Curs 11: Lanţuri Markov. Definiţie, proprietăţi, simulare

1.1 Lanţuri Markov: generalităţi

Lanţurile Markov se folosesc în modelarea şi simularea unor sisteme în care se produc anumite evenimente la momente discrete de timp $t = 0, 1, 2, \ldots$. Astfel de lanţuri se folosesc în design-ul algoritmilor de rutare, al protocoalelor pentru reţele wireless, în controlul mişcării roboţilor, ranking-ul paginilor WEB şi al echipelor sportive, în analiza protocoalelor de management a memoriei, studiul performanţei serverelor, în algoritmi randomizaţi, în machine learning (de exemplu, în computational advertising) etc. Există chiar şi un limbaj de modelare pentru sisteme hardware/software în timp real, numit *POOSL, Parallel Object Oriented Specification Language*, care generează un lanţ Markov, ca model al sistemului.

Un sistem este reprezentat de o mulţime finită de stări (noduri) $S = \{1, 2, \ldots, m\}$ sau infinit numărabilă $S = \mathbb{N}$. Mulţimea S se numeşte spaţiul stărilor sau al nodurilor unei reţele. Schimbările de stare se produc la întâmplare, la momente discrete de timp $t = 0, 1, 2, \ldots, n, \ldots$. În continuare vom discuta preponderent exemple în care la momente discrete de timp se produce o transmitere de informaţie de la un nod din reţea spre altul sau ”un călător virtual” trece de la un nod spre altul. Nodul spre care se transmite informaţia (sau nodul în care trece călătorul) depinde de diverse circumstanţe şi astfel informaţia (călătorul) are o traiectorie aleatoare în reţea.

Fiecărui moment de timp $n \in \mathbb{N}$ i se asociază o variabilă aleatoare X_n ce ia valori în mulţimea nodurilor:

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

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

Definiţia 1.1.1 Un şir de variabile aleatoare (X_n), $n \in \mathbb{N}$, definite pe acelaşi spaţiu de probabilitate $(\Omega, \mathcal{K}, P)$ cu valori în mulţimea stărilor (nodurilor) S, defineşte un *lanţ Markov* dacă probabilitatea ca informaţia să treacă la momentul $n + 1$ în nodul j ştiind că în momentele de timp anterioare se afla respectiv în nodurile $s_0, s_1, \ldots, s_{n-1}, i$ este

$$P(X_{n+1} = j | X_0 = s_0, X_1 = s_1, \ldots, X_{n-1} = s_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i). \quad (1)$$

Probabilitatea $P(X_{n+1} = j | X_n = i)$ se citeşte "probabilitatea ca informaţia să treacă în nodul j la momentul $n+1$, ştiind că se află în nodul i la momentul n". Relaţia (1) se numeşte *proprietate markoviană*. Aceasta caracterizează "lipsa parţială de memorie" a lanţului Markov: cunoscând succesiunea de noduri $s_0, s_1, \ldots, s_{n-1}, i$, prin care informaţia a trecut până la momentul prezent n, doar nodul i, în care se află în prezent, influenţează probabilitatea de trecere în momentul următor într-un alt nod, nu şi drumul parcurs până la momentul curent n (cu alte cuvinte doar istoria recentă, nu şi cea trecută, influenţează evoluţia viitoare).

În mod normal probabilităţile $P(X_{n+1} = j | X_n = i)$ depind de n, adică

$$P(X_{n+1} = j | X_n = i) = p_{ij}(n).$$

Un lanţ Markov (X_n) se numeşte *omogen*, dacă probabilităţile $P(X_{n+1} = j | X_n = i)$ nu depind de n.

În continuare ne referim la lanţuri Markov omogene cu o mulţime finită de stări, iar în loc de trecerea informaţiei sau a călătorului virtual de la un nod la altul, spunem trecerea lanţului Markov.

Pentru un lanţ Markov omogen, notăm cu $p_{ij} = P(X_{n+1} = j | X_n = i)$, $i, j = \overline{1, m}$, probabilitatea ca la momentul $n+1$ lanţul Markov să treacă în nodul j, ştiind că la momentul n se afla în nodul i. p_{ij} se numeşte probabilitate de trecere într-un singur pas din nodul i în nodul j, iar matricea Q, de elemente $Q(i, j) := p_{ij}$, $i, j = \overline{1, m}$, se numeşte *matricea de tranziţie a lanţului Markov*.

În concluzie, un lanţ Markov defineşte o lege de mişcare la întâmplare pe mulţimea nodurilor.

Matricea de tranziţie $Q = (p_{ij})_{i,j \in S}$ are proprietăţile:

1. $p_{ij} \geq 0, \forall\ (i, j) \in S \times S$;

2. $\sum\limits_{j=1}^{m} p_{ij} = 1$, $\forall\ i \in S$, adică suma elementelor de pe fiecare linie este 1.

O astfel de matrice se numeşte *matrice stochastică*, iar liniile ei se numesc *vectori stochastici* sau *vectori probabilişti*. Elementele de pe linia i, $i = \overline{1, m}$, a matricei Q, $p_{i1}, p_{i2}, \ldots, p_{im}$, indică probabilităţile ca din starea i lanţul Markov să treacă respectiv în stările $1, 2, \ldots, m$.

Proprietăţi ale matricelor stochastice

- Notăm cu $\mathbf{e}$ vectorul care are toate coordonatele egale cu 1, adică $\mathbf{e} = [1, 1, \ldots, 1]^T$.

Astfel produsul $Q\mathbf{e}$ este:

$$\begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ & & \vdots & \\ p_{k1} & p_{k2} & \dots & p_{km} \\ & & \vdots & \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} + p_{12} + \dots + p_{1m} \\ \vdots \\ p_{k1} + p_{k2} + \dots + p_{km} \\ \vdots \\ p_{m1} + p_{m2} + \dots + p_{mm} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Această relaţie ne permite exprimarea concentrată a proprietăţii lui Q de a avea toate liniile vectori stochastici:

$$Q\mathbf{e} = \mathbf{e}.$$

Relaţia de mai sus exprimă şi faptul că $\mathbf{e}$ este vector propriu al matricei Q, corespunzător valorii proprii $\lambda = 1$. În continuare o vom folosi ca relaţie de definiţie a unei matrice stochastice (subînţelegând că elementele ei sunt mai mari sau egale cu zero).

• Produsul a două matrice stochastice P, Q este o matrice stochastică, deoarece $P\mathbf{e} = \mathbf{e}$ şi $Q\mathbf{e} = \mathbf{e}$ implică $(PQ)\mathbf{e} = P(Q\mathbf{e}) = P\mathbf{e} = \mathbf{e}$.

Ca o consecinţă a acestei proprietăţi rezultă că dacă Q este matricea de tranziţie a unui lanţ Markov, atunci Q^n este matrice stochastică, $\forall\, n \in \mathbb{N}$.

• Dacă P, Q sunt matrice stochastice şi $\alpha \in (0, 1)$, atunci combinaţia convexă

$$M = \alpha P + (1 - \alpha)Q$$

este o matrice stochastică.

Mulţimea nodurilor unui lanţ Markov şi matricea de tranziţie definesc un graf orientat. Există arc orientat de la i la j, dacă probabilitatea p_{ij} este nenulă. Graful astfel asociat se numeşte *graf de tranziţie al lanţului Markov*.

Exemplul 1. Fie $S = \{1, 2, 3\}$ mulţimea nodurilor unei reţele şi Q matricea de tranziţie de la un nod la altul:

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

Graful asociat este vizualizat în Fig. 1.

1.1.1 Simularea unui lanţ Markov

O realizare a lanţului Markov (X_n) sau o observaţie asupra lanţului este un şir de noduri ce pot fi vizitate de lanţ, $(s_0, s_1, \dots, s_n, \dots)$, $s_k \in S$, şi se numeşte **traiectorie a lanţului Markov**.

Pentru a putea analiza şi simula un lanţ Markov, trebuie precizată **distribuţia iniţială de probabilitate**, care dă probabilitatea ca traiectoria aleatoare a lanţului să pornească

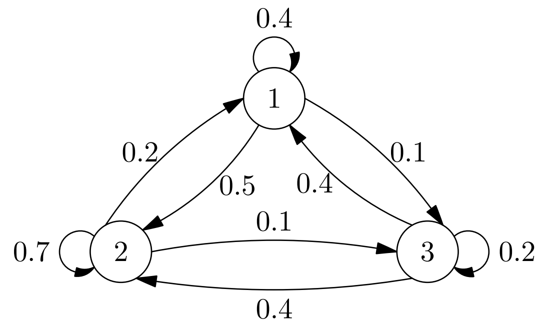

Fig.1: Graful de tranziţie al unui lanţ Markov. Pe fiecare arc este indicată probabilitatea de trecere între nodurile conectate de arc.

dintr-un nod i. Mai precis, distribuţia iniţială de probabilitate este un vector probabilist (vector cu coordonatele în [0,1] şi suma coordonatelor egală cu 1):

$$\pi_0 = [\pi_0(1), \pi_0(2), \ldots, \pi_0(m)]^T, \quad \pi_0(k) \geq 0, \quad \sum_{k=1}^{m} \pi_0(k) = 1,$$

unde $\pi_0(k) = P(X_0 = k)$ este probabilitatea ca la momentul $t = 0$ lanţul să pornească din nodul k. Cu alte cuvinte, distribuţia iniţială de probabilitate a lanţului Markov este distribuţia de probabilitate a variabilei aleatoare discrete X_0:

$$X_0 = \begin{pmatrix} 1 & 2 & \ldots & m \\ \pi_0(1) & \pi_0(2) & \ldots & \pi_0(m) \end{pmatrix}. \tag{2}$$

Dacă distribuţia iniţială de probabilitate este, de exemplu, $\pi_0 = [0, 1, 0, \ldots, 0]^T$, atunci spunem că lanţul porneşte sigur (adică cu probabilitatea 1) din nodul 2.

Dacă pentru lanţul Markov din Exemplul 1 distribuţia iniţială de probabilitate este

$$X_0 = \begin{pmatrix} 1 & 2 & 3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix},$$

atunci înseamnă că probabilitatea ca un mers (drum) aleator în mulţimea $S = \{1, 2, 3\}$ să pornească, de exemplu, din starea 2 este 0.5.

Un lanţ Markov este simulat în mod iterativ. **Algoritmul de simulare a unui LM** este prototip pentru *clasa algoritmilor iterativi aleatori*. Având dat spaţiul stărilor $S = \{1, 2, \ldots, m\}$, distribuţia iniţială de probabilitate $\pi_0 = [\pi_0(1), \pi_0(2), \ldots, \pi_0(m)]^T$ şi matricea de tranziţie $Q = (p_{ij})$, $i, j = \overline{1, m}$, pentru un lanţ Markov (X_n), putem genera o traiectorie aleatoare $s_0, s_1, \ldots, s_n$ astfel:

- Construim simulatorul unei variabile aleatoare discrete arbitrare:

$$Y = \begin{pmatrix} 1 & 2 & \ldots & m \\ \mathtt{p_1} & \mathtt{p_2} & \ldots & \mathtt{p_m} \end{pmatrix},$$

pe care îl notăm simbolic cu

$$\texttt{simulator}(1, 2, \ldots, \mathtt{m}; \mathbf{p}),$$

unde $\mathbf{p}$ este vectorul de probabilitate al variabilei Y. Prin $\mathtt{j} = \texttt{simulator}(1, 2, \ldots, \mathtt{m}; \mathbf{p})$ simbolizăm faptul că simulatorul generează numărul (starea) j.

• Se generează starea iniţială s_0, din care porneşte traiectoria, simulând variabila aleatoare X_0 definită în (2).

Probabilităţile de trecere din starea $i = s_0$ în una din stările sistemului sunt date de elementele din linia i a matricei de tranziţie $Q = (p_{ij})$. Astfel starea s_1 la momentul $t = 1$ este o observaţie asupra variabilei aleatoare discrete

$$T_1 = \begin{pmatrix} 1 & 2 & \ldots & m \\ Q_{i1} & Q_{i2} & \ldots & Q_{im} \end{pmatrix},$$

ce ia valorile $\{1, 2, \ldots, m\}$ cu probabilităţile din linia i a matricei de tranziţie etc.

Algoritmul de generare a segmentului de traiectorie $\mathtt{s_0}, \mathtt{s_1}, \ldots, \mathtt{s_n}$ este:

```
function LantMarkov(m, π0, Q, n)
    s0 = simulator(1, 2, ..., m; π0);
    i = s0;
    for k = 1 : n
        p = Q[i, :]; //Q[i, :] = [linia i din matricea Q];
        sk = simulator(1, 2, ..., m; p);
        i = sk;
    end for
    return s0, s1, ..., sn;
end function
```

1.1.2 Analiza unui lanţ Markov

Pe lângă tranziţia într-un singur pas a unui lanţ Markov, suntem interesaţi şi de tranziţia dintr-un nod în altul în n paşi. Fie

$$P(X_n = j | X_0 = i), \quad i, j \in S,$$

probabilitatea ca lanţul să treacă din nodul iniţial i în nodul j după n paşi. Să arătăm că această probabilitate este dată de elementul din poziţia (i, j) a matricei de tranziţie Q la puterea n, notat cu $Q^n(i, j)$.

Propoziţia 1.1.1 *Are loc:*

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

Demonstraţie: Vom arăta mai întâi câ $P(X_2 = j|X_0 = i) = Q^2(i,j)$. Evenimentul $(X_2 = j|X_0 = i)$ se poate scrie astfel:

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

deci

$$\begin{aligned} P(X_2 = j|X_0 = i) &= \frac{\sum_{k=1}^{m} P(X_0 = i, X_1 = k, X_2 = j)}{P(X_0 = i)} \\ &= \frac{\sum_{k=1}^{m} P(X_0 = i)P(X_1 = k|X_0 = i)\underbrace{P(X_2 = j|X_0 = i, X_1 = k)}_{P(X_2=j|X_1=k)}}{P(X_0 = i)} \\ &= \sum_{k=1}^{m} Q(i,k)Q(k,j) = Q^2(i,j). \end{aligned}$$

Presupunem că egalitatea $P(X_{n-1} = j|X_0 = i) = Q^{n-1}(i,j)$ este adevărată. Pentru a o demonstra şi în cazul n, exprimăm evenimentul

$$(X_n = j|X_0 = i) = \bigcup_{k=1}^{m} (X_n = j, X_{n-1} = k|X_0 = i).$$

Procedând la fel ca în cazul $n = 2$ se obţine concluzia. □

Are loc:

$$Q^n(i,j) = P(X_n = j|X_0 = i) = P(X_{n+k} = j|X_k = i), \text{ pentru orice } k \in \mathbb{N}.$$

Deci, probabilitatea de a trece în n paşi din nodul i în nodul j nu depinde de momentul în care lanţul este în nodul i.

Pentru a putea face predicţii asupra traiectoriei aleatoare definite de lanţ să calculăm câteva probabilităţi ale unor evenimente de interes.

Propoziţia 1.1.2 *Probabilitatea ca lanţul să evolueze pe traiectoria $s_0, s_1, s_2, \ldots, s_n$ este:*

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

Demonstraţie: Din formula condiţionării iterate şi a proprietăţii markoviene (1) avem

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

□

Exemplul 2. Considerăm lanţul Markov din Exemplul 1, având distribuţia iniţială de probabilitate $\pi_0 = [0.2, 0.5, 0.3]^T$. Să se calculeze probabilitatea ca lanţul să evolueze din starea $s_0 = 2$ în starea $s_4 = 1$ pe traiectoria $2, 1, 3, 2, 1$.

Conform formulei deduse, probabilitatea ca lanţul să aibă traiectoria $2, 1, 3, 2, 1$ este

$$P = \pi_0(2)Q(2,1)Q(1,3)Q(3,2)Q(2,1) = 0.5 \times 0.2 \times 0.1 \times 0.4 \times 0.2 = 0.0008.$$

Distribuţia de probabilitate a variabilei de stare la momentul n

În definiţia unui lanţ Markov (X_n) nu se precizează şi distribuţia de probabilitate a variabilei aleatoare de stare X_n. Variabila aleatoare X_n ia valorile $\{1, 2, \ldots, m\}$ şi evenimentul $(X_n = j)$ este evenimentul ca la momentul n traiectoria aleatoare să ajungă în nodul $j \in S$. Notăm cu $\pi_n(j) = P(X_n = j)$.

Vom arăta că dacă se cunoaşte distribuţia iniţială de probabilitate π_0 şi matricea de tranziţie Q a lanţului Markov (X_n), atunci putem determina şi distribuţia de probabilitate $\pi_n = [\pi_n(1), \pi_n(2), \ldots, \pi_n(m)]^T$ a variabilei aleatoare X_n, $\forall\, n \in \mathbb{N}^*$.

Propoziţia 1.1.3 *Distribuţia de probabilitate a stării la momentul n este:*

$$\pi_n^T = \pi_0^T Q^n \tag{3}$$

sau detaliat:

$$\begin{bmatrix} \pi_n(1) & \pi_n(2) & \ldots & \pi_n(m) \end{bmatrix} = \begin{bmatrix} \pi_0(1) & \pi_0(2) & \ldots & \pi_0(m) \end{bmatrix} Q^n.$$

Observaţie Vectorii probabilişti π_0, π_n sunt matrice coloană, deci transpusele lor sunt matrice linie

Demonstraţie: Notăm cu A evenimentul $(X_n = j)$ şi cu H_i evenimentele (ipoteze) $H_i = (X_0 = i)$, $i = 1, 2, \ldots, m$. Evident că $H_1, H_2, \ldots, H_m$ constituie o descompunere a evenimentului sigur în m evenimente mutual exclusive două câte două. Conform formulei probabilităţii totale avem:

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i).$$

Rescriem formula probabilităţii totale înlocuind A cu $(X_n = j)$ şi H_i cu $(X_0 = i)$:

$$P(X_n = j) = \sum_{i=1}^{m} P(X_0 = i)P(X_n = j|X_0 = i),$$

adică

$$P(X_n = j) = \sum_{i=1}^{m} \pi_0(i)Q^n(i,j).$$

Rezultă astfel că

$$\pi_n(j) = \sum_{i=1}^{m} \pi_0(i) Q^n(i,j),$$

ceea ce ne conduce la relaţia matriceală

$$[\ \pi_n(1) \quad \pi_n(2) \quad \ldots \quad \pi_n(m)\ ] = [\ \pi_0(1) \quad \pi_0(2) \quad \ldots \quad \pi_0(m)\ ] Q^n$$

sau concentrat

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

□

Din relaţia (3) rezultă că distribuţia de probabilitate X_n a stării la momentul n se poate calcula recursiv pornind de la distribuţia iniţială π_0:

$$\begin{aligned} \pi_1^T &= \pi_0^T Q \\ \pi_2^T &= \pi_0^T Q^2 = \pi_1^T Q \\ &\ \vdots \\ \pi_n^T &= \pi_{n-1}^T Q, \ n \in \mathbb{N}^*. \end{aligned} \tag{4}$$

Exemplul 3. Pentru lanţul Markov dat în Exemplul 1, considerând distribuţia iniţială de probabilitate $\pi_0 = [0.2, 0.35, 0.45]^T$, distribuţiile de probabilitate ale variabilelor aleatoare (stărilor) $X_1, X_2, \ldots, X_{10}$, calculate conform relaţiei de recurenţă de mai sus, sunt:

$$\begin{aligned}
\pi_1 &= [0.33000000000 \quad 0.525000000000 \quad 0.145000000000]^T \\
\pi_2 &= [0.29500000000 \quad 0.590500000000 \quad 0.114500000000]^T \\
\pi_3 &= [0.28190000000 \quad 0.606650000000 \quad 0.111450000000]^T \\
\pi_4 &= [0.27867000000 \quad 0.610185000000 \quad 0.111145000000]^T \\
\pi_5 &= [0.27796300000 \quad 0.610922500000 \quad 0.111114500000]^T \\
\pi_6 &= [0.27781550000 \quad 0.611073050000 \quad 0.111111450000]^T \\
\pi_7 &= [0.27778539000 \quad 0.611103465000 \quad 0.111111145000]^T \\
\pi_8 &= [0.27777930700 \quad 0.611109578500 \quad 0.111111114500]^T \\
\pi_9 &= [0.27777808430 \quad 0.611110804250 \quad 0.111111111450]^T \\
\pi_{10} &= [0.27777783915 \quad 0.611111049705 \quad 0.111111111145]^T.
\end{aligned}$$

Faptul că matricei de tranziţie Q şi distribuţiei iniţiale de probabilitate π_0 li se asociază un şir de vectori probabilişti (π_n), ne conduce la întrebările:
• În ce condiţii este şirul (π_n) convergent?
• Dacă şirul (π_n) este convergent la π, ce reprezintă limita sa π?
• Fiecare distribuţie iniţială π_0 defineşte un alt şir (π_n) prin $\pi_n^T = \pi_0^T Q^n$. Ne aşteptăm ca în caz de convergenţă să obţinem mereu altă limită?

Pentru a răspunde acestor întrebări detaliem câteva particularităţi ale matricei Q şi a transpusei sale Q^T, precum şi ale şirului (π_n).

Propoziţia 1.1.4 *Dacă şirul (π_n) al distribuţiilor de stare este convergent, atunci limita sa este un vector probabilist π, adică un vector de coordonate mai mari sau egale cu zero şi suma coordonatelor este 1.*

Demonstraţie: Fie $\pi_n = [\pi_n(1), \pi_n(2), \ldots, \pi_n(m)]^T$ cu $\pi_n(i) \in [0, 1]$ şi $\sum_{i=1}^{m} \pi_n(i) = 1$. Presupunem că şirul vectorial (π_n) este convergent la vectorul $\pi = [\pi_1, \pi_2, \ldots, \pi_m]^T$. Se ştie din analiză că $\pi_n \to \pi$ dacă şi numai dacă $\pi_n(i) \to \pi_i$, $\forall\ i = 1, 2, \ldots, m$. Astfel, şirul $S_n = \pi_n(1) + \pi_n(2) + \cdots + \pi_n(m) \to \pi_1 + \pi_2 + \cdots + \pi_m$. Dar $S_n = 1$ pentru orice n, deci şi limita sa este 1, adică $\pi_1 + \pi_2 + \cdots + \pi_m = 1$ (ceea ce evidenţiază că π este vector probabilist). □

Propoziţia 1.1.5 *Dacă şirul (π_n) al distribuţiilor de stare la momentul n converge la π, atunci are loc:*

$$\pi^T = \pi^T Q. \tag{5}$$

Demonstraţie: Dacă $\lim_{n\to\infty} \pi_n = \pi$, atunci şi $\lim_{n\to\infty} \pi_{n-1} = \pi$. Trecând la limită când $n \to \infty$ în relaţia

$$\pi_n^T = \pi_{n-1}^T Q,$$

obţinem (5). □

Definiţia 1.1.2 O distribuţie de probabilitate π, pe spaţiul nodurilor unui lanţ Markov, cu proprietatea că $\pi^T = \pi^T Q$ se numeşte *distribuţie invariantă, staţionară sau distribuţie de echilibru.*

Ce înseamnă asta?

Din faptul că π este limita şirului (π_n) rezultă că pentru orice $\varepsilon > 0$ există un rang $N(\varepsilon)$ astfel încât pentru orice $n \geq N(\varepsilon)$ avem $\|\pi_n - \pi\| < \varepsilon$, adică dacă se calculează π_n, pentru n suficient de mare, cu ajutorul relaţiei $\pi_n^T = \pi_0^T Q^n$, atunci π_n aproximează destul de bine limita π şi deci de la un anumit rang avem că $\pi_{n+1}^T = \underbrace{\pi_n^T}_{\approx \pi^T} Q \approx \pi^T$ etc.

Dacă există distribuţia de echilibru $\pi = [\pi(1), \pi(2), \ldots, \pi(m)]^T$ (ca limită a şirului π_n), atunci $\pi(j)$ reprezintă şansa asimptotică de a fi vizitat nodul j. Cu alte cuvinte, dacă mişcarea aleatoare pe S continuă indefinit şi şirul (π_n) este convergent, atunci de la un moment dat mişcarea aleatoare se stabilizează şi vizitează fiecare nod $j \in \{1, 2, \ldots, m\}$ cu aceeaşi frecvenţă $\pi(j)$.

În cazul Exemplului 3 dacă veţi calcula distribuţiile de probabilitate $\pi_{50}, \pi_{51}, \ldots, \pi_{100}$ şi afişaţi coordonatele cu 15 zecimale, veţi obţine

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

Observăm că $P(X_{50} = k) = \cdots = P(X_{100} = k)$, pentru $k = 1, 2, 3$. Orice distribuţie de probabilitate π_n, cu $n \geq 50$, am calcula cu 15 zecimale am obţine distribuţii identice cu π_{50}, adică după momentul $n = 50$ distribuţia este staţionară, nu se mai modifică în primele 15 zecimale şi deci sistemul a ajuns într-un echilibru.

De-a lungul oricărei traiectorii începând cu momentul $n = 50$, $s_{50}, s_{51}, \ldots, s_{50+N}$, de lungime $N + 1$, starea 1 este "vizitată" de lanţul Markov în proporţie de $100\,\pi(1)\% = 27.7777777778\%$, starea 2 în proporţie de $100\,\pi(2)\% = 61.1111111111\%$, iar starea 3 în proporţie de $100\,\pi(3)\% = 11.1111111111\%$. Mai mult, observăm că pentru $n > 50$ norma diferenţei dintre două distribuţii consecutive $\pi_n - \pi_{n-1}$ este mai mică decât 10^{-15}, adică $\|\pi_n - \pi_{n-1}\| < 10^{-15}$.

1.2 Lanţuri Markov ireductibile şi aperiodice

Ne întrebăm acum în mod natural, ce anume proprietăţi ale lanţului Markov asigură convergenţa şirului de distribuţii (π_n)? Dacă lanţul Markov ar fi definit pe nodurile grafului WEB (ce constă din paginile WEB) şi mişcarea aleatoare s-ar face alegând din fiecare pagină cu o anumită probabilitate paginile către care există linkuri, atunci existenţa distribuţiei de echilibru ar permite caracterizarea popularităţii paginilor WEB cu ajutorul distribuţiei π, $\pi(j)$ reprezentând frecvenţa asimptotică cu care un navigator aleator pe WEB ar vizita pagina j, adică popularitatea paginii j.

În continuare prezentăm condiţiile pe care trebuie să le îndeplinească graful de tranziţie al unui lanţ Markov pentru ca şirul distribuţiilor (π_n) să fie convergent. Particularităţile pe care le evidenţiem sunt cele care au inspirat modul de definire a navigării aleatoare de către Larry Page şi Serghei Brin, fondatorii motorului de căutare Google şi autorii PageRank-ului.

Definiţia 1.2.1 Un lanţ Markov pe $S = \{1, 2, \ldots, m\}$ se numeşte *lanţ 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 lanţul Markov poate trece într-un număr de paşi din nodul i în nodul j.

Practic lanţul este ireductibil dacă şi numai dacă graful de tranziţie este tare conex (există drum de arce între orice două noduri).

Dacă matricea de tranziţie are toate elementele $Q(i, j) > 0$, $\forall\, i, j \in \overline{1, m}$, atunci lanţul Markov este ireductibil.

Exemplul 4. Lanţul Markov având spaţiul stărilor $S = \{1, 2, 3, 4\}$ şi matricea de tranziţie:

$$Q = \begin{pmatrix} 0.35 & 0.65 & 0 & 0 \\ 0.45 & 0.55 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}$$

nu este un lanţ ireductibil, aşa cum se poate observa mai simplu din graful asociat. Mulţimile de stări $\{1, 2\}$ şi $\{3, 4\}$ nu comunică între ele.

Observaţia 1.2.1 *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 mulţi paşi, adică s-ar putea ca $Q^n(i,j) > 0$, pentru un $n > 1$.*

Un lanţ Markov poate avea şi *traiectorii periodice* (Fig. 2), adică traiectorii în care succesiunea de noduri $s_0, s_1, \ldots, s_{T-1}, s_T = s_0$ se repetă indefinit într-o traiectorie a lanţului.

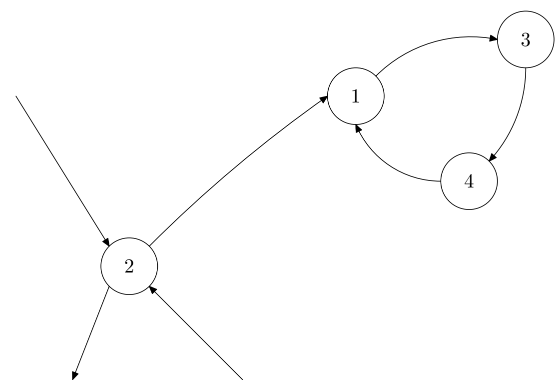

Fig.2: Graful de tranziţie al unui lanţ Markov ce are traiectoria periodică (1,3,4,1)

Din analiza grafului din figură rezultă că probabilităţile $Q^3(1,1) > 0$, $Q^6(1,1) > 0$ şi în general $Q^{3k}(1,1) > 0$, $\forall\, k \in \mathbb{N}^*$, deci cu o probabilitate pozitivă, o traiectorie ce porneşte din 1 se reîntoarce în 1 după 3 paşi, 6 paşi sau mai general un multiplu de 3 paşi. Această proprietate indică faptul că traiectoria ce porneşte din 1 este periodică. Analog pentru 3 şi 4.

Definiţia 1.2.2 *Perioada unui nod $i \in S$* este numărul

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

În cazul exemplului nostru cel mai mare divizor comun al numerelor de forma $3k$, $k \in \mathbb{N}^*$, este 3, deci nodul 1 este periodic de perioadă 3 şi analog nodurile 3 şi 4.

Un nod i a cărui perioadă este 1, adică nodul pentru care cel mai mare divizor comun al numerelor $n \in \mathbb{N}^*$ cu proprietatea că $Q^n(i,i) > 0$ este 1, se numeşte *nod aperiodic*, iar un lanţ Markov care are toate nodurile aperiodice se numeşte *lanţ Markov aperiodic*. Cel mai simplu exemplu de nod aperiodic este un nod i pentru care $Q(i,i) > 0$.

Evident că dacă matricea de tranziţie are toate elementele de pe diagonala principală strict pozitive, $Q(i,i) > 0$, atunci lanţul este aperiodic.

Remarcăm faptul că proprietăţile de ireductibilitate şi aperiodicitate ale unui lanţ Markov sunt proprietăţi ale matricei de tranziţie.

Se poate demonstra (algebric) următoarea proprietate:

Propoziţia 1.2.1 *Un lanţ Markov ireductibil are toate nodurile de aceeaşi perioadă.*

Prin urmare, dacă un nod al unui lanţ ireductibil este aperiodic, atunci toate nodurile sunt aperiodice. Astfel, pentru a arăta că un lanţ ireductibil este aperiodic este suficient să identificăm un nod i pentru care $Q(i,i) > 0$ (probabilitatea de trecere de la starea i la ea însăşi este nenulă), pentru a concluziona că nodul i este aperiodic şi deci toate stările lanţului sunt aperiodice. Dar există lanţuri ireductibile care nu au nici un nod i astfel încât $Q(i,i) > 0$, adică graful asociat nu conţine nici o buclă, şi totuşi lanţul este aperiodic.

De exemplu pentru lanţul ce are matricea de tranziţie:

$$Q = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

graful asociat nu are nici o buclă. Lanţul este ireductibil şi nodul 4 aparţine unui ciclu de lungime 2 şi unuia de lungime 3. Cum cel mai mare divizor comun dintre 2 şi 3 este 1, adică $(2,3) = 1$, rezultă că nodul 4 este aperiodic şi cum lanţul este ireductibil, atunci toate nodurile sunt aperiodice.

Se poate demonstra următorul rezultat, pe baza căruia s-a fundamentat şi algoritmul PageRank–Google:

Propoziţia 1.2.2 *Dacă un lanţ Markov pe S este ireductibil şi aperiodic, atunci oricare ar fi distribuţia iniţială de probabilitate π_0, şirul distribuţiilor de probabilitate la momentul n asociat, (π_n), este convergent şi limita acestuia este un vector probabilist π care nu depinde de distribuţia iniţială de probabilitate (deci indiferent de distribuţia iniţială de probabilitate, şirurile asociate converg la aceeaşi limită π). Această limită este unica distribuţie de echilibru a lanţului Markov. Mai mult, şirul (Q^n) este convergent la o matrice de rang 1 având fiecare linie egală cu π^T:*

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

Cum distribuţia de echilibru π verifică relaţia $\pi^T = \pi^T Q$, ceea ce este echivalent, aplicând transpunerea, cu

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

rezultă că π **este vector propriu probabilist al matricei Q^T, corespunzător valorii proprii 1**.

Acest rezultat matematic a influenţat modul de definire al navigării aleatoare pe graful WEB, ca un lanţ Markov ireductibil şi aperiodic pe mulţimea paginilor WEB (vezi Material Suplimentar).