Catene di Markov

Anno 2007/08

S. Bonaccorsi

November 12, 2007

Outline

Tempo e Incertezza

Osserviamo un fenomeno che evolve nel tempo: quando è impossibile fare una previsione deterministica sull'evoluzione del sistema, possiamo associare una variabile casuale ad ogni istante futuro di osservazione, e nel dedurre dalle caratteristiche di questa le previsioni che ci interessano. Introdurremo quindi una famiglia di variabili aleatorie

$$\{X_t,\ t \in T\}, \qquad T \text{ è un sottoinsieme di } \mathbb{R}_+,$$

a valori in un sottoinsieme E di $\mathbb{R}$, che viene chiamato *spazio delle fasi*.

Classificazione dei processi stocastici

Prima classificazione dei processi stocastici:

- processi stocastici discreti nel tempo e con spazio delle fasi discreto;
- processi stocastici continui nel tempo e con spazio delle fasi discreto;
- processi stocastici discreti nel tempo e con spazio delle fasi continuo;
- processi stocastici continui nel tempo e con spazio delle fasi continuo.

Seconda classificazione dei processi stocastici:

- Modelli statici:
  - Il valore delle variabili non cambia nel tempo e il processo è costituito da una successione di variabili aleatorie indipendenti;
- Modelli dinamici:
  - Il valore delle variabili cambia nel tempo
  - Lo stato corrente e l'evoluzione futura dipendono dalla storia, cioè dalla particolare successione di valori già ottenuti.

Outline

Introduzione ai processi di Markov

- "realizzazione del processo":
  - insieme dei valori rilevati in una successione di osservazioni: $\{(t_0, x_{t_0}); (t_1, x_{t_1}); \ldots; (t_n, x_{t_n})\}$.
- Problema: prevedere lo stato del sistema al tempo t_{n+1}, ossia assegnare la variabile $X_{t_{n+1}}$
  - Se il sistema ha un insieme discreto di stati, basta assegnare le probabilità $\mathbb{P}(X_{t_{n+1}} = x_j)$, $j = 1, 2, 3, \ldots$

  Sapendo la realizzazione del processo fino al tempo t_n, scriveremo le *probabilità di transizione* dallo stato x_{t_n} NOTO ad un certo stato x:

$$\mathbb{P}(X_{t_{n+1}} = x \mid X_{t_0} = x_0; X_{t_1} = x_1; \ldots; X_{t_n} = x_n) \qquad (1)$$

per indicare la probabilità che passi allo stato x, sapendo che in t_0 si è trovato in x_0, in t_1 si è trovato in x_1, ecc.

Catene di Markov

Processi di Markov: l'evoluzione futura dipende solo dallo stato del sistema all'istante attuale; il passato non ha alcuna influenza. Per i processi di Markov la (1) si semplifica in $\mathbb{P}(X_{t_{n+1}} = x \mid X_{t_n} = x_n)$.

Noi ci limitiamo a considerare processi di Markov discreti nel tempo e nello spazio. Lo spazio degli eventi E è dato da

- $E = \{1, \dots, M\}$ un insieme finito;
- $E = \{1, 2, \dots\} = \mathbb{N}$ un insieme infinito;
- $E = \{\dots, -2, -1, 0, 1, 2, \dots\} = \mathbb{Z}$ ancora un insieme infinito (tutti gli interi con segno).

Senza perdita di generalità, i tempi di osservazione saranno sempre $t = 0, 1, 2, \dots$.

- L'evoluzione del sistema è completamente determinata prendendo in considerazione le transizioni in ordine:

$$\mathbb{P}(X_{t_{n+1}} = x_j \mid X_{t_n} = x_i) = p_{ij}(n)$$

dove la coppia di indici "*ij*" indica che si considera la probabilità della transizione $x_i \to x_j$, mentre l'indice "*n*" tra parentesi sta ad indicare che si tratta della *n*-esima transizione.

- Il caso più studiato è quello caratterizzato da probabilità di transizione che non dipendono dal numero d'ordine della transizione considerata. Questa condizione si esprime anche dicendo che la catena è *omogenea* nel tempo. In tal caso potremo semplificare la notazione eliminando l'indice "(n)" dagli elementi p_{ij}.

Scriviamo le probabilità di transizione come matrice

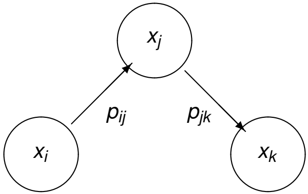
$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1M} \\ p_{21} & p_{22} & \dots & p_{2M} \\ \dots & \dots & \dots & \dots \\ p_{M1} & p_{M2} & \dots & p_{MM} \end{pmatrix}$$

La matrice di transizione P gode delle proprietà seguenti.

- L'elemento $p_{ij}(n)$ è un numero reale, ≥ 0 e ≤ 1.
- La i-esima riga è formata dalle probabilità di transizione dallo stato i ad un qualunque stato j (compreso i stesso): per il principio delle probabilità totali, la somma degli elementi di ogni riga è 1: $\sum_{j=1}^{M} p_{ij}(n) = 1$.
- La matrice è quadrata, perché le righe corrispondono agli stati iniziali e le colonne agli stati finali della transizione, che evidentemente sono uguali.

Rappresentazione grafica

Una matrice stocastica P è rappresentabile graficamente da un grafo. Ogni nodo rappresenta uno stato e l'arco $x_i \rightarrow x_j$ rappresenta la probabilità di transizione p_{ij}



Esempio (1)

- Supponiamo che un'industria produca due tipi di Cola. Una persona che, all'acquisto precedente, ha comprato Cola1, per il 90% di possibilità comprerà ancora Cola1. Chi ha comprato invece Cola2, per l'80% di possibilità comprerà ancora Cola2.
- Matrice di transizione:

$$P = \begin{pmatrix} 0.90 & 0.10 \\ 0.20 & 0.80 \end{pmatrix}$$

Esempio (2)

- Si consideri un sistema di stoccaggio nel quale la sequenza di eventi lungo ogni periodo è la seguente:
  - si osserva il livello i di magazzino all'inizio del periodo; se $i \leq 1$, vengono ordinate $(4 - i)$ unità, mentre se $i \geq 2$ non viene emesso nessun ordine. Le consegne degli ordini sono immediate.
  - la domanda d segue la seguente distribuzione di probabilità:
    - con probabilità 1/3, $d = 0$;
    - con probabilità 1/3, $d = 1$;
    - con probabilità 1/3, $d = 2$.
  - si osserva quindi il livello di magazzino all'inizio del prossimo periodo.
- Descrivere la matrice di transizione del sistema.

La relazione di Chapman-Kolmogorov

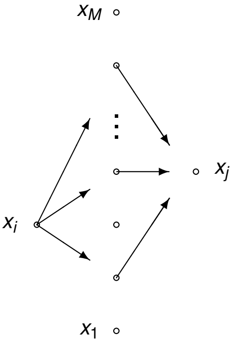
Consideriamo una catena di Markov omogenea, associata ad una matrice di transizione, che racchiude tutte le informazioni sulla evoluzione del sistema.

- Domanda: qual è la probabilità di transizione dallo stato x_i allo stato x_j in m passi, $p_{ij}^{(m)}$?
- in due transizioni si ha

$$p_{ij}^{(2)} = \sum_{k \in E} p_{ij} p_{kj}. \tag{2}$$

La matrice $P^{(2)}$ il cui termine generico è $p_{ij}^{(2)}$ si ottiene come prodotto di P con se stessa:

$$P^{(2)} = \left(p_{ij}^{(2)} \right) = P \cdot P = P^2.$$



Infatti l'evento che stiamo considerando è del tipo $x_i \to x_k \to x_j$, dove $k \in E$. Fissato k, per il principio della probabilità composta, $p_{ik}p_{kj}$ è la probabilità della transizione attraverso il particolare stato intermedio k. Poiché questo può essere qualunque, e al variare di k gli eventi sono mutualmente esclusivi, basterà applicare il principio delle probabilità totali per ottenere la (2).

L’osservazione è facilmente generalizzabile: nota la matrice $P^{(2)}$, la probabilità di transizione in tre passi dallo stato x_i allo stato x_j vale $p_{ij}^{(3)} = \sum_{k \in E} p_{ik}^{(2)} p_{kj}$ e quindi la matrice $P^{(3)}$ i cui elementi sono $p_{ij}^{(3)}$ si ottiene come prodotto tra il quadrato di P e P, cioè $P^{(3)} = P^2 \cdot P = P^3$.

In generale si ha $p_{ij}^{(n)} = \sum_{k=1}^{r} p_{ik}^{(n-1)} p_{kj}$,

$P^{(n)} = \left(p_{ij}^{(n)}\right) = P^{(n-1)} \cdot P = P^{n-1} \cdot P = P^n$ e

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)}. \tag{3}$$

L’ultima relazione prende il nome di equazione di Chapman - Kolmogorov ed è caratteristica dei processi di Markov.

Esempio (1-continua)

- Se una persona usualmente compra Cola2, qual è la probabilità che compri Cola1 dopo *due* acquisti?

$$\mathbb{P}(X_2 = 1 | X_0 = 2) = P_{21}(2)$$

- Calcoliamo

$$P^2 = \begin{pmatrix} 0.90 & 0.10 \\ 0.20 & 0.80 \end{pmatrix} \begin{pmatrix} 0.90 & 0.10 \\ 0.20 & 0.80 \end{pmatrix} = \begin{pmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{pmatrix}$$

- La probabilità cercata è 0.34

Determinazione dello stato del sistema

- Fissato lo spazio degli stati E, poiché le variabili aleatorie assumono valori in E, la legge di X_n è individuata dai numeri $\pi_1, \pi_2, \ldots$ dove, per $k \in E$:

$$\pi_k = \pi_k(n) = \mathbb{P}(X_n = x_k).$$

  $\pi(n) = (\pi_1(n), \ldots, \pi_r(n))$, $\pi(n)$ è un vettore riga di lunghezza = cardinalità di E (eventualmente infinita!).
- Se X_0 ha legge $\pi(0)$, calcoliamo la legge di X_n:

$$\pi_k(n) = \sum_{j \in E} \mathbb{P}(X_n = x_k \mid X_0 = x_j)\mathbb{P}(X_0 = x_j) = \sum_{j \in E} p_{jk}^{(n)} \pi_j(0)$$

  cioè i vettori $\pi(0)$ e $\pi(n)$ sono legati dalla relazione $\pi(n) = \pi(0)P^n$.
- Caso particolare: se X_0 assume il valore x_j con probabilità 1, diremo che la catena di Markov *parte* dallo stato x_j e la distribuzione di $\pi(n)$ coincide con la riga j-esima di P^n.

Classificazione degli stati

- Uno stato j è **raggiungibile** da uno stato i se esiste un cammino che da i arriva a j.
- Due stati i e j si dice che **comunicano** se j è raggiungibile da i e viceversa.
- Un insieme di stati S in una catena di Markov è un **insieme chiuso** se nessuno stato fuori S è raggiungibile dagli stati in S.
- Uno stato i si definisce stato **assorbente** se $p_{ii} = 1$.

Classificazione degli stati

- Uno stato i si definisce stato **transiente** se esiste uno stato j raggiungibile da i, ma i non è raggiungibile da j.
- Uno stato che non è transiente viene definito stato **ricorrente**.
- L'**orbita** dello stato i è l'insieme di tutti i tempi n per cui è positiva la probabilità di transizione in n passi da i in sé: $O(x_i) = \{n : p_{ii}^{(n)} > 0\}$.
- Uno stato i è **periodico** di periodo $k > 1$ se k è il massimo comun denominatore dell'insieme $O(x_i)$, ossia più piccolo numero tale che tutti i cammini che dallo stato i ritornano ad i hanno una lunghezza che è un multiplo di k.
- Se uno stato non è periodico si definisce **aperiodico**.
- Se tutti gli stati in una catena sono ricorrenti, aperiodici e comunicano tra loro, la catena si definisce **regolare**.

Esempio (3)

Consideriamo un gioco d'azzardo.

- Con probabilità p vinco 1\$ e con probabilità $1 - p$ perdo 1\$.
- L'obiettivo è quello di incrementare il proprio capitale fino a 4\$.
- Possiamo definire X_t come il valore del mio capitale (somma in mio possesso) dopo aver giocato t volte.
- La matrice di transizione è

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Gli stati 0 e 4 sono stati assorbenti, gli stati $\{1, 2, 3\}$ formano la classe degli stati transitori.

Tempo di arrivo

Sia A una classe di stati in E; definiamo

$$T_A = \min\{n \geq 1 \,:\, X_n \in A\}.$$

se A è formato dal solo stato a, useremo la notazione T_a invece di $T_{\{a\}}$. Osserviamo la relazione

$$p_{x,y}(n) = \sum_{m=1}^{n} \mathbb{P}^x(T_y = m) p_{y,y}(n-m).$$

Se a è uno stato assorbente allora

$$p_{x,a}(n) = \mathbb{P}^x(T_a \leq n)$$

è la probabilità di essere assorbiti in a entro n passi, partendo da x. Osserviamo ancora la relazione

$$\mathbb{P}^x(T_y = n+1) = \sum_{z \neq y} p_{x,z} \mathbb{P}^z(T_y = n).$$

Stati ricorrenti e transienti

Definiamo per ogni coppia di stati $x, y \in E$ la probabilità di arrivare, in un qualche momento, in y partendo da x

$$\rho_{x,y} = \mathbb{P}^x(T_y < \infty) = \sum_{m=1}^{\infty} \mathbb{P}^x(T_y = m).$$

Possiamo classificare gli stati in base alla probabilità di ritornare in sé

$$\begin{cases} \rho_{x,x} = 1 & x \text{ è ricorrente} \\ \rho_{x,x} < 1 & x \text{ è transiente} \end{cases} \tag{4}$$

Esempio (3-continua)

Calcoliamo la probabilità di essere assorbiti in 0 partendo da uno degli stati transitori. Scriviamo il sistema

$$\begin{cases} \rho_{1,0} = p_{1,0} + \sum_{j=1}^{3} p_{1,j}\rho_{j,0} \\ \rho_{2,0} = p_{2,0} + \sum_{j=1}^{3} p_{2,j}\rho_{j,0} \\ \rho_{3,0} = p_{3,0} + \sum_{j=1}^{3} p_{3,j}\rho_{j,0} \end{cases}$$

Da cui

$$\begin{cases} \rho_{1,0} = (1-p)[1-p(1-p)]/[1-2p(1-p)] & \text{se } p = 1/2 : 3/4 \\ \rho_{2,0} = (1-p)^2/[1-2p(1-p)] & \text{se } p = 1/2 : 1/2 \\ \rho_{3,0} = (1-p)^3/[1-2p(1-p)] & \text{se } p = 1/2 : 1/4 \end{cases}$$

Numero di passaggi

Indichiamo con $N(y)$ il numero di volte che una traiettoria passa per lo stato y, ossia

$$N(y) = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n=y\}}, \quad \mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} p_{x,y}(n).$$

Possiamo caratterizzare l'evento $(N(y) \geq 1)$ osservando che coincide con $(T_y < \infty)$, e quindi $\mathbb{P}_x(N(y) \geq 1) = \rho_{x,y}$; analogamente,

$$\mathbb{P}_x(N(y) \geq k) = \rho_{x,y}\rho_{y,y}^{k-1},$$
$$\mathbb{P}_x(N(y) = k) = \rho_{x,y}\rho_{y,y}^{k-1}(1 - \rho_{y,y}).$$

Dalla distribuzione di $N(y)$ possiamo calcolare la media

$$\mathbb{E}_x[N(y)] = \sum_{k=1}^{\infty} k\, \rho_{x,y}\rho_{y,y}^{k-1}(1-\rho_{y,y}) = \frac{\rho_{x,y}}{1-\rho_{y,y}}.$$

Allora, ricordando la definizione (4), si ottiene che, posto x lo stato iniziale, uno stato transitorio y è caratterizzato da essere visitato, in media, un numero finito di volte; uno stato ricorrente, viceversa, o non è mai raggiunto, oppure viene visitato, in media, infinite volte.

Esempio (3-continua)

Calcoliamo la probabilità di essere ritorno in uno stato transitorio $\rho_{x,x}$ per $x = 1, 2, 3$. Per $x = 2$ si ha che T_2 è necessariamente 2 oppure infinito; allora $\rho_{2,2} = \mathbb{P}^2(T_2 = 2) = 1/2$. Per $x = 1$ posso tornare in un numero pari di passi, dopo avere oscillato tra 2 e 3:

$$\mathbb{P}^1(T_1 = 2m) = [p(1-p)]^m, \qquad \rho_{1,1} = \sum_{m=1}^{\infty} [p(1-p)]^m = \frac{p(1-p)}{1-p(1-p)}.$$

Per simmetria vale lo stesso per lo stato 3. Il numero medio di passaggi in x partendo dal capitale iniziale x è:

$$\mathbb{E}_1[N(1)] = \mathbb{E}_3[N(3)] = \frac{p(1-p)}{1-2p(1-p)}, \quad \mathbb{E}_2[N(2)] = 1.$$

Classificazione degli stati

Theorem

Una catena di Markov con stati finiti ha almeno uno stato ricorrente.

Dimostrazione. Se uno stato y è transitorio, allora $\mathbb{E}_x[N(y)] < \infty$, dunque la serie $\sum_n p_{x,y}(n)$ è convergente, ed il termine n-esimo converge a zero. Se tutti gli stati fossero transitori, allora la precedente osservazione dovrebbe per ogni $y = 1, \ldots, M$; esiste allora $\bar{n}$ grande per cui $p_{x,y}(n) < \frac{1}{2M}$ per ogni y e per ogni $n > \bar{n}$; ma allora

$$1 = \sum_{y \in E} p_{x,y}(n) < \sum_{y \in E} \frac{1}{2(d+1)} < \frac{1}{2},$$

assurdo. Allora è assurda l'ipotesi che tutti gli stati siano transitori, quindi deve esistere almeno uno stato ricorrente.

□

Classi irriducibili.

Consideriamo la relazione di equivalenza $x \sim y$ sullo spazio degli stati definita da $x \sim y$ se $x \to y \to x$. Le classi di stati che vengono in questo modo individuate si dicono irriducibili. In altre parole: una classe è irriducibile se ogni coppia di elementi x, y scelti nella classe verifica $x \sim y$.

Theorem

Se x è uno stato ricorrente e $x \to y$, allora y è ricorrente, $y \to x$, $\rho_{x,y} = \rho_{y,x} = 1$.

Classi chiuse.

Una classe è chiusa se non si può uscire dalla classe: per ogni $x \in C$, $y \in E \setminus C$, per ogni $n \geq 1$ risulta $p_{x,y}(n) = 0$. Il prossimo risultato è una conseguenza del teorema 1.

Theorem

Se C è una classe chiusa, irriducibile, finita, allora C è ricorrente (ossia, tutti gli elementi di C sono ricorrenti).

Outline

Distribuzioni limite

Il Problema Sia $\pi(0)$ una generica distribuzione iniziale. Si ottiene $\pi(n) = \pi(0)P^{(n)} = \pi(0)P^n$, dove P^n è la potenza n-esima della matrice P. Per studiare il comportamento asintotico (per $n \to \infty$) del sistema sarà necessario studiare il limite di P^n per $n \to \infty$.

- Diremo che una distribuzione di probabilità su E $\pi = (\pi_x)_{x \in E}$ è **invariante** per la catena di Markov associata alla matrice P se $\pi P = \pi$.
- Se la legge dello stato iniziale è una misura invariante π, allora tutti gli stati seguenti hanno la stessa legge:

$$\pi(n) = \pi P^n = (\pi P)P^{n-1} = \pi P^{n-1} = \cdots = \pi.$$

Esistenza della distribuzione invariante

- Esistenza della distribuzione invariante.
    - **Teorema di Markov-Kakutani**. Una matrice di transizione su uno spazio finito di stati ha sempre almeno una probabilità invariante.
- Osserviamo che la probabilità invariante non è necessariamente unica. Anzi, se π^0 e π^1 sono due probabilità invarianti, allora per ogni $\lambda \in (0, 1)$, la legge $\pi^\lambda = \lambda\pi^1 + (1 - \lambda)\pi^0$ è invariante per P.
- Vale tuttavia il seguente risultato:
    - *se una catena è irriducibile, allora possiede al più una probabilità invariante*.

Catene di Markov regolari

- Diremo **regolare** una matrice di transizione che verifica:
  - per un qualche valore di n gli elementi della matrice P^n hanno tutti i valori strettamente positivi.
- Osserviamo subito che se per un qualche n la matrice P^n ammette tutti gli elementi > 0, allora ciò è vero anche per tutte le potenze seguenti. Infatti

$$p_{ij}^{(n+1)} = \sum_{k=1}^{r} p_{ik} p_{kj}^{(n)} \tag{5}$$

  e, per ipotesi, almeno un addendo è maggiore di zero; quindi, tutti gli elementi $p_{ij}^{(n+1)}$ sono positivi.
- Per dimostrare che una matrice è regolare si può utilizzare il seguente criterio:
  - se tutti gli stati comunicano tra loro ed inoltre esiste j tale che $p_{jj} > 0$, allora la catena è regolare.

Comportamento al limite per catene regolari

- Nel caso di matrici *regolari* vale il seguente risultato:
  - **Teorema di Markov**: Se la matrice P è regolare, allora

$$\lim_{n\to\infty} p_{ij}^{(n)} = p_j,$$

ossia i limiti delle probabilità $p_{ij}(n)$ esistono e sono indipendenti dall'indice i.

- La distribuzione di probabilità $\pi = (p_j)_{j\in E}$ è l'unica distribuzione invariante per la catena: infatti se il vettore $\pi(0)$ è una distribuzione invariante, risulta $\pi(0) = \pi(0)P^n$ per ogni n quindi anche al limite; allora

$$\pi(0)_j = \lim_{n\to\infty} (\pi(0)P^n)_j = \sum_{i\in E} \pi_i p_{ij}^{(n)} = \sum_{i\in E} \pi_i p_j = p_j.$$

- Una matrice regolare è **ergodica**: dopo un numero elevato di transizioni la distribuzione del sistema tende a coincidere con la distribuzione stazionaria, indipendentemente dalla distribuzione iniziale.

Calcolo della distribuzione invariante

Una distribuzione invariante verifica $\pi P = \pi$; allora possiamo scrivere

$$p_j = \sum_{k \in E} p_k p_{kj}, \qquad j = 1, 2, \ldots, r \tag{6}$$

Il sistema (6), associato alla condizione ovvia

$$\sum_{j \in E} p_j = 1,$$

permette di determinare la/e distribuzione/i invariante/i.
Nel caso di una catena regolare, la distribuzione invariante è anche unica e ergodica.

Esempio (1-continua)

Dalla matrice del consumo della Cola $P = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$ cerchiamo la distribuzione ergodica $\pi = (a;\ b)$:

$$\begin{cases} a = 0.9a + 0.2b \\ b = 0.1a + 0.8b \\ a + b = 1 \end{cases}$$

da cui

$$\begin{cases} 2a = b \\ a + b = 1 \end{cases}$$

e quindi $a = 2/3$, $b = 1/3$. Possiamo dire (la catena è ergodica) che a lungo termine, la prima bibita venderà il doppio della seconda!

Esempio (2-continua)

Abbiamo visto la matrice P per la gestione del magazzino

$$\begin{pmatrix} 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Calcoliamo la distribuzione invariante $\pi = (a\ b\ c\ d\ e)$: il sistema si scrive

$$\begin{cases} \pi = \pi P \\ a + b + c + d + e = 1 \end{cases}$$

Esclusa l'ultima, una riga si può cancellare (è combinazione lineare delle altre). Il risultato è

$$\pi = (1/9; \quad 5/27; \quad 1/3; \quad 2/9; \quad 4/27)$$

Numero di ritorni e distribuzione ergodica

Il seguente teorema lega (anche nel caso di una catena a infiniti stati) il numero di ritorni in uno stato con la distribuzione invariante. Poniamo N_x^n il numero di ritorni a partire dallo stato x effettuati entro il passo n-esimo.

Theorem

Per una catena irriducibile si ha

$$\lim_{n\to\infty} \frac{1}{n} N_x^n = \frac{1}{m_x},$$

Inoltre, se $m_x < \infty$ e poniamo $\pi_x = \frac{1}{m_x}$, allora $\pi = (\pi_1, \pi_2, \dots)$ è l'unica distribuzione invariante della catena. Viceversa, se $m_x = +\infty$ per un qualche $x \in E$, allora non esiste alcuna distribuzione invariante.

Catene periodiche

- Per ogni punto $x \in E$, costruiamo l'orbita

$$O(x) = \{\, n : p^{(n)}_{x,x} > 0 \,\}$$

Indichiamo con d_x il massimo comun denominatore dell'orbita $O(x)$.

- ◦ Se P è una matrice irriducibile, allora $d_x = d$ è costante per ogni $x \in E$.
- Per una catena irriducibile, il valore d è detto il *periodo* della catena, e una catena irriducibile con periodo $d = 1$ è detta **aperiodica**.
- Si dimostra che una catena irriducibile aperiodica è regolare, quindi ergodica; una catena periodica di periodo $d \geq 2$ non è regolare né ergodica (quindi, la distribuzione limite $\lim_{n\to\infty} \pi^{(n)} = \lim_{n\to\infty} \pi(0)P^n$ dipende dalla distribuzione iniziale $\pi(0)$.

Esercizio

Consideriamo il seguente sistema: sia $E = \{a, b, c, d\}$ lo spazio degli stati e sia P la matrice di transizione

$$P = \begin{pmatrix} 0 & q & 0 & p \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ q & 0 & p & 0 \end{pmatrix}.$$

Sia $\pi(0) = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$ lo stato iniziale del sistema.

(a) Descrivere il sistema.

(b) Determinare la distribuzione invariante del sistema.

(c) Determinare la distribuzione $\pi(10)$ del sistema a partire dallo stato $\pi(0)$.

Soluzione

Osserviamo che tutti gli stati del sistema comunicano: ad esempio, la traiettoria $a \to b \to c \to d \to a$ ha probabilità $\frac{1}{64}$.
Quindi tutti gli stati appartengono alla stessa classe, ossia il sistema è irriducibile.
Il sistema è periodico di periodo 2. Basta infatti osservare che a partire dallo stato 1, le possibilità di ritorno in 1 sono

$$p_{1,1}^{(2n)} = q, \qquad p_{1,1}^{(2n+1)} = 0.$$

Vogliamo determinare la distribuzione invariante; a tal fine, è necessario risolvere il sistema

$$\begin{cases} \pi = \pi P \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \end{cases}$$

che ha soluzione $\pi = \begin{pmatrix} \frac{q}{2} & \frac{q}{2} & \frac{p}{2} & \frac{p}{2} \end{pmatrix}$.
Il sistema non è ergodico, tuttavia partendo dallo stato $\pi(0)$ si ottiene $\pi(2n) = \begin{pmatrix} 0 & q & 0 & p \end{pmatrix}$ mentre $\pi(2n+1) = \begin{pmatrix} q & 0 & p & 0 \end{pmatrix}$.