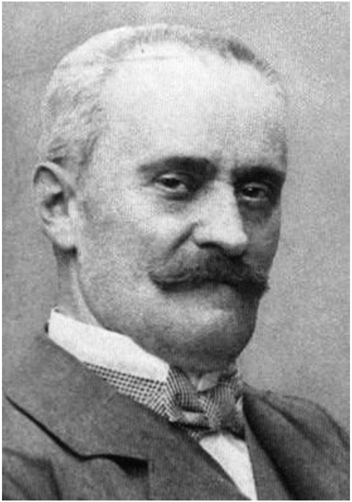
Farkas-lemma

Hajnal Péter

2021. tavasz

Farkas Gyula: Fizikus, egyetemi tanár



Sárosd, 1847. március 28. — Pestszentlőrinc, 1930. december 27.

Farkas Gyula: https://tudosnaptar.kfki.hu/historia

- A pesti egyetemen jogot hallgatott, később Jedlik Ányos hatására természettudományi szakot választott.
- 1870-ben a székesfehérvári reáliskola fizika-kémiatanára volt. 1876-ban megszerezte a tanári képesítést.
- 1880-ban doktorált. 1887-től a kolozsvári egyetem fizikatanszékén rendkívüli, 1888-tól rendes tanár. Ekkor kezdte meg kutatásait. 1907–1908-ban az egyetem rektora volt. Érdeklődése az elméleti fizikának csaknem minden ágára kiterjedt. Ő hívta Kolozsvárra Schlesinger Lajost, Fejér Lipótot, Riesz Frigyest és Haar Alfrédot. 1915-ben Farkas Gyula nyugalomba vonult. Az 1920-as években Farkas Gyula a szegedi Matematikai Szemináriumnak ajándékozta saját könyvtárát.
- Különösen a termodinamika axiomatikus megalapozásával és a lineáris egyenlőtlenségek elméletével foglalkozott. Minkowskitól függetlenül felfedezte az elmélet alaptételét, az ún. Farkas-Minkowski-tételt. A Farkas-lemma a lineáris programozás fontos tétele.
- 1898-tól az MTA tagja.

Az M. Tud. Akadémia III. osztályának 1898. október 17.-én tartott üléséből

A FOURIER-FÉLE MECHANIKAI ELV ALKALMAZÁSÁNAK ALGEBRAI ALAPJA.

FARKAS GYULA l. tagtól.

Már két közleményben következtettem ezt. Most ebben a harmadikban aránylag igen egyszerű módon teszem, a melynél egyszerűbbet találni aligha lehetséges, a mennyiben t. i. számon tartjuk, hogy akárhány egymástól független homogen, lineáris egyenlőtlenség rendelhető össze, mihelyt a vezérmennyiségek száma kettőnél nagyobb. Nem számol ezzel az eshetőséggel egy újabb időben közölt e tárgyú német dolgozat sem (HENNEBERG, Crelle Jour. 1894). Pedig leginkább ennek az általánosságnak a szempontjából hasznos az alkalmazás általános analytikai módszere.

Legyen

$$\begin{array}{l} A_{11}u_1+A_{12}u_2+\cdots+A_{1n}u_n\equiv\theta_1\geqq 0 \\ A_{21}u_1+A_{22}u_2+\cdots+A_{2n}u_n\equiv\theta_2\geqq 0 \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \end{array} \qquad (1)$$

egyenlőtlenségi rendszerünk, és ennek minden megoldásában teljesüljön:

$$A_1u_1+A_2u_2+\cdots+A_nu_n\equiv\vartheta\geqq 0. \qquad (2)$$

A FOURIER-féle mechanikai elv alkalmazásának algebrai alapját a következő tantétel képezi: mindig léteznek olyan nem negativ, az u vezérmennyiségektől független multiplicatorok λ, hogy

$$\vartheta\equiv\lambda_1\theta_1+\lambda_2\theta_2+\cdots \qquad (3)$$

I. alternatíva forma

Farkas-lemma, I. alternatíva forma

Legyen $Ax \preceq b$ egy egyenlőtlenségrendszer, ahol $A \in \mathbb{R}^{k\times n}$, $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ és $b \in \mathbb{R}^k$. Ekkor a következő két állítás közül pontosan egy teljesül:

(i) Az egyenlőtlenségrendszer megoldható, azaz alkalmas $x_0 \in \mathbb{R}^n$ szám n-esre $Ax_0 \preceq b$.

(ii) Alkalmas $0 \preceq \lambda \in \mathbb{R}^k$ nemnegatív szám k-asra $\lambda^{\mathsf{T}} A = 0^{\mathsf{T}}$ és $\lambda^{\mathsf{T}} b = -1$.

Bizonyítás

- Azt kell igazolni, hogy az egyenlőtlenség nem megoldhatósága ekvivalens (ii)-vel.
- A megoldhatóságot a Fourier—Motzkin-eliminációval dönthetjük el.
- Következtetésekkel vezetünk le újabb és újabb egyenlőtlenségeket, amíg a minden változót eliminálunk. Ekkor a bal oldalon 0 áll.
- A nem megoldhatóság pontosan akkor teljesül, ha a jobb oldalon (az egyenlőtlenség nagyobb oldalán) negatív szám áll. Pozitív számmal szorozva elérhetjük, hogy jobb oldalom -1 álljon. Levezettük a $0 \leq -1$ egyenlőtlenséget.
- A levezetéseknél látott univerzalitás alapján az pontosan akkor teljesül, ha az eredeti egyenlőtlenségekből nem-negatív együthatókkal kikombinálható a $0 \leq -1$ egyenlőtlenség.
- Ez éppen a bizonyítandó.

Megjegyzések a bizonyításhoz

- A (ii) lehetőség látványosan mutatja az egyenlőtlenség rendszer nem megoldhatóságát:

Példa

$$\left\{\begin{aligned} x-y+z &\leq 0 \\ -2x+y-z &\leq -2 \\ y-z &\leq 1 \end{aligned}\right.$$

Mátrix alakban

$$\begin{pmatrix} 1 & -1 & 1 \\ -2 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

Azaz

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -2 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

Megjegyzések a bizonyításhoz (folytatás)

Példa (folytatás)

Legyen $\lambda^\mathsf{T} = \left(1, \frac{1}{2}, \frac{1}{2}\right)$. Ekkor

$$\left(1, \tfrac{1}{2}, \tfrac{1}{2}\right)\begin{pmatrix} 1 & -1 & 1 \\ -2 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} = (0, 0, 0), \quad \left(1, \tfrac{1}{2}, \tfrac{1}{2}\right)\begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} = -1.$$

Azaz az eredeti egyenlőtlenségrendszer nem megoldható.
Hagyományos írásmóddal

$$\begin{aligned} x - y + z &\leq 0 \qquad / \cdot 1 \\ -2x + y - z &\leq -2 \qquad / \cdot \frac{1}{2} \\ y - z &\leq 1 \qquad / \cdot \frac{1}{2} \\ \text{összegezve:} \quad 0 \cdot x + 0 \cdot y + 0 \cdot z &= -1 \end{aligned}$$

Azaz nincs megoldás.

II. alternatíva forma

Farkas-lemma, II. alternatíva forma

Legyen $\begin{cases} Ax = b \\ x \succeq 0 \end{cases}$ egy előjel feltételes egyenletrendszer, ahol

$A \in \mathbb{R}^{\ell \times n}$, $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ és $b \in \mathbb{R}^{\ell}$. Ekkor a következő két állítás

közül pontosan egy teljesül:

(i) Az előjel feltételes egyenletrendszer megoldható, azaz alkalmas $0 \preceq x_0 \in \mathbb{R}^n$ szám n-esre $Ax_0 = b$.

(ii) Alkalmas $\lambda \in \mathbb{R}^k$ szám ℓ-esre $\lambda^\mathsf{T} A \succeq 0^\mathsf{T}$ és $\lambda^\mathsf{T} b = -1$.

Bizonyítás

- Egyenletrendszerünk pontosan akkor megoldható, ha a

$$\begin{cases} Ax \preceq b \\ -Ax \preceq -b \\ -x \preceq 0 \end{cases} \quad \text{egyenlőtlenségrendszer megoldható.}$$

- Ezen egyenlőtlenségrendszer NEM megoldhatósága ekvivalens azzal, hogy alkalmas $\mu, \nu \in \mathbb{R}_+^\ell$ és $\rho \in \mathbb{R}_+^n$ súlyvektorok esetén

$$\mu^\mathsf{T} A - \nu^\mathsf{T} A - \rho^\mathsf{T} I = 0^\mathsf{T}$$
$$\mu^\mathsf{T} b - \nu^\mathsf{T} b - \rho^\mathsf{T} 0 = -1$$

- Ekvivalens módon: alkalmas $\mu, \nu \in \mathbb{R}_+^\ell$ és $\rho \in \mathbb{R}_+^n$ súlyvektorok esetén

$$(\mu - \nu)^\mathsf{T} A = \rho^\mathsf{T} I \succeq 0^\mathsf{T}$$
$$(\mu - \nu)^\mathsf{T} b = -1$$

- $\lambda = \mu - \nu$ jelölés bevezetésével kapjuk a bizonyítandót.

Megjegyzések a bizonyításhoz

- A (ii) lehetőség látványosan mutatja az egyenlőtlenség rendszer nem megoldhatóságát:

Példa

$$\begin{cases} 2x + 3y = 0 \\ x - 3y = 0 \\ -x + y = 1 \\ x, y \geq 0. \end{cases}$$

Mátrix alakban

$$\begin{pmatrix} 2 & 3 \\ 1 & -3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Azaz

$$A = \begin{pmatrix} 1 & 3 \\ 1 & -3 \\ -1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Megjegyzések a bizonyításhoz (folytatás)

Példa (folytatás)

Legyen $\lambda^{\mathsf{T}} = \big(2, 1, -1\big)$. Ekkor

$$(2, 1, -1) \begin{pmatrix} 1 & 3 \\ 1 & -3 \\ -1 & 1 \end{pmatrix} = (6, 2) \succeq (0, 0)\,, \quad (2, 1, -1) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -1.$$

Azaz az eredeti egyenlőtlenségrendszer nem megoldható.
Hagyományos írásmóddal

$$\begin{aligned} 2x + 3y &= 0 & \quad / \cdot 2 \\ x - 3y &= 0 & \quad / \cdot 1 \\ -x + y &= 1 & \quad / \cdot (-1) \\ \text{összegezve:} \quad 6x + 2y &= -1 & \end{aligned}$$

Azaz nem lehetséges nem-negatív megoldás.

Szünet

Logika

- A Farkas-lemma állítása logikai szempontból is érdekes. Egy egyenlőtlenségrendszer megoldhatósága egy létezés állítása. Úgy nevezett egzisztenciális állítás. Formalizálva a $\exists$ egzisztenciális kvantort használnánk.

- Egy egyenlőtlenségrendszer nem megoldhatósága a megoldhatóság tagadása. Egy egzisztenciális állítás definíció szerinti tagadása egy univerzális állítás. Egy egyenlőtlenségrendszer nem megoldhatósága azt jelenti, hogy bármit helyettesítsünk is be, nem teljesülhet a rendszer mindegyik feltétele. Leírva egy $\forall$ univerzális kvantort használnánk.

- A Farkas-lemma egy egzisztenciális állítás tagadását szintén egzisztenciális állításként fogalmazza meg. Ez egy nagyon meglepő és mély állítás.

Farkas-lemma: Logikai forma I

Emlékeztető

Legyen $A \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^k$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$.

$$Ax \preceq b \quad \vdash \quad c^{\mathsf{T}}x \leq d \qquad \text{esetén} \qquad Ax \preceq b \quad \models \quad c^{\mathsf{T}}x \leq d$$

Az állítás nagyon egyszerű. Lényegében általános iskolai tudásunkból adódik. Az állítás megfordítható. Ez már egyetemi anyag.

Farkas-lemma: Logikai forma I

$$Ax \preceq 0 \quad \vdash \quad c^{\mathsf{T}}x \leq 0$$

akkor és csak akkor, ha

$$Ax \preceq 0 \quad \models \quad c^{\mathsf{T}}x \leq 0.$$

Farkas-lemma: Logikai forma I: Bizonyítás

- Azt kell igazolnunk, hogy a logikai következmények le is vezethetők. Azaz $Ax \preceq 0 \quad \models \quad c^{\mathsf{T}}x \leq 0$ esetén $c^{\mathsf{T}}x \leq 0$ le is vezethető logikai szabályaink alapján.

- $Ax \preceq 0 \quad \models \quad c^{\mathsf{T}}x \leq 0$ esetén $\begin{cases} Ax & \preceq 0 \\ -c^{\mathsf{T}}x & \leq -\varepsilon \end{cases}$ nem megoldható tetszőleges ε pozitív értékre.

- A Farkas-lemma alapján alkalmas $\mu_\varepsilon \in \mathbb{R}_+^k$, $\nu_\varepsilon \in \mathbb{R}_+$ számokra

$$\mu_\varepsilon^{\mathsf{T}} A - \nu_\varepsilon c^{\mathsf{T}} = 0^{\mathsf{T}}, \qquad \mu_\varepsilon^{\mathsf{T}} 0 - \nu_\varepsilon \varepsilon = -1.$$

- Azaz $\nu_\varepsilon = \frac{1}{\varepsilon} \in \mathbb{R}_{++}$, továbbá

$$\frac{1}{\nu_\varepsilon} \mu_\varepsilon^{\mathsf{T}} A = c^{\mathsf{T}}.$$

Farkas-lemma: Logikai forma II

Farkas-lemma: Logikai forma II

$$Ax \preceq b \quad \models \quad c^{\mathsf{T}}x \leq d$$

akkor és csak akkor, ha

$$Ax \preceq b \quad \vdash \quad c^{\mathsf{T}}x \leq d, \quad \text{vagy} \quad Ax \preceq b \quad \vdash_{+} \quad c^{\mathsf{T}}x < d.$$

Szünet

Markov-láncok: Valószínűségszámítási szemlélet

• Egy kombinatorikus Markov-lánc egy sztochasztikus folyamat. A folyamat egy rendszert ír le, amely diszkrét $t = 0, 1, 2, 3, \ldots$ időskálán változik. Minden $t_0 \in \mathbb{N}$ időpillanatban egy véges S állapothalmaz ($S = \{s_1, s_2, \ldots, s_n\}$) egy elemében van. $t_0 \to t_0 + 1$ „lépésnél” ha az s_i állapotban van, akor $p_{i,j}$ valószínűséggel kerül s_j állapotba. Ezt egy $P \in \mathbb{R}^{S \times S} \simeq \mathbb{R}^{n \times n}$ átmeneti mátrix írja le, ami minden t_0 időpillanatban ugyanaz.

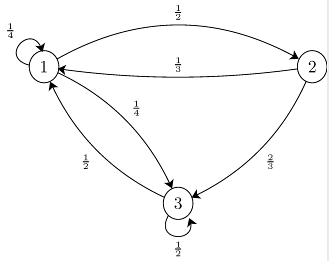
Észrevétel

Egy Markov-lánc átmeneti mátrixa egy $\mathbb{R}_+^{n \times n}$-beli mátrix, amely mindegyik sorösszege 1 (az ilyeneket sztochasztikus mátrixnak nevezzük). Ez meg is fordítható: Egy $\mathbb{R}_+^{n \times n}$-beli mátrix, amely mindegyik sorösszege 1 az leír egy Markov-láncot.

Markov-láncok: Kombinatorikus szemlélet

- Adott egy P átmeneti mátrix.
- S-en mint egy csúcshalmazon egy párhuzamos élek nélküli irányított $\overrightarrow{G}$ gráfot definiálunk. $\overrightarrow{s_i s_j}$ él akkor és csak akkor szerepel a gráfunkban, ha $p_{i,j} > 0$. A véletlen folyamat felfogható mint egy véletlen séta $\overrightarrow{G}$ (egy v csúcsból kivezető véletlen lépés egy a v-ből kivezető éleken lévő valószínűségi eloszlással adott).

Példa (internetről letöltve)



Markov-láncok: Alapok

Észrevétel

A véletlen séta kiindulópontját $\pi \in \mathbb{R}^S$ eloszlással választjuk. Majd egy P átmeneti mátrix által leírt véletlen sétálásba kezdünk.

(i) Az egy lépés utáni helyünk egy véletlen csúcs, amely eloszlása

$$\pi^\mathsf{T} P.$$

(ii) L lépés utáni helyünk egy véletlen csúcs, amely eloszlása

$$\pi^\mathsf{T} P^L.$$

Definíció: Stacionárius eloszlás

A csúcshalmaz feletti π eloszlás stacionárius, ha egy lépés megtétele után ugyanaz az eloszlás írja le helyünket.

Azaz π akkor és csak akkor stacionárius, ha $\pi^\mathsf{T} P = \pi^\mathsf{T}$.

Markov-láncok: A tétel

Tétel

Legyen P egy sztochasztikus mátrix. Ekkor az általa leírt Markov-láncnak van stacionárius eloszlása.

Az állítás egy lineáris egyenlőtlenségrendszer megoldhatóságaként fogalmazható meg, amely rendszer:

$$x^{\mathsf{T}}P = x^{\mathsf{T}}$$
$$x^{\mathsf{T}}1 = 1$$
$$x^{\mathsf{T}} \succeq 0$$

Ekvivalens módon:

$$(P^{\mathsf{T}} - I)x = 0$$
$$1^{\mathsf{T}}x = 1$$
$$x \succeq 0$$

Markov-láncok: A bizonyítás

• A Farkas-lemma alapján ki kell zárnunk a második alternatívát. Azaz be kell látnunk, hogy nem létezhet $\lambda \in \mathbb{R}^n, \rho \in \mathbb{R}$, hogy $(\lambda^\mathsf{T}, \rho) \begin{pmatrix} P^\mathsf{T} - I \\ 1^\mathsf{T} \end{pmatrix} \succeq 0^\mathsf{T}$ és $(\lambda^\mathsf{T}, \rho) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1$

• Azaz nem lehet, hogy teljesüljön, hogy

$$(P - I|1) \begin{pmatrix} \lambda \\ \rho \end{pmatrix} \succeq 0 \qquad \text{és} \qquad (0^\mathsf{T}|1) \begin{pmatrix} \lambda \\ \rho \end{pmatrix} = -1.$$

• Azaz nem lehet, hogy teljesüljön, hogy

$$(P - I)\lambda \succeq 1, \qquad \text{azaz} \qquad P\lambda \succeq I\lambda + 1.$$

• Valóban: P sorai valószínűségi eloszlások. $P\lambda$ koordinátái a λ komponenseinek egy-egy várható értéke. Speciálisan nem haladhatják meg λ legnagyobb komponensének értékét. Emiatt nem állhat fenn az egyenlőség a jobb oldallal.

Vége van!

Köszönöm a figyelmet!