BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

Alkalmazott algebra

BMETE90MX57 (FELSŐBB MATEMATIKA INFORMATIKUSOKNAK)

Nemnegatív mátrixok

PRIMITÍV, IRREDUCIBILIS, REDUCIBILIS MÁTRIX, PERRON–FROBENIUS-TÉTEL

Wettl Ferenc
ALGEBRA TANSZÉK

Tartalom

Ismeretek, képességek, célok

- nemnegatív mátrixok osztályozása: pozitív, primitív, irreducibilis, reducibilis,
- Perron- és Perron–Frobenius-tételek, a tételbeli mátrixhatványok kiszámítása, Collatz–Wielandt-tétel,
- reducibilitás és primitívség eldöntése,
- sztochasztikus mátrixok, Frobenius–Kőnig-tétel, Birkhoff-tétel
- sztochasztikus folyamatok, Markov-láncok

Pozitív mátrixok

Mátrixok összehasonlítása

D $\mathbf{A} > \mathbf{B}$, ha $a_{ij} > b_{ij}$ $\forall$ i és j esetén

pozitív: $\mathbf{A} > \mathbf{O}$ (nemnegatív: $\mathbf{A} \geqslant \mathbf{O}$), azaz ha $a_{ij} > 0$ ($a_{ij} \geqslant 0$)

Á Néhány észrevétel:

- $\mathbf{A} \geqslant \mathbf{O} \Leftrightarrow \mathbf{Ax} \geqslant \mathbf{0}$ minden $\mathbf{x} \geqslant 0$ vektorra,
- $\mathbf{A} > \mathbf{O} \Leftrightarrow \mathbf{Ax} > \mathbf{0}$ minden $\mathbf{x} \geqslant 0$, $\mathbf{x} \neq \mathbf{0}$ vektorra,
- $\mathbf{A} \geqslant \mathbf{O}$, és $\mathbf{x} \geqslant \mathbf{y} \geqslant \mathbf{0} \Rightarrow \mathbf{Ax} \geqslant \mathbf{Ay}$.

D A nemnegatív mátrixok pozitivitásának 4 fokozata ($a_{ij}^{(k)} = [\mathbf{A}^k]_{ij}$):

$\mathbf{A}$ pozitív:	$\forall i, j$	$a_{ij} > 0$	
$\mathbf{A}$ primitív:	$\exists k\ \forall i, j$	$a_{ij}^{(k)} > 0$	(nem primitív: imprimitív)
$\mathbf{A}$ irreducibilis:	$\forall i, j\ \exists k$	$a_{ij}^{(k)} > 0$	
$\mathbf{A}$ reducibilis:	$\exists i, j\ \forall k$	$a_{ij}^{(k)} = 0$	

Példák

- $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ pozitív
- $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ primitív, mert $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$
- $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ irreducibilis, mert $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ reducibilis, mert $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

Perron-tétel

T **Perron-tétel: pozitív sajátérték és sajátvektor**

Ha $\mathbf{A} > \mathbf{O}$, *akkor*

1. $r > 0$, *(*$r = \varrho(\mathbf{A})$ *a spektrálsugár),*
2. r *sajátérték egy pozitív sajátvektorral,*
3. $\mathbf{A}$*-nak e pozitív sajátvektor skalárszorosain kívül nincs más nemnegatív sajátvektora.*

D $\mathbf{p}$ (jobb) Perron-vektor és $\mathbf{q}$ bal Perron-vektor, ha

$$\mathbf{A}\mathbf{p} = r\mathbf{p}, \ \sum_{i=1}^{n} p_i = 1, \qquad \mathbf{q}^\mathsf{T}\mathbf{A} = r\mathbf{q}^\mathsf{T}, \ \sum_{i=1}^{n} q_i = 1$$

T **Perron-tétel: egyszeres és domináns sajátérték**

Ha $\mathbf{A} > \mathbf{O}$, *akkor*

4. *az* r *sajátérték algebrai multiplicitása* 1*,*
5. r *domináns, azaz minden további* λ *sajátértékre* $|\lambda| < r$.

Nem negatív mátrixok

A Perron-tétel nem áll a nemnegatív mátrixokra

A Perron-tétel pl. ezekre nem áll:

- a $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \geqslant \mathbf{O}$, mindkét sajátértéke 0, ezért spektrálsugara is 0,
- az $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \geqslant \mathbf{O}$ mátrix spektrálsugara 1, de az 1 kétszeres sajátérték, és több lineárisan független pozitív sajátvektor is tartozik hozzá,
- a $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ mátrix sajátértékei 1 és -1, így spektrálsugara ugyancsak 1, de a spektrálkörön több különböző sajátértéke is van,
- az $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ mátrixnak nincs pozitív sajátvektora.

Ugyanakkor pl.

- az $\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \geqslant \mathbf{O}$, sajátértékei 2, -1, spektrálsugara 2, ami egyszeres sajátérték, a spektrálkörön ez az egyetlen sajátérték, a hozzá tartozó $(1, 1)$ sajátvektor pozitív, és ennek konstansszorosait kivéve más pozitív sajátvektor nincs, mert a -1-hez tartozó sajátvektor $(1, -2)$.

Mi igaz minden nemnegatív mátrixra

T **Perron–Frobenius-tétel – gyenge változat**

Ha $\mathbf{A} \geqslant \mathbf{O}$*, akkor az* $r = \varrho(\mathbf{A})$ *spektrálsugár sajátértéke* $\mathbf{A}$*-nak, melyhez tartozik nemnegatív sajátvektor.*

B Alapötlet: $\mathbf{A}_k = \mathbf{A} + \frac{1}{k}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}, \quad k \in \mathbb{N}.$

T **Collatz–Wielandt-tétel**

Az $\mathbf{A} \geqslant \mathbf{O}$ *mátrix* r *spektrálsugarára*

$$r = \max_{\substack{\mathbf{x} \\ \mathbf{0} \neq \mathbf{x} \geqslant \mathbf{0}}} \min_{\substack{1 \leqslant i \leqslant n \\ x_i \neq 0}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i}, \text{ másként fogalmazva } r = \max_{\substack{\mathbf{x} \\ \mathbf{0} \neq \mathbf{x} \geqslant \mathbf{0}}} \max_{c\mathbf{x} \leqslant \mathbf{A}\mathbf{x}} c$$

B Ötlet: $\mathbf{A} > \mathbf{O} \ \rightsquigarrow \ c\mathbf{x} \leqslant \mathbf{A}\mathbf{x} \ \rightsquigarrow \ c\mathbf{q}^\mathsf{T}\mathbf{x} \leqslant \mathbf{q}^\mathsf{T}\mathbf{A}\mathbf{x} = r\mathbf{q}^\mathsf{T}\mathbf{x} \ \rightsquigarrow \ c \leqslant r.$
Másrészt az $\mathbf{x} = \mathbf{p}$ vektorra $r\mathbf{p} = \mathbf{A}\mathbf{p} \rightsquigarrow \max c = r.$

Mi igaz minden nemnegatív mátrixra

K **Nemnegatív mátrixok spektrálsugarának becslése**

Ha $\mathbf{A} \geqslant \mathbf{O}$*, akkor a spektrálsugár a sorösszegek minimuma és maximuma, illetve az oszlopösszegek minimuma és maximuma közé esik, azaz*

$$\min_i \left\{ \sum_{j=1}^{n} a_{ij} \right\} \leqslant \varrho(\mathbf{A}) \leqslant \max_i \left\{ \sum_{j=1}^{n} a_{ij} \right\}$$

$$\min_j \left\{ \sum_{i=1}^{n} a_{ij} \right\} \leqslant \varrho(\mathbf{A}) \leqslant \max_j \left\{ \sum_{i=1}^{n} a_{ij} \right\}$$

K **Konstans sorösszeg vagy oszlopösszeg** Ha $\mathbf{A} \geqslant \mathbf{O}$, és minden sorösszeg c, akkor a spektrálsugár c. (oszlopösszegre is)

- Becsüljük meg az alábbi mátrixok spektrálsugarát a sor és oszlopösszegekkel!

$$\mathsf{A} = \begin{bmatrix} 3 & 3 & 1 & 3 \\ 2 & 3 & 2 & 2 \\ 3 & 3 & 2 & 3 \\ 3 & 3 & 1 & 1 \end{bmatrix} \qquad \mathsf{D} = \begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

M A sorösszegei: 10, 9, 11, 8, oszlopösszegei: 11, 12, 6, 9, így $8 \leq \rho(\mathsf{A}) \leq 11$, a sajátértékek 9.366, -1.242, $0.438 \pm 0.569i$, $\rho(\mathsf{A}) \approx 9.366$.

M Dürer híres rajzának bűvös négyzetében minden sorösszeg 34, így $\rho(\mathsf{A}) = 34$. (Számtalan más összeg is 34, a sajátértékek pedig 34, 8, 0, -8.)

Irreducibilis mátrixok

T **Reducibilitás szükséges és elégséges feltétele**

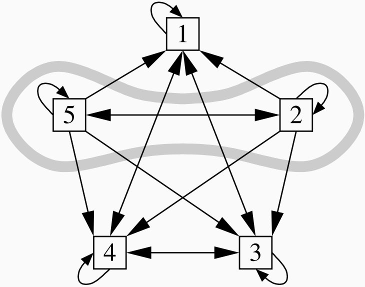
Az $\mathbf{A} \geqslant \mathbf{O}$ *mátrix pontosan akkor reducibilis, ha a sorok és oszlopok azonos permutációjával* $\begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{O} & \mathbf{Z} \end{bmatrix}$ *alakra hozható, ahol* $\mathbf{X}$ *és* $\mathbf{Z}$ *négyzetes mátrixok (létezik olyan* $\mathbf{P}$ *permutáló mátrix, hogy* $\mathbf{PAP}^\mathsf{T}$ *a fenti alakú).*

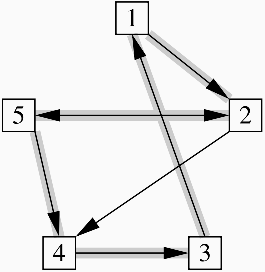
B csak az számít, hogy pozitív vagy 0, a nagyság nem.

- G irányított gráf: van $i \to j$ él $\iff a_{ij} > 0$,
- $\mathbf{G}$ szomszédsági mátrixa $\mathbf{A}$-ból: a pozitív elemeket 1-re cseréljük
- $[\mathbf{G}^k]_{ij} > 0 \iff$ az i-edikből megy k-hosszú irányított út a j-edikbe
- $\mathbf{A}$ pontosan akkor irreducibilis, ha bármely két csúcs között vezet irányított út, azaz ha a gráf erősen összefüggő
- sorszámozhatjuk a pontokat úgy, hogy a tételbeli alakja legyen (átsorszámozás = azonos sor és oszloppermutációk)

P reducibilis, vagy irreducibilis?

$$\mathsf{A} = \begin{bmatrix} 11 & 0 & 13 & 14 & 0 \\ 21 & 22 & 23 & 24 & 25 \\ 31 & 0 & 33 & 34 & 0 \\ 41 & 0 & 43 & 44 & 0 \\ 51 & 52 & 53 & 54 & 55 \end{bmatrix}, \quad \mathsf{B} = \begin{bmatrix} 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 24 & 25 \\ 31 & 0 & 0 & 0 & 0 \\ 0 & 0 & 43 & 0 & 0 \\ 0 & 52 & 0 & 54 & 0 \end{bmatrix}.$$





Permutáló mátrix

- A reducibilis: az első és utolsó sorok és oszlopok cseréje megfelel

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad PAP^T = \left[\begin{array}{cc|ccc} 55 & 52 & 53 & 54 & 51 \\ 25 & 22 & 23 & 24 & 21 \\ \hline 0 & 0 & 33 & 34 & 31 \\ 0 & 0 & 43 & 44 & 41 \\ 0 & 0 & 13 & 14 & 11 \end{array}\right].$$

Más megoldás: az $1 \to 3 \to 4 \to 5 \to 2 \to 1$ csere is megteszi:

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad PAP^T = \left[\begin{array}{cc|ccc} 22 & 25 & 21 & 23 & 24 \\ 52 & 55 & 51 & 53 & 54 \\ \hline 0 & 0 & 11 & 13 & 14 \\ 0 & 0 & 31 & 33 & 34 \\ 0 & 0 & 41 & 43 & 44 \end{array}\right].$$

- sorok és oszlopok azonos permutációja, nem elemi sorműveletek!

Irreducibilis mátrixok

- Összefoglalás:

A	algebrai feltétel		gráfelméleti feltétel
pozitív:	$\forall i, j$	$a_{ij} > 0$	irányított teljes gráf hurokélekkel
primitív:	$\exists k\ \forall i, j$	$a_{ij}^{(k)} > 0$	$\forall$ két csúcs között fut k-hosszú út
irreducibilis:	$\forall i, j\ \exists k$	$a_{ij}^{(k)} > 0$	erősen összefüggő
reducibilis:	$\exists i, j\ \forall k$	$a_{ij}^{(k)} = 0$	nem erősen összefüggő

Perron–Frobenius-tétel 1.

T

Perron–Frobenius-tétel 1. – sajátérték/sajátvektor

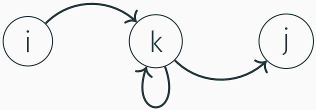
Ha az $\mathbf{A} \geqslant \mathbf{O}$ *irreducibilis, akkor*

1. $r > 0$,
2. *r sajátértéke* $\mathbf{A}$*-nak, melyhez tartozik pozitív sajátvektor,*
3. $\mathbf{A}$*-nak e pozitív sajátvektor skalárszorosain kívül nincs más nemnegatív sajátvektora,*
4. *r egyszeres sajátérték.*

m A Perron-tétel állításai közül nem maradt igaz az irreducibilis nemnegatív mátrixokra az, hogy a spektrálkörön csak egyetlen sajátérték van.

T Elégséges feltétel mátrix primitivitására: Ha $\mathbf{A} \geqslant \mathbf{O}$ irreducibilis és főátlójában van pozitív elem, akkor primitív.

B A gráfon!



Primitív és imprimitív mátrixok

P Döntsük el, hogy melyik mátrix primitív!

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 & 6 \\ 7 & 0 & 8 \\ 0 & 9 & 0 \end{bmatrix}$$

M A reducibilis $\rightsquigarrow$ nem primitív. (a többi irred.)

- B pozitív $\rightsquigarrow$ primitív.
- $C^3 = I$ $\rightsquigarrow$ nem primitív.
- $D^2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ reducibilis $\rightsquigarrow$ D^{2m} is $\rightsquigarrow$ D nem primitív.

Primitív és imprimitív mátrixok

- $\mathbf{E}$ irreducibilis és a főátlóján van pozitív elem $\rightsquigarrow$ primitív.
- Az $\mathbf{F}$ primitív, mivel $\mathbf{F}^5 = \begin{bmatrix} 27216 & 20412 & 31104 \\ 36288 & 54432 & 57348 \\ 23814 & 46656 & 54432 \end{bmatrix} > \mathbf{O}$,
- Egyszerűbb: $0 \mapsto 0$, pozitív $\mapsto 1$, szorzás $\mapsto$ AND, összeadás $\mapsto$ OR:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Sőt, csak négyzetreemelésekkel még gyorsabb:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Eszerint $\mathbf{F}^8 > \mathbf{O}$, tehát $\mathbf{F}$ primitív.

T

Perron–Frobenius-tétel – sajátértékek a spektrálkörön

Ha az $\mathbf{A} \geqslant \mathbf{O}$ *irreducibilis, akkor*

1. *a spektrálkör határára eső sajátértékek* 1 *multiplicitásúak, és* $\{r, r\varepsilon, \dots, r\varepsilon^{k-1}\}$ *alakba írhatók, ahol* $\varepsilon = \mathrm{e}^{2\pi\mathrm{i}/k}$,
2. $\mathbf{A}$ *primitív* $\iff$ $\forall\lambda \neq r$ *sajátértékére* $|\lambda| < r$,
3. $\mathbf{A}$ *pontosan akkor primitív, ha létezik a* $\lim_{k\to\infty}(\mathbf{A}/r)^k$ *határérték. Ekkor e határérték megegyezik az* $\mathbf{A}$ *spektrálfelbontásában szereplő, az* r *sajátértékhez tartozó vetítő mátrixszal (ahol* $\mathbf{p}$ *a Perron és* $\mathbf{q}$ *a bal Perron vektor), azaz*
$$\lim_{k\to\infty}(\mathbf{A}/r)^k = \frac{\mathbf{p}\mathbf{q}^\mathsf{T}}{\mathbf{q}^\mathsf{T}\mathbf{p}} > \mathbf{0},$$
4. *Ha* $\mathbf{A}$ *imprimitív, akkor*
$$\lim_{k\to\infty}\frac{\mathbf{I} + (\mathbf{A}/r) + (\mathbf{A}/r)^2 + \dots + (\mathbf{A}/r)^{k-1}}{k} = \frac{\mathbf{p}\mathbf{q}^\mathsf{T}}{\mathbf{q}^\mathsf{T}\mathbf{p}} > \mathbf{0}.$$

Sztochasztikus mátrixok

D A nemnegatív vektort sztochasztikusnak nevezzük, ha koordinátáinak összege 1 (azaz 1-normája 1). A nemnegatív $\mathbf{A}$ mátrix sztochasztikus, ha minden oszlopvektora sztochasztikus. (Sorsztochasztikus, ha minden sora).

- Ha $\mathbf{A}$ és $\mathbf{v}$ sztochasztikus, akkor $\mathbf{u} = \mathbf{A}\mathbf{v}$ is:

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_j = \sum_{j=1}^{n} v_j \sum_{i=1}^{n} a_{ij} = \sum_{j=1}^{n} v_j = 1.$$

- sztochasztikus mátrixok szorzata sztochasztikus mátrix.
- Az $\mathbf{A}$ mátrix pontosan akkor sztochasztikus, ha $\mathbf{A}$-nak az $\mathbf{1} = (1, 1, \dots, 1)$ vektor bal sajátvektora 1 sajátértékkel.

T Ha $\mathbf{S}$ sztochasztikus mátrix, akkor

1. $\lambda = 1$ egy sajátérték,
2. a spektrálsugara 1, és
3. ha $\mathbf{S}$ primitív, akkor $\lambda \neq 1$ esetén $|\lambda| < 1$.

Duplán sztochasztikus mátrixok*

D Duplán sztochasztikus, ha sor- és oszlopsztochasztikus is

- Duplán sztochasztikus mátrixok szorzata is duplán sztochasztikus.
- Minden permutáló mátrix duplán sztochasztikus.
- Ha $\mathbf{U} = [u_{ij}]$ unitér, akkor az $\mathbf{A} = [|u_{ij}|^2]$ mátrix duplán sztochasztikus, ugyanis $\sum_{i=1}^{n} |u_{ij}|^2 = \sum_{j=1}^{n} |u_{ij}|^2 = 1$.

T Frobenius–Kőnig-tétel: Az n-edrendű $\mathbf{A}$ mátrixban pontosan akkor eleme minden kígyónak a 0, ha $\mathbf{A}$ részmátrixai közt van olyan $s \times t$ méretű zérusmátrix, hogy $s + t = n + 1$.

B Ötlet:

$$\begin{array}{cc}
\begin{array}{c} \\ \\ \\ s \\ \\ \end{array} & \left[\begin{array}{ccc|cc} \ast & \ast & \ast & ? & ? \\ \ast & \ast & \ast & ? & ? \\ \hline 0 & 0 & 0 & \ast & \ast \\ 0 & 0 & 0 & \ast & \ast \\ 0 & 0 & 0 & \ast & \ast \end{array}\right] \\
 & t
\end{array}$$

K Minden duplán sztochasztikus mátrixban van legalább egy kígyó, melynek minden eleme pozitív.

T Birkhoff-tétel: Minden n-edrendű duplán sztochasztikus mátrix előáll permutáló mátrixok konvex lineáris kombinációjaként, azaz a duplán sztochasztikus mátrixok az $\mathbb{R}^{n\times n}$ térben olyan konvex poliédert alkotnak, melynek csúcsai a permutáló mátrixok. Azaz

$$\mathbf{A} = \sum_{i=1}^{k} c_i \mathbf{P}_i, \text{ ahol } c_1 + c_2 + \ldots + c_k = 1,\ c_i \geqslant 0.$$

P Például:

$$\begin{bmatrix} .3 & .4 & .3 \\ .2 & .5 & .3 \\ .5 & .1 & .4 \end{bmatrix}$$

$$= .3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + .3\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} + .2\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} + .1\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + .1\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Markov-láncok

Markov-lánc

- Egy rendszer következő állapota csak a pillanatnyi állapot függvénye, a múlté nem.
- Populációk fejlődése, bizonyos kémiai, termodinamikai, gazdasági folyamatok, tömegkiszolgálási és sorbanállási rendszerekben, statisztika, web-oldalak rangsorolása,...

D Legyen $\mathcal{S}$ egy megszámlálható halmaz (pl. $\mathcal{S} = \{1, 2, \ldots, N\}$, vagy $\mathcal{S} = \mathbb{N}$). Az $\mathcal{S}$-értékű valószínűségi változók egy X_0, X_1, X_2,..., X_n,... sorozata diszkrét paraméterű homogén Markov-lánc, ha

$$\mathbb{P}(X_{n+1} = i \mid X_n = j, X_{n-1} = k, \ldots, X_0 = \ell) = \mathbb{P}(X_{n+1} = i \mid X_n = j) \text{ és}$$
$$\mathbb{P}(X_{n+1} = i \mid X_n = j) = \mathbb{P}(X_1 = i \mid X_0 = j) = p_{ij},$$

Az $\mathcal{S}$ halmazt a Markov-lánc állapotterének nevezzük.

Markov-lánc lineáris algebrai modellje

- $\mathbf{p}_0 = (p_1, p_2, \dots)$ a kezdeti valószínűségeloszlás vektora ($p_i = \mathbb{P}(X_0 = i)$, $\mathbf{p} \geqslant \mathbf{0}$, $\sum_i p_i = 1$).
- A $\mathbf{P} = [p_{ij}]$ egy $|\mathcal{S}| \times |\mathcal{S}|$-es mátrix, az átmenetvalószínűségek mátrixa, vagy átmenetmátrix.
- A kezdeti állapotból az i-be való jutás valószínűsége

$$\mathbb{P}(X_1 = i) = \sum_j \mathbb{P}(X_1 = i \mid X_0 = j)\mathbb{P}(X_0 = j) = \sum_j p_{ij} p_j = [\mathbf{P}\mathbf{p}_0]_i,$$

- tehát a második állapot eloszlásvektora $\mathbf{P}\mathbf{p}_0$
- az n-edik állapot valószínűségeloszlása $\mathbf{p}_n = \mathbf{P}^n \mathbf{p}_0$ (általában $\mathbb{P}(X_{n+m} = i \mid X_n = j) = [\mathbf{P}^m]_{ij}$).

T Ha $\mathcal{S}$ egy megszámlálható halmaz, $\mathbf{p}$ egy valószínűségeloszlás $\mathcal{S}$-en, és $\mathbf{P}$ egy $|\mathcal{S}| \times |\mathcal{S}|$ méretű sztochasztikus mátrix, akkor létezik olyan $\mathcal{S}$ állapotterű Markov-lánc, melynek kezdeti eloszlása $\mathbf{p}$, és átmenetmátrixa $\mathbf{P}$.

Markov-lánc gráfos modellje: bolyongás a gráfon

- Minden Markov-lánc modellezhető egy súlyozott élű irányított gráfon való bolyongással.
- A gráf csúcsai az állapotok, és a j-edik csúcsból akkor vezet egy p_{ij} súlyú él az i-edikbe, ha $\mathbb{P}(X_1 = i \mid X_0 = j) = p_{ij}$, azaz a j-edik állapotot p_{ij} valószínűséggel követi az i-edik.
- A bolyongót letesszük a gráf egyik csúcsára a kezdeti $\mathbf{p}$ valószínűségeloszlás szerint, az időegységenként körbenéz, és a kifutó élekre írt valószínűségeknek megfelelően véletlenül választ közülük
- Mivel $\mathbf{P}$ sztochasztikus, e gráf minden csúcsából kifutó élek súlyainak összege 1.

Két kérdés

- Létezik-e a $\lim_{m \to \infty} \mathsf{p}_m$ határérték?
- Ha ez nem létezik, létezik-e a

$$\lim_{m \to \infty} \frac{\mathsf{p}_0 + \mathsf{p}_1 + \ldots \mathsf{p}_{m-1}}{m}$$

határérték függetlenül p_0 értékétől?

Markov-láncok

Néhány példa

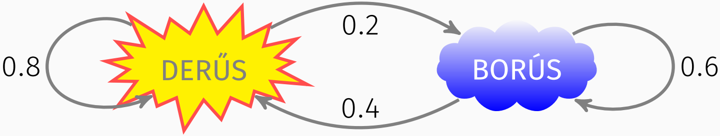
Időjárás

P Derűs napot 80% eséllyel derűs, míg borúst 60% eséllyel borús nap követ.

M Az átmenetmátrix

$$P = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$$

A folyamat gráfja:



Csön-csön gyűrű

P Páros sok gyerek körben ül, egyikük kezében rejtve egy gyűrű. Ritmusra mindenki úgy tesz, mintha egyik szomszédja kezébe adná a gyűrűt. Tfh minden játékos fix valószínűséggel, véletlenül választva adja át a gyűrűt.
A Markov-lánc állapota az, hogy kinél van a gyűrű.
A Markov-lánc átmenetmátrixában legyen $a_{i+1,i} = p_i$, $a_{i-1,i} = 1 - p_i$, ahol $p_i \in [0,1]$, és $i = 0, 1, \ldots, n-1$, számolás mod n, azaz

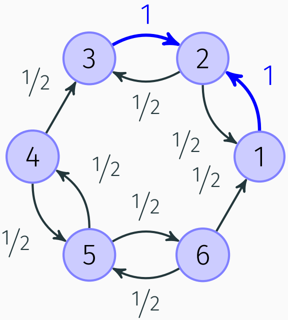
$$\mathbf{P} = \begin{bmatrix} 0 & 1-p_2 & 0 & \ldots & 0 & p_n \\ p_1 & 0 & 1-p_3 & \ldots & 0 & 0 \\ 0 & p_2 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 1-p_1 & 0 & 0 & \ldots & p_{n-1} & 0 \end{bmatrix}$$

Csön-csön gyűrű

m Mivel a résztvevők száma páros, ezért minden lépésben változik a Markov-lánc állapotának paritása, így a $\mathbf{p}_m$ vektorok határértéke nem létezik.

P egy 6-fős játék mátrixa:

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$



ha a gyűrű az $\{1, 2, 3\}$ halmazba kerül, onnan többé nem jut ki, ha egyszer elhagyja a $\{4, 5, 6\}$ halmazt, oda többé nem tér vissza.

Ki nevet a végén?

P A táblán a Starttól a Célig öt további mező van. A játékos dob, majd annyit lép a Cél felé, amennyi a dobás eredménye, de ha nagyobbat dob, mint amennyi a célba éréshez szükséges, vissza kell fordulnia. Akkor ér a Célba, ha épp ott fejezi be a lépéseket.

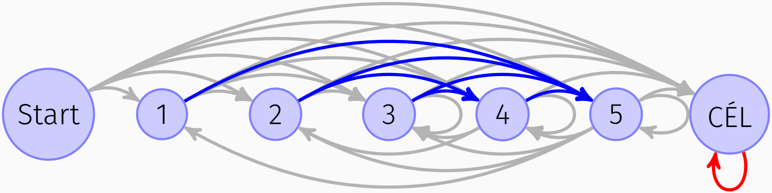
Start	1	2	3	4	5	CÉL

M A játékhoz tartozó átmenetmátrix

$$\mathbf{P} = \frac{1}{6}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 2 & 1 & 1 & 0 \\ 1 & 2 & 2 & 2 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 6 \end{bmatrix}$$

A kezdeti eloszlás $(1, 0, 0, 0, 0, 0, 0)$, és a Start-ba sosem jutunk

Ki nevet a végén?



A szürke élekhez $1/6$, a kékekhez $2/6$, míg a piroshoz 1 valószínűség tartozik.

Azt sejtjük, hogy a játékos 1 valószínűséggel véges időn belül CÉL-ba ér, ezért az állapotvektorok határértéke $(0, 0, 0, 0, 0, 0, 1)$.

Markov-láncok

Osztályozás

Az állapotok osztályozása

D A j állapotból az i elérhető (jelölése $j \to i$), ha van olyan $n \geqslant 0$ egész, hogy $\mathbb{P}(X_n = i \mid X_0 = j) > 0$. (algebra: van olyan n, hogy $[\mathbf{P}^n]_{ij} > 0$; gráf: van irányított út j-ből i-be)

D i és j állapotok érintkeznek (közlekednek, $i \leftrightarrow j$), ha $i \to j$ és $j \to i$

Á E reláció osztályozza az állapotokat (ekvivalencia-reláció).

P A „Ki nevet a végén?"-ben három osztály (Start, Cél, többi), a „Csön-csön gyűrű"-ben két osztály ($\{1, 2, 3\}$, $\{4, 5, 6\}$) van.

D Egy Markov-lánc irreducibilis, ha egyetlen osztályból áll.

T A Markov-lánc pontosan akkor irreducibilis, ha átmenetmátrixa irreducibilis, azaz $\forall\ (i, j)$ párhoz $\exists\ m$, hogy $[\mathbf{P}^m]_{ij} > 0$.

P Az „Időjárásmodell" irreducibilis, a „Csön-csön gyűrű" lehet reducibilis és irreducibilis is, a megadott konkrét esetben reducibilis, a „Ki nevet a végén?" reducibilis.

Az állapotok osztályozása

D Az i állapot d_i periódusa azon kísérletek sorszámának legnagyobb közös osztója, amelyekben a Markov-lánc az i állapotból indulva visszatérhet i-be, azaz $d_i = \mathrm{lnko}\,\{\, n > 0 : \mathbb{P}(X_n = i \mid X_0 = i) > 0 \,\}$.

P Például a „Csön-csön gyűrű” játék mindegyik állapotának 2 a periódusa.

D Az állapot aperiodikus, ha $d_i = 1$. A Markov-lánc aperiodikus, ha minden állapota aperiodikus.

P Az „Időjárásmodell” és a „Ki nevet a végén?” aperiodikus.

D Az i állapot visszatérő, ha a Markov-lánc az i-ből indulva 1 valószínűséggel visszatér az i-be, azaz $\mathbb{P}(\exists n > 0 : X_n = i \mid X_0 = i) = 1$.

D Egy állapot átmeneti, ha nem visszatérő.

P A „Csön-csön gyűrű” $\{\,1, 2, 3\,\}$-beli állapotai visszatérők, a $\{\,4, 5, 6\,\}$-beliek átmenetiek.

Az állapotok osztályozása

m Általában is igaz, hogy a visszatérés, az átmenetiség és a periódus ún. osztálytulajdonság, azaz egy osztály minden elemére azonos.

Á Egy véges állapotterű Markov-láncban egy osztály pontosan akkor átmeneti, ha gráfján vezet ki belőle él, és pontosan akkor visszatérő, ha nem. Ha a Markov-lánc elhagy egy átmeneti osztályt, akkor oda többé nem jut vissza, ha belép egy visszatérő osztályba, akkor onnan többé nem tud kijönni. Minden Markov-lánc állapottere diszjunkt átmeneti és visszatérő osztályok uniója.

P A „Csön-csön gyűrű” (csupa pozitív valószínűség esetén) és az Időjárámodell állapotai egyetlen visszatérő osztályt alkotnak, A 6-fős változata egy visszatérő és egy átmeneti osztályból áll.

P A „Ki nevet a végén?” játék két átmeneti és egy visszatérő osztály uniója.

Irreducibilis Markov-láncok

m A továbbiakban kizárólag csak véges állapotterű Markov-láncokkal foglalkozunk.

D A $\mathbf{P}$ átmenetmátrixú véges Markov-lánc állapotterén értelmezett valamely $\boldsymbol{\pi}$ eloszlásvektor stacionárius, ha $\mathbf{P}\boldsymbol{\pi} = \boldsymbol{\pi}$.

m Primitív mátrixok hatványainak határértéke megegyezik a jobb és bal Perron-vektor diadikus és skaláris szorzatának hányadosával. Mivel egy $n \times n$-es átmenetmátrix bal Perron-vektora $\frac{1}{n}\mathbf{1}$, ahol $\mathbf{1}$ a csupa-1 vektor, ezért ha $\boldsymbol{\pi}$ jelöli a jobb Perron-vektort, akkor

$$\lim_{m\to\infty} \mathbf{P}^m = \frac{\boldsymbol{\pi}(\mathbf{1}/n)^\mathsf{T}}{\boldsymbol{\pi}^\mathsf{T}(\mathbf{1}/n)} = \boldsymbol{\pi}\mathbf{1}^\mathsf{T}.$$

K $\lim\limits_{m\to\infty} \mathbf{p}_m = \boldsymbol{\pi}$, ugyanis tetszőleges $\mathbf{p}_0$ eloszlásvektorra $\mathbf{1}^\mathsf{T}\mathbf{p}_0 = 1$, így

$$\lim_{m\to\infty} \mathbf{p}_m = \lim_{m\to\infty} \mathbf{P}^m\mathbf{p}_0 = \boldsymbol{\pi}\mathbf{1}^\mathsf{T}\mathbf{p}_0 = \boldsymbol{\pi}.$$

Irreducibilis Markov-láncok

P Az Időjárásmodell esetén a $\mathsf{P} = \left[\begin{smallmatrix} .8 & .4 \\ .2 & .6 \end{smallmatrix}\right]$ átmenetmátrix primitív, az 1 sajátértékhez tartozó jobb sajátvektora, s vele a stacionárius eloszlás $\boldsymbol{\pi} = (2/3, 1/3)$, vagyis a napoknak 2/3-a derűs. Másrészt

$$\lim_{m\to\infty} \mathsf{P}^m = \begin{bmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix}.$$

P A „Ki nevet a végén?” ugyan reducibilis, de K átmenetmátrixának az 1 egyszeres multiplicitású sajátértéke és nincs több sajátérték a spektrálkörén, így létezik a $\lim_{n\to\infty} \mathsf{K}^n$ határérték, a jobb sajátvektora fejben számolva is ellenőrizhetően $\boldsymbol{\pi} = (0, 0, 0, 0, 0, 0, 1) = \mathbf{e}_7$, így az állapotvektorok határértéke $\mathbf{e}_7$, vagyis valóban a CÉL-ban végzünk (1 valószínűséggel).

Irreducibilis és imprimitív Markov-láncok

- Ha $\mathbf{P}$ irreducibilis, de imprimitív (ekkor több sajátérték van a spektrálkörön, pl. a csupa pozitív valószínűséggel definiált „Csön-csön gyűrű” ilyen, sajátértékei 1 és -1), akkor létezik ugyan stacionárius megoldás, de az nem az állapotvektorok határértéke. A stacionárius vektor i-edik koordinátája megadja, hogy a Markov-lánc „idejének” átlagosan hányad részét tölti az i-edik állapotban (mint a primitív esetben).
- Az állapotvektoroknak nincs határértékük, de átlaguknak igen, a stacionárius vektor, ugyanis

$$\lim_{m\to\infty} \frac{\mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^{m-1}}{m} = \boldsymbol{\pi}\mathbf{1}^\mathsf{T},$$

amiből

$$\lim_{m\to\infty} \frac{\mathbf{p}_0 + \mathbf{p}_1 + \cdots + \mathbf{p}_{m-1}}{m} = \boldsymbol{\pi}.$$

Reducibilis Markov-láncok

- A korábbi 6 fős „Csön-csön gyűrű” reducibilis ugyan, de a $\boldsymbol{\pi} = (1/4, 1/2, 1/4, 0, 0, 0)$ stacionárius vektor még kifejezi, hogy a Markov-lánc „idejének” átlagosan hányad részét tölti az *i*-edik állapotban (az átmeneti osztályban töltött idő „elenyészik”).

P Ha azonban több visszatérő osztály is van, akkor már ez sem igaz. P. a következő esetben két stacionárius vektor is van: $(1/2, 1/2, 0, 0, 0, 0)$, $(0, 0, 1/4, 1/2, 1/4, 0)$:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$