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 staţionaritate a surselor, dar şi condiţiile în care aceasta este atinsă.

2.2 Introducere teoretică

2.2.1 Definiţii

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

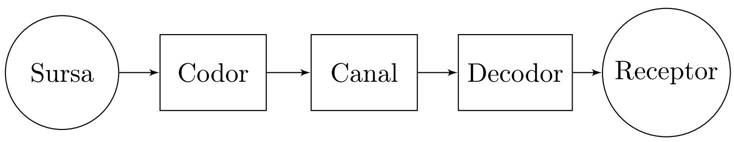


Figura 2.1: Schemă simplificată a unui sistem de comunicaţie.

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

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

Sursele de informaţie sunt împărţite în două clase:

1. *surse discrete* – sunt surse a căror ieşire este o secvenţă de simboluri dintr-un alfabet discret finit $\mathcal{X} = \{x_1, \ldots, x_D\}$. De exemplu, caracterele alfanumerice, caracterelele 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 mulţime infinită.

Sursele discrete de informaţie pot fi, la rândul lor, împărţite 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 şir de i simboluri transmise anterior. În acest caz, probabilitatea de transmisie a simbolului x_m depinde de $\{x_{m-1}, \ldots, x_{m-i}\}$ astfel:
$$P(X_m = x_m | X_{m-1} = x_{m-1}, \ldots, X_{m-i} = x_{m-i}).$$
Dacă emiterea unui simbol de ieşire este condiţionată numai de starea precedentă, sursa se numeşte sursă Markov de ordinul I, ea putând fi modelată cu ajutorul unui lanţ Markov finit.

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

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

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

O *sursă ergodică* este o sursă staţionară cu memorie finită la care toate şirurile de simboluri sunt şiruri tipice. Un şir A_ϵ^n, de lungime n, este tipic dacă conţine $n_1 = np_1$ simboluri x_1, $n_2 = np_2$ simboluri x_2, ş.a.m.d. ($n = \sum_{k\geq 1} n_k$) astfel încât probabilitatea mulţimii A_ϵ^n tinde spre 1. Formal, $P(A_\epsilon^n) > 1 - \epsilon$, pentru n suficient de mare.

2.2.2 Lanţuri Markov

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

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

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

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

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. Probabilităţile 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 *staţionar* dacă probabilitatea de tranziţie dintr-o stare în alta nu depinde de momentul de timp ales, adică:

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

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

Un lanţ Markov staţionar cu D stări ($\text{card}(\mathcal{X}) = D$) este, aşadar, caracterizat de o *matrice de tranziţie* T:

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

ce respectă următoarea condiţie:

suma elementelor de pe fiecare linie este 1

adică,

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

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:

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

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

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

Folosind ultimele două relaţii, putem scrie:

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

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

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

Dacă stările lanţului Markov sunt parcurse într-un număr finit de paşi atunci lanţul Markov se numeşte *ireductibil*. Dacă lungimea oricărui drum de la o stare la ea însăşi este mai mică sau egală cu 1, atunci spunem despre lanţul Markov că este *neperiodic*.

2.2.3 Staționaritatea surselor Markov

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

$$\omega = \lim_{n\to\infty} P^{(n)} \tag{2.10}$$

se numeşte *distribuţia staţionară (de echilibru)* corespunzătoare procesului Markov. Pentru un lanţ Markov finit, ireductibil şi neperiodic, distribuţia staţionară este unică. Mai mult, în acest caz, staţionaritatea este atinsă pornind de la orice vector de probabilităţi iniţiale, $P^{(0)}$.

Distribuţia de echilibru $\omega = [\omega_1, \ldots, \omega_D]$ respectă relaţia:

$$\sum_{i=1}^{D} \omega_i = 1. \tag{2.11}$$

În funcţie 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 numă relativ mic de paşi) atunci această stare se numeşte *tranzitorie*. Din contră, o stare ce nu poate fi părăsită sau este părăsită doar după un număr relativ mare de paşi, se numeşte *absorbantă*.

În funcţie de viteza cu care se atinge distribuţia de echilibru, convergenţa lanţului Markov poate fi *lentă* sau *rapidă*.

De asemenea, convergenţa unui lanţ Markov poate fi *monotonă* sau *oscilantă* în funcţie de modul de convergenţa a vectorului de probabilităţi $P^{(n)}$ către distribuţia de echilibru, ω.

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

- matricea de tranziţie
- vectorul de distribuție la staționaritate ω.

1. *Fixarea matricii de tranziție*

   Vectorul de distribuţie la echilibru ω se determină folosind următorul sistem de ecuaţii:

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

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

2. *Fixarea vectorului de distribuție la staționaritate*

   În cazul impunerii vectorului de distribuţie la staționaritate ω, determinarea matricei de tranziţii T implică sistemul de ecuaţii:

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

   unde a doua ecuaţie provine din condiţia de a avea suma pe fiecare linie a matricei T egală cu 1.

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

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

Dacă considerăm vectorul de distribuţie la echilibru:

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

Determinarea numerică a elementelor matricei de tranziţie T va conduce la formarea unui sistem de ecuaţii compatibil nederminat, ce va avea soluţii de forma: $\{T|2q = p\}$. Verificaţi prin calcul direct acest lucru.

2.3 Desfăşurarea lucrării

Aplicaţie*: Sursa Markov cu 2 stări*

Un lanţ Markov cu 2 stări $\{S_1, S_2\}$ este reprezentat în Fig. 2.2 şi este caracterizat de matricea de tranziţie:

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

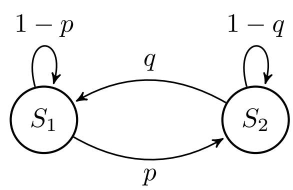


Figura 2.2: Lanţ Markov cu 2 stări

1. Scrieţi o funcţie în Python care să calculeze vectorul de probabilităţi asociat sursei la un moment de timp n ($P^{(n)}$), considerând cunoscut vectorul de probabilităţi la momentul anterior de timp ($P^{(n-1)}$).

2. Scrieţi o funcţie compactă care să calculeze vectorul de probabilităţi asociat sursei la un moment de timp n ($P^{(n)}$).

3. Reprezentaţi grafic cele 2 componente ale vectorului de probabilităti pentru toate momentele de timp $\{0, 2, ..., n-1\}$. 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 convergenţa lanţului Markov în următoarele cazuri:

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

   iar p şi q egale cu:

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

Soluţie Python - aplicaţia 1

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

   Observaţie: Vectorii de probabilităţi sunt de tipul 1×2!

```
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
```

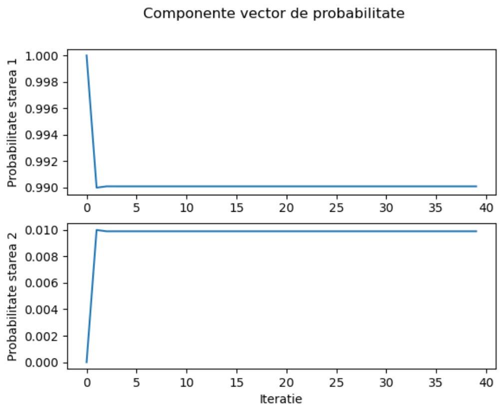
2. Calculul $P^{(n)}$ a sursei Markov cu 2 stări in funcţie de $P^{(0)}$

```
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)
```

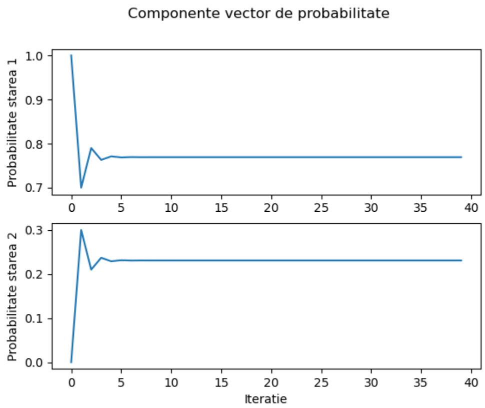
3. Grafice

   Se foloseşte o buclă *for*:

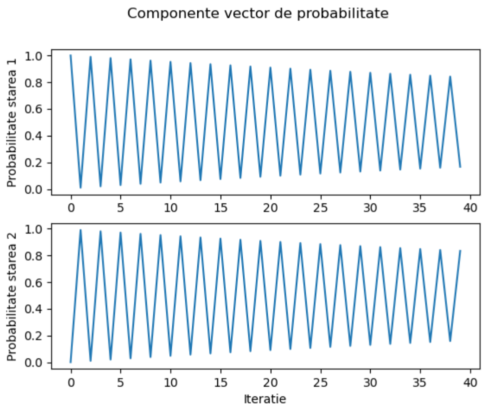
```
P = [P_0]
for i in range(n - 1):
    P.append(compute_next_state(P[-1], T))
P = np.asanyarray(P)

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()
```



(a) $P_0 = \begin{bmatrix}1 & 0\end{bmatrix}$, $p = 0.01$, $q = 1$



(b) $P_0 = \begin{bmatrix}1 & 0\end{bmatrix}$, $p = 0.3$, $q = 1$



(c) $P_0 = \begin{bmatrix}1 & 0\end{bmatrix}$, $p = 0.99$, $q = 1$

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 transmisiuni (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 Exerciţii

1. Fie matricea de tranziţie:
$$T = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$
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 probablități inițiale $P^{(0)} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ și matricea de tranziţie:
$$T = \begin{bmatrix} p & 1-2p & p \\ 1-2p & p & p \\ p & p & 1-2p \end{bmatrix}$$
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)