

## 5. Processi di Markov

Introduciamo i processi di Markov analizzando il sistema  $U3 - U2$ , già usato per studiare macrostati e microstati, e la probabilità associata alle diverse configurazioni. In particolare, se si studia il sistema attraverso il macrostato di colore  $n_3$  (numero di biglie nere in  $U3$ ), allora:

- Il sistema ha 3 possibili configurazioni, identificabili dal valore di  $n_3 = 1, 2, 3$
- Ogni altro macrostato è ricavabile dallo stato  $n_3$ , in virtù di una "proprietà di conservazione" del colore

Inoltre, la probabilità **a priori** di ciascuna configurazione è:  $P(n_3 = 1) = 3/10$ ,  $P(n_3 = 2) = 6/10$ ,  $P(n_3 = 3) = 1/10$ . E' quindi possibile calcolare il valore medio di  $n_3$ , che è uguale a  $M[n_3] = \sum_{i=1}^3 i \cdot P(n_3 = i) = 1.8$

A questo punto, è possibile trasformare questo gioco di urne al fine di realizzare un **processo stocastico**. Una *definizione operativa* di questo concetto è che occorre implementare una sequenza di configurazioni che si realizzano come conseguenza di eventi stocastici. In particolare:

1. Inizializziamo in modo casuale lo stato del sistema
2. Estraiamo casualmente una biglia da  $U_3$  e una da  $U_2$ , e le scambiamo di posto
3. Iteriamo il passo precedente per un certo tempo  $T$

**Attenzione.** Notiamo subito che, indipendentemente dal numero di iterazioni, lo stato relativo a  $T = k$  dipende unicamente dallo stato  $T = k - 1$ . Questo tipo di trasformazioni prendono il nome di **catene di Markov**.

A questo punto è possibile calcolare la media aritmetica della variabile  $n_3$  e verificare che corrisponda alla media della distribuzione di probabilità a priori. I risultati empirici hanno mostrato che la frequenza delle configurazioni del sistema si avvicina alla distribuzione a priori.

In altre parole, possiamo dire che ad ogni istante, la probabilità di passare da configurazione  $j$  (qualsiasi) ad una configurazione  $i$  (qualsiasi) è la stessa. La notazione  $P(i \leftarrow j)$  rappresenta la probabilità di transire dalla configurazione  $j$  alla configurazione  $i$ .

Nel caso precedente, dal momento che  $j$  e  $i$  possono assumere 3 valori (il sistema ha infatti 3 possibili configurazioni), si hanno  $3 \times 3 = 9$  probabilità di transizioni (ovvero 9 numeri).

Ad esempio, per calcolare  $P(3 \leftarrow 2)$ , ovvero  $P(n_3 = 3 \leftarrow n_3 = 2)$ , è necessario calcolare la probabilità di estrarre la probabilità di estrarre una biglia nera da  $U_2$  e una bianca da  $U_3$ , e dunque

$$P(3 \leftarrow 2) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

**Nota.** Alcune transizioni, sebbene siano "numericamente sensate", non sono possibili, ad esempio  $P(3 \leftarrow 1) = 0$ , mentre una è certa, infatti  $P(2 \leftarrow 1) = 1$ .

Introduciamo ora la **matrice di Markov** o **matrice stocastica**, che raccoglie le probabilità di transizione tra configurazioni, che rappresentiamo come  $W_{ij} = P(i \leftarrow j)$ .

Ad esempio, la matrice di Markov del gioco di urne è

$$W = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{2}{3} & \frac{1}{2} & 1 \\ 0 & \frac{1}{6} & 0 \end{pmatrix}$$

Questa matrice ha due proprietà fondamentali:

1.  $\forall_{i,j=1}^3 W_{i,j} \geq 0$ , tutte le sue voci sono non negative
2.  $\forall_{j=1}^3 \sum_{i=1}^3 W_{i,j} = 1$ , la somma delle probabilità per ciascuna colonna è pari a 1

**Nota.** La proprietà 2 deriva dal fatto che  $\sum_i P(i \leftarrow j) = 1$

In particolare, ogni matrice  $W$  quadrata ( $N \times N$ ) che possieda queste proprietà è una matrice di Markov (o matrice stocastica).

Il processo stocastico definito sopra può quindi essere descritto tramite  $W$ , in particolare:

a) L'inizializzazione casuale equivale a definire una  $P^{(0)}$ , da intendere come  $P^{(t=0)}$ , ovvero  $P_i^{(t=0)}$  è la probabilità di essere in  $n_3 = i$  al tempo  $t = 0$ .  $P^{(t=0)}$  è un valore a 3 entrate, descritto tramite un vettore colonna.

b) A questo punto, è possibile calcolare il valore di  $P^{(i)}$ , dove  $P^{(1)} = P^{(t=1)}$ , e  $P_i^{(1)}$  è la probabilità di  $n_3 = i$  al tempo  $t = 1$ .

Vale inoltre la regola  $P_i^{(i)} = \sum_{j=1}^3 P_j^{(0)} P(i \leftarrow j)$ , ovvero occorre sommare a tutte le possibili  $j$  di partenza (al tempo  $t = 0$ ) la probabilità di avere  $n_3 = j$  al tempo  $t = 0$ , moltiplicata per la probabilità di transire da  $n_3 = j \rightarrow n_3 = i$ , che denotiamo con  $P(i \leftarrow j)$ . Siccome  $W_{i,j} = P(i \leftarrow j)$  vale la relazione matriciale  $P^{(1)} = W P^{(0)}$ , che in componenti è equivalente a  $P_i^{(1)} = \sum_{j=1}^3 W_{i,j} P_j^{(0)}$ .

Ovviamente vale che  $P_i^{(2)} = \sum_{j=1}^3 P_j^{(1)} P(i \leftarrow j)$ , ovvero  $P_i^{(2)} = \sum_{j=1}^3 W_{i,j} P_j^{(1)}$  ma da quanto visto

$$\begin{aligned} P_i^{(2)} &= \sum_{j=1}^3 W_{i,j} \sum_{k=1}^3 W_{j,k} P_k^{(0)} = \\ &= \sum_{k=1}^3 \left( \sum_{j=1}^3 W_{i,j} W_{j,k} \right) P_k^{(0)} = \\ &= \sum_{k=1}^3 W_{i,k}^2 P_k^{(0)} \end{aligned}$$

ovvero

$$P^{(2)} = W \cdot P^{(1)} = W^2 \cdot P^{(0)}$$

Questa proprietà può essere generalizzata dicendo che

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

Tutto ciò risulta banale se si ragiona sul significato di  $W^2$ , infatti

$$W_{i,j}^2 = \sum_k W_{i,k} W_{k,j} = \sum_k P(k \leftarrow j) P(i \leftarrow k)$$

Ovvero,  $W_{i,j}^2$  rappresenta il prodotto della probabilità di transire (prima) da  $j$  a  $k$  per la probabilità di transire (poi) da  $k$  a  $i$ , sommato su tutte le configurazioni intermedie.

**Nota.** Ovviamente, se  $W$  è una matrice stocastica, anche  $W^2$  è una matrice stocastica.

Concludiamo osservando che, dopo un certo numero di iterazioni, la matrice di Markov elevata alla  $n$ -esima potenza assume una forma in cui tutte le colonne sono identiche, riflettendo le frequenze (asintotiche) di visita alle configurazioni. Questo risultato implica che, asintoticamente, la probabilità di trovarsi in una configurazione non dipende dalla configurazione di partenza, ma solo dalla configurazione stessa, indipendentemente dal percorso compiuto.

## Lo spettro delle matrici di Markov

Il calcolo degli autovalori e degli autovettori per una matrice stocastica consiste nel determinare uno scalare  $\lambda$  (autovalore) e un vettore non nullo  $v$  (autovettore) tali che

$$W \cdot v^{(\lambda)} = \lambda \cdot v^{(\lambda)}$$

dove  $v^{(\lambda)}$  indica che l'autovettore  $v$  è associato all'autovalore  $\lambda$ . Il problema del calcolo degli autovalori e autovettori per una matrice stocastica si affronta distinguendo due casi principali:

1. **Caso**  $\lambda = 1$ . L'autovettore associato,  $v^{(1)}$ , può essere riscalato e interpretato come una distribuzione di probabilità.
2. **Caso**  $\lambda \neq 1$ . In questo caso,  $\lambda$  deve avere modulo minore o uguale a 1, ovvero  $|\lambda| \leq 1$ . Gli autovettori associati a questi autovalori non possono essere riscalati per rappresentare una distribuzione di probabilità.

In particolare, l'autovettore corrispondente all'autovalore 1 è detto probabilità **asintotica** (o **stazionaria**) del processo.

Inoltre, se tutti gli autovalori diversi da  $\lambda = 1$  hanno modulo strettamente minore di 1, il sistema è garantito a convergere, per tempi evolutivi tendenti all'infinito, verso la distribuzione stazionaria, indipendentemente dalla distribuzione di probabilità iniziale. In altre

parole, le configurazioni del processo saranno visitate in accordo alla distribuzione stazionaria.

**Nota.** Calcolare lo spettro di una matrice di Markov significa trovare i suoi autovalori, che indicano come il sistema evolve nel tempo. Questi autovalori ci dicono se lo stato del sistema si stabilizza con il passare del tempo.