PAC - 29 MAGGIO 2003

E 13. Catena di Markov: 3 colori

Supponiamo di colorare i vertici A, B, C, D di un quadrato nel modo seguente: disponiamo di 3 colori e due vertici adiacenti non possono avere lo stesso colore. Sia $\sigma = \{\sigma(A), \sigma(B), \sigma(C), \sigma(D)\}$ la configurazione di colori sui vertici A, B, C, D e sia Ω lo spazio delle configurazioni σ realizzabili con la suddetta regola. Sia ν la misura di probabilità uniforme su Ω.

- Costruire una catena di Markov $\{\sigma_1, \sigma_2, \dots\}$ su Ω che ha come misura invariante ν.
- Simulare l'evoluzione temporale della catena di Markov per una fissata condizione iniziale di colori σ_1. Verificare la convergenza della catena alla misura invariante ν calcolando la frequenza di un'arbitraria configurazione di riferimento $\xi \in \Omega$. Indicare il tempo necessario per poter parlare di "perdita di memoria" della catena.
- Sia Ω_0 il sottoinsieme di Ω in cui vertici opposti hanno lo stesso colore:
$$\Omega_0 = \{\sigma \in \Omega \ : \ \ \sigma(A) = \sigma(C) \ \text{ e } \ \sigma(B) = \sigma(D)\} \, .$$
Sia h_n il numero di volte in cui la catena si è trovata in Ω_0 fino al tempo n:
$$h_n = \sum_{k=1}^{n} \mathbb{I}(\sigma_k \in \Omega_0) \, .$$
Verificare graficamente la convergenza per un valore arbitrario di σ_1:
$$\lim_{n\to\infty} \frac{h_n}{n} = \nu(\Omega_0) = \frac{1}{3} \, .$$
- Estendere l'analisi al caso di un esagono o altro poligono regolare.

caputo@mat.uniroma3.it.

Suggerimenti

E13. *Spazio degli stati.* Identifichiamo i tre colori (giallo, rosso, nero) con i numeri $(0,1,2)$. Sia $\widetilde{\Omega} = \{0,1,2\}^4$. Lo spazio delle configurazioni (o stati) allora è

$$\Omega = \{\sigma \in \widetilde{\Omega} \,:\; \sigma(A) \neq \sigma(B)\,,\; \sigma(B) \neq \sigma(C)\,,\; \sigma(C) \neq \sigma(D)\,,\; \sigma(D) \neq \sigma(A)\}$$

Un semplice conto mostra che $|\Omega| = 18$. Infatti, fissati i colori di A e B rimangono solo 3 modi diversi di colorare C e D. Poiché i colori di A e B possono essere scelti in 6 modi diversi si ottengono 18 configurazioni. La misura uniforme su Ω è data da $\nu(\sigma) = 1/18$ per ogni $\sigma \in \Omega$.

Probabilità di transizione. Il meccanismo della transizione da una configurazione $\sigma \in \Omega$ a una nuova $\eta \in \Omega$ può essere definito come segue. Scegliamo a caso uno dei 4 vertici. Supponiamo di aver scelto C. Se i due vertici adiacenti B e D hanno lo stesso colore, cioè $\sigma(B) = \sigma(D)$ allora ci sono due possibilità per colorare il vertice scelto C. Tiriamo una moneta (equa) e scegliamo una delle due. In questo caso $\eta(C) = \sigma(C)$ con prob. $1/2$ e $\eta(C) = 3 - \sigma(C) - \sigma(B)$ con prob. $1/2$[1]. Se invece $\sigma(B) \neq \sigma(D)$ non resta che una possibilità per colorare C, quindi $\eta(C) = \sigma(C)$ e la configurazione non cambia. Osserviamo che con questa scelta la nuova configurazione η è o uguale a σ oppure differisce da σ in un solo vertice.

Volendo essere più precisi possiamo scrivere la probabilità di transizione nel modo seguente. Chiamiamo $\Gamma_A(\sigma)$ l'insieme di configurazioni che differiscono da σ solo per il colore nel vertice A:

$$\Gamma_A(\sigma) = \{\eta \in \Omega \,:\; \eta(A) \neq \sigma(A)\,,\; \eta(B) = \sigma(B)\,,\; \eta(C) = \sigma(C)\,,\; \eta(D) = \sigma(D)\}$$

e analogo per $\Gamma_B(\sigma)$ ecc. Siano σ, η due configurazioni tali che $\eta \in \Gamma_C(\sigma)$. Allora poniamo

$$p(\sigma,\eta) = \frac{1}{4}\begin{cases} 0 & \sigma(B) \neq \sigma(D) \\ \frac{1}{2} & \sigma(B) = \sigma(D)\,,\; \eta(C) = \sigma(C) \\ \frac{1}{2} & \sigma(B) = \sigma(D)\,,\; \eta(C) = 3 - \sigma(C) - \sigma(B) \end{cases} \qquad \big(\eta \in \Gamma_C(\sigma)\big)$$

Analogamente definiamo $p(\sigma,\eta)$ per $\eta \in \Gamma_A(\sigma)$, $\Gamma_B(\sigma)$, $\Gamma_D(\sigma)$. Se η e σ hanno colori diversi in più di un vertice allora poniamo $p(\sigma,\eta) = 0$. Infine se $\eta = \sigma$ poniamo

$$p(\sigma,\sigma) = 1 - \sum_{\eta \in \Omega:\; \eta \neq \sigma} p(\sigma,\eta)$$

Si può verificare per esempio che $p(\sigma,\sigma) = \frac{1}{2}$ se $\sigma \in \Omega_0$.

Misura invariante. Verifichiamo che la misura uniforme ν è invariante per la catena, cioè che per ogni $\sigma \in \Omega$ si ha

$$\nu(\sigma) = \sum_{\eta \in \Omega} \nu(\eta) p(\eta,\sigma)$$

Poiché ν è uniforme basterà quindi mostrare che $\sum_{\eta \in \Omega} p(\eta,\sigma) = 1$. Notiamo che per costruzione si ha

$$\sum_{\eta \in \Omega} p(\sigma,\eta) = 1$$

Quindi l'identità cercata segue una volta che mostriamo che $p(\cdot,\cdot)$ è simmetrica: $p(\sigma,\eta) = p(\eta,\sigma)$, $\sigma,\eta \in \Omega$. Ma questo segue immediatamente dalla definizione data: se per es. $\eta \in \Gamma_A(\sigma)$ allora $\sigma \in \Gamma_A(\eta)$ e $p(\sigma,\eta) = p(\eta,\sigma) = \frac{1}{8}\,\mathbb{I}(\sigma(B) = \sigma(D))$.

Ergodicità. Sia $p^n(\sigma,\eta)$ la matrice di transizione su n–passi, definita per ricorrenza dalle equazioni:

$$p^{n+1}(\sigma,\eta) = \sum_{\xi \in \Omega} p(\sigma,\xi) p^n(\xi,\eta)\,, \quad p^1(\sigma,\eta) = p(\sigma,\eta)\,.$$

E' possibile verificare che la catena è *regolare*: esiste un intero m tale che $p^m(\sigma,\eta) > 0$ per ogni $\sigma,\eta \in \Omega$. Da un teorema di Markov segue che ν è l'unica misura invariante, e soddisfa

$$\lim_{n \to \infty} p^n(\sigma,\eta) = \nu(\eta)\,, \qquad \sigma,\eta \in \Omega \tag{0.1}$$

Notare l'indipendenza del limite dalla condizione iniziale σ. La convergenza (0.1) può essere verificata come segue: fissato il valore di σ lasciamo evolvere M "storie" indipendenti della catena fino al tempo n (tutte con inizio in σ). Chiamiamo $\eta^1, \ldots, \eta^M$ le configurazioni finali. Scegliamo un $\xi \in \Omega$ qualunque, per es. $\xi(A) = 0, \xi(B) = 1, \xi(C) = 0, \xi(D) = 1$. Contando le volte che $\eta^i = \xi$ e dividendo per M si ottiene una stima $\bar{p}_n(\sigma,\xi)$ per $p^n(\sigma,\xi)$. Fissiamo M molto grande (per es. $M = 10^5$) e grafichiamo la frequenza $\bar{p}_n(\sigma,\xi)$ per valori crescenti di n ($n = 1,2,3,\ldots$).

[1] Stiamo usando il fatto che dati due numeri differenti x, y in $\{0,1,2\}$ il terzo numero z in $\{0,1,2\}$ differente da entrambi è dato da $3 - x - y$.

Si osserverà che $\bar{p}_n(\sigma, \xi)$ approssima bene il valore limite $\nu(\xi) = 1/18$ a partire da un certo n_0. Diremo che n_0 è il tempo necessario per avere "perdita di memoria".

Se $\sigma_1, \sigma_2, \ldots$ denotano gli stati successivi di una *singola* "storia" della catena si può inoltre dimostrare che per ogni funzione $\varphi : \Omega \to \mathbb{R}$ si ha la convergenza:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \varphi(\sigma_k) = \mathbb{E}_\nu(\varphi) = \sum_{\sigma\in\Omega} \nu(\sigma)\varphi(\sigma) \tag{0.2}$$

L'interpretazione corretta della fromula precedente è la seguente: data una qualunque configurazione iniziale σ, l'insieme delle "storie" con inizio in σ per cui vale la convergenza (0.2) ha probabilità 1. Il caso $\varphi = \mathbb{I}(\Omega_0)$ fornisce il limite di h_n/n proposto nel testo.