Probabilità e Statistica[1]

Isadora Antoniano-Villalobos
isadora.antoniano@unive.it

Laurea in Informatica
(Data science/ Tecnologie e scienze dell'informazione)
Università Ca' Foscari di Venezia

Anno academico 2023/2024

[1]Materiale didattico redatto da: Isadora Antoniano-Villalobos & Federica Giummolè

Catene di Markov

Introduzione

Fino ad ora abbiamo considerato successioni di variabili aleatorie indipendenti e identicamente distribuite.

Spesso però nelle applicazioni capita di incontrare **sequenze di variabili dipendenti le une dalle altre**. Basti pensare a modelli probabilistici per fenomeni che si evolvono nel tempo, dove la variabile di interesse al tempo presente dipende da ciò che si è verificato nel passato.

- Le **catene di Markov** sono un modello in cui le variabili sono legate da un particolare tipo di dipendenza.

Definizioni

Catena di Markov (omogenea)

Sia $X_0, X_1, X_2, \ldots$ una successione di variabili casuali (discrete) a valori in un insieme finito (o numerabile) $S = \{1, 2, \ldots, M\}$, detto **spazio degli stati**.

$X = \{X_n\}_{n \geq 0}$ è una **catena di Markov** (omogenea) se

$$P(X_{n+1} = j \quad | \quad X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \\ = P(X_{n+1} = j | X_n = i) = p_{ij}.$$

➔ $P = (p_{ij})_{ij}$ è la **matrice di transizione** e i suoi elementi p_{ij} sono le **probabilità di transizione**.

La matrice di transizione

La matrice di transizione

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{pmatrix}$$

è tale che

- $p_{ij} \geq 0$, $i, j = 1 \ldots, M$,
- $\sum_{j=1}^{M} p_{ij} = 1$, $i = 1 \ldots, M$.

➔ Conoscere la matrice di transizione e la funzione di probabilità dello stato iniziale, $\pi^{(0)}$, permette di calcolare probabilità condizionate, congiunte e marginali della catena.

La matrice di transizione

Esempio: il mondo di Oz

Nel mondo di Oz ci sono tre possibili situazioni meteorologiche:
1=pioggia, 2=sole, 3=neve.

Inoltre:

- non ci sono mai due giorni consecutivi di sole;
- se oggi c'è sole, domani nevica o piove con la stessa probabilità;
- se nevica o piove, con probabilità 0.5 domani rimane invariato e con probabilità 0.5 domani cambia a caso.

➔ Se oggi c'è il sole, come sarà il tempo fra due giorni?

La matrice di transizione

☞ Il modello che descrive bene la situazione è una catena di Markov con spazio degli stati $S = \{1, 2, 3\}$ e matrice di transizione

$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

Lo stato iniziale è $X_0 = 2$ (sole) e ci domandiamo quanto vale $P(X_2 = j | X_0 = 2)$, $j = 1, 2, 3$.

Transizione a due passi

In generale,

$$
\begin{aligned}
& P(X_2 = j | X_0 = i) \\
& \quad = \sum_{k \in S} P(X_2 = j | X_1 = k) P(X_1 = k | X_0 = i) \\
& \quad = \sum_{k \in S} p_{kj} p_{ik} = p_{ij}^{(2)},
\end{aligned}
$$

dove $p_{ij}^{(2)}$ è l'elemento ij della matrice $P^2 = P \cdot P$.

➔ La matrice P^2 è chiamata la **matrice di transizione a due passi** della catena di Markov.

Transizione a due passi

Esempio: il mondo di Oz fra due giorni

Nel nostro esempio siamo interessati alla seconda righa di P^2:

$$p_{21}^{(2)} = P(X_2 = 1 | X_0 = 2) = \begin{pmatrix} 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 1/4 \end{pmatrix} = \frac{3}{8},$$

$$p_{22}^{(2)} = P(X_2 = 2 | X_0 = 2) = \begin{pmatrix} 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/4 \\ 0 \\ 1/4 \end{pmatrix} = \frac{1}{4},$$

$$p_{23}^{(2)} = P(X_2 = 3 | X_0 = 2) = \begin{pmatrix} 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/4 \\ 1/2 \\ 1/2 \end{pmatrix} = \frac{3}{8}.$$

Transizione a n passi

Reiterando l'operazione a due passi, si ottiene la **matrice di transizione a n passi**:

$$P^n = P \cdot P \cdot ... \cdot P,$$

i cui elementi sono le probabilità condizionate a n passi

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Distribuzioni marginali

Le **distribuzioni marginali** della catena si ottengono dalla funzione di probabilità iniziale della catena, $\pi^{(0)}$, e dalla matrice di transizione a n passi:

$$\begin{aligned}\pi_i^{(n)} &= P(X_n = i) = \sum_{k \in S} P(X_n = i | X_0 = k) P(X_0 = k) \\ &= \sum_{k \in S} p_{ki}^{(n)} \pi_k^{(0)}.\end{aligned}$$

Perciò,

$$\pi^{(n)} = \pi^{(0)} \cdot P^n.$$

Distribuzioni marginali

Esempio: il mondo di Oz fra due giorni

Qual è la distribuzione del tempo fra due giorni?

👉 Oggi c'è il sole! e $\Rightarrow\ \pi_2^{(0)} = 1$

$$
\begin{aligned}
\pi^{(2)} &= \pi^{(0)} \cdot P^2 \\
&= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 7/16 & 3/16 & 3/8 \\ 3/8 & 2/8 & 3/8 \\ 3/8 & 3/16 & 7/16 \end{pmatrix} \\
&= \begin{pmatrix} 3/8 & 2/8 & 3/8 \end{pmatrix}
\end{aligned}
$$

Catene regolari

Catene regolari

Una catena di Markov si dice **regolare** se esiste un indice n per cui P^n ha tutti elementi strettamente positivi.

- **Esempio (mondo di Oz):** P ha uno zero, ma P^2 no! Perciò la catena è regolare.
- **Esempio:** $P = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix}$ non è regolare (provatelo!).

Distribuzione stazionaria

Se P è regolare allora esiste $\pi = (\pi_1, \ldots, \pi_M)$ tale che

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \pi_1 & \pi_2 & \ldots & \pi_M \\ \pi_1 & \pi_2 & \ldots & \pi_M \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \ldots & \pi_M \end{pmatrix}$$

Distribuzione stazionaria

Si dimostra che π è l'unica **distribuzione stazionaria** della catena, ovvero tale che

$$\pi P = \pi, \qquad \sum_{i=1}^{M} \pi_i = 1.$$

Catena stazionaria

Catena stazionaria

Se la distribuzione iniziale di X_0 è la distribuzione stazionaria π, allora

$$\pi P^n = \pi, \qquad \forall n,$$

e tutte le distribuzioni marginali della catena sono uguali a π.
Si dice allora che la catena di Markov è una **catena stazionaria**.

Esempio: distribuzione stazionaria nel mondo di Oz

☞ Provate a dimostrare che

$$\lim_n P^n = \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \end{pmatrix}.$$

Allora $\pi = (0.4, 0.2, 0.4)$ è la distribuzione stazionaria della catena.

Catena stazionaria

Infatti, la condizione $\pi P = \pi$, con $\sum_{i=1}^{3} \pi_i = 1$, porta al sistema

$$\begin{cases} 4\pi_1 = 2\pi_1 + 2\pi_2 + \pi_3 \\ 4\pi_2 = \pi_1 + \pi_3 \\ 4\pi_3 = \pi_1 + 2\pi_2 + 2\pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \dots \begin{cases} \pi_3 = 2\pi_2 \\ \pi_1 = 2\pi_2 \\ \pi_1 + \pi_2 + \pi_3 = 1, \end{cases}$$

che dà come soluzione proprio $\pi = (2/5, 1/5, 2/5)$.

La passeggiata aleatoria

La passeggiata aleatoria

Una **passeggiata aleatoria** (*random walk* in inglese) è una catena di Markov con spazio degli stati $S = \mathbb{Z}$ che ad ogni istante si muove di un passo a destra o a sinistra con probabilità rispettivamente p e $1 - p$:

$$p_{ii+1} = p, \qquad p_{ii-1} = 1 - p, \qquad p_{ij} = 0 \text{ altrimenti.}$$

In queste condizioni,

- se $p > 1/2$ il sistema andrà verso $+\infty$;
- se $p < 1/2$ il sistema andrà verso $-\infty$;
- se $p = 1/2$ il suo andamento è meno prevedibile (ma per la legge dei grandi numeri, se parte da 0 tende a 0... pensateci!).

Simulazione della passeggiata aleatoria

Per simulare una **traiettoria** della passeggiata aleatoria:

1. si simulano n valori $x_1, \ldots, x_n$ da una variabile aleatoria di Bernoulli, $X \sim \text{Be}(p)$;
2. si trasformano i valori simulati in $y_i = 2x_i - 1$ (tutti valori ± 1);
3. si calcolano le somme cumulate parziali del vettore $y_1, \ldots, y_n$.

La passeggiata aleatoria con barriere

Si possono aggiungere alla passeggiata aleatoria due **barriere non assorbenti**, in modo che, una volta raggiunte, il sistema rimbalzi allo stato precedente.

- In questo caso $S = \{-L, \ldots, -2, -1, 0, 1, 2, \ldots, L\}$, con $p_{-L,\,-L+1} = p_{L,\,L-1} = 1$:

$$P = \begin{pmatrix} 0 & 1 & \ldots & \ldots & \ldots & 0 \\ 1-p & 0 & p & 0 & \ldots & 0 \\ 0 & 1-p & 0 & p & 0 & \ldots \\ \vdots & \ldots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1-p & 0 & p \\ 0 & \ldots & \ldots & \ldots & 1 & 0 \end{pmatrix}$$

La passeggiata aleatoria con barriere

Alternativamente, si possono aggiungere alla passeggiata aleatoria due **barriere assorbenti**, in modo che, una volta raggiunte, il sistema non si muova più da lì.

- In questo caso $S = \{-L, \ldots, -2, -1, 0, 1, 2, \ldots, L\}$, con $p_{-L,-L} = p_{L,L} = 1$:

$$P = \begin{pmatrix} 1 & 0 & \ldots & \ldots & \ldots & 0 \\ 1-p & 0 & p & 0 & \ldots & 0 \\ 0 & 1-p & 0 & p & 0 & \ldots \\ \vdots & \ldots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1-p & 0 & p \\ 0 & \ldots & \ldots & \ldots & 0 & 1 \end{pmatrix}$$