

Appunti di:
Processi stocastici

A.A. 2024/2025

Sapienza Università di Roma
Dipartimento di Scienze Matematiche per l'intelligenza artificiale



Autore: Carboni Francesco

Contents

1	Catene di Markov omogenee	2
1.1	Stati assorbenti e transienti	4
1.2	Distribuzione invariante	5

1 Catene di Markov omogenee

Sia \mathcal{S} un insieme finito di stati e X_1, X_2, \dots una successione di variabili aleatorie tali che

$$X_i : \Omega \rightarrow \mathcal{S}$$

Ogni X_i dipende solo dalla variabile aleatoria che la precede nella sequenza, ovvero X_{i-1} , questo si traduce in:

$$\mathbb{P}(X_k = x_k | X_1 = x_1, X_2 = x_2, \dots, X_{k-1} = x_{k-1}) = \mathbb{P}(X_k = x_k | X_{k-1} = x_{k-1})$$

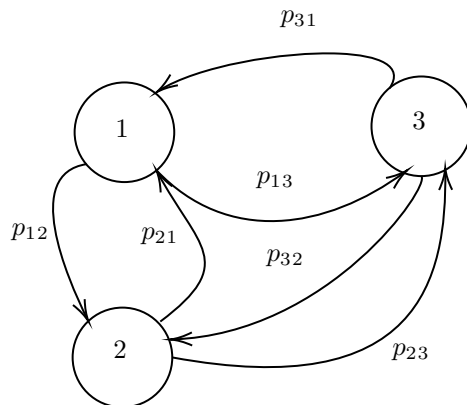
Per alleggerire la notazione scriveremo $p_{ij} = \mathbb{P}(X_k = i | X_{k-1} = j)$ per indicare la **probabilità di transizione**, possiamo immaginare le *MC* come una serie di stati in \mathcal{S} , dove p_{ij} rappresenta la probabilità di passare dallo stato j allo stato i . Per semplificare lo studio delle Catene di Markov possiamo aggiungere l'ipotesi, almeno per il momento, che la probabilità di transizione non dipenda dal tempo k , ovvero p_{ij} è uguale per ogni X_k . Catene di Markov di questo tipo vengono chiamate **omogenee** e possono essere descritte da una matrice $P \in \mathfrak{M}_{n,n}(I)$ chiamata matrice di transizione:

$$\begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{n1} & \dots & \dots & p_{nn} \end{pmatrix}$$

dove per ogni riga vale $\sum_j p_{ij} = 1$. Lo stato iniziale X_0 della catena può essere definito in due modi:

- si sceglie in modo deterministico $X_0 = i$.
- Si utilizza un vettore di probabilità $\mathbf{q} = (q_1, q_2 \dots q_n)$, e si sceglie in modo aleatorio lo stato iniziale.

Una catena di Markov quindi è formata da un'insieme di stati e da una funzione di probabilità che regola i passaggi da uno stato all'altro, ed è proprio in questo che differisce da una macchina a stati deterministica, potendo essere pensata come una macchina a stati stocastica. Oltre alla matrice di transizione una catena di Markov omogenea trova una rappresentazione anche in un grafo orientato, dove i vertici sono gli stati di \mathcal{S} e gli archi sono pesati con p_{ij} .



Esempio 1.1. Consideriamo un sistema di due bit collegati ai lanci di una moneta con la seguente legge: se la moneta da testa viene flipato il primo bit, altrimenti viene flipato il secondo. Assumiamo inoltre che la moneta non sia truccata e quindi

$$\mathbb{P}(T) = \mathbb{P}(C) = \frac{1}{2}.$$

La matrice di transizione che ne deriva è

$$P = \begin{matrix} & \begin{matrix} [00] & [01] & [10] & [11] \end{matrix} \\ \begin{matrix} [00] \\ [01] \\ [10] \\ [11] \end{matrix} & \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \end{matrix}$$

$\mathbf{q} = (q_{00}, q_{01}, q_{10}, q_{11})$ è il vettore di probabilità iniziale, dove q_{00} indica la probabilità di iniziare da $[00]$. Calcoliamo la probabilità che da un $[ij]$ stato di partenza ottenga tutte le altre configurazioni al primo passo:

- $\mathbb{P}(X_1 = [00]) = \frac{1}{2}(q_2 + q_3)$
- $\mathbb{P}(X_1 = [01]) = \frac{1}{2}(q_1 + q_4)$
- $\mathbb{P}(X_1 = [10]) = \frac{1}{2}(q_1 + q_4)$
- $\mathbb{P}(X_1 = [11]) = \frac{1}{2}(q_2 + q_3)$

Quindi l'analisi del primo passo del processo sarà:

$$\text{passo 1} \rightarrow \left(\frac{1}{2}(q_2 + q_3), \frac{1}{2}(q_1 + q_4), \frac{1}{2}(q_1 + q_4), \frac{1}{2}(q_2 + q_3)\right)$$

Il calcolo svolto per ottenere il vettore probabilità del primo passo non è altro che il prodotto vettore per matrice $\mathbf{q} \cdot P$. Da questa osservazione segue che possiamo ridurre la computazione di ogni passo ad un prodotto del vettore risultante dal passo precedente per la matrice di transizione, che rimane invariata.

$$\begin{aligned} \mathbf{q}_1 &= \mathbf{q} \cdot P \\ \mathbf{q}_2 &= \mathbf{q}_1 \cdot P = \mathbf{q} \cdot P^2 \\ &\vdots \\ \mathbf{q}_n &= \mathbf{q}_{n-1} \cdot P = \mathbf{q} \cdot P^n \end{aligned}$$

Osservazione 1.2. Se la matrice P è diagonalizzabile, ovvero è simile ad una matrice diagonale, allora P può essere scritta come $P = U^{-1}DU$, dove D è la matrice diagonale. Grazie alla diagonalizzazione abbiamo un modo più semplice per calcolare le potenze di matrici, infatti

$$P^n = (U^{-1}DU)(U^{-1}DU) \dots (U^{-1}DU) = U^{-1}D^nU$$

ma la potenza della matrice diagonale è $D^n = (a_{ij}^n)$.

Definizione 1.3. Sia P una matrice quadrata, p_{ij} gli elementi della matrice dove $p_{ij} \in [0, 1]$ e $\sum_j p_{ij} = 1$, allora P è detta *matrice stocastica*.

Esempio 1.4. Consideriamo il processo dell'esempio precedente, consideriamo solo i passi pari del processo, per analizzare i passi del processo dobbiamo calcolare la probabilità di andare da $i \rightarrow j$ in due passi.

$$\sum_k p_{ik}p_{kj} \rightarrow \text{probabilità di andare da } i \text{ a } j \text{ in due passi}$$

La nuova matrice di transizione $P' = (\sum_k p_{ik}p_{kj})$ è stocastica.

Esempio 1.5. Consideriamo ora due MC sullo stesso \mathcal{S} :

$P \rightarrow$ passi pari

$Q \rightarrow$ passi dispari

Possiamo rappresentare il processo come un'alternanza delle matrici stocastiche P e Q

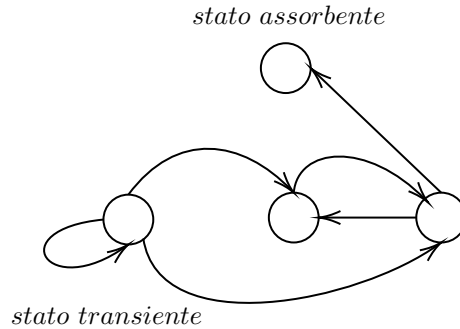
$$qPQPQP \dots$$

Allo stesso modo possiamo considerare $PQ = R$ con $r_{ij} = \sum_k p_{ik}q_{kj}$, ottenendo comunque una matrice stocastica R , quindi il prodotto di matrici stocastiche è una matrice stocastica.

1.1 Stati assorbenti e transienti

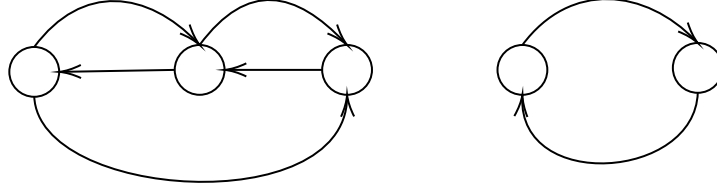
Una serie di risultati interessanti nascono dall'ipotesi di eseguire la MC per molto tempo, e chiedersi quale sia la probabilità che si finisca in un determinato stato.

Definizione 1.6. Uno stato i è detto assorbente se $p_{ii} = 1$, ovvero una volta entrato in quello stato non potrà più uscirne. Uno stato j è detto *transiente*, se una volta che il processo lo abbandona, questo non vi ritornerà più.



Ovviamente essendo una macchina stocastica non può esistere la definizione di stati su cui passeremo un numero finito di volte, se continuiamo il processo per infinito tempo. Da queste semplici definizioni osserviamo, che se una MC ha uno o più stati assorbenti e continuiamo il processo per un tempo indeterminato, allora o si finisce in uno degli stati assorbenti, o ci sono più stati che vengono visitati infinite volte.

Esempio 1.7. Consideriamo la MC rappresentata dal grafo non connesso qui sotto:



formalmente questa è un'unica catena di Markov ma se ne scriviamo la matrice di transizione, sarà una matrice divisa a blocchi.

$$\left(\begin{array}{c|c} MC_1 & 0 \\ \hline 0 & MC_2 \end{array} \right)$$

Sia $\mathcal{S} = \{1, \dots, k, k+1, \dots, n\}$ l'insieme degli stati, dove gli stati da 1 a k sono rappresentati nel sottografo di sinistra, mentre quelli da $k+1$ ad n a destra. Non essendo i due grafi connessi, una volta che finiremo in uno dei due la probabilità di arrivare ad uno qualunque degli stati dell'altro sarà 0, per questo possiamo rinormalizzare il vettore delle probabilità solo su quelle che riguardano il singolo sottografo:

$$\text{per } 1 \leq i \leq k \quad \mathbf{q}' = \left(\frac{q_i}{\sum_{i=1}^k q_i} \right) \quad \text{per } k+1 \leq i \leq n \quad \mathbf{q}'' = \left(\frac{q_i}{\sum_{i=k+1}^n q_i} \right)$$

1.2 Distribuzione invariante

Sia $p^{(n)} = \mathbf{q}P^n$, supponiamo che $\lim_{n \rightarrow \infty} \mathbf{q}P^n = \boldsymbol{\pi}$, quindi se lasciamo proseguire il processo per un tempo illimitato questo si stabilizza e il limite è proprio $\boldsymbol{\pi}$. Inoltre se $\boldsymbol{\pi}$ è il limite otteniamo che

$$\lim_{n \rightarrow \infty} \mathbf{q}P^n = \lim_{n \rightarrow \infty} \mathbf{q}P^{n+1} = \left(\lim_{n \rightarrow \infty} \mathbf{q}P^n \right) P = \boldsymbol{\pi}P$$

$$\boldsymbol{\pi} = \boldsymbol{\pi}P.$$

Allora $\boldsymbol{\pi}$ è un autovettore sinistro di P , e viene chiamato **distribuzione invariante**. In altre parole, la distribuzione $\boldsymbol{\pi}$ resta invariata sotto l'evoluzione della catena di Markov. In questo caso, $\boldsymbol{\pi}$ rappresenta una situazione stazionaria in cui, anche se il sistema evolve nel tempo, la distribuzione complessiva degli stati rimane costante.

Teorema 1.8 (Frobenius). *Una matrice stocastica ha sempre un autovalore uguale ad 1.*