

## Lucrarea 2

# Surse discrete Markov cu memorie de ordinul întâi

## 2.1 Obiectivul lucrării

În această lucrare, se studiază sursele discrete Markov, punându-se accent pe sursele discrete Markov de ordinul 1. Totodată, lucrarea prezintă conceptul de staționaritate a surselor, dar și condițiile în care aceasta este atinsă.

## 2.2 Introducere teoretică

### 2.2.1 Definiții

O schemă simplificată generală a unui sistem de comunicație este prezentată în Fig. 2.1.

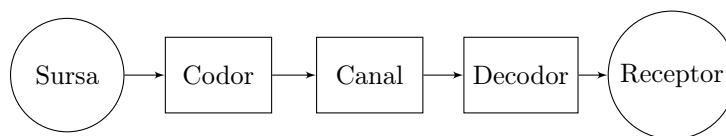


Figura 2.1: Schemă simplificată a unui sistem de comunicație.

În sistemele de comunicație, în particular, în cazul surselor, se utilizează următoarea terminologie:

- *simbol/literă* – elementul fundamental, ireductibil care conține informație, ce corespunde unei realizări particulare a sursei de informație; de exemplu, o sursă de tensiune care generează  $n$  nivele de tensiune. Transmiterea unui nivel de tensiune corespunde unei realizări a evenimentului  $\{X = k\}_{k=1,\dots,n}$ .
- *alfabet* – totalitatea simbolurilor/literelor; În exemplul anterior, alfabetul este alcătuit din  $\mathcal{X} = \{1, 2, \dots, n\}$ .
- *cuvânt* – succesiune finită de simboluri. Un cuvânt poate conține o singură literă;
- *limbă* – totalitatea cuvintelor care pot fi formate cu ajutorul unui alfabet.

Sursele de informație sunt împărțite în două clase:

1. *surse discrete* – sunt surse a căror ieșire este o secvență de simboluri dintr-un alfabet discret finit  $\mathcal{X} = \{x_1, \dots, x_D\}$ . De exemplu, caracterele alfanumerice, caracterele unei tastaturi, caracterele binare, etc. Această lucrare de laborator se axează pe aceste clase de surse.
2. *surse continue* – sunt surse care emit semnale continue dintr-o mulțime infinită.

Sursele discrete de informație pot fi, la rândul lor, împărțite în:

1. *surse cu memorie* – surse pentru care probabilitatea de a transmite un simbol depinde de simbolul care a fost transmis anterior sau de un șir de  $i$  simboluri transmise anterior. În acest caz, probabilitatea de transmisie a simbolului  $x_m$  depinde de  $\{x_{m-1}, \dots, x_{m-i}\}$  astfel:

$$P(X_m = x_m | X_{m-1} = x_{m-1}, \dots, X_{m-i} = x_{m-i}).$$

Dacă emiterea unui simbol de ieșire este condiționată numai de starea precedentă, sursa se numește sursă Markov de ordinul I, ea putând fi modelată cu ajutorul unui lanț Markov finit.

2. *surse fără memorie* – surse a căror probabilitate de transmisie a unui simbol nu depinde de simbolurile transmise anterior.

O *sursă discretă staționară (sau omogenă)* generează simboluri ale căror probabilități nu depinde de originea timpului, ci doar de pozițiile lor relative. Mai precis,

$$P(X_{t_k} = a) = P(X_{t_k + \tau} = a), \quad \forall \tau. \quad (2.1)$$

O *sursă ergodică* este o sursă staționară cu memorie finită la care toate șirurile de simboluri sunt șiruri tipice. Un șir  $A_\epsilon^n$ , de lungime  $n$ , este tipic dacă conține  $n_1 = np_1$  simboluri  $x_1$ ,  $n_2 = np_2$  simboluri  $x_2$ , ș.a.m.d. ( $n = \sum_{k \geq 1} n_k$ ) astfel încât probabilitatea mulțimii  $A_\epsilon^n$  tinde spre 1. Formal,  $P(A_\epsilon^n) > 1 - \epsilon$ , pentru  $n$  suficient de mare.

## 2.2.2 Lanțuri Markov

Un exemplu simplu de procese stochastice este cel al lanțurilor (proceselor) Markov, în care emiterea unui simbol depinde numai de simbolul transmis anterior. Mai precis,

$$\begin{aligned} p(X_m = x_m | X_{m-1} = x_{m-1}, \dots, X_1 = x_1) \\ = p(X_m = x_m | X_{m-1} = x_{m-1}). \end{aligned} \quad (2.2)$$

pentru orice  $x_1, x_2, \dots, x_m \in \mathcal{X}$ . În acest caz, folosind regula lui Bayes, pentru o sursă staționară, se poate scrie:

$$\begin{aligned} p(x_1, x_2, \dots, x_m) &= p(x_m | x_{m-1}, \dots, x_1) p(x_{m-1}, \dots, x_1) \\ &= p(x_m | x_{m-1}) p(x_{m-1}, \dots, x_1) \\ &= \dots \\ &= p(x_m | x_{m-1}) p(x_{m-1} | x_{m-2}) \dots p(x_2 | x_1) p(x_1). \end{aligned}$$

Notând cu  $p(x_i^{(n)}) = P(X_n = x_i)$  este probabilitatea de a transmite simbolul  $x_i \in \mathcal{X}$  la momentul de timp  $n$ .

În fiecare moment de timp, sursa Markov trece într-o nouă stare asociată emiterii unui nou simbol din alfabetul utilizat. Probabilitățile de trecere dintr-o stare în alta sunt notate cu:

$$p_{ij} = p(x_j|x_i) = P(X_j = x_j|X_i = x_i).$$

Un proces Markov este *staționar* dacă probabilitatea de tranziție dintr-o stare în alta nu depinde de momentul de timp ales, adică:

$$P(X_{T+2} = b|X_{T+1} = a) = P(X_2 = b|X_1 = a) = P(b|a) = P_{a,b}, \quad (2.3)$$

pentru orice  $T > 0$  și  $a, b \in \mathcal{X}$ .

Un lanț Markov staționar cu  $D$  stări ( $\text{card}(\mathcal{X}) = D$ ) este, așadar, caracterizat de o *matrice de tranziție*  $T$ :

$$T = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1D} \\ p_{21} & p_{22} & \dots & p_{2D} \\ \dots & \dots & \dots & \dots \\ p_{D1} & p_{D2} & \dots & p_{DD} \end{bmatrix} = [p_{ij}]_{i,j \in \{1, \dots, D\}}, \quad (2.4)$$

ce respectă următoarea condiție:

$$\text{suma elementelor de pe fiecare linie este } 1$$

adică,

$$\sum_{j=1}^D p_{ij} = 1, \quad \forall i \in \{1, 2, \dots, D\}. \quad (2.5)$$

Dacă probabilitatea de a transmite simbolul  $x_i \in \mathcal{X}$  la momentul de timp  $n$  este notată cu  $p(x_i^{(n)})$ , atunci probabilitatea de a transmite un simbol la momentul de timp  $n+1$  este egală cu:

$$p(x_j^{(n+1)}) = \sum_{i=1}^D p(x_i^{(n)})p_{ij}, \quad \forall j \in \{1, \dots, D\}. \quad (2.6)$$

Distribuția de probabilitate a tuturor stărilor la momentul de timp  $n+1$  este dată de matricea:

$$P^{(n+1)} = \left[ p(x_1^{(n+1)}), p(x_2^{(n+1)}), \dots, p(x_D^{(n+1)}) \right]. \quad (2.7)$$

Folosind ultimele două relații, putem scrie:

$$P^{(n+1)} = P^{(n)} \cdot T. \quad (2.8)$$

Notând cu  $P^{(0)}$  vectorul probabilităților stărilor inițiale, atunci rezultă:

$$P^{(n)} = P^{(0)} \cdot T^n. \quad (2.9)$$

Dacă stările lanțului Markov sunt parcurse într-un număr finit de pași atunci lanțul Markov se numește *ireductibil*. Dacă lungimea oricărui drum de la o stare la ea însăși este mai mică sau egală cu 1, atunci spunem despre lanțul Markov că este *neperiodic*.

### 2.2.3 Staționaritatea surselor Markov

Dacă vectorul probabilităților stărilor la momentul  $n + 1$  este egal cu vectorul probabilităților stărilor la momentul  $n$ ,  $P^{(n)} = P^{(n+1)}$ , atunci,

$$\omega = \lim_{n \rightarrow \infty} P^{(n)} \quad (2.10)$$

se numește *distribuția staționară (de echilibru)* corespunzătoare procesului Markov. Pentru un lanț Markov finit, ireductibil și neperiodic, distribuția staționară este unică. Mai mult, în acest caz, staționaritatea este atinsă pornind de la orice vector de probabilități inițiale,  $P^{(0)}$ .

Distribuția de echilibru  $\omega = [\omega_1, \dots, \omega_D]$  respectă relația:

$$\sum_{i=1}^D \omega_i = 1. \quad (2.11)$$

În funcție de modul în care se trece printr-o anumită stare, aceasta poate fi *tranzitorie* sau *absorbantă*. Astfel, dacă într-o stare a sursei, nu se poate rămâne (decât un număr relativ mic de pași) atunci această stare se numește *tranzitorie*. Din contră, o stare ce nu poate fi părăsită sau este părăsită doar după un număr relativ mare de pași, se numește *absorbantă*.

În funcție de viteza cu care se atinge distribuția de echilibru, convergența lanțului Markov poate fi *lentă* sau *rapidă*.

De asemenea, convergența unui lanț Markov poate fi *monotonă* sau *oscilantă* în funcție de modul de convergență a vectorului de probabilități  $P^{(n)}$  către distribuția de echilibru,  $\omega$ .

În continuare, vom determina condițiile de staționaritate ale unei surse Markov în două cazuri, fixând pe rând:

- matricea de tranziție
- vectorul de distribuție la staționaritate  $\omega$ .

#### 1. Fixarea matricii de tranziție

Vectorul de distribuție la echilibru  $\omega$  se determină folosind următorul sistem de ecuații:

$$\begin{cases} \omega = \omega T \\ \sum_{i=1}^D \omega_i = 1 \end{cases} \quad (2.12)$$

întrucât  $\omega = \lim_{n \rightarrow \infty} P^{(n)}$ .

#### 2. Fixarea vectorului de distribuție la staționaritate

În cazul impunerii vectorului de distribuție la staționaritate  $\omega$ , determinarea matricii de tranziții  $T$  implică sistemul de ecuații:

$$\begin{cases} \omega = \omega T \\ \sum_{i=1}^D p_{ij} = 1, \quad \forall i \in \{1, \dots, D\} \end{cases} \quad (2.13)$$

unde a doua ecuație provine din condiția de a avea suma pe fiecare linie a matricii  $T$  egală cu 1.

De exemplu, în cazul simplu în care avem doar două stări, lanțul Markov este caracterizat de matricea de tranziție:

$$T = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$

Dacă considerăm vectorul de distribuție la echilibru:

$$\omega = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Determinarea numerică a elementelor matricei de tranziție  $T$  va conduce la formarea unui sistem de ecuații compatibil nedeterminat, ce va avea soluții de forma:  $\{T | 2q = p\}$ . Verificați prin calcul direct acest lucru.

## 2.3 Desfășurarea lucrării

**Aplicație:** Sursa Markov cu 2 stări

Un lanț Markov cu 2 stări  $\{S_1, S_2\}$  este reprezentat în Fig. 2.2 și este caracterizat de matricea de tranziție:

$$T = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}. \quad (2.14)$$

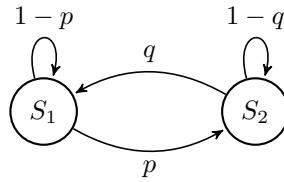


Figura 2.2: Lanț Markov cu 2 stări

1. Scrieți o funcție în Python care să calculeze vectorul de probabilități asociat sursei la un moment de timp  $n$  ( $P^{(n)}$ ), considerând cunoscut vectorul de probabilități la momentul anterior de timp ( $P^{(n-1)}$ ).
2. Scrieți o funcție compactă care să calculeze vectorul de probabilități asociat sursei la un moment de timp  $n$  ( $P^{(n)}$ ).
3. Reprezentați grafic cele 2 componente ale vectorului de probabilități pentru toate momentele de timp  $\{0, 2, \dots, n-1\}$ . La primul moment de timp (0), se consideră că sursa transmite simboluri cu probabilități egale cu elementele vectorului  $P^{(0)}$ .

Cu ajutorul reprezentărilor grafice, studiați convergența lanțului Markov în următoarele cazuri:

- $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- $\begin{bmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{bmatrix}$
- $\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$

iar  $p$  și  $q$  egale cu:

- $p = 0.01, q = 1$
- $p = 0.99, q = 1$
- $p = 0.01, q = 0$
- $p = 0.7, q = 1$

### Soluție Python - aplicația 1

1. Calculul  $P^{(n)}$  a sursei Markov cu 2 stări în funcție de  $P^{(n-1)}$

*Observație:* Vectorii de probabilități sunt de tipul  $1 \times 2!$

```
def compute_next_state(P_old, T):
    P_old = np.asanyarray(P_old)
    T = np.asanyarray(T)

    if np.any(T < 0) or np.any(T > 1) or not np.allclose(np.sum(T, ...
axis=1), 1):
        raise Exception('T nu este o matrice de tranzitie valida.')

    if np.any(P_old < 0) or np.any(P_old > 1) or not ...
np.allclose(np.sum(P_old), 1):
        raise Exception('P_old nu este un vector de probabilitati valid.')

    return P_old @ T
```

2. Calculul  $P^{(n)}$  a sursei Markov cu 2 stări în funcție de  $P^{(0)}$

```
def compute_nth_state(P_start, T, n):
    P_start = np.asanyarray(P_start)
    T = np.asanyarray(T)
    n = int(n)

    if np.any(T < 0) or np.any(T > 1) or not np.allclose(np.sum(T, ...
axis=1), 1):
        raise Exception('T nu este o matrice de tranzitie valida.')

    if np.any(P_start < 0) or np.any(P_start > 1) or not ...
np.allclose(np.sum(P_start), 1):
        raise Exception('P_start nu este un vector de probabilitati valid.')

    if n <= 0:
        raise Exception('n nu este un numar natural nenul.')

    return P_start @ np.linalg.matrix_power(T, n)
```

3. Grafice

Se folosește o buclă *for*:

```
P = [P_0]
for i in range(n - 1):
    P.append(compute_next_state(P[-1], T))
P = np.asanyarray(P)

fig, (ax1, ax2) = plt.subplots(2)
fig.suptitle('Vectori de probabilitate')
ax1.plot(range(n), P[:, 0])
ax1.set_ylabel('Probabilitate starea 1')
ax2.plot(range(n), P[:, 1])
ax2.set_xlabel('Iteratie')
ax2.set_ylabel('Probabilitate starea 2')
plt.show()
```

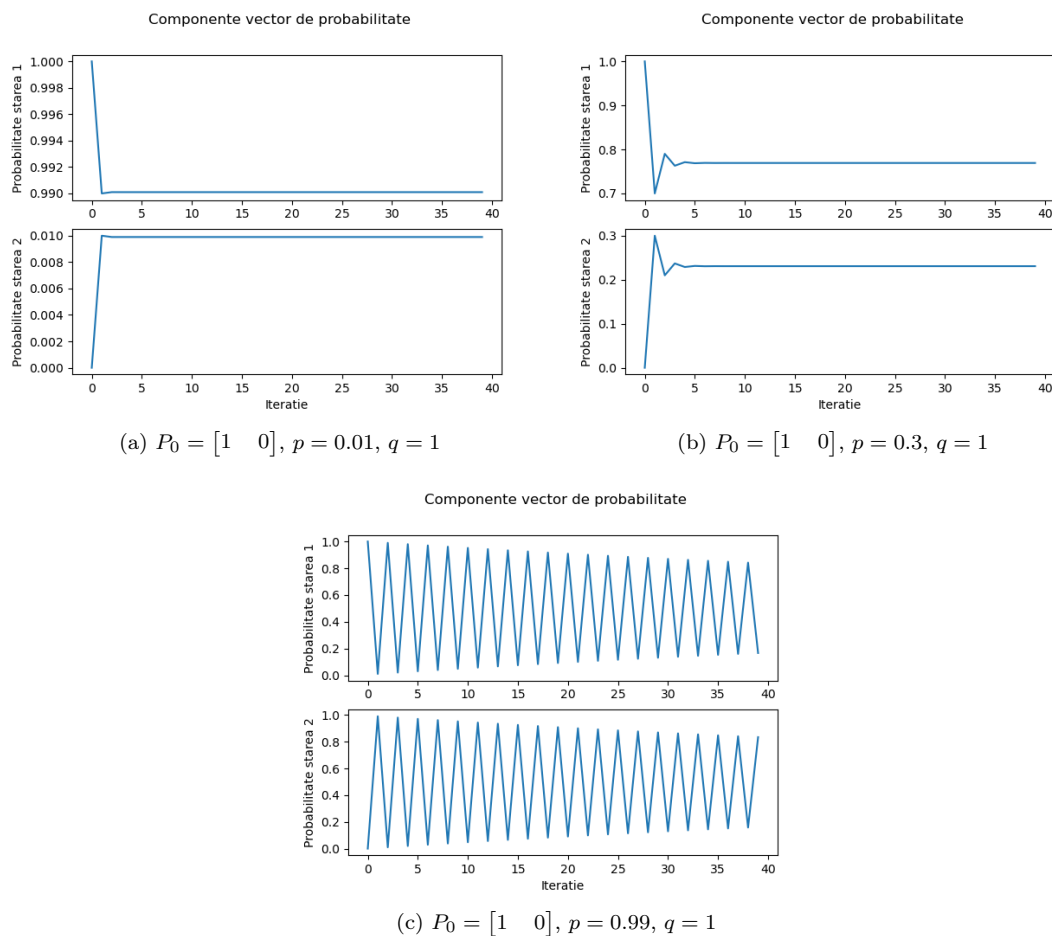


Figura 2.3: Vectori de probabilitate pentru o sursă Markov cu 2 stări

În figura 2.3, se poate observa modul în care variază probabilitățile de transmitere a două simboluri emise de o sursă Markov cu 2 stări în funcție de probabilitățile de tranziție. Se presupune că, la prima iterație, se va emite primul simbol întotdeauna. În primul caz, se ajunge la convergență brusc întrucât una din probabilitățile de tranziție dintr-o stare în alta este mică. Drept urmare, după un număr mic de transmisii (iterații), sursa va emite primul simbol cu probabilitatea 0.99, în timp ce al doilea va fi transmis cu probabilitatea 0.01. De asemenea, se poate ajunge la convergență chiar dacă anterior au existat anumite oscilații în probabilitățile de transmitere a simbolurilor (cazul b). Observați faptul că, în unele situații, nu se ajunge la convergență într-un număr finit de pași (cazul c).

## 2.4 Exerciții

1. Fie matricea de tranziție:

$$T = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

Scrieți un program care afișează vectorul de probabilități după 7 iterații, considerând un vector de probabilități inițiale  $P^{(0)}$  citit de la tastatură.

2. Fie un vector de probabilități inițiale  $P^{(0)} = [1 \ 0 \ 0]$  și matricea de tranziție:

$$T = \begin{bmatrix} p & 1-2p & p \\ 1-2p & p & p \\ p & p & 1-2p \end{bmatrix}$$

Scrieți un program care afișează cu ce probabilitate a fost transmis al doilea simbol după 5 iterații, considerând variabila  $p$  citită de la tastatură.

3. Fie matricea de tranziție

$$T = \begin{bmatrix} 0.01 & 0.98 & 0.01 \\ 0.98 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$$

afișați o listă cu probabilitățile corespunzătoare stării 3 pentru 8 iterații consecutive, considerând un vector de probabilități inițiale  $P^{(0)}$  citit de la tastatură.

4. Fie o matrice de tranziție  $T$  de dimensiune  $K \times K$ . Considerând probabilitatea inițială a fiecărui element ca fiind  $\frac{1}{K}$ , afișați dacă se obține convergența după  $N$  iterații. Se consideră că aceasta a fost atinsă dacă fiecare element al diferenței în modul dintre vectorul de probabilități la 2 momente consecutive este mai mic decât  $10^{-3}$ .

Date de intrare:

Pe prima linie 2 numere întregi separate prin spațiu:  $N \ K$

Pe următoarele  $K$  linii  $K$  numere reale separate prin spațiu, liniile matricii  $T$ .

Date de ieșire:

Pe prima linie True / False (dacă se atinge sau nu convergența după  $N$  iterații)