

Capitolul 2

Surse Markov

2.1 Breviar teoretic

Fie o sursă având alfabetul $[X] = [x_1 x_2 \dots x_D]$. Sursa este **Markov de ordinul k** dacă se respectă relația:

$$p(x_{i_n}/x_{i_{n-1}}x_{i_{n-2}}) \dots x_{i_0}) = p(x_{i_n}/x_{i_{n-1}}x_{i_{n-2}}) \dots x_{i_{n-k}})) \quad (2.1)$$

unde cu $p(x_{i_n})$ am notat probabilitatea ca la momentul n să se emită simbolul x_i .

O sursă Markov de ordinul k este, în fapt, o sursă discretă cu memorie unde șansa de a emite un simbol la momentul curent depinde ultimele k simboluri emise.

În practică, și, în acest material, de interes sunt numai sursele Markov de ordinul 2 și, mai ales de ordinul 1.

În continuare, dacă nu vom specifica altfel, ne vom referi numai la surse Markov de ordinul 1.

Se spune că o sursă se găsește la momentul n în starea i , dacă la momentul $n - 1$ a emis simbolul x_i .

O sursă Markov este **omogenă** dacă probabilitatea de tranziție între două stări i și j nu depinde de momentul de timp la care se petrece tranziția. Deci, pentru o sursă omogenă, tranzițiile dintre două momente de timp se caracterizează de același set de probabilități.

Considerând momentul prezent ca fiind n atunci avem setul de relații:

$$p(x_{i_n}) = \sum_{j=1}^D p(x_{i_n}/x_{j_{n-1}}) \cdot p(x_{j_{n-1}}), (\forall) i = 1, n \quad (2.2)$$

Dacă sursa este omogenă, atunci, notăm:

$$p(x_{j_n}/x_{i_{n-1}}) = p(S_j/S_i) = p_{ij}$$

Dat fiind faptul că tot ce se află într-o stare la un moment de timp trebuie să se întâmple ceva cu el la momentul următor avem relația:

$$\sum_{j=1}^D p_{ij} = 1, (\forall) i \quad (2.3)$$

Tranziția unei surse între două stări poate fi caracterizată în două moduri: printr-un graf de tranziție al stărilor sau prin matricea de tranziție.

La un moment de timp, n , o sursă poate fi caracterizată printr-un vector (coloană) de stare:

$$P_n = \begin{bmatrix} p(x_{1n}) \\ p(x_{2n}) \\ \dots \\ p(x_{Dn}) \end{bmatrix}$$

Relația 2.2 poate fi rescrisă matricial:

$$\begin{aligned} \begin{bmatrix} p(x_{1n}) \\ p(x_{2n}) \\ \dots \\ p(x_{Dn}) \end{bmatrix} &= \begin{bmatrix} p(x_{1n}/x_{1n-1}) & p(x_{1n}/x_{2n-1}) & \dots & p(x_{1n}/x_{Dn-1}) \\ p(x_{2n}/x_{1n-1}) & p(x_{2n}/x_{2n-1}) & \dots & p(x_{2n}/x_{Dn-1}) \\ \dots & \dots & \dots & \dots \\ p(x_{Dn}/x_{1n-1}) & p(x_{Dn}/x_{2n-1}) & \dots & p(x_{Dn}/x_{Dn-1}) \end{bmatrix} \cdot \begin{bmatrix} p(x_{1n-1}) \\ p(x_{2n-1}) \\ \dots \\ p(x_{Dn-1}) \end{bmatrix} \\ &\iff P_n = \begin{bmatrix} p_{11} & p_{21} & \dots & p_{D1} \\ p_{12} & p_{22} & \dots & p_{D2} \\ \dots & \dots & \dots & \dots \\ p_{1D} & p_{2D} & \dots & p_{DD} \end{bmatrix} \cdot P_{n-1} \\ &\iff P_n = T^t \cdot P_{n-1} \end{aligned} \quad (2.4)$$

unde cu T am notat **matricea de tranziție** :

$$T = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1D} \\ p_{21} & p_{22} & \dots & p_{2D} \\ \dots & \dots & \dots & \dots \\ p_{D1} & p_{D2} & \dots & p_{DD} \end{bmatrix}$$

Matricea de tranziție este stohastică, adică are suma elementelor de pe linii 1, ceea ce, în faptm rezulta direct din relația 2.3. Dacă notăm cu P_0 vectorul stărilor la mometul 0, atunci obținem:

$$P_n = (T^t)^n \cdot P_0 \quad (2.5)$$

Dacă o sursă omogenă, cu D stări, evoluează pe baza matricei de tranziție:

$$T = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1D} \\ p_{21} & p_{22} & \dots & p_{2D} \\ \dots & \dots & \dots & \dots \\ p_{D1} & p_{D2} & \dots & p_{DD} \end{bmatrix} \quad (2.6)$$

atunci trecerea ei de la o stare la alta poate fi descrisă, într-un mod mai intuitiv, cu ajutorul unui graf al stărilor. Acesta conține în noduri stările sursei, iar pe muchii se notează probabilitatea de tranziție din starea inițială în starea finală. Un exemplu de asemenea graf este în figura 2.1. De notat că relația 2.3 înseamnă, în graf, că suma probabilităților asociate muchiilor care ies din graf trebuie să fie 1.

Se poate arăta că orice sursă evoluează după un număr de tranziții (momente de timp) într-o stare staționară. O stare staționară este caracterizată de un vector al stărilor constant de la o tranziție la alta. Convergența către starea staționară poate fi:

- monotonă (strict crescătoare, respectiv descrescătoare) către valoarea de echilibru;
- oscilantă, cu o perioadă finită, în jurul valorii de echilibru;

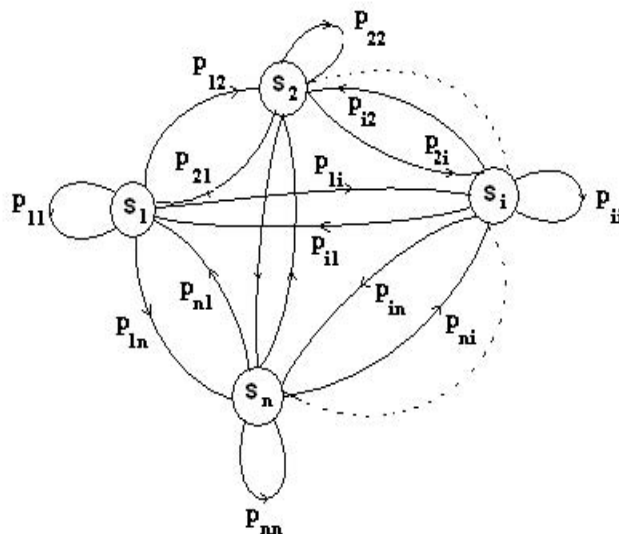


Figura 2.1: Exemplu: graful de tranziție al stărilor asociat matricii de tranziție definită de relația 2.6

2.2 Probleme rezolvate

1. [1] [7] Fie o sursă Markov de ordinul 1. Sursa este omogenă, iar matricea de tranziție asociată este cea dată de relația 2.6. Să se arate ca sursa acceptă stări staționare.

Rezolvare:

După cum am precizat starea staționară este caracterizată de un set de probabilități ale stărilor care nu se mai modifică de la o tranziție la alta. Adică:

$$P_{st} = T^t \cdot P_{st} \quad (2.7)$$

Ecuția 2.7 ne aduce aminte de problema vectorilor și valorilor proprii și anume: dacă λ și V sunt o valoare proprie și, respectiv, vectorul propriu corespunzător, al unei matrice A , atunci există relația:

$$A \cdot V = \lambda V$$

Revenind la problema noastră este suficient să demonstrăm că matricea de tranziție T acceptă $\lambda = 1$ ca valoare proprie, caz în care ecuația VVP (cu vectori și valori proprii) se scrie sub forma

$$A \cdot V = V$$

.

După cum se știe valorile proprii ale unei matrice se definesc a fi soluțiile ecuației caracteristice:

$$\det(A - \lambda \cdot I) = 0$$

În cazul matricei de tranziție dacă arătam că 1 este soluție a ecuației caracteristice problema este rezolvată. Deci:

$$\begin{aligned} & \det(T^t - \lambda \cdot I) = 0 \\ \Leftrightarrow & \begin{vmatrix} p_{11} - \lambda & p_{21} & \dots & p_{D1} \\ p_{12} & p_{22} - \lambda & \dots & p_{D2} \\ \dots & \dots & \dots & \dots \\ p_{1D} & p_{2D} & \dots & p_{DD} - \lambda \end{vmatrix} = 0 \end{aligned}$$

Adunând primele $D - 1$ linii la ultima, se obține relația echivalentă:

$$\begin{vmatrix} p_{11} - \lambda & p_{21} & \dots & p_{D1} \\ p_{12} & p_{22} - \lambda & \dots & p_{D2} \\ \dots & \dots & \dots & \dots \\ \left(\sum_{i=1}^D p_{1i}\right) - \lambda & \left(\sum_{i=1}^D p_{2i}\right) - \lambda & \dots & \left(\sum_{i=1}^D p_{Di}\right) - \lambda \end{vmatrix} = 0$$

Aplicând relația 2.3, obținem pe ultima linie $1 - \lambda$ ceea ce înseamnă că 1 este într-adevar valoare proprie a matricei de tranziție. De altfel, rezultatul este unul cunoscut, și anume că: orice matrice stohastică are o valoare proprie egală cu unitatea.

Se poate arăta că celelalte valori proprii sunt de modul cel mult unitar: $|\lambda| \leq 1$.

Se poate face un studiu al convergenței surselor Markov spre staționaritate în funcție de valorile proprii ale matricei de tranziție:

- dacă valorile proprii sunt pozitive, ceea ce corespunde cazului când stările păstrează mai mult decât dau, ($p_{ii} \geq p_{ij}, i \neq j$) sursa are o convergență monotonă spre staționaritate.
- dacă valorile proprii sunt negative, ceea ce corespunde cazului când stările dau mai mult decât păstrează ($p_{ii} \geq p_{ij}, i \neq j$), sursa are o convergență oscilantă spre staționaritate.
- dacă valorile proprii sunt 1, ceea ce corespunde unei matrice de tranziție cu 1 pe diagonala principală (nu se schimbă nimic între stări), sursa este staționară de la început, în poziția inițială.
- dacă valorile proprii sunt -1 , ceea ce corespunde unei matrice de tranziție cu 1 în alte poziții față de diagonala principală (stările se schimbă total); sursa este permanent oscilantă.

2. [7] O sursă Markov are un alfabet binar, $[x_1 x_2]$. Graful de tranziție este reprezentat în figura 2.2. Vectorul probabilităților inițiale este:

$$a. P_{0a} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix} \text{ sau } b. P_{0b} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 3 \end{bmatrix}$$

- (a) Care este probabilitatea ca sursa să genereze simbolul x_1 după 4 tranziții?
- (b) Care este vectorul de probabilități ce caracterizează starea staționară?
- (c) Cum trebuie modificate probabilitățile p_{21} și p_{22} pentru sursa sa devină staționară de la momentul 1 cu $P_{st} = P_{0a}$?
- (d) Care este entropia sursei staționare? Când această entropie este maximă?

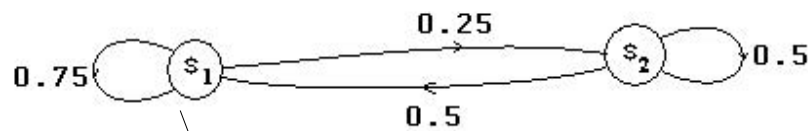


Figura 2.2: Graful de tranziție al stărilor.

Rezolvare: Matricea de tranziție asociată grafului este:

$$T = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

- (a) Relatia 2.5 ne spune că este suficient să aflăm $(T^t)^4$ pentru a calcula probabilitatea cerută.

$$\begin{aligned}
 T^t &= \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \\
 (T^t)^2 &= \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{11}{16} & \frac{10}{16} \\ \frac{5}{16} & \frac{6}{16} \end{bmatrix} \\
 (T^t)^4 &= \begin{bmatrix} \frac{11}{16} & \frac{10}{16} \\ \frac{5}{16} & \frac{6}{16} \end{bmatrix} \cdot \begin{bmatrix} \frac{11}{16} & \frac{10}{16} \\ \frac{5}{16} & \frac{6}{16} \end{bmatrix} = \begin{bmatrix} \frac{171}{256} & \frac{85}{256} \\ \frac{84}{256} & \frac{86}{256} \end{bmatrix} \\
 P_4 = (T^t)^4 \cdot P_0 &= \begin{bmatrix} \frac{341}{512} \\ \frac{171}{512} \end{bmatrix} \\
 \implies p(x_{14}) &= \frac{341}{512}
 \end{aligned}$$

- (b) Rezultatul de la problema 1 precizează că starea staționară nu depinde de starea inițială decât dacă valorile proprii ale matricei de tranziție sunt de modul 1. Nu este cazul în problema curentă. Să verificăm totuși acest lucru.

În cazul în care $P_0 = P_{0a}$ obținem:

$$\begin{aligned}
 P_{1a} = T^t \cdot P_{0a} &= \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0.625 \\ 0.375 \end{bmatrix} \\
 P_{2a} = \begin{bmatrix} 0.656 \\ 0.344 \end{bmatrix} & P_{3a} = \begin{bmatrix} 0.664 \\ 0.336 \end{bmatrix} & P_{4a} = \begin{bmatrix} 0.666 \\ 0.334 \end{bmatrix} & P_{5a} = \begin{bmatrix} 0.666 \\ 0.334 \end{bmatrix} \dots P_{st} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}
 \end{aligned}$$

În cazul în care $P_0 = P_{0b}$ obținem:

$$P_{1b} = T^t \cdot P_{0b} = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0.583 \\ 0.417 \end{bmatrix}$$

$$P_{2b} = \begin{bmatrix} 0.656 \\ 0.344 \end{bmatrix} \quad P_{3b} = \begin{bmatrix} 0.645 \\ 0.345 \end{bmatrix} \quad P_{4b} = \begin{bmatrix} 0.662 \\ 0.338 \end{bmatrix} \quad P_{5b} = \begin{bmatrix} 0.666 \\ 0.334 \end{bmatrix} \quad \dots P_{st} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Evoluția sursei în timp, în cele două cazuri, poate fi văzută și în figura 2.3.

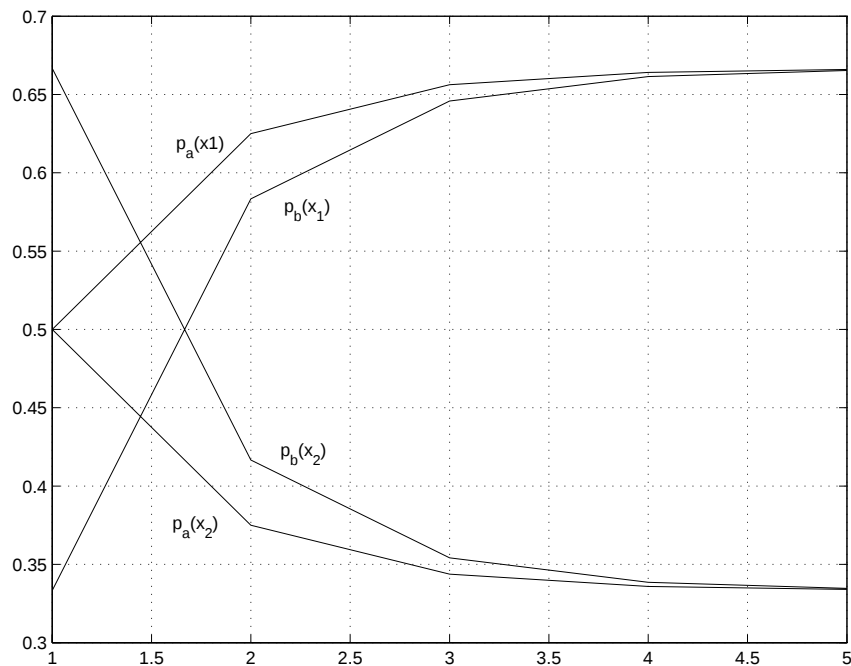


Figura 2.3: Starea staționară a sursei nu depinde de starea inițială.

(c) Notăm $p_{21} = x$. Atunci $p_{22} = 1 - x$ și matricea de tranzit devine:

$$T = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ x & 1 - x \end{bmatrix}$$

Cerința presupune că $P_1 = P_{0a} = T^t \cdot P_{0a}$.

$$\Rightarrow \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & x \\ \frac{1}{4} & 1 - x \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Se obțin două ecuații:

$$\begin{aligned} \frac{3}{4} \frac{1}{2} + \frac{x}{2} &= \frac{1}{2} \\ \frac{1}{4} \frac{1}{2} + \frac{1-x}{2} &= \frac{1}{2} \end{aligned}$$

Ambele ecuații acceptă soluția $x = \frac{1}{4}$, pentru care matricea de tranzit devine:

$$T = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

- (d) Entropia unei surse discrete fără memorie era ușor de calculat pentru că, la orice moment de timp sursa emitea simbolurile cu aceleași probabilități, și se putea scrie o formulă stabilă. Pentru o sursă discretă cu memorie este mult mai dificil, deoarece la momente de timp diferite simbolurile sunt emise cu probabilități diferite și, pe lângă medierea probabilistică ar mai trebui o medire temporală sau o jurnalizare a probabilităților la toate momentele de timp. Totuși, dacă sursa este staționară problema este aceeași cu cea de la sursele discrete fără memorie. Entropia sursei staționare (mai ales în cazul de la punctul (c)) se calculează cu formula:

$$H(X_{st}) = - \sum_{i=1}^n p_{i_{st}} \log p_{i_{st}} = 1 \frac{\text{bit}}{\text{simbol}}$$

Entropia maximă se obține pentru probabilități egale (ceea ce s-a și întâmplat).

3. [4] *Fie o sursă Markov de ordinul 2. Sursa este binară și complet determinată de probabilitățile condiționate:*

$$p(0|00) = p(1|11) = 0.8$$

$$p(1|00) = p(0|11) = 0.8$$

$$p(0|01) = p(0|11) = p(1|01) = p(1|10) = 0.5$$

Să se traseze graficul de tranziție al stărilor. Să se arate că sursa este staționară și să se calculeze vectorul staționar.

Rezolvare:

Am precizat că, în cazul unei surse Markov de ordinul 1 starea este funcție de ultimul simbol transmis. De data aceasta sursa este de ordinul 2 și atunci starea este funcție de ultimii doi biți transmiși. Fiind o sursă binară rezultă un alfabet al stărilor cu 4 elemente. Convenim asupra următoarei codificări:

- starea S_0 înseamnă că la momentul $n - 2$ s-a transmis 0, iar momentul $n - 1$ tot 0;
 $S_0 = (00)$
- starea S_1 înseamnă că la momentul $n - 2$ s-a transmis 0, iar momentul $n - 1$, 1;
 $S_1 = (01)$
- $S_2 = (10)$
- $S_3 = (11)$

Ce înseamnă stările în cazul unei transmisiuni? Avem succesiunea de simboluri transmise de tipul $x_1x_2x_3$. Simbolul x_3 a fost transmis cu probabilitatea $p(x_3/x_1x_2)$. Simbolurile x_1x_2 codifică starea la momentul $n - 1$, iar x_2x_3 starea la momentul n . Analizând setul de probabilități ce caracterizează sursa rezultă:

- s-a transmis succesiunea de simboluri 00 (starea S_0). După ele poate urma un simbol de :
 - 0 cu probabilitate 0.8; succesiunea este 000, iar starea la momentul curent S_0 deci $p(S_0/S_0) = 0.8$

- 1 cu probabilitate 0.2; succesiunea este 001, iar starea la momentul curent S_1 deci $p(S_1/S_0) = 0.2$

Atunci $p(S_2/S_0) = p(S_3/S_0) = 0$

- s-a transmis succesiunea de simboluri 01 (starea S_1). După ele poate urma un simbol de :

- 0 cu probabilitate 0.5; succesiunea este 010, iar starea la momentul curent S_2 deci $p(S_2/S_1) = 0.5$

- 1 cu probabilitate 0.5; succesiunea este 011, iar starea la momentul curent S_3 deci $p(S_3/S_1) = 0.5$

Atunci $p(S_0/S_1) = p(S_1/S_1) = 0$

- s-a transmis succesiunea de simboluri 10 (starea S_2). După ele poate urma un simbol de :

- 0 cu probabilitate 0.5; succesiunea este 100, iar starea la momentul curent S_0 deci $p(S_0/S_2) = 0.5$

- 1 cu probabilitate 0.5; succesiunea este 101, iar starea la momentul curent S_1 deci $p(S_1/S_2) = 0.5$

Atunci $p(S_2/S_2) = p(S_3/S_2) = 0$

- s-a transmis succesiunea de simboluri 11 (starea S_3). După ele poate urma un simbol de :

- 0 cu probabilitate 0.2; succesiunea este 110, iar starea la momentul curent S_2 deci $p(S_2/S_3) = 0.2$

- 1 cu probabilitate 0.8; succesiunea este 111, iar starea la momentul curent S_3 deci $p(S_3/S_3) = 0.8$

Atunci $p(S_0/S_3) = p(S_1/S_3) = 0$

Folosind notația

$$p(S_j/S_i) = p_{ij} \text{ și } Q = [p_{ij}]$$

matricea de tranziție Q este:

$$Q = \begin{bmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0.5 & 0.5 & 0 & 0.2 \\ 0 & 0 & 0.2 & 0.8 \end{bmatrix}$$

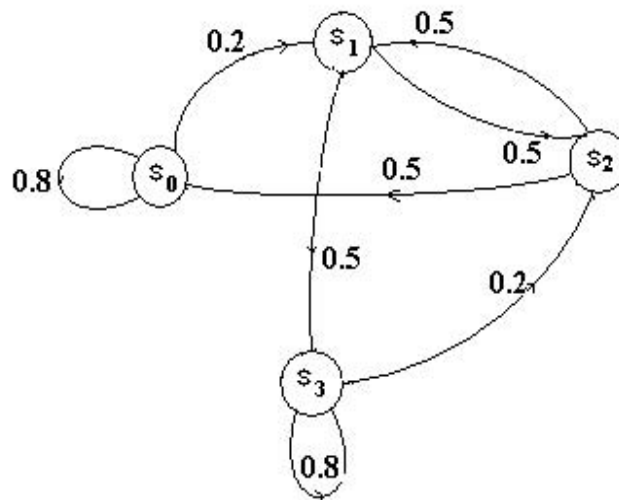
Graful de tranziție al stărilor asociat este prezentat în figura 2.4

Dat fiind faptul că și această matrice este stochastică (are suma elementelor pe linie egale cu 1), rezultatul de la problema precedentă este valabil (admite stare staționară). Aflarea stării staționare presupune rezolvarea sistemului 5 ecuații și 4 necunoscute:

$$P_{st} = Q^t \cdot P_{st}$$

unde

$$P_{st} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \text{ și } \alpha + \beta + \gamma + \delta = 1$$

Figura 2.4: Graful de tranziție al stărilor asociat matricei Q

Adică

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.5 & 0 \\ 0.2 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.2 \\ 0 & 0.5 & 0 & 0.8 \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$$

$$\alpha + \beta + \gamma + \delta = 1$$

Rezultatul obținut la prima problema ne asigură că acest sistem este determinat; în fapt, relația care dă suma liniilor matricei de tranziție 1 presupune că una dintre cele patru ecuații este redundantă și, în fapt, avem un sistem de patru ecuații cu patru necunoscute.

Sistemul se rezolvă prin metoda substituțiilor (Gauss). Soluția obținută este:

$$P_{st} = \begin{bmatrix} \frac{5}{14} \\ \frac{1}{7} \\ \frac{1}{7} \\ \frac{5}{14} \end{bmatrix}$$

Această soluție verifică relația de staționariate.

2.2.1 Probleme propuse

1. Să se studieze convergența (sunt oscilante respectiv convergente monotone sau oscilante) sursele Markov de ordinul 1 caracterizate de matricea de convergență și stare inițială:

(a)

$$T = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \quad P_0 = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$$

(b)

$$T = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix} \quad P_0 = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$$

(c)

(d)

$$T = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \quad P_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(e)

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad P_0 = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$$

(f)

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad P_0 = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$$

2. Fie o sursa Markov de ordinul 1 cu alfabet binar. Ea este caracterizată de matricea de tranziție T . În cazurile discutate până acum problema a fost de determinat vectorul staționar dacă se știe matricea de tranziție. De data aceasta se cere să se determine matricea de tranziție dacă se știe vectorul staționar: $P_{st} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$

Indicație Sistemul inițial este subdeterminat și de aceea trebuie reduse gradele de libertate prin prefixarea unor valori ale matricei de tranziție.