

PS

(curs 11 - 511)

Lanțuri Markov

- **spațiul stărilor sau al nodurilor unei rețele:** un sistem reprezentat de o mulțime finită de stări (noduri) $S = \{1, 2, \dots, m\}$ sau infinit numărabilă $S = \mathbb{N}$;
- Schimbările de stare se produc la întâmplare, la momente discrete de timp $t = 0, 1, 2, \dots, n, \dots$
- Fiecărui moment de timp $n \in \mathbb{N}$ i se asociază o variabilă aleatoare X_n ce ia valori în mulțimea nodurilor:

$$X_n = \begin{pmatrix} 1 & 2 & \dots & m \\ \pi_n(1) & \pi_n(2) & \dots & \pi_n(m) \end{pmatrix},$$

unde $\pi_n(i)$ este probabilitatea ca la momentul n informația să ajungă în nodul $i \in S$.

Definiție

Un șir de variabile aleatoare (X_n) , $n \in \mathbb{N}$, definite pe același spațiu de probabilitate (Ω, \mathcal{K}, P) cu valori în mulțimea stărilor (nodurilor) S , definește un **lanț Markov** dacă

$$P(X_{n+1} = j | X_0 = s_0, \dots, X_{n-1} = s_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i)$$

Proprietatea markoviană: "lipsa parțială de memorie" a lanțului Markov: doar istoria recentă, nu și cea trecută, influențează evoluția viitoare.

Un lanț Markov definește o lege de mișcare la întâmplare pe mulțimea nodurilor.

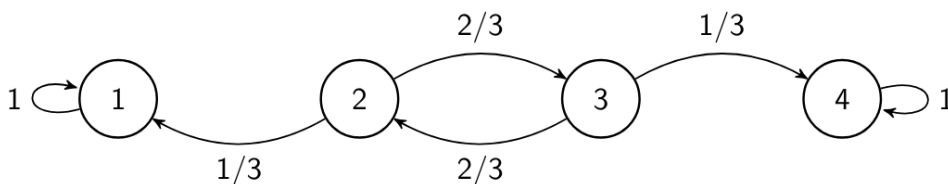
Lanțuri Markov omogene

Definiție

Lanț Markov (X_n) **omogen:** $P(X_{n+1} = j | X_n = i)$ nu depinde de n .

p_{ij} - probabilitate de trecere într-un singur pas din nodul i în nodul j

$$p_{ij} = P(X_{n+1} = j | X_n = i), \forall i, j \in \{1, m\}$$



Matricea de tranziție

Definim **Q-matricea de tranziție a lanțului Markov**-cu elementele

$$Q(i, j) := p_{ij}, i, j \in \{1, m\}. \text{ Avem: } Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- 1 Matrice stohastică: $p_{ij} \geq 0$, $(i, j) \in S \times S$; și suma elementelor de pe fiecare linie este 1: $\sum_{j=1}^m p_{ij} = 1$, $i \in S$;
- 2 Linia i = *vector stohastic/probabilist*, iar elementele ei sunt probab. ca lanțul Markov să treacă din starea i respectiv în stările $1, 2, \dots, m$.

Proprietățile matricelor stohastice

- $Qe = e$, unde $e = [1, 1, \dots, 1]^T$, deci e este vector propriu al matricei Q , corespunzător valorii proprii $\lambda = 1$.
- Produsul a două matrice stohastice P, Q este o matrice stohastică;
- dc. Q este matricea de tranziție a unui lanț Markov, atunci Q^n -stohastică, $n \in \mathbb{N}$.

Mulțimea nodurilor unui lanț Markov și matricea de tranziție definesc un graf orientat: **graf de tranziție al lanțului Markov**. Există arc orientat de la i la j , dacă probabilitatea p_{ij} este nenulă.

Exemplu: Fie $S = \{1, 2, 3\}$ mulțimea nodurilor unei rețele și Q matricea de tranziție de la un nod la altul:

$$Q = \begin{pmatrix} 0.4 & 0.5 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}.$$

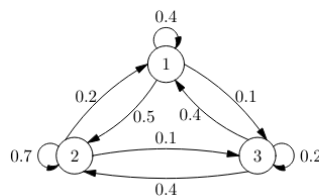


Figure: Graful de tranziție al unui lanț Markov. Pe fiecare arc este indicată probabilitatea de trecere între nodurile conectate de arc.

Probabilitatea ca lanțul să treacă din nodul inițial i în nodul j după n pași:

$$P(X_n = j | X_0 = i) = Q^n(i, j).$$

$(X_2 = j | X_0 = i)$ se poate scrie astfel:

$$(X_2 = j | X_0 = i) = \bigcup_{k=1}^m (X_2 = j, X_1 = k | X_0 = i),$$

$$\text{deci } P(X_2 = j | X_0 = i) = \frac{\sum_{k=1}^m P(X_0 = i, X_1 = k, X_2 = j)}{P(X_0 = i)}$$

$$= \frac{\sum_{k=1}^m P(X_0 = i)P(X_1 = k | X_0 = i)P(X_2 = j | X_0 = i, X_1 = k)P(X_2=j|X_1=k)}{P(X_0 = i)} =$$

$$\sum_{k=1}^m Q(i, k)Q(k, j) = Q^2(i, j).$$

Probabilitatea ca lanțul să evolueze pe traiectoria $s_0, s_1, s_2, \dots, s_n$ este:

$$P(X_0 = s_0, X_1 = s_1, \dots, X_n = s_n) = \pi_0(s_0)Q(s_0, s_1)Q(s_1, s_2) \cdots Q(s_{n-1}, s_n).$$

Demo:

$$\begin{aligned} P(X_0 = s_0, X_1 = s_1, \dots, X_{n-1} = s_{n-1}, X_n = s_n) \\ &= P(X_0 = s_0)P(X_1 = s_1|X_0 = s_0)P(X_2 = s_2|X_0 = s_0, X_1 = s_1) \cdots \\ &\quad \cdots P(X_n = s_n|X_0 = s_0, X_1 = s_1, \dots, X_{n-1} = s_{n-1}) \\ &= \pi_0(s_0)P(X_1 = s_1|X_0 = s_0)P(X_2 = s_2|X_1 = s_1) \cdots P(X_n = s_n|X_{n-1} = s_{n-1}) \\ &= \pi_0(s_0)Q(s_0, s_1)Q(s_1, s_2) \cdots Q(s_{n-1}, s_n). \end{aligned}$$

O realizare a lanțului Markov (X_n) sau o observație asupra lanțului este un șir de noduri ce pot fi vizitate de lanț, $(s_0, s_1, \dots, s_n, \dots)$, $s_k \in S$, și se numește **traiectorie a lanțului Markov**.

Pentru a putea analiza și simula un lanț Markov, trebuie:

- precizată **distribuția inițială de probabilitate**: distribuția de probabilitate a variabilei aleatoare discrete X_0 :
$$X_0 = \begin{pmatrix} 1 & 2 & \dots & m \\ \pi_0(1) & \pi_0(2) & \dots & \pi_0(m) \end{pmatrix}.$$
- simulat în mod iterativ.
- **Algoritmul de simulare a unui LM** este prototip pentru *clasa algoritmilor iterativi aleatori*.

Algoritmul de generare a segmentului de traiectorie s_0, s_1, \dots, s_n

```
begin algorithmic[1]
```

```
FunctionLantMarkov(m,  $\pi_0$ , Q, n)
```

```
State  $s_0$  = simulator(1, 2, ..., m;  $\pi_0$ );
```

```
State i =  $s_0$ ;
```

```
For k = 1 : n
```

```
State p = Q[i, :]; //Q[i, :] = [linia i din matricea Q];
```

```
State  $s_k$  = simulator(1, 2, ..., m; p);
```

```
State i =  $s_k$ ;
```

```
EndFor
```

```
State return  $s_0, s_1, \dots, s_n$ ;
```

```
EndFunction1
```

Distribuția de probabilitate la momentul n

Ne interesează variabila aleatoare X_n , adică distribuția ei de probabilitate, notată cu $\pi_n = [\pi_n(1), \pi_n(2), \dots, \pi_n(m)]^T$, $n \in \mathbb{N}^*$.

- ia valorile $\{1, 2, \dots, m\}$;
- evenimentul $(X_n = j)$ este evenimentul ca la momentul n traiectoria aleatoare să ajungă în nodul $j \in S$;
- notăm cu $\pi_n(j) = P(X_n = j)$;
- este suficient a se cunoaște distribuția inițială de probabilitate π_0 și matricea de tranziție Q a lanțului

$$\pi_n^T = \pi_0^T Q^n$$

sau detaliat:

$$[\pi_n(1) \ \pi_n(2) \ \dots \ \pi_n(m)] = [\pi_0(1) \ \pi_0(2) \ \dots \ \pi_0(m)] Q^n.$$

Deci, distribuția de probabilitate X_n a stării la momentul n se poate calcula recursiv pornind de la distribuția inițială π_0 :

$$\begin{aligned} \pi_1^T &= \pi_0^T Q \\ \pi_2^T &= \pi_0^T Q^2 = \pi_1^T Q \\ &\vdots \\ \pi_n^T &= \pi_{n-1}^T Q, \quad n \in \mathbb{N}^*. \end{aligned} \tag{1}$$

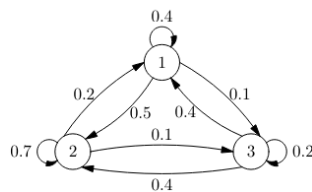


Figure: Graful de tranziție al unui lanț Markov

Dacă considerăm distribuția inițială de probabilitate

$\pi_0 = [0.2, 0.35, 0.45]^T$, distribuțiile de probabilitate ale variabilelor aleatoare (stărilor) X_1, X_2, \dots, X_{10} , calculate conform relației de recurență de mai sus, sunt:

$$\begin{aligned} \pi_1 &= [0.3300000000 \quad 0.5250000000 \quad 0.1450000000]^T \\ \pi_2 &= [0.2950000000 \quad 0.5905000000 \quad 0.1145000000]^T \\ \pi_3 &= [0.2819000000 \quad 0.6066500000 \quad 0.1114500000]^T \\ \pi_4 &= [0.2786700000 \quad 0.6101850000 \quad 0.1111450000]^T \\ \pi_5 &= [0.2779630000 \quad 0.6109225000 \quad 0.1111145000]^T \\ \pi_6 &= [0.2778155000 \quad 0.6110730500 \quad 0.1111114500]^T \\ \pi_7 &= [0.2777853900 \quad 0.6111034650 \quad 0.1111111450]^T \\ \pi_8 &= [0.2777793070 \quad 0.6111095785 \quad 0.1111111145]^T \\ \pi_9 &= [0.2777780843 \quad 0.6111108042 \quad 0.1111111114]^T \\ \pi_{10} &= [0.2777778391 \quad 0.6111110497 \quad 0.1111111114]^T. \end{aligned}$$

Dacă se continuă calculul distribuțiilor de probabilitate $\pi_{50}, \pi_{51}, \dots, \pi_{100}$ și se afișează coordonatele cu 15 zecimale, se obține

$$\pi_{50} = \pi_{51} = \dots = \pi_{100} = [0.277777777778 \quad 0.611111111111 \quad 0.111111111111]^T.$$

- Avem $P(X_{50} = k) = \dots = P(X_{100} = k)$, pentru $k = 1, 2, 3$
- Orice distribuție de probabilitate π_n , cu $n \geq 50$, am calcula cu 15 zecimale am obține distribuții identice cu π_{50} , adică după momentul $n = 50$ distribuția este staționară, nu se mai modifică în primele 15 zecimale și deci sistemul a ajuns într-un echilibru.
- De-a lungul oricărei traiectorii începând cu momentul $n = 50$, $s_{50}, s_{51}, \dots, s_{50+N}$, de lungime $N + 1$ avem
 - starea 1 este "vizitată" de lanțul Markov în proporție de $100 \pi(1)\% = 27.7777777778\%$,
 - starea 2 în proporție de $100 \pi(2)\% = 61.1111111111\%$,
 - starea 3 în proporție de $100 \pi(3)\% = 11.1111111111\%$.
- pentru $n > 50$ avem $\|\pi_n - \pi_{n-1}\| < 10^{-15}$.

Distribuție de echilibru



Universitatea Politehnică Timișoara

- Dacă șirul (π_n) al distribuțiilor de stare este convergent, atunci limita sa este un vector probabilist π , adică un vector de coordonate mai mari sau egale cu zero și suma coordonatelor este 1.
- Dacă șirul (π_n) al distribuțiilor de stare la momentul n converge la π , atunci are loc:

$$\pi^T = \pi^T Q.$$

Definiție

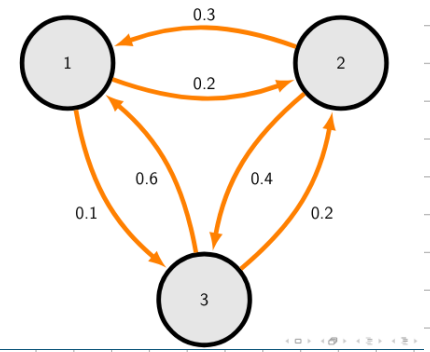
O distribuție de probabilitate π , pe spațiul nodurilor unui lanț Markov, cu proprietatea că $\pi^T = \pi^T Q$ se numește *distribuție invariantă*, *staționară* sau *distribuție de echilibru*.

- Dacă există distribuția de echilibru $\pi = [\pi(1), \pi(2), \dots, \pi(m)]^T$ (ca limită a șirului π_n), atunci $\pi(j)$ reprezintă șansa asimptotică de a fi vizitat nodul j .
- Deci, dacă mișcarea aleatoare pe S continuă indefinit și șirul (π_n) este convergent, atunci de la un moment dat mișcarea aleatoare se stabilizează și vizitează fiecare nod $j \in \{1, 2, \dots, m\}$ cu aceeași frecvență $\pi(j)$.

Definiție

Un lanț Markov pe $S = \{1, 2, \dots, m\}$ se numește *lanț ireductibil*, dacă pentru oricare două noduri $i, j \in S$ există $n = n(i, j) \in \mathbb{N}^*$ astfel încât $Q^n(i, j) > 0$, adică cu probabilitate nenulă lanțul Markov poate trece într-un număr de pași din nodul i în nodul j .

- Practic, lanțul este ireductibil dacă și numai dacă graful de tranziție este tare conex (există drum de arce între orice două noduri).
- Dacă matricea de tranziție are toate elementele $Q(i, j) > 0$, $i, j \in 1, m$, atunci lanțul Markov este ireductibil.
- Simplul fapt că $Q(i, j) = 0$ nu asigură că i nu comunică cu j . Cele două noduri nu comunică într-un pas, dar pot comunica în mai mulți pași, adică s-ar putea ca $Q^n(i, j) > 0$, pentru un $n > 1$.



Un lanț Markov poate avea și *traectorii periodice*, adică traectorii în care succesiunea de noduri $s_0, s_1, \dots, s_{T-1}, s_T = s_0$ se repetă indefinit într-o traectorie a lanțului.

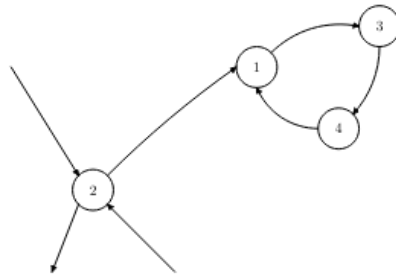


Figure: Graful de tranziție al unui lanț Markov ce are traiectoria periodică (1,3,4,1)

Definiție

Perioada unui nod $i \in S$ este numărul

$$\tau_i = c.m.m.d.c. \{n \in \mathbb{N}^* \mid Q^n(i, i) > 0\}.$$

- Un nod i a cărui perioadă este 1 se numește **nod aperiodic**
- un lanț Markov care are toate nodurile aperiodice se numește **lanț Markov aperiodic**.
- Cel mai simplu exemplu de nod aperiodic este un nod i pentru care $Q(i, i) > 0$.
- Dacă matricea de tranziție are toate elementele de pe diagonala principală strict pozitive, $Q(i, i) > 0$, atunci lanțul este aperiodic.

Observatie: Proprietățile de ireductibilitate și aperiodicitate ale unui lanț Markov sunt proprietăți ale matricei de tranziție.

- Un lanț Markov ireductibil are toate nodurile de aceeași perioadă.
- Dacă un nod al unui lanț ireductibil este aperiodic, atunci toate nodurile sunt aperiodice.
- Pentru a arăta că un lanț ireductibil este aperiodic este suficient să identificăm un nod i pentru care $Q(i, i) > 0$.

Dacă un lanț Markov pe S este ireductibil și aperiodic, atunci oricare ar fi distribuția inițială de probabilitate π_0 , șirul distribuțiilor de probabilitate la momentul n asociat, (π_n) , este convergent.

- Limita acestui șir este un vector probabilist π care nu depinde de distribuția inițială de probabilitate (deci indiferent de distribuția inițială de probabilitate, șirurile asociate converg la aceeași limită π).
- Această limită este unica distribuție de echilibru a lanțului Markov.
- Mai mult, șirul (Q^n) este convergent la o matrice de rang 1 având fiecare linie egală cu π^T :

$$\lim_{n \rightarrow \infty} Q^n = \begin{pmatrix} \pi(1) & \pi(2) & \dots & \pi(m) \\ \pi(1) & \pi(2) & \dots & \pi(m) \\ \vdots & \vdots & \dots & \vdots \\ \pi(1) & \pi(2) & \dots & \pi(m) \end{pmatrix}.$$

Cum determinăm pe π ?

π este vector propriu probabilist al matricei Q^T , corespunzător valorii proprii 1 pentru ca verifică

$$Q^T \pi = \pi,$$

(distribuția de echilibru π verifică relația $\pi^T = \pi^T Q$)

Acest rezultat matematic a influențat modul de definire al navigării aleatoare pe graful WEB, ca un lanț Markov ireductibil și aperiodic pe mulțimea paginilor WEB (vezi Material Suplimentar).