíció második feltétele ekvivalens azzal, hogy $\exists u = (u_i) \in \mathbb{R}^n$, melyre $u_i > 0$, $1 \leq i \leq n$ és

$$Bu > 0.$$

Továbbá azt is tudjuk, hogy egy M-mátrix diagonális elemei mindig pozitívak.

22. Tétel. *Legyen $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{n \times n}$ és $\mathcal{A}_{ij} = \langle a_{ij}, r_{ij} \rangle$, $1 \leq i, j \leq n$. Továbbá legyen $B = (b_{ij}) \in \mathbb{R}^{n \times n}$, melyre*

$$b_{ij} := \begin{cases} |a_{ii}| - r_{ii}, & \text{ha } i = j \\ -|\mathcal{A}_{ij}| & \text{különben.} \end{cases}$$

Ha B M-mátrix, akkor a Gauss-elimináció elvégezhető $\mathbf{A}$ intervallummátrixra sor- és oszlopcserék nélkül.

Bizonyítás: Mivel B M-mártix, ezért $\exists u = (u_i) \in \mathbb{R}^n$, melyre $u_i > 0$, $1 \leq i \leq n$ és $Bu > 0$. Ez azt jelenti, hogy

$$(|a_{ii}| - r_{ii})u_i > \sum_{j=1, j\neq i}^{n} |\mathcal{A}_{ij}|u_j,$$

$1 \leq i \leq n$. Mivel a bal oldal nemnegatív és $u_i > 0$, ezért $i = 1$-re $|a_{11}| - r_{11} > 0$, amiből az következik, hogy $0 \notin \mathcal{A}_{11}$. Tehát a Gauss-elimináció első lépését el lehet végezni, és így megkapjuk az $\mathbf{A}' = (\mathcal{A}'_{ij})$ intervallummátrixot. Ha megmutatjuk, hogy a tétel feltételei fennállnak az $\widetilde{\mathbf{A}}' = (\widetilde{\mathcal{A}}'_{ij}) \in I(\mathbb{R})^{(n-1)\times(n-1)}$-re, melyre

$$\widetilde{\mathcal{A}}'_{ij} = \mathcal{A}'_{ij} = \langle a'_{ij}, r'_{ij}\rangle, \qquad 2 \leq i, j \leq n,$$

akkor teljes indukcióval kész vagyunk.
Legyen $i \geq 2$, ekkor

$$\sum_{j=2, j\neq i}^{n} |\mathcal{A}'_{ij}|u_j = \sum_{j=2, j\neq i}^{n} \left|\mathcal{A}_{ij} - \mathcal{A}_{1j}\frac{\mathcal{A}_{i1}}{\mathcal{A}_{11}}\right| u_j \leq$$

$$\leq \sum_{j=2, j\neq i}^{n} |\mathcal{A}_{ij}|u_j + |\mathcal{A}_{i1}| \left|\frac{1}{\mathcal{A}_{11}}\right| \sum_{j=2, j\neq i}^{n} |\mathcal{A}_{1j}|u_j.$$

Tekintsük ismét a bizonyítás elején szereplő egyenlőtlenséget $i = 1$-re és a szumma i-edik tagját vigyük át a másik oldalra. Ekkor

$$\sum_{j=2, j\neq i}^{n} |\mathcal{A}_{1j}|u_j < (|a_{11}| - r_{11})u_1 - |\mathcal{A}_{1i}|u_i.$$

Továbbá

$$\left|\frac{1}{\mathcal{A}_{11}}\right| = \left|\left\langle \frac{a_{11}}{a_{11}^2 - r_{11}^2}, \frac{r_{11}}{a_{11}^2 - r_{11}^2}\right\rangle\right| = \frac{|a_{11}|}{a_{11}^2 - r_{11}^2} + \frac{r_{11}}{a_{11}^2 - r_{11}^2} = \frac{1}{|a_{11}| - r_{11}}.$$

Ezeket felhasználva

$$\sum_{j=2, j\neq i}^{n} |\mathcal{A}'_{ij}|u_j \leq \sum_{j=2, j\neq i}^{n} |\mathcal{A}_{ij}|u_j + |\mathcal{A}_{i1}|\frac{1}{|a_{11}| - r_{11}}((|a_{11}| - r_{11})u_1 - |\mathcal{A}_{1i}|u_i).$$

Ha a zárójelet felbontjuk, és az első tagját egyszerűsítjük $|a_{11}| - r_{11}$-gyel, akkor azt be tudjuk vinni a szummába, és az alábbi becslést kapjuk

$$\sum_{j=2, j\neq i}^{n} |\mathcal{A}'_{ij}|u_j \leq \sum_{j=1, j\neq i}^{n} |\mathcal{A}_{ij}|u_j - \frac{|\mathcal{A}_{i1}||\mathcal{A}_{1i}|}{|a_{11}| - r_{11}}u_i.$$

Erre megint alkalmazhatjuk az első egyenlőtlenséget, ekkor

$$\sum_{j=2,j\neq i}^{n} |\mathcal{A}'_{ij}|u_j < u_i \left(|a_{ii}| - r_{ii} - \frac{|\mathcal{A}_{i1}||\mathcal{A}_{1i}|}{|a_{11}| - r_{11}} \right).$$

Végül ha az előző lemmát az $\mathcal{A} = \mathcal{A}_{ii}$, $\mathcal{B} = \mathcal{A}_{i1}$, $\mathcal{C} = \mathcal{A}_{1i}$ és $\mathcal{D} = \mathcal{A}_{11}$ intervallumokra alkalmazzuk, akkor

$$\mathcal{Z} = \mathcal{A} - \frac{1}{\mathcal{D}}\mathcal{B}\mathcal{C} = \mathcal{A}_{ii} - \frac{1}{\mathcal{A}_{11}}\mathcal{A}_{i1}\mathcal{A}_{1i} = \mathcal{A}'_{ii},$$

és így

$$|a_{ii}| - r_{ii} - \frac{|\mathcal{A}_{i1}||\mathcal{A}_{1i}|}{|a_{11}| - r_{11}} \leq |a'_{ii}| - r'_{ii}.$$

Ezzel tovább tudunk becsülni, és a következőre jutunk:

$$\sum_{j=1,j\neq i}^{n} |\mathcal{A}'_{ij}|u_j < (|a'_{ii}| - r'_{ii})u_i,$$

és ezt kellett belátnunk. □

Az intervallummátrixok egy igen fontos osztálya teljesíti az előző tétel feltételeit.

22. Definíció. *Legyen* $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{n\times n}$ *és* $\mathcal{A}_{ij} = \langle a_{ij}, r_{ij}\rangle$, $1 \leq i,j \leq n$. *Az* $\mathbf{A}$ *intervallummátrix szigorúan diagonálisan domináns, ha*

$$|a_{ii}| - r_{ii} > \sum_{j=1,j\neq i}^{n} |\mathcal{A}_{ij}|, \qquad 1 \leq i \leq n.$$

A definícióból rögtön következik, hogy egy szigorúan diagonálisan domináns $\mathbf{A}$ intervallummátrix diagonális elemei nem tartalmazhatják a 0-t. Továbbá az is látszik, hogy minden valós $\widehat{A} = (\widehat{a}_{ij}) \in \mathbf{A}$ mátrix esetén

$$|\widehat{a}_{ii}| > \sum_{j=1,j\neq i}^{n} |\widehat{a}_{ij}|, \qquad 1 \leq i \leq n.$$

Azaz minden valós $\widehat{A} \in \mathbf{A}$ mátrix szigorúan diagonálisan domináns a hagyományos értelemben, ezáltal nemszinguláris.
Egy szigorúan diagonálisan domináns $\mathbf{A}$ intervallummátrix teljesíti az előző tétel feltételét is, $u = (u_i)$, $u_i = 1$, $1 \leq i \leq n$ választással. Tehát kimondhatjuk a következő következményt.

6. Következmény. *Legyen* $\mathbf{A}$ *szigorúan diagonálisan domináns intervallummátrix. Ekkor a Gauss-elimináció elvégezhető az* $\mathbf{A}$ *intervallummátrixra sor- és oszlpocserék nélkül.*

3.3. Gauss-elimináció tridiagonális intervallummátrixokra

Legyen

$$\mathbf{A} = \begin{pmatrix} \mathcal{A}_1 & \mathcal{C}_1 & & & \\ \mathcal{B}_2 & \mathcal{A}_2 & \mathcal{C}_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & \ddots & \ddots & \mathcal{C}_{n-1} \\ & & & \mathcal{B}_n & \mathcal{A}_n \end{pmatrix}.$$

23. Tétel. *Legyen az* $\mathbf{A}$ *intervallummátrix tridiagonális, és tegyük fel, hogy*

$$\begin{aligned} \mathcal{A}_i &= \langle a_i, r_i \rangle, & 1 \le i \le n, \\ \mathcal{B}_i &= \langle b_i, s_i \rangle \neq 0, & 2 \le i \le n, \\ \mathcal{C}_i &= \langle c_i, t_i \rangle \neq 0, & 1 \le i \le n-1. \end{aligned}$$

Továbbá tegyük fel, hogy

$$\begin{aligned} |a_1| - r_1 &> |\mathcal{C}_1|, \\ |a_i| - r_i &\ge |\mathcal{B}_i| + |\mathcal{C}_i|, \quad 2 \le i \le n-1, \ . \\ |a_n| - r_n &> |\mathcal{B}_n|. \end{aligned}$$

Ekkor a Gauss-elimináció elvégezhető $\mathbf{A}$ *intervallummátrixra sor- és oszlpocserék nélkül.*

Bizonyítás: Írjuk fel az előző tételbeli B mátrixot ebben az esetben.

$$B = \begin{pmatrix} |a_1| - r_1 & -|\mathcal{C}_1| & & & \\ -|\mathcal{B}_2| & |a_2| - r_2 & -|\mathcal{C}_2| & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & \ddots & \ddots & -|\mathcal{C}_{n-1}| \\ & & & -|\mathcal{B}_n| & |a_n| - r_n \end{pmatrix}.$$

Tehát B olyan diagonálisan domináns tridiagonális valós mátrix, melyre teljesül, hogy az első és az utolsó sorban szigorú egyenlőtlenség van, azaz M-mátrix és az előző tétel alkalmazható $\mathbf{A}$-ra. □

3.4. Gauss-elimináció nem szigorúan diagonálisan domináns mátrixokra

Ebben a fejezetben megnézzük, hogy mit lehet tenni abban az esetben, ha a lineáris egyenletrendszer $\mathbf{A}$ mátrixa nem szigorúan diagonálisan domináns. Az ötlet az, hogy alkalmazunk egy olyan transzformációt a rendszerre, ami szigorúan diagonálisan dominánssá transzformálja az $\mathbf{A}$ mátrixot.

Legyen $\mathbf{A} = (\mathcal{A}_{ij}) \in I(\mathbb{R})^{n\times n}$ és $\mathcal{A}_{ij} = \langle a_{ij}, r_{ij}\rangle$, $1 \le i, j \le n$. Tegyük fel továbbá, hogy minden $A \in \mathbf{A}$ valós mátrix esetén létezik A^{-1}. Legyen $A_c := (a_{ij}) \in \mathbb{R}^{n\times n}$. Ez invertálható, hiszen $A_c \in \mathbf{A}$. Szorozzuk be az egyenlet mindkét oldalát A_c^{-1}-zel, ekkor az

$$\widetilde{\mathbf{A}} := A_c^{-1}\mathbf{A}$$

és

$$\widetilde{\mathbf{b}} := A_c^{-1}\mathbf{b}$$

jelöléseket használva az új egyenletrendszerünk

$$\widetilde{\mathbf{A}}x = \widetilde{\mathbf{b}}.$$

Ekkor

$$\{x : Ax = b, A \in \mathbf{A}, b \in \mathbf{b}\} \subseteq \{y : \widetilde{A}y = \widetilde{b}, \widetilde{A} \in \widetilde{\mathbf{A}}, \widetilde{b} \in \widetilde{\mathbf{B}}\}.$$

Ugyanis legyen az x egy eleme a baloldali halmaznak, azaz létezik $A \in \mathbf{A}$ és $b \in \mathbf{b}$, hogy

$$Ax = b.$$

Ekkor

$$A_c^{-1}Ax = A_c^{-1}b,$$

és mivel

$$A_c^{-1}A \in \widetilde{\mathbf{A}}, \qquad A_c^{-1}b \in \widetilde{\mathbf{b}},$$

az állítást beláttuk.

Ha az $\mathbf{A}$ mátrix elemei nem túl szélesek, akkor az $\widetilde{\mathbf{A}}$ erősen diagonálisan domináns és a Gauss-elimináció elvégezhető. Ha ugyanis $d(\mathbf{A}) = 0$, akkor $\widetilde{\mathbf{A}} = I$ és ekkor $\widetilde{\mathbf{A}}$ persze erősen diagonálisan domináns. Ha az $\mathbf{A}$ elemeinek szélessége nem túl nagy, akkor $\widetilde{\mathbf{A}}$ nem sokban fog eltérni az egységmátrixtól.

Azonban az $\widetilde{\mathbf{A}}$ intervallummátrix erősen diagonális dominanciája nem csak az $\mathbf{A}$ mátrix elemeinek szélességétől függ. Legyen

$$\mathcal{A}_{ij} = \langle a_{ij}, r_{ij} \rangle, \quad 1 \leq i, j \leq n,$$

$$\mathbf{D} = (\mathcal{D}_{ij}), \quad \mathcal{D}_{ij} = \langle 0, r_{ij} \rangle, \quad 1 \leq i, j \leq n.$$

Ekkor

$$\widetilde{\mathbf{A}} = A_c^{-1}\mathbf{A} = A_c^{-1}(A_c + \mathbf{D}) =$$

$$= I + A_c^{-1}\mathbf{D} = I + \mathbf{H},$$

ahol $\mathbf{H} = A_c^{-1}\mathbf{D}$. Mivel

$$\|\mathbf{H}\| \leq \|A_c^{-1}\| \cdot \|\mathbf{D}\| = \frac{1}{2}\|A_c^{-1}\| \cdot \|d(\mathbf{A})\| =$$

$$= \frac{1}{2}\|A_c^{-1}\| \cdot \|A_c\| \cdot \frac{\|d(\mathbf{A})\|}{\|A_c\|} = \frac{1}{2} cond(A_c) \frac{\|d(\mathbf{A})\|}{\|A_c\|},$$

ezért az $\widetilde{\mathbf{A}}$ annál inkább diagonálisan domináns, minél kisebb az A_c kondíciószáma.

4. fejezet

Intervallum-együtthatós lineáris egyenletrendszerek megoldáshalmazának behatárolása reguláris esetben

Mint azt az előző fejezetben láttuk, a Gauss-eliminációt olyan intervallummátrixok esetén lehet jól használni, melyekben az elemek viszonylag keskenyek. Ebben a fejezetben két olyan eljárást ismertetünk, ami abban az esetben hatékony, amikor ezek az intervallumok viszonylag szélesek. Viszont a hátrányuk az, hogy több számolással járnak, mint a Gauss-elimináció. Először E. R. Hansen eredményét közöljük a [3] cikk alapján. Itt a bizonyításokra nem térünk ki, mivel a második módszer, melyet J. Rohn közölt a [4] cikkben, lényegében ugyanarra az eredményre jut, mint a Hansen-féle, de $2n$ db lineáris egyenletrendszer megoldása helyett csak egy mátrix invertálása szükséges.

4.1. E. R. Hansen eredménye

Legyen $\mathbf{A} \in I(\mathbb{R})^{n\times n}$ és $\mathbf{b} \in I(\mathbb{R})^n$.

23. Definíció. *Egy intervallumot/intervallumvektort/intervallummátrixot centráltnak nevezünk, ha a centruma a* 0 *szám/vektor/mátrix. Egy intervallummátrixot az identitás körül centráltnak nevezünk, ha a centruma az identitásmátrix.*

Tegyük fel, hogy $\forall A \in \mathbf{A}$ reguláris. Ekkor az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrendszer megoldáshalmaza a következőképpen adható meg:

$$\Sigma = \{x = A^{-1}b : A \in \mathbf{A}, b \in \mathbf{b}\}.$$

A pontos megoldáshalmaz helyett most is a legszűkebb olyan intervallumvektort keressük, ami azt tartalmazza.

Ha elvégezzük az intervallum-együtthatós lineáris egyenletrendszeren az előző fejezetben ismertetett transzformációt, akkor — mint azt láttuk — ha $\mathbf{A}$ és $\mathbf{b}$ elemei viszonylag szűkek, akkor csak kis mértékben növeli a megoldáshalmazt, ha viszont szélesek, akkor nagyon megnövelheti azt. Enélkül viszont az intervallumok szélessége általában nagyon gyorsan nő a megoldás során és a végső eredmény kevéssé használható lesz. Így az eredeti intervallum-együtthatós lineáris egyenletrendszer helyett tekinsük az

$$\widetilde{\mathbf{A}}x = \widetilde{\mathbf{b}}$$

intervallum-együtthatós lineáris egyenletrendszert, ahol

$$\widetilde{\mathbf{A}} := A_c^{-1}\mathbf{A}$$

és

$$\widetilde{\mathbf{b}} := A_c^{-1}\mathbf{b}.$$

Láttuk, hogy $\widetilde{\mathbf{A}}$ az identitás körül centrált, így

$$\widetilde{\mathbf{A}} = [I - \Delta, I + \Delta], \qquad \widetilde{\mathbf{b}} = [b_c - \delta, b_c + \delta].$$

24. Tétel. *Tegyük fel, hogy $\forall A \in \widetilde{\mathbf{A}}$ mátrix szigorúan diagonálisan domináns. (Ekkor $\widetilde{\mathbf{A}}$ nem tartalmaz szinguláris mátrixot.) Ekkor az alábbiak teljesülnek a megoldáshalmazt tartalmazó $\mathbf{x}$ intervallumvektorra.*

1. *$\mathbf{x}_i$ maximális értéke, ha az nemnegatív:*

$$\overline{\mathbf{x}}_i = e_i^T (I - \Delta)^{-1} s^{(i)},$$

ahol

$$s_j^{(i)} = \begin{cases} (b_c + \delta)_i, & \text{ha } j = i \\ \max\{-(b_c - \delta)_j, (b_c + \delta)_j\}, & \text{ha } j \neq i. \end{cases}$$

2. $\mathbf{x}_i$ *minimális értéke, ha az nemnegatív:*

$$\underline{\mathbf{x}}_i = \frac{1}{2((I-\Delta)^{-1})_{ii}-1} e_i^T (I-\Delta)^{-1} t^{(i)},$$

ahol

$$t_j^{(i)} = \begin{cases} (b_c - \delta)_i, & \text{ha } j = i \\ \min\{(b_c-\delta)_j, -(b_c+\delta)_j\}, & \text{ha } j \neq i. \end{cases}$$

3. $\mathbf{x}_i$ *maximális értéke, ha az negatív:*

$$\overline{\mathbf{x}}_i = \frac{1}{2((I-\Delta)^{-1})_{ii}-1} e_i^T (I-\Delta)^{-1} s^{(i)}.$$

4. $\mathbf{x}_i$ *minimális értéke, ha az negatív:*

$$\underline{\mathbf{x}}_i = e_i^T (I-\Delta)^{-1} t^{(i)}.$$

Megjegyezzük, hogy $\mathbf{x}_i$ maximális értéke csak úgy lehet negatív, ha $(b_c+\delta)_i < 0$, ugyanis az $\widetilde{\mathbf{A}}x = \widetilde{\mathbf{b}}$ intervallum-együtthatós lineáris egyenletrendszer megoldáshalmaza tartalmazza a $\widetilde{\mathbf{b}}$ intervallumvektort, mivel $I \in \widetilde{\mathbf{A}}$. Ezért ha $(b_c+\delta)_i \geq 0$, akkor $\overline{\mathbf{x}}_i \geq 0$.

Azt is megjegyezzük, hogy $s^{(i)}$ és $t^{(i)}$ kiszámítható elágazás nélkül, ugyanis

$$\max\{-(b_c-\delta)_j, (b_c+\delta)_j\} = (b_c)_i + \delta_i,$$

és

$$\min\{(b_c-\delta)_j, -(b_c+\delta)_j\} = -\max\{-(b_c-\delta)_j, (b_c+\delta)_j\}.$$

4.2. J. Rohn eredménye

Ugyanazokat a jelöléseket használjuk, mint az előző részben. Az előző tételben a szigorúan diagonális dominancia volt a regularitás elégséges feltétele. A [12] cikk alapján az $[I-\Delta, I+\Delta]$ intervallummátrix akkor és csak akkor reguláris, ha

$$\varrho(\Delta) < 1,$$

ahol $\varrho(\Delta)$ a Δ spektrálsugara. Ebből az következik, hogy az

$$M = (I-\Delta)^{-1} = (m_{ij})$$

mátrix létezik és nemnegatív. Legyen

$$\underline{x}_i := \min_{x\in X} x_i,$$

$$\overline{x}_i := \max_{x\in X} x_i,$$

ahol X az

$$[I-\Delta, I+\Delta]x = [b_c-\delta, b_c+\delta]$$

intervallum-együtthatós lineáris egyenletrendszer megoldáshalmaza.

25. Tétel. *Tegyük fel, hogy $\varrho(\Delta) < 1$. Ekkor $\forall i = 1,2,...,n$-re*

$$\underline{x}_i = \min\left\{-(M(|b_c|+\delta))_i + m_{ii}(b_c+|b_c|)_i, -\frac{1}{2m_{ii}-1}((M(|b_c|+\delta))_i + m_{ii}(b_c+|b_c|)_i)\right\},$$

$$\overline{x}_i = \max\left\{(M(|b_c|+\delta))_i + m_{ii}(b_c-|b_c|)_i, \frac{1}{2m_{ii}-1}((M(|b_c|+\delta))_i + m_{ii}(b_c-|b_c|)_i)\right\}.$$

Bizonyítás: A tétel bizonyítása három részből áll. Először azt látjuk be, hogy minden $x\in X$ esetén

$$x_i \le \max\{\widetilde{x}_i, \nu_i\widetilde{x}_i\},$$

ahol

$$\widetilde{x}_i = (M(|b_c|+\delta))_i + m_{ii}(b_c-|b_c|)_i$$

és

$$\nu_i = \frac{1}{2m_{ii}-1}.$$

Majd megmutatjuk, hogy $\widetilde{x}_i = x'_i$ és $\nu_i\widetilde{x}_i = x''_i$ valamely $x', x'' \in X$-re. Ebből az állítás második része következik. Végül megmutatjuk az állítás első felét.

1. • Először megmutatjuk, hogy

$$M\Delta = \Delta M = M - I. \tag{4.1}$$

Ugyanis

$$M - I = (I-\Delta)^{-1} - (I-\Delta)^{-1}(I-\Delta) = (I-\Delta)^{-1}(I-(I-\Delta)) =$$
$$= (I-\Delta)^{-1}(I-I+\Delta) = M\Delta,$$

és

$$M - I = (I-\Delta)^{-1} - (I-\Delta)(I-\Delta)^{-1} = (I-(I-\Delta))(I-\Delta)^{-1} =$$
$$= (I-I+\Delta)(I-\Delta)^{-1} = \Delta M.$$

- Megmutatjuk, hogy $\nu_i \in (0,1]$. Mivel $m_{ii} \geq 1$, ezért $2m_{ii} - 1 \geq 1$, és így

$$\frac{1}{2m_{ii}-1} = \nu_i \in (0,1]. \tag{4.2}$$

- Legyen D diagonális mátrix, $j = 1, 2, ..., n$,

$$D_{jj} := \begin{cases} 1, & \text{ha } j \neq i \text{ és } (b_c)_j \geq 0, \\ -1, & \text{ha } j \neq i \text{ és } (b_c)_j < 0, \\ 1, & \text{ha } j = i. \end{cases}$$

És legyen

$$\widehat{b} := Db_c + \delta = \begin{pmatrix} |(b_c)_1| \\ \vdots \\ |(b_c)_{i-1}| \\ (b_c)_i \\ |(b_c)_{i+1}| \\ \vdots \\ |(b_c)_n| \end{pmatrix} + \delta.$$

Ekkor

$$\widetilde{x}_i = (M(|b_c| + \delta))_i + m_{ii}(b_c - |b_c|)_i = (M\widehat{b})_i. \tag{4.3}$$

- Legyen $x \in X$ tetszőleges, azaz $\exists A \in [I - \Delta, I + \Delta]$ és $b \in [b_c - \delta, b_c + \delta]$, hogy $Ax = b$. Továbbá legyen

$$x' = Dx = \begin{pmatrix} |x_1| \\ \vdots \\ |x_{i-1}| \\ x_i \\ |x_{i+1}| \\ \vdots \\ |x_n| \end{pmatrix}.$$

Ekkor

$$M(x' - |x|) + |x| \leq M\widehat{b}, \tag{4.4}$$

ugyanis

$$x'_i = x_i = b_i + ((I - A)x)_i \leq (b_c + \delta)_i + (\Delta|x|)_i = (\widehat{b} + \Delta|x|)_i, \tag{4.5}$$

és $j \neq i$ esetén

$$x'_j = |x_j| \leq |b_j| + |((I-A)x)_j| \leq |b_c|_j + \delta_j + (\Delta|x|)_j = (\widehat{b} + \Delta|x|)_j. \tag{4.6}$$

(4.5) és (4.6) alapján

$$x' \leq \widehat{b} + \Delta|x|.$$

Az egyenlet mindkét oldalát balról M-mel szorozva

$$Mx' \leq M\widehat{b} + M\Delta|x|.$$

(4.1) alapján

$$Mx' \leq M\widehat{b} + (M-I)|x|.$$

Ha $(M-I)|x|$-t átvisszük a másik oldalra megkapjuk (4.4)-t.

- Két eset van.
  - Ha $x_i \geq 0$, akkor $x' = |x|$, és így (4.4) miatt

    $$x_i = |x_i| \leq (M\widehat{b})_i = \widetilde{x}_i.$$

  - Ha $x_i < 0$, akkor $x'_i = x_i$ és $|x_i| = -x_i$. (4.4) miatt

    $$(M(x' - |x|))_i + |x_i| = 2m_{ii}x_i - x_i = (2m_{ii} - 1)x_i \leq (B\widehat{b})_i = \widetilde{x}_i.$$

    Ezért $x_i \leq \nu_i \widetilde{x}_i$, ami bizonyítja az első részt.

2. Legyen $x' := DM\widehat{b}$ és $x'' := DM(\widehat{b} - 2\nu_i\widetilde{x}_i\Delta e_i)$. Megmutatjuk, hogy $x', x'' \in X$ és, hogy $x'_i = \widetilde{x}_i$ és $x''_i = \nu_i\widetilde{x}_i$.

   - Először nézzük x'-t. (4.1) miatt

     $$(I - D\Delta D)x' = (I - D\Delta D)DM\widehat{b} = DM\widehat{b} - D\Delta M\widehat{b} = DM\widehat{b} - D(M-I)\widehat{b} =$$

     $$= DM\widehat{b} - DM\widehat{b} + D\widehat{b} = D\widehat{b} = D(Db_c + \delta) = b_c + D\delta.$$

     Azaz

     $$(I - D\Delta D)x' = b_c + D\delta. \tag{4.7}$$

     Mivel

     - $I - D\Delta D \in [I - \Delta, I + \Delta]$ és

– $b_c + D\delta \in [b_c - \delta, b_c + \delta]$,

ezért (4.7) miatt $x' \in X$ teljesül.

- Most nézzük x''-t. Legyen D' diagonális mátrix, ahol $D'_{ii} = -1$ és $D'_{jj} = D_{jj}$. Ekkor (4.1) miatt

$$(I - D\Delta D')DM = DM - D\Delta D'DM = DM - D\Delta(I - 2e_ie_i^T)M =$$

$$= DM - D\Delta M + D\Delta 2e_ie_i^T M = DM - D(M - I) + D\delta 2e_ie_i^T M =$$

$$= DM - DM + D + 2D\Delta e_ie_i^T M = D + 2D\Delta e_ie_i^T M.$$

Ezt felhasználva

$$(I - D\Delta D')x'' = (I - D\Delta D')DM(\widehat{b} - 2\nu_i\widetilde{x}_i\Delta e_i) =$$

$$= (D + 2D\Delta e_ie_i^T M)(\widehat{b} - 2\nu_i\widetilde{x}_i\Delta e_i) =$$

$$= D\widehat{b} - 2\nu_i\widetilde{x}_i D\Delta e_i + 2D\Delta e_ie_i^T M\widehat{b} - 4\nu_i\widetilde{x}_i D\Delta e_ie_i^T M\Delta e_i =$$

$$= D\widehat{b} + 2\widetilde{x}_i D\Delta e_i(-\nu_i + 1 - 2\nu_i e_i^T M\Delta e_i) =$$

$$= D\widehat{b} + 2\widetilde{x}_i D\Delta e_i\left(-\frac{1}{2m_{ii} - 1} + 1 - \frac{2(m_{ii} - 1)}{2m_{ii} - 1}\right) = D\widehat{b} = b_c + D\delta.$$

Azaz

$$(I - D\Delta D')x'' = b_c + D\delta. \tag{4.8}$$

Mivel

– $I - D\Delta D' \in [I - \Delta, I + \Delta]$ és

– $b_c + D\delta \in [b_c - \delta, b_c + \delta]$,

ezért (4.8) miatt $x'' \in X$ teljesül.

- A második pont igazolásához még azt kell belátni, hogy $x'_i = \widetilde{x}_i$ és $x''_i = \nu_i\widetilde{x}_i$.

  – Mivel $e_i^T D = e_i$, ezért

$$x'_i = e_i^T DM\widehat{b} = e_i^T M\widehat{b} = (M\widehat{b})_i = \widetilde{x}_i.$$

  – $e_i^T D = e_i$ és (4.1) miatt

$$x''_i = (DM\widehat{b})_i - (2\nu_i\widetilde{x}_i DM\Delta e_i)_i = \widetilde{x}_i - 2\nu_i\widetilde{x}_i e_i^T D(M - I)e_i =$$

$$= \widetilde{x}_i - 2\nu_i\widetilde{x}_i(m_{ii} - 1) = \widetilde{x}_i - \frac{2\widetilde{x}_i(m_{ii} - 1)}{2m_{ii} - 1} = \nu_i\widetilde{x}_i.$$

Ezzel beláttuk a tétel maximumra vonatkozó állítását.

3. Tekintsük az $[I - \Delta, I + \Delta]x = [-b_c - \delta, -b_c + \delta]$ intervallum-együtthatós lineáris egyenletrendszer $X_0 = -X$ megoldáshalmazát. Ha az imént belátottakat erre alkalmazzuk, akkor megkapjuk a minimumra vonatkozó állítást. $\square$

5. fejezet

Intervallum-együtthatós lineáris egyenletrendszerek megoldáshalmazának behatárolása általános esetben

Ebben a fejezetben összefoglaljuk a [7], [8], [5], [6] cikkek legfontosabb állításait, majd az algoritmusok leírása után pár példán bemutatjuk, hogy azokat MATLAB-ban leprogramozva milyen eredményeket kaptunk.

5.1. Elméleti háttér

Egy általános módszert írunk le, mely megadja egy tetszőleges intervallum-együtthatós lineáris egyenletrendszer megoldáshalmazát tartalmazó legszűkebb intervallumvektort, vagy ad egy szinguláris mátrixot, mely eleme a rendszer baloldali mátrixának. Tehát most az

$$\mathbf{A} = [A_c - \Delta, A_c + \Delta] \in I(\mathbb{R})^{n\times n}$$

és a

$$\mathbf{b} = [b_c - \delta, b_c + \delta] \in I(\mathbb{R})^n$$

intervallummátrixról és vektorról nem teszünk fel semmit.

A következőkben az alábbi jelöléseket használjuk.

24. Definíció. *Legyen $x \in \mathbb{R}^n$ tetszőleges vektor, ekkor*

$$(sgn(x))_i := \begin{cases} 1, & ha \quad x_i \geq 0, \\ -1, & ha \quad x_i < 0 \end{cases} \qquad (i = 1, ..., n).$$

25. Definíció. *Jelölje*

$$\mathbb{R}^n_z := \{x : T_z x \geq 0\},$$

ahol $T_z = diag(z_1, ..., z_n)$ és $z \in Y_n$ előre rögzített vektor.

26. Definíció. *Legyen $z, z' \in Y_n$. Ekkor azt mondjuk, hogy z és z' szomszédosak, ha pontosan egy koordinátájukban különböznek.*

Az $\mathbf{A}x = \mathbf{b}$ intervallum-együtthatós lineáris egyenletrendszer megoldáshalmazát továbbra is Σ-val jelöljük, azaz

$$\Sigma = \{x : \exists A \in \mathbf{A} \wedge \exists b \in \mathbf{b}, Ax = b\}.$$

Az Oettli-Prager-tétel [9] szerint ez a megoldáshalmaz a következőképpen írható le:

$$\Sigma = \{x : |A_c x - b_c| \leq \Delta|x| + \delta\}.$$

Ismeretes, hogy ha $\mathbf{A}$ reguláris, akkor Σ kompakt és összefüggő halmaz, ellenkező esetben pedig Σ minden komponense (azaz nemüres összefüggő részhalmaza, ami a tartalmazásra nézve maximális) nemkorlátos. A megoldáshalmaz általában egy bonyolult nemkonvex struktúra, ezért most is az őt tartalmazó legszűkebb intervallumvektort keressük, melyet $\mathbf{x}(\mathbf{A}, \mathbf{b})$-vel jelölünk. Azaz

$$\mathbf{x}(\mathbf{A}, \mathbf{b}) = [\underline{x}, \overline{x}],$$

ahol

$$\underline{x}_i = \min\{x_i : x \in \Sigma\},$$
$$\overline{x}_i = \max\{x_i : x \in \Sigma\},$$

$(i = 1, ..., n)$. Ha $\mathbf{A}$ szinguláris, akkor Σ vagy üres, vagy nemkorlátos, ezért ebben az esetben $\mathbf{x}(\mathbf{A}, \mathbf{b})$-t nem definiáljuk.

A megoldáshalmazt tartalmazó legszűkebb intervallumvektor megadásáról szóló fő tétel előtt kimondjuk az ezt az eredményt megalapzó három egymásra épülő tételt.

26. Tétel. *Legyen* $\mathbf{A} \in I(\mathbb{R})^{n\times n}$ *és* $\mathbf{b} \in I(\mathbb{R})^n$*, és legyen* $Z \subseteq Y_n$ *melyre a következők teljesülnek:*

1. $sgn(x_0) \in Z$ *valamely* $x_0 \in \Sigma$ *esetén,*

2. $\Sigma \cap \mathbb{R}^n_z$ *korlátos halmaz minden* $z \in Z$ *esetén,*

3. *ha* $z \in Z$ *és* $y \in Y_n$ *szomszédosak és* $\Sigma \cap \mathbb{R}^n_z \cap \mathbb{R}^n_y \neq 0$*, akkor* $y \in Z$.

Ekkor $\mathbf{A}$ *reguláris és*

$$\Sigma \subseteq \bigcup_{z\in Z} \mathbb{R}^n_z.$$

Tehát a tétel ad egy szükséges feltételt az $\mathbf{A}$ intervallummátrix regularitására, és a megoldáshalmazba tartozó vektorok előjeleit korlátozza a Z halmazra. A következő tételben kicsit változtatunk a Z halmaz tulajdonságain, és így egy Σ-t tartalmazó intervallumvektort tudunk adni, ami persze még nem biztos, hogy a legszűkebb.

27. Tétel. *Legyen* $\mathbf{A} \in I(\mathbb{R})^{n\times n}$ *és* $\mathbf{b} \in I(\mathbb{R})^n$*, és legyen* $Z \subseteq Y_n$ *melyre a következők teljesülnek:*

1. $sgn(x_0) \in Z$ *valamely* $x_0 \in \Sigma$ *esetén,*

2. *minden* $z \in Z$*-re, melyre* $\Sigma \cap \mathbb{R}^n_z \neq 0$*, létezik egy* $[\underline{x}_z, \overline{x}_z]$ *intervallumvektor, melyre* $\Sigma \cap \mathbb{R}^n_z \subseteq [\underline{x}_z, \overline{x}_z]$*,*

3. *ha* $z \in Z$*,* $\Sigma \cap \mathbb{R}^n_z \neq 0$ *és* $(\underline{x}_z)_j(\overline{x}_z)_j \leq 0$ *valamely* j *esetén, akkor* $z - 2z_je_j \in Z$.

Ekkor $\mathbf{A}$ *reguláris és*

$$\Sigma \subseteq \bigcup_{z\in Z_0} [\underline{x}_z, \overline{x}_z],$$

ahol

$$Z_0 = \{z \in Z : \Sigma \cap \mathbb{R}^n_z \neq 0\}.$$

A következő tételben egy abszolútértékes egyenlőtlenségrendszer megoldására vezetjük vissza a problémát, melynek megoldására később még visszatérünk. Ismét változtatunk a Z halmaz tulajdonságain, amivel az előzőnél egy jobban használható eredményre jutunk.

28. Tétel. *Legyen* $\mathbf{A} = [A_c - \Delta, A_c + \Delta] \in I(\mathbb{R})^{n\times n}$ *és* $\mathbf{b} = [b_c - \delta, b_c + \delta] \in I(\mathbb{R})^n$, *és legyen* $Z \subseteq Y_n$ *melyre a következők teljesülnek:*

1. $sgn(x_0) \in Z$ *valamely* $x_0 \in \Sigma$ *esetén,*

2. *minden* $z \in Z$*-re az alábbi egyenlőtlenségeknek*

$$(QA_c - I)T_z \geq |Q|\Delta \tag{5.1}$$

$$(QA_c - I)T_{-z} \geq |Q|\Delta \tag{5.2}$$

   létezik Q_z *és* Q_{-z} *megoldása,*

3. *ha* $z \in Z$, $Q_{-z}b_c - |Q_{-z}|\delta \leq Q_z b_c + |Q_z|\delta$ *és* $(Q_{-z}b_c - |Q_{-z}|\delta)_j(Q_z b_c + |Q_z|\delta)_j \leq 0$ *valamely* j *esetén, akkor* $z - 2z_j e_j \in Z$.

Ekkor $\mathbf{A}$ *reguláris és*

$$\Sigma \subseteq \bigcup_{z\in Z_1} [Q_{-z}b_c - |Q_{-z}|\delta, Q_z b_c + |Q_z|\delta] \subseteq$$

$$\subseteq [\min_{z\in Z_1}(Q_{-z}b_c - |Q_{-z}|\delta), \max_{z\in Z_1}(Q_z b_c + |Q_z|\delta)],$$

ahol

$$Z_1 = \{z \in Z : Q_{-z}b_c - |Q_{-z}|\delta \leq Q_z b_c + |Q_z|\delta\}.$$

Legyen mostantól

$$\overline{x}_z := Q_z b_c + |Q_z|\delta,$$

$$\underline{x}_z := Q_{-z}b_c - |Q_{-z}|\delta.$$

Tehát ha a tétel feltételei teljesülnek, akkor

$$\Sigma \subseteq [\min_{z\in Z_1} \underline{x}_z, \max_{z\in Z_1} \overline{x}_z] \tag{5.3}$$

A következő tétel azt mondja ki, hogy ha az (5.1), (5.2) abszolútértékes egyenlőtlenségeket egyenlőséggel oldjuk meg, akkor az (5.3)-béli tartalmazó intervallum legszűkebb tartalmazó intervallummá válik.

29. Tétel. *Legyen* $\mathbf{A} = [A_c - \Delta, A_c + \Delta] \in I(\mathbb{R})^{n\times n}$ *és* $\mathbf{b} = [b_c - \delta, b_c + \delta] \in I(\mathbb{R})^n$, *és legyen* $Z \subseteq Y_n$ *melyre a következők teljesülnek:*

1. $sgn(x_0) \in Z$ *valamely* $x_0 \in \Sigma$ *esetén,*

2. *minden* $z \in Z$*-re az alábbi egyenlőségeknek*

$$QA_c - |Q|\Delta T_z = I \tag{5.4}$$

$$QA_c - |Q|\Delta T_{-z} = I \tag{5.5}$$

   létezik Q_z *és* Q_{-z} *megoldása,*

3. *ha* $z \in Z$, $\underline{x}_z \leq \overline{x}_z$ *és* $(\underline{x}_z)_j(\overline{x}_z)_j \leq 0$ *valamely* j *esetén, akkor* $z - 2z_je_j \in Z$.

Ekkor $\mathbf{A}$ *reguláris és*

$$\mathbf{x}(\mathbf{A}, \mathbf{b}) = [\min_{z\in Z_1} \underline{x}_z, \max_{z\in Z_1} \overline{x}_z], \tag{5.6}$$

ahol

$$Z_1 = \{z \in Z : \underline{x}_z \leq \overline{x}_z\}.$$

Tehát a fenti tétel segítségével meg tudjuk adni egy tetszőleges intervallum egyenletrendszer megoldáshalmazát tartalmazó legszűkebb intervallumvektort, ha van ilyen.

Most térjünk rá az abszolútértékes egyenlet (5.4), (5.5) megoldására. Legyen

$$x^T = Q_{i.}, \quad i \in \{1, 2, ..., n\}.$$

Ekkor x vektor az

$$x^T A_c - |x|^T \Delta T_z = e_i^T \tag{5.7}$$

megoldása, és így

$$A_c^T x - T_z \Delta^T |x| = e_i, \tag{5.8}$$

ami

$$Ax + B|x| = b \tag{5.9}$$

alakban van, ahol $A, B \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n$. A megoldás minket abban az esetben érdekel, ha nem létezik olyan S szinguláris mátrix, melyre

$$|S - A| \leq |B|, \tag{5.10}$$

hiszen ha létezik ilyen S, akkor ez eleme az $\mathbf{A}$ intervallummátrixnak, és így az szinguláris.

A következőkben felsoroljuk azokat az állításokat, melyeket (5.9) egyenletrendszer megoldása során felhasználunk.

19. Állítás. *Legyen* $A \in \mathbb{R}^{n\times n}$ *reguláris,* $b, c \in \mathbb{R}^n$ *és* $\alpha = 1 + c^T A^{-1} b$*. Ekkor*

1. $\det(A + bc^T) = \alpha \det(A)$,

2. *ha* $\alpha = 0$, *akkor* $A + bc^T$ *szinguláris,*

3. *ha* $\alpha \neq 0$, *akkor* $(A + bc^T)^{-1} = A^{-1} - \frac{1}{\alpha} A^{-1} bc^T A^{-1}$.

20. Állítás. *Legyen* $\mathbf{A} = [A-|B|, A+|B|] \in I(\mathbb{R})^{n\times n}$*.* $\mathbf{A}$ *akkor és csak akkor szinguláris, ha* $|Ax| \leq |B||x|$ *egyenlőtlenségnek létezik nemtriviális megoldása.*

21. Állítás. *Legyen* $\mathbf{A} = [A - |B|, A + |B|] \in I(\mathbb{R})^{n\times n}$ *reguláris és*

$$(A + BT_{z'})x' = (A + BT_{z''})x''$$

valamely $z', z'' \in Y_n$, $x' \neq x''$ *esetén. Ekkor létezik olyan* j *index, melyre* $z'_j z''_j = -1$ *és* $x'_j x''_j > 0$.

22. Állítás. *Legyen*

$$(A + BT_{z'})x' = (A + BT_{z''})x''$$

valamely $z', z'' \in Y_n$ *esetén és* $x' \neq x''$ *olyan, hogy minden* l *indexre* $z'_l z''_l = -1$ *esetén* $x'_l x''_l \leq 0$*. Továbbá legyen* $x = x' - x''$,

$$y_j = \begin{cases} (Ax)_j/(|B||x|)_j, & ha(|B||x|)_j > 0 \\ 1, & ha(|B||x|)_j = 0 \end{cases} \qquad (j = 1, ..., n) \tag{5.11}$$

és

$$z = sgn(x). \tag{5.12}$$

Ekkor

$$S = A - T_y|B|T_z \tag{5.13}$$

szinguláris mátrix, melyre $|S - A| \leq |B|$ *és* $Sx = 0$.

5.2. Algoritmusok

Először azt az algoritmust írjuk le, amely vagy megoldja az (5.9) abszolútértékes egyenletrendszert, vagy ad egy S szinguláris mátrixot, melyre $|S - A| \leq |B|$.

1. Ha A szinguláris, akkor $S = A$ és kész vagyunk.
2. Legyen $z = sgn(A^{-1}b)$.
3. Ha $A + BT_z$ szinguláris, akkor $S = A + BT_z$ és kész vagyunk.
4. Legyen $x = (A + BT_z)^{-1}b$ és $C = -(A + BT_z)^{-1}B$.
5. Legyen $i = 0$, $r = 0 \in \mathbb{R}^n$, $X = 0 \in \mathbb{R}^{n \times n}$.
6. Amíg $z_j x_j < 0$ valamely j-re
   (a) Legyen $i = i + 1$ és $k = \min\{j : z_j x_j < 0\}$.
   (b) Ha $1 + 2z_k C_{kk} \leq 0$, akkor $S = A + B(T_z + (1/C_{kk})e_k e_k^T)$ és kész vagyunk.
   (c) Ha ($k < n$ és $r_k > \max_{j>k} r_j$) vagy ($k = n$ és $r_n > 0$), akkor
      i. $x = x - X_{.k}$.
      ii. Ha $(|B||x|)_j > 0$, akkor legyen $y_j = (Ax)_j/(|B||x|)_j$ egyébként legyen $y_j = 1$ $(j = 1, 2, ..., n)$.
      iii. Legyen $z = sgn(x)$ és $S = A - T_y|B|T_z$ és kész vagyunk.
   (d) Legyen $r_k = i$, $X_{.k} = x$, $z_k = -z_k$ és $\alpha = 2z_k/(1 - 2z_k C_{kk})$.
   (e) Legyen $x = x + \alpha x_k C_{.k}$ és $C = C + \alpha C_{.k} C_{k.}$.

Tehát a fenti algoritmussal $A = A_c^T$, $B = -T_z \Delta^T$, $b = e_i$, $(i = 1, 2, ..., n)$ választással Q_z illetve Q_{-z} sorait ki tudjuk számítani.

Most térjünk rá arra az algoritmusra, amely egy intervallum-együtthatós lineáris egyenletrendszerhez megadja a megoldáshalmazát tartalmazó legszűkebb intervallumvektort, ha ilyen létezik. Ellenkező esetben megad egy olyan szinguláris S mátrixot, ami benne van az egyenletrendszer együttható intervallummátrixában.

1. Ha A_c szinguláris, akkor $S = A_c$, és kész vagyunk.

2. Legyen $x_c = A_c^{-1} b_c$, $z = sgn(x_c)$, $\underline{x} = \overline{x} = x_c$, $Z = \{z\}$ és $D = \emptyset$.

3. Amíg $Z \neq \emptyset$:

   (a) Választunk egy $z \in Z$-t, $Z = Z \backslash \{z\}$ és $D = D \cup \{z\}$.

   (b) A fenti algoritmussal kiszámítjuk Q_z-t és Q_{-z}-t, ha léteznek. Ha valamelyik nem létezik, akkor az algoritmus ad egy S szinguláris mátrixot, és kész vagyunk.

   (c) Legyen $\overline{x}_z = Q_z b_c + |Q_z|\delta$ és $\underline{x}_z = Q_{-z} b_c - |Q_{-z}|\delta$.

   (d) Ha $\underline{x}_z \leq \overline{x}_z$, akkor

      i. Legyen $\underline{x} = \min\{\underline{x}, \underline{x}_z\}$ és $\overline{x} = \max\{\overline{x}, \overline{x}_z\}$.

      ii. Válasszunk egy tetszőleges z-vel szomszédos z'-t, és legyen j az az index, amelyre $z'_j = -z_j$. Ha $(\underline{x})_j (\overline{x})_j \leq 0$ és $z' \notin Z \cup D$, akkor legyen $Z = Z \cup \{z'\}$. Ezt addig ismételjük, amíg z összes szomszédját meg nem vizsgáltuk.

4. $\mathbf{x}(\mathbf{A}, \mathbf{b}) = [\underline{x}, \overline{x}]$.

A fent leírt algoritmusokat leprogramoztuk, melynek MATLAB-kódjai a mellékelt CD-n találhatók. A programunkat számos példán kipróbáltuk, melyek közül néhány m-file-ja szintén megtalálható a CD-n. A következőkben röviden összefoglaljuk ezeknek az eredményeit.

- Először pontmátrixokon próbáltuk ki az algoritmust, mely mind reguláris, mind szinguláris esetben a várt eredményt adta.

- Ezek után megnéztük, hogy a [3]-ban szereplő mintafeladatra milyen eredményt ad. Itt ugyanazt az eredményt kaptuk, mint a cikkben.

- Végül különböző szerkezetű és méretű intervallummátrixokra teszteltük az algoritmust.

6. fejezet

Összefoglalás

- A dolgozatban az
$$\mathbf{A}x = \mathbf{b}$$
úgynevezett intervallum-együtthatós lineáris egyenletrendszerekkel foglalkoztunk.

- A munka elején a tárgyaláshoz szükséges fogalmakat definiáltuk, és az ehhez kapcsolódó legfontosabb állításokat írtuk le.

- Először egy elméleti kérdést vizsgáltunk, mégpedig az intervallum-együtthatós lineáris egyenletrendszerek megoldhatóságát.

- Ezek után leírtuk, hogy milyen feltételek mellett lehet alkalmazni reguláris baloldal esetén a Gauss-eliminációt az intervallumos feladatokra.

- Majd E. R. Hansen illetve J. Rohn eredményeivel foglalkoztunk, mely szintén reguláris $\mathbf{A}$ mátrix esetén alkalmazható.

- Ezt követően egy olyan algoritmust mutattunk be, amelynél már nem kell feltenni a regularitást, így tetszőleges $\mathbf{A}$ mátrix esetén, ha $\mathbf{A}$ szinguláris, akkor ad egy szinguláris részmátrixot, ellenkező esetben pedig megadja a megoldásokat tartalmazó legszűkebb intervallumvektort. Ez az algoritmus szintén J. Rohn nevéhez köthető.

- Ez utóbbi algoritmushoz MATLAB-programot is készítettünk, melyet különböző típusú egyenletrendszerekre teszteltünk.

- A munkához a kapcsolódó irodalom aktuális eredményeit használtuk fel.

Számomra az intervallumaritmetika egy teljesen ismeretlen terület volt, a téma felfedezése sok izgalmat tartogatott. A felkészülés során számos cikket olvastam és szeretnék a továbbiakban olyan ehhez kapcsolódó témával foglalkozni, amibe még nem ástam bele magam, például speciális tulajdonságú (szimmetrikus, pozitív definit) **A** mátrixok esetén milyen a megoldáshalmaz, illetve lehet-e az algoritmusokat egyszerűsíteni. Nagy lelkesedéssel tölt el annak a lehetősége, hogy diplomám megszerzése után is foglalkozhassak ezzel a kutatási területtel.

Ezúton szeretném megragadni az alkalmat, hogy köszönetet mondjak témavezetőmnek, Gergó Lajos Tanár Úrnak, hogy figyelmembe ajánlotta a témát, és a dolgozat megírása folyamán rengeteg segítséget nyújtott.

Irodalomjegyzék

[1] G. Alefeld and J. Herzberger, Introduction to Interval Computations, Academic Press, NewY ork, 1983.

[2] J. Rohn, Solvability of Systems of Linear Interval Equations, *SIAM J. MATRIX ANAL. APPL.*, Vol. 25, No. 1, pp. 237-245, 2003.

[3] E. R. Hansen, Bounding the Solution of Interval Linear Equations, *SIAM J. NUMER. ANAL.*, Vol. 29, No. 5, pp. 1493-1503, October 1992.

[4] J. Rohn, Cheap and tight bounds: The recent result by E. Hansen can be made more efficient, *Interval Comput.*, 4 (1993), pp. 13-21.

[5] J. Rohn, An algorithm for solving the absolute value equation, *Electronic Journal of Linear Algebra*, 18 (2009), pp. 589-599.

[6] J. Rohn, An algorithm for solving the absolute value equation: An improvement, Technical Report 1063, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, January 2010.

[7] J. Rohn, A general method for enclosing solutions of interval lin- ear equations, Technical Report 1067, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, March 2010.

[8] J. Rohn, An Algorithm for Computing the Hull of the Solution Set of Interval Linear Equations, Technical report 1074, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, April 2010.

[9] W. Oettli and W. Prager, Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides, *Numer. Math.*, 6 (1964), pp. 405-409.

[10] J. Rohn, An existence theorem for systems of linear equations, *Linear Multilinear Algebra*, 29 (1991), pp. 141-144.

[11] S. Poljak and J. Rohn, Checking robust nonsingularity is NP-hard, *Math. Control Signals Systems*, 6 (1993), pp. 1-9.

[12] J. Rohn, Systems of linear interval equations, *Lin. Alg. Appls.* 126 (1989), 39-78