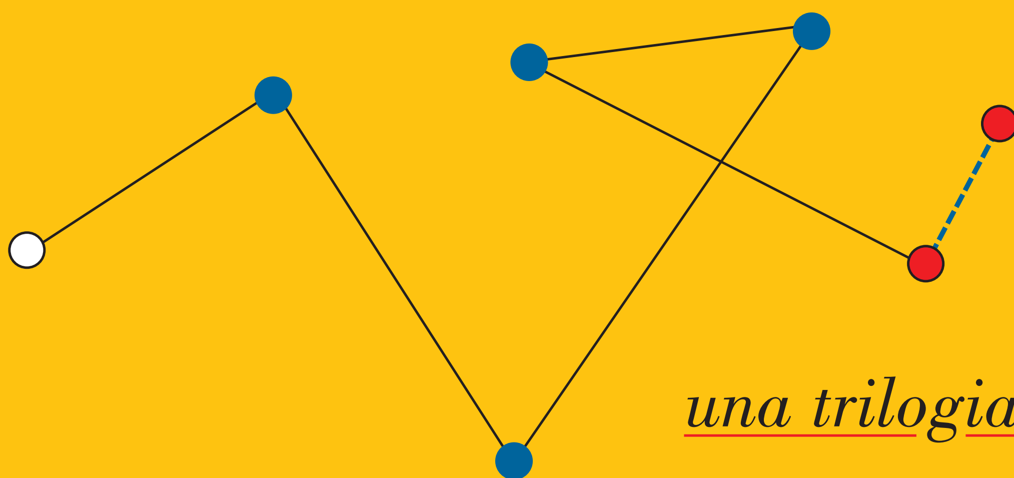
Nicolò Pintacuda

Invito alle Catene di Markov

edizione elettronica



una trilogia

Un corso minimo sulle catene di Markov

Passeggiata a caso in una rete

Variazioni su tema di Markov

Edizioni ETS

www.edizioniets.com

isbn: 978-884672367-3

variazioni su tema di markov

di nicolò pintacuda

Ai lettori più tenaci del mio *Corso minimo sulle catene di Markov* è rivolto questo lavoro, una sorta di ampio vassoio di *desserts* markoviani, preparati con cura artigianale e tecniche povere, solo scavando nella materia con curiosità e pazienza.

I temi trattati sono disparati per importanza, livello di difficoltà e originalità. Né io mi sono sforzato di costringerli entro uno schema sistematico, inevitabilmente fittizio.

Al *Corso* faccio costante riferimento, anche per la terminologia; occasionalmente, mi accadrà di rimandare al mio volume *Catene di Markov*, Edizioni ETS Pisa, 2000, immodestamente citato come il *Libro*.

1. probabilità di visita

Nelle prime pagine del *Corso* si accenna alla probabilità di visitare un dato insieme di stati, riguardata come funzione dello stato di partenza.

Nel primo dei due esempi colà illustrati essa vale sempre 1, nell'altro no, e anzi prende valori arbitrariamente piccoli.

Va detto che i due comportamenti esemplificati sono i soli possibili.

(1.1) Proposizione. *Se la probabilità di visitare un dato insieme di stati è discosta da zero*[†]*, allora vale* 1 *dappertutto.*

Dim. La prova si basa sul principio che "tanto va la gatta al lardo": per dirla in modo assai grossolano, se esiste sempre una consistente probabilità che qualcosa accada, prima o poi accadrà.

Siano H l'insieme di stati, $u_H(x)$ la probabilità che una catena X che parte da x visiti H; sia ancora T l'istante (aleatorio) della prima visita di X in H (che può essere $+\infty$). Abbreviamo u_H in u per comodità.

L'assenza di memoria delle catene di Markov (*Corso*, pag. 2) assicura che, condizionalmente al fatto di trovarsi in y a un dato istante (k) e di non

[†] cioè: minorata da un numero strettamente positivo

aver visitato H, la probabilità di non visitarlo più coincide con quella di non visitarlo quando si parte da y, quindi con $1-u(y)$; si ha insomma l'uguaglianza

$$P\big(T > k, X_k = y, (X_k, X_{k+1}, \dots) \text{ non visita } H\big) =$$

$$= P(T > k, X_k = y)\,\big(1 - u(y)\big)$$

Supposto $u \geq q > 0$, l'espressione si maggiora con

$$(1-q)\,P(T > k, X_k = y)$$

donde, sommando per tutti gli y, la relazione

$$1 - u(x) = P\big(T > k, (X_k, X_{k+1}, \dots) \text{ non visita } H\big) \leq (1-q)\,P(T > k)$$

Facendo tendere k all'infinito, gli eventi $\{T > k\}$ tendono decrescendo verso l'evento $\{T = +\infty\} = \{X \text{ non visita } H\}$ e si ottiene

$$1 - u(x) \leq (1-q)\,\big(1 - u(x)\big)$$

che comporta $u(x) = 1$. Poiché x era arbitrario, la prova è conclusa. $\diamondsuit$

La netta dicotomia espressa dalla Proposizione (1.1) si riformula come segue: i soli valori possibili per l'estremo inferiore della funzione probabilità di visitare un dato insieme di stati sono 0 e 1.

Analoga sorprendente alternativa sussiste per la probabilità di compiere *infinite* visite in un insieme dato.

Non senza rischi di confusione tipografica, indico con $v_H(x)$ la probabilità di compiere infinite visite in H, partendo da x (brevemente $v(x)$ quando H sia sottointeso). Va da sé che si ha $v_H \leq u_H$.

Esaminando le transizioni possibili a partire da x e sfruttando l'assenza di memoria, si trova la relazione $v_H(x) = \sum_y \mathrm{P}(x,y)\; v_H(y)$: come si suol dire, la funzione v_H è *armonica*. Per dirla tutta:

(1.2) LEMMA. *La funzione v_H è la massima minorante armonica di u_H.*

DIM. Abbreviamo u_H in u e v_H in v per comodità.

Supponiamo che f sia armonica maggiorata da u e facciamo vedere che è anche maggiorata da v. Serve all'uopo un artificio simile a quello adoperato per ricavare la formula (2.4) del *Corso*.

Si tratta di far entrare in gioco la probabilità $g_n(x)$ che una catena X partita da x e *osservata dall'istante n in poi*, visiti H.

In simboli,

$$g_n(x) = P(A_n) \quad \text{con} \quad A_n = \{(X_n, X_{n+1}, \dots) \text{ visita } H\}$$

Si vede che gli eventi A_n sono decrescenti e la loro intersezione è l'evento

$$\{X \text{ visita } H \text{ infinite volte}\}$$

Perciò le funzioni g_n tendono (decrescendo) verso v.

È d'altra parte evidente che $g_0 = u$; ragionando sulla prima transizione e sfruttando l'assenza di memoria, si stabilisce poi la relazione ricorrente

$$g_{n+1}(x) = \sum_y \mathrm{P}(x,y)\, g_n(y)$$

Dal momento che f è armonica e maggiorata da $u = g_0$, si deve avere $f \leq g_n$ per ogni n; facendo tendere n all'infinito, si ottiene $f \leq v$. $\diamondsuit$

L'omologa della Proposizione (1.1) è la seguente

(1.3) Proposizione. *Se la probabilità di visitare infinite volte un dato insieme è discosta da 1*[†]*, allora essa è identicamente nulla.*

Dim. Sia $u(x)$ la probabilità di visitare H per una catena X che parte da x; la probabilità $v(x)$ di infinite visite sia maggiorata da s, con $0 < s < 1$.

Affinché X visiti H infinite volte, occorre intanto che lo visiti una prima volta, e poi che vi compia ulteriori infinite visite.

Sfruttando questa volta la forma *forte* dell'assenza di memoria al tempo aleatorio T del primo ingresso di X in H (*Corso*, Appendice 10.1 pag. 21), $v(x)$ sarà il prodotto della probabilità di visita $u(x)$ per quella di infinite visite a partire dal punto (aleatorio) di H nel quale avviene il primo ingresso di X, questa comunque maggiorata da s per ipotesi[‡].

Se ne trae la relazione $v \leq su$; vi si legge che la funzione (armonica) v/s è maggiorata da u, e dunque da v (Lemma 1.2), e si ottiene $v \leq sv$, disuguaglianza che forza v ad annullarsi dappertutto. $\diamondsuit$

Facciamo sosta per raccogliere un po' di commenti.

Se l'estremo inferiore della funzione u_H è nullo, lo è *a fortiori* quello della funzione v_H, che ne è maggiorata; se v_H vale dappertutto 1, lo stesso accade per u_H. Reciprocamente, se la funzione u_H è identicamente 1, lo è anche v_H: dopo ogni visita in H, la catena si trova infatti in luoghi da cui rivisita l'insieme H qc, e questo indefinitamente.

Ne discende l'alternativa seguente:

Ci sono tre possibilità per la funzione probabilità di infinite visite:

[†] cioè: maggiorata da un numero strettamente minore di 1

[‡] discorso intuitivamente chiaro, servirebbe però maggior rigore; il lettore saprà precisare i dettagli come nella prova della (1.1)

(a) *la funzione è identicamente nulla;*
(b) *la funzione vale identicamente* 1*;*
(c) *la funzione è* "divaricata"*: ha estremo inferiore nullo e estremo superiore uguale a* 1

Il caso (a) si manifesta ad esempio se H si riduce a un singolo stato transitorio, il caso (b) per ogni catena connessa ricorrente (*Corso*, pag. 5 e pag. 6). Del caso (c) non è arduo trovare esempi; qui sorvolo, sul tema avrò modo di tornare più avanti; mi piace solo battezzare *polidromo* un insieme H se la probabilità di infinite visite in H è non costante (e pertanto "divaricata").

Nella prova della (1.3), basta invero che v_H sia discosta da 1 *sull'insieme* H; se non è identicamente nulla, il suo estremo superiore preso su H deve dunque valere 1.

Per visitare infinite volte un insieme, il fatto di trovarsi già nell'insieme non è un vantaggio: nell'esempio in Appendice la funzione probabilità di infinite visite ha estremo inferiore nullo nell'insieme e estremo superiore uguale a 1 nel complementare !

Osservazione importante: se la catena è connessa e l'insieme H polidromo, la funzione v_H *non ha massimo né minimo*.

È infatti armonica; se si annulla nello stato a, deve fare altrettanto in ogni stato raggiungibile da a con una transizione; passo dopo passo, si raggiunge qualsiasi stato in virtù della connessità, e dunque la funzione si annulla dappertutto. Analogo ragionamento per il massimo.

2. eremitaggio

Prendiamo una catena connessa transitoria. La probabilità $u_a(x)$ di visitare lo stato a partendo da x non si annulla mai (connessità) né vale 1 identicamente (transitorietà); per effetto della Proposizione (1.1), deve avere estremo inferiore nullo.

Detto in modo colloquiale, esistono posti da cui visitare lo stato a riesce arbitrariamente improbabile.

La cosa notevole è che la dinamica della catena la spinge automaticamente verso questi luoghi, lontana dal mondo come un eremita.

Più precisamente:

(2.1) PROPOSIZIONE. *Se X è una catena connessa transitoria, allora quasi certamente la successione $u_a(X_n)$ tende a zero per ogni a.*

DIM. Di nuovo sfrutteremo il principio della gatta e del lardo, che usammo nella prova della Proposizione (1.1), ma in una variante più raffinata.

Faremo vedere che la catena non può muoversi indefinitamente in una regione da cui rimane consistente la probabilità di visitare a, senza dover prima o poi visitarlo, e infinite volte; cosa questa incompatibile con la natura transitoria di a.

Poniamo $G^q = \{u_a > q\}$, con $0 < q < 1$; siano $S_1, S_2, \ldots$ i tempi delle successive visite della catena X_n in G^q, e sia T il tempo della sua prima visita in a (tutti aleatori, che possono anche valere $+\infty$).

Proviamo intanto che ha probabilità nulla l'evento

$$\Omega^* = \{X \text{ visita } G^q \text{ infinite volte e non visita } a\}$$

Per assenza di memoria (forte) si può scrivere

$$P\big(S_k < \infty, T > S_k, X_{S_k} = y, (X_{S_k}, X_{S_k+1}, \ldots) \text{ non visita } a\big) =$$

$$= P(S_k < \infty, T > S_k, X_{S_k} = y)\ \big(1 - u_a(y)\big) \leq$$
$$\leq (1-q)\ P(S_k < \infty, T > S_k, X_{S_k} = y)$$

da cui sommando per tutti i valori di y (in G^q) si ottiene

$$P(S_k < \infty, T = \infty) \leq (1-q)\ P(S_k < \infty, T > S_k)$$

Quando k tende all'infinito, gli eventi al primo e quelli al secondo membro tendono decrescendo verso Ω^* e si ha

$$P(\Omega^*) \leq (1-q)\, P(\Omega^*)$$

da cui $P(\Omega^*) = 0$.

È dunque provato che la probabilità di infinite visite in G^q è maggiorata da u_a, quindi (per effetto del Lemma 1.2), dalla probabilità di infinite visite in a, che peraltro è nulla identicamente.

Se ne deduce che (qc) la catena visita G^q un numero finito di volte e poi ne esce definitivamente; facendo intervenire valori sempre più piccoli di q, si vede che (qc) la successione $u_a(X_n)$ tende a zero; per riunione numerabile di insiemi trascurabili, la conclusione vale qc per ogni a. ♢

Nell'esempio (2.3) pag. 3 del *Corso* (la passeggiata asimmetrica), la catena tende qc a $+\infty$, in accordo con la Proposizione.

Se si vuole sapere tutta la verità, nel *Libro* (Teorema 4.23, pag. 50) è dimostrato il seguente

(2.2) TEOREMA. *Ciascuna delle successioni $u_H(X_n)$ e $v_H(X_n)$ tende qc verso l'indicatrice*[†] *dell'evento* $\{X$ visita H *infinite volte*$\}$.

[†] la variabile aleatoria che vale 1 se l'evento è verificato e zero altrimenti

Ma la prova è tecnicamente più difficile.

3. priorità

In molte situazioni (si pensi in particolare a modelli di tipo biologico oppure economico) ha speciale interesse considerare la probabilità che una catena visiti un dato stato (vantaggioso) senza averne prima visitato un determinato altro (dannoso).

È spontaneo voler dedurre queste probabilità dalle probabilità di visita "non vincolate", trattate nel capitolo 1.

Se la catena è (connessa) ricorrente, le funzioni probabilità di visita sono tutte identicamente uguali a 1, non possono quindi dare informazione alcuna; diverso e assai più ricco si presenta lo scenario se la catena è transitoria.

Nel caso transitorio, infatti, la probabilità $u_{a,b}(x)$ che una catena partita da x visiti b senza prima esser passata da a si sa esprimere come combinazione lineare delle funzioni u_a e u_b (probabilità di visitare a , b rispettivamente).

Il trucco sta nel trattare simultaneamente la probabilità "speculare" $u_{b,a}$; usando l'assenza (forte) di memoria al tempo (aleatorio) d'ingresso in b, si scrive la relazione

$$u_a(x) = u_{b,a}(x) + u_{a,b}(x)\, u_a(b)$$

cui corrisponde l'analoga che si ottiene scambiando a con b.

Se ne ricava

$$u_{b,a} = \frac{u_a - u_a(b)\, u_b}{1 - u_a(b)\, u_b(a)} \tag{3.1}$$

e l'espressione simmetrica per $u_{a,b}$.

Notiamo che il sistema lineare è risolubile, poiché il prodotto $u_a(b)\, u_b(a)$ rappresenta la probabilità di partire da a, visitare b, ripartirne e visitare a, di uno cioè dei modi di tornare in a partendo da a; questo prodotto è pertanto maggiorato dalla probabilità di ritorno in a, sicuramente minore di 1 dato che lo stato è transitorio.

Il lettore si guarderà dal ritenere che la somma $u_{a,b} + u_{b,a}$ debba sempre valere 1 (la catena potrebbe non visitare a e nemmeno b.)

Se visitare uno stato è più probabile che visitarne un altro, non per questo la probabilità di visitare il primo senza passare dal secondo dev'essere maggiore di quella (speculare) di visitare il secondo senza passare dal primo.

In simboli: $u_a(c) > u_b(c)$ *non* implica $u_{b,a}(c) > u_{a,b}(c)$.

Alla smentita si presta l'esempio (2.3) del *Corso*, modificato ponendo $\mathrm{P}(x, x+1) = 3/5$ e $\mathrm{P}(x, x-1) = 2/5$; facilmente si ottiene $u_a(x) = 1$ per

$x \le a$ e $u_a(x) = (2/3)^{(x-a)}$ per $x > a$; se si prende $a = 3$, $b = 0$, $c = 1$, si trova allora

$$u_3(1) = 1 > \frac{2}{3} = u_0(1) \qquad u_{0,3}(1) = \frac{9}{19} < \frac{10}{19} = u_{0,3}(1)$$

4. tempi di scambio

Cambiamo scenario, per collocarci nell'ambito delle catene connesse *veloci*.

In queste ipotesi, il tempo di attesa perché una catena partita dallo stato r compia la sua prima visita in a ha valor medio finito $t(r, a)$; è il Teorema 6.1 a pag. 10 del *Corso*.

Esaminando con cura la prova di quel teorema, vediamo che il secondo membro della maggiorazione (6.2) rappresenta il tempo aleatorio per la prima visita in r dopo che la catena ha visitato a.

Per la forma forte di assenza di memoria, il valor atteso di questo secondo membro è dunque la somma $s(r, a) = t(r, a) + t(a, r)$, che diremo *tempo di scambio* tra i due stati.

Intuitivamente, è il tempo medio per scambiare un messaggio tra i due stati tramite un messaggero aleatorio.

Introduciamo la probabilità $q(r, a)$ che una catena partita da r visiti a senza prima esser tornata nello stato di partenza r (la diremo probabilità *di escursione* in a); il numero q che appare nella (6.3) della dimostrazione del *Corso* è $1 - q(r, a)$ e la somma della serie geometrica vale $q(r, a)^{-1}$.

Veniamo al valor medio $E(R_1)$ del tempo di ritorno in r; detta π la distribuzione invariante di probabilità (unica per la Proposizione 5.5 pag. 9 del *Corso*), esso coincide con il reciproco di $\pi(r)$ (ultime righe del capitolo 7).

Così rimesso a nuovo, il Teorema (7.1) diventa

(4.1) TEOREMA. *In una catena connessa veloce con distribuzione invariante π di probabilità, il tempo di scambio tra due stati distinti r e a è dato dalla relazione*

$$s(r, a) = \frac{1}{\pi(r)\, q(r, a)} \tag{4.2}$$

dove $q(r, a)$ indica la probabilità di escursione da r in a.

Un facile corollario del teorema è la relazione di reciprocità

$$\pi(x)\, q(x, y) = \pi(y)\, q(y, x) \qquad (x \ne y) \tag{4.3}$$

che rivela l'esistenza di una simmetria inaspettata; di qui l'ovvia maggiorazione

$$q(x,y) \leq \frac{\pi(y)}{\pi(y)} \tag{4.4}$$

5. ritorno a casa

Farò vedere in questo capitolo che per tornare da un posto molto lontano si fa assai più in fretta che per andarci.

Quest'affermazione, ermetica e paradossale, ha evidente bisogno di precisazioni.

Che vuol dire "lontano", dal momento che non è data alcuna distanza nello spazio degli stati ?

Aggiriamo l'ostacolo definendo una nozione di convergenza per una funzione sullo spazio degli stati: la funzione positiva f *tende a zero*† se, per ogni reale $c > 0$, l'insieme $\{f > c\}$ è finito.

Intuitivamente, una funzione tende a zero se diventa arbitrariamente piccola fuori da un insieme finito abbastanza grande ("lontano"); va da sé che supporremo infinito lo spazio degli stati, ché altrimenti ogni funzione tenderebbe a zero banalmente.

Riformulato con precisione e con le notazioni del capitolo precedente, il mio enunciato è il seguente:

(5.1) Teorema. *In una catena connessa veloce, per ogni a fissato, la funzione*

$$f(x) = \frac{t(x,a)}{t(a,x)} \quad \text{per } x \neq a \qquad f(a) = 0$$

tende a zero.

Dim. Sia S il tempo di ritorno in a della catena X che parte da a, e ne sia $T = T_x$ il tempo d'ingresso in x (con $x \neq a$); indichiamo come al solito con $q(a,x)$ la probabilità di escursione da a in x.

Un quoziente f/g tende a zero se e solo se lo fa $f/(f+g)$; si tratta dunque di mostrare che tende a zero la funzione

$$\frac{t(x,a)}{s(a,x)}$$

chiaramente maggiorata, in virtù della (4.2), da $h(x) = q(a,x)\, t(x,a)$.

† secondo il filtro delle parti cofinite, direbbero i matematici

Facciamo vedere che quest'ultima funzione tende a zero.

Ancora una volta giocheremo sulla forma forte dell'assenza di memoria.

I lettori sono ormai abbastanza smaliziati per verificare che T è un tempo d'arresto (*Corso*, punto a) dell'Appendice 10.1 pag. 21), e assicurarsi poi che l'evento $A = \{T < S\}$ è "anteriore a T" (punto b) di quell'Appendice).

Condizionalmente all'evento A, la catena $(X_T, X_{T+1}, \ldots)$ ha dunque la stessa legge di una catena che parte da x; d'altra parte, $S - T$ è il suo tempo di primo ingresso in a.

Pertanto $t(x,a)$ è il valore atteso di $(S-T)$ *condizionato all'evento* A, cioè calcolato rispetto alla probabilità condizionata $P_A(\cdot) = P(\cdot \cap A)/P(A)$; lo indicheremo con $E_A(S-T)$.

Per ogni intero k vale l'ovvia disuguaglianza

$$S - T \le S \le S I_{\{S>k\}} + k$$

da cui si trae

$$t(x,a) \le E_A\big(S I_{\{S>k\}}\big) + k \le \frac{E\big(S I_{\{S>k\}}\big)}{P(A)} + k$$

Moltiplicando per $P(A) = q(a,x)$ e tenendo conto della (4.4), perveniamo alla relazione

$$h(x) = q(a,x)\, t(x,a) \le E\big(S I_{\{S>k\}}\big) + k\, \frac{\pi(x)}{\pi(a)}$$

Sia $c > 0$; poiché $S I_{\{S>k\}}$ tende decrescendo a zero quando k tende all'infinito, per un opportuno valore di k si avrà $E\big(S I_{\{S>k\}}\big) < c/2$; l'insieme $\{h > c\}$ ne risulterà incluso nell'insieme (finito) $\{x : \pi(x) > c\, \pi(a)/2k\}$.

Il teorema è provato. ♢

6. saliscendi

Esaminiamo ora un modello di catena che in un certo senso generalizza la passeggiata degli esempi di pag. 3 del *Corso*.

Gli stati sono i numeri naturali, le transizioni sono possibili solo da uno stato a quelli adiacenti, oppure a se stesso; a differenza però dagli esempi citati, le probabilità di transizione sono in generale diverse da punto a punto.

Più precisamente, sono date da

$$\mathrm{P}(x, x+1) = p_x > 0, \qquad \mathrm{P}(x,x) = r_x, \qquad \mathrm{P}(x, x-1) = q_x > 0 \text{ per } x > 0$$

Con la convenzione $q_0 = 0$, si ha $p_x + q_x + r_x = 1$ per ogni x.

Catene di tal genere sono chiaramente connesse, e offrono un modello ricco e sufficientemente flessibile da prestarsi a rappresentare svariate situazioni interessanti†.

I citati esempi del *Corso*, opportunamente modificati perché qui non ci sono gli interi negativi, mostrano che esistono saliscendi ricorrenti e saliscendi transitori.

Fermiamo l'attenzione sul caso transitorio.

La catena abbandona prima o poi definitivamente ogni prefissato insieme finito di stati, e dunque tende all'infinito; ma poiché non sa scavalcare nessuno stato, visita tutti quelli alla sua destra.

Si deve pertanto avere sempre $u_a(x) = 1$ per $x \le a$; per $x > a$, di contro, dev'essere $u_a(x) < 1$, altrimenti troveremmo due stati a e b per i quali il prodotto $u_a(b)\, u_b(a)$ vale 1, e questo non può accadere‡.

Un uso (ormai disinvolto) dell'assenza (forte) di memoria porta poi a stabilire, per $0 < a < x$, la relazione

$$u_0(x) = u_a(x)\, u_0(a) \tag{6.1}$$

Alla luce della discussione precedente, essa mostra che la probabilità u_0 di visitare l'*origine* è una funzione strettamente decrescente; perciò tende verso il proprio estremo inferiore, che è zero in virtù della dicotomia (1.1).

Reciprocamente

(6.2) PROPOSIZIONE. *Data una funzione che vale* 1 *nell'origine e decresce strettamente verso zero, esiste un saliscendi per il quale è la probabilità di visitare l'origine.*

DIM. Sia g la funzione assegnata, e introduciamo le probabilità di transizione

$$r_x = 0 \quad (x \ge 0) \qquad p_0 = 1 \qquad p_x = \frac{g(x-1) - g(x)}{g(x-1) - g(x+1)} \quad (x \ge 1)$$

Un calcolo elementare permette di verificare che per $x > 0$ vale l'identità $p_x g(x+1) + q_x g(x-1) = g(x)$; in altre parole, la funzione g è armonica fuori dell'origine, come dev'essere la probabilità u_0 di visitare l'origine (Proposizione 2.5 pag. 4 del *Corso*).

Basta questo per concludere che $u_0 = g$? A priori no, ma consideriamo la differenza $f = u_0 - g$; è una funzione nulla nell'origine, armonica fuori e nulla all'infinito.

---

† in letteratura sono generalmente note con il nome, un tantino drammatico, di catene *di nascita e morte*

‡ vedi osservazione dopo la (3.1)

Supponiamo per assurdo che esista un punto c nel quale $f(c) > 0$; per ipotesi, f sarà minore di $f(c)$ fuori da un insieme finito F, e pertanto assumerà un massimo $M > 0$ in un punto a di quest'insieme (diverso dall'origine).

Giocando sull'armonicità, la funzione varrà M in tutti i punti raggiungibili da a con una transizione e, passo dopo passo, nell'origine, dove però si annulla.

L'assurdo prova che $f \leq 0$; applicando quanto sopra alla funzione $-f$, si conclude che f è identicamente nulla, e $u_0 = g$. ♢

Torniamo ora (6.1); da questa relazione emerge che tutte le probabilità di visita si possono ottenere a partire dalla probabilità di visitare l'origine, e precisamente si ha

$$u_a = 1 \wedge \frac{u_0}{u_0(a)} \tag{6.3}$$

Ne risulta, sostituendo nella (3.1), un'espressione semplice per la probabilità $u_{b,a}$ di visitare a senza aver prima visitato b ($a < b$), nella quale compare solo la funzione probabilità di visitare l'origine:

$$u_{b,a}(x) = \frac{u_0(x) - u_0(b)}{u_0(a) - u_0(b)} \qquad (a \leq x \leq b) \tag{6.4}$$

(ovviamente si trova 1 per $x < a$ e zero per $x > b$).

La flessibilità del modello del saliscendi è ulteriormente testimoniata dalla seguente

(6.5) PROPOSIZIONE. *Data una misura positiva che non si annulli in alcun singoletto, esiste un saliscendi per il quale è una misura invariante.*

DIM. Sia σ la misura (*distribuzione*, nel linguaggio del *Corso*), individuata dai valori $\sigma(x)$ nei singoletti.

Definiamo per ricorrenza le probabilità di transizione attraverso le condizioni:

$$q_o = 0 \qquad 0 < p_0 < \frac{\sigma(1)}{\sigma(0)} \wedge (1 - q_0)$$

$$q_{x+1} = \frac{p_x\, \sigma(x)}{\sigma(x+1)} \qquad 0 < p_{x+1} < \frac{\sigma(x+2)}{\sigma(x+1)} \wedge (1 - q_{x+1})$$

Lascio al lettore il compito, facile ma fastidioso, di controllare che le condizioni sono compatibili e che per ogni x si ha $p_x + q_x < 1$, così che si può correttamente definire $r_x = 1 - p_x - q_x$.

Importa notare che la relazione $q_{x+1}\, \sigma(x+1) = p_x\, \sigma(x)$ vale per ogni x, sicché sommandola a quella scritta con $x - 1$ al posto di x si ottiene

$$\sigma(x) - r_x\, \sigma(x) = (p_x + q_x)\, \sigma(x) = q_{x+1}\, \sigma(x+1) + p_{x-1}\, \sigma(x-1)$$

che sancisce l'invarianza di σ. $\diamondsuit$

Se si prende una distribuzione di probabilità, si vede che esiste un saliscendi che la ammette come invariante, dunque un saliscendi ricorrente veloce (Proposizione 7.3 pag. 14 del *Corso*).

Altre notizie sulle catene di questo capitolo si trovano alle pagg. 136–140 del *Libro*.

7. ricorrenza sugli interi relativi

La passeggiata semplice (*Corso*, pag. 3) è un esempio di catena "omogenea", nel senso che la probabilità di una transizione dallo stato x allo stato y dipende solo dalla differenza $y - x$.

S'intuisce che, a cagione dell'omogeneità, gli stati saranno tutti transitori oppure tutti ricorrenti.

Vi è una distribuzione di probabilità $\lambda(z)$ che regola gli spostamenti, le probabilità di transizione sono date da $\mathrm{P}(x, x+z) = \lambda(z)$.

Nella passeggiata semplice questa distribuzione (che chiamerò *legge di spostamento*) è data da $\lambda(-1) = \lambda(+1) = 1/2$.

Che succede con una legge di spostamento più complicata ? Supponiamo per esempio che sia

$$\lambda(0) = 3/8, \qquad \lambda(-2) = \lambda(+2) = 1/16, \qquad \lambda(-1) = \lambda(+1) = 1/4 \tag{7.1}$$

Ebbene, anche in questo caso la catena è ricorrente; userò una tecnica che consente anzi di stabilire la ricorrenza per una classe ben più ampia di leggi di spostamento.

Premetto un lemma, che si riduce a un calcolo banale.

(7.2) LEMMA. *Se la funzione g è armonica, $E(g(X_n))$ non dipende da n.*

DIM. Senza parole,

$$E\big(g(X_{n+1})\big) = \sum_y P(X_{n+1} = y)\, g(y) = \sum_{x,y} P(X_n = x, X_{n+1} = y)\, g(y) =$$

$$= \sum_{x,y} P(X_n = x)\ \mathrm{P}(x,y)\, g(y) = \sum_x P(X_n = x) \sum_y \mathrm{P}(x,y) g(y) =$$

$$= \sum_x P(X_n = x)\, g(x) = E\big(g(X_n)\big)$$

$\diamondsuit$

Notiamo intanto che la legge (7.1) è simmetrica rispetto all'origine; quello che più conta, è *centrata*: lo spostamento aleatorio ha media nulla, in simboli si ha $\sum_z \lambda(z)\, z = 0$. Equivale a dire che la funzione $f(x) = x$ è armonica:

$$\sum_z \mathrm{P}(x, x+z) g(x+z) = \sum_z \lambda(z)(x+z) = x + \sum_z \lambda(z) z = x = f(x)$$

Tolto allora un intervallo finito F opportuno (in questo caso basta prendere gli interi da -2 a $+2$), la funzione $g(x) = |x|$ è armonica: per $a \notin F$, infatti, le transizioni portano sempre nella stessa semiretta di a, dove la funzione g coincide con x oppure con $-x$ (a seconda del segno di a).

Per $a \notin F$, sia N un intero maggiore di $|a|$ e poniamo $G = \{g > N\}$.

Voglio ragionare sulla probabilità che la catena X che parte da a visiti G senza aver prima visitato F.

All'uopo farò una piccola operazione di chirurgia (anti)estetica che consiste nel modificare le probabilità di transizione ponendo $\mathrm{P}(x, x) = 1$ per $x \in F \cup G$; la catena ne sarà stravolta, ma la probabilità che mi interessa non ne risulterà alterata.

In compenso, la funzione g, già armonica fuori di F, diverrà armonica *dappertutto*, e avremo $E(g(X_n)) = E(g(X_0)) = g(a)$ per ogni n.

Supponiamo per assurdo che la catena sia transitoria; essa visita qc l'insieme $F \cup G$ (che ha complementare finito); indichiamo con T il tempo della prima visita.

A motivo della chirurgia, la catena si ferma all'ingresso in $F \cup G$, sicché la differenza $g(X_n) - g(X_T)$ si annulla sull'evento $\{T \le n\}$ e, per costruzione, non supera mai $2N$ in valore assoluto. Pertanto

$$\big|g(a) - E\big(g(X_T)\big)\big| = \big|E\big(X_n) - g(X_T)\big)\big| \le 2N\, P(T > n)$$

Se n tende all'infinito, $P(T > n)$ tende a zero, e quindi $g(a) = E(g(X_T))$; sull'evento $A = \{X \text{ non visita } F\}$, la posizione X_T sta in G, e $g(X_T) \ge N$; passando alla scrittura integrale, che ha meno parentesi,

$$g(a) = \int g(X_T) \ge \int_A g(X_T) \ge N\, P(A) = N\,\big(1 - u_F(a)\big)$$

dove u_F indica come al solito la funzione probabilità di visitare F.

Dato che N era arbitrario, ne segue $u_F(a) = 1$; ma anche a era arbitrario in F^c, perciò u_F vale identicamente 1, contraddicendo l'ipotesi che la catena sia transitoria (vedi pag. 3).

La catena con legge di spostamento (7.1) è davvero ricorrente. ♢

Il ragionamento precedente resta valido per una qualsiasi legge di spostamento centrata e con supporto finito. Ma l'ipotesi di supporto finito non è realmente necessaria, vedi *Libro*, Teorema (11.10) pag. 171.

8. convergenza in legge

Uno dei teoremi più emozionanti del *Corso* è senza dubbio quello della convergenza alla stazionarietà (capitolo 5).

Con opportune ipotesi, esso stabilisce che la probabilità di trovarsi nei vari stati tende con il tempo al valore di una distribuzione limite di probabilità, la quale dipende solo dalla dinamica delle probabilità di transizione ma non dalle condizioni di partenza.

Se si vuole formulare in modo più suggestivo questo risultato, è bene guardare alle distribuzioni e alla loro convergenza con occhio più globale.

Diciamo intanto che una legge (o misura) di probabilità μ andrebbe riguardata, come s'usa tra gente bene educata, come una funzione *d'insieme*; il valore $\mu(A)$ si ottiene semplicemente sommando per tutti gli elementi $x \in A$ i valori della *distribuzione* $\mu(x)$.

Per esprimere in modo pulito l'avvicinarsi delle leggi a un limite, occorre saper dire di quanto "differiscono" due leggi; una ragionevole misura di difformità appare l'estremo superiore, per tutti i possibili insiemi A, del valore assoluto delle differenze tra i valori assunti in A dalle due leggi; in simboli:

$$\delta(\mu, \nu) = \sup_A \big|\mu(A) - \nu(A)\big| \tag{8.1}$$

Facilmente si controlla che la (8.1) è una vera propria *distanza* tra leggi, con tutte le proprietà formali delle distanze; è sempre maggiorata da 1.

La differenza $\mu(A) - \nu(A)$ cambia segno se si passa dall'insieme A al complementare, ragion per cui possiamo tranquillamente sopprimere il valore assoluto nella definizione (8.1), e scrivere

$$\delta(\mu, \nu) = \sup_A \big|\mu(A) - \nu(A)\big| = \sup_A \big(\mu(A) - \nu(A)\big) \tag{8.2}$$

Operiamo ora un piccolo ritocco nella dimostrazione di pag. 9 del *Corso*; sommando per $x \in A$ l'uguaglianza $P(T \le n, X_n = x) = P(T \le n, Y_n = x)$ otteniamo $P(T \le n, X_n \in A) = P(T \le n, Y_n \in A)$; da cui

$$P(X_n \in A) - P(Y_n \in A) = P(T > n, X_n \in A) - P(T > n, Y_n \in A) \le$$

$$\le P(T > n)$$

Con le ipotesi del teorema, $P(X_n \in A)$ è il valore in A della legge μ_n di X_n (come a pag. 6 del *Corso*), mentre $P(Y_n \in A) = \pi(A)$.

Si ha dunque

$$\mu_n(A) - \pi(A) \leq P(T > n)$$

Prendendo l'estremo superiore su A e rammentando la definizione (8.2), perveniamo all'elegante e concisa enunciazione ("globale")

$$\delta(\mu_n, \pi) \leq P(T > n) \to 0 \tag{8.3}$$

Un esempio di convergenza alla stazionarietà, del tutto banale, si ha quando le probabilità di transizione non dipendono dallo stato di partenza, esiste cioè una legge λ per la quale

$$\mathrm{P}(x, y) = \lambda(y) \tag{8.4}$$

In questo caso è elementare calcolare

$$\mu\mathrm{P}(x) = \sum_y \mu(y)\, \mathrm{P}(y, x) = \sum_y \mu(y)\, \lambda(x) = \lambda(x)$$

da cui concludere che la legge λ è invariante e accorgersi, grazie alla (4.1) di pag. 6 del *Corso*, che $\mu_1 = \mu_2 = \ldots = \lambda$.

In altri termini, la convergenza alla stazionarietà è istantanea (già al primo passo).

Vediamo un esempio meno brutale; sia π una legge invariante per le probabilità di transizione P, e poniamo

$$\overline{\mathrm{P}}(x, y) = r\, \mathrm{P}(x, y) + (1 - r)\, \pi(y) \tag{8.5}$$

Affinché la (8.5) definisca delle probabilità di transizione, occorre e basta che sia $0 \leq r \leq 1$; per $r = 0$ si riottiene l'esempio precedente, per $r = 1$ di nuovo P; quindi supporremo $0 < r < 1$.

È immediato riconoscere che π è invariante anche per le probabilità di transizione $\overline{\mathrm{P}}$.

Prima di studiare questo caso, occorre osservare che se si trasformano due leggi mediante probabilità di transizione (come nella 4.1 del *Corso*), la distanza tra le leggi trasformate non supera mai quella tra le leggi di partenza; in simboli, e con notazione concisa:

$$\delta(\mu\mathrm{P}, \nu\mathrm{P}) \leq \delta(\mu, \nu) \tag{8.6}$$

La verifica non è difficile:

$$\mu\mathrm{P}(A)-\nu\mathrm{P}(A)=\sum_{x\in A}\big(\mu\mathrm{P}(x)-\nu\mathrm{P}(x)\big)=\sum_{x\in A}\sum_{y}\big(\mu(y)-\nu(y)\big)\,\mathrm{P}(y,x)\le$$

$$\le\sum_{x\in A}\sum_{y}\big(\mu(y)-\nu(y)\big)^{+}\,\mathrm{P}(y,x)\le\sum_{y}\big(\mu(y)-\nu(y)\big)^{+}=\sum_{y\in B}\big(\mu(y)-\nu(y)\big)=$$

$$=\mu(B)-\nu(B)\le\delta(\mu,\nu)$$

dove l'esponente $+$ denota com'è usuale la parte positiva di un numero[†] e B è l'insieme $\{y:\mu(y)\ge\nu(y)\}$.

Prendendo l'estremo superiore su A, la (8.6) è dimostrata. $\diamondsuit$

Torniamo all'esempio (8.5); per ogni legge μ vale l'identità

$$\mu\overline{\mathrm{P}}(A)=r\,\mu\mathrm{P}(A)+(1-r)\,\pi(A)$$

da cui, tenendo conto della (8.6), si trae

$$\mu\overline{\mathrm{P}}(A)-\pi(A)=r\,\big(\mu\mathrm{P}(A)-\pi\mathrm{P}(A)\big)\le r\,\delta(\mu\mathrm{P},\pi\mathrm{P})\le r\,\delta(\mu,\pi)$$

e in definitiva

$$\delta(\mu\overline{\mathrm{P}},\pi)\le r\,\delta(\mu,\pi)$$

Per iterazione, si avrà allora

$$\delta(\mu_{n+1},\pi)=\delta(\mu_n\overline{\mathrm{P}},\pi)\le r\,\delta(\mu_n,\pi)$$

e dunque

$$\delta(\mu_n,\pi)\le r^n \tag{8.7}$$

La maggiorazione (8.7) mostra in modo chiaro che le leggi μ_n tendono alla legge invariante π.

Brillanti quanto si vuole, gli esempi (8.4) e (8.6) riescono però leggermente fuorvianti. In entrambi i casi, infatti, la convergenza alla legge limite è molto veloce: istantanea nel primo, nel secondo è più veloce di una progressione geometrica (di ragione r).

In molte situazioni in cui sussiste convergenza, questa può rivelarsi in realtà assai lenta. Arbitrariamente lenta, oserei dire. Vale infatti la

[†] il numero stesso se positivo, zero altrimenti

(8.8) PROPOSIZIONE. *Data una successione di numeri reali che decresce strettamente verso zero, esiste una catena che converge in legge più lentamente della successione.*

DIM. Non si perde in generalità supponendo $t_0 < 1$; la scrittura

$$\pi\{n+1, n+2, \ldots\} = t_n$$

definisce una legge di probabilità sull'insieme dei naturali.

Sappiamo dal teorema (6.5) che esiste una catena di tipo "saliscendi" per la quale π è invariante, dunque ricorrente e veloce; dalla dimostrazione di quel teorema emerge che i numeri r_x non si annullano, per cui la catena (connessa) è ultraconnessa (*Corso*, Proposizione 5.4, pag. 9).

Supponiamo che la catena parta da 0; ci sono le ipotesi del Teorema di convergenza del *Corso*, e dunque si ha convergenza in legge verso π, ossia le distanze $\delta(\mu_n, \pi)$ tendono a zero.

Ma, per costruzione, all'istante n la catena deve trovarsi entro l'insieme $A_n = \{0, 1, \ldots, n\}$; di conseguenza $\mu_n(A_n^c) = 0$ e

$$\delta(\pi, \mu_n) \geq \pi(A_n^c) - \mu_n(A_n^c) = \pi(A_n^c) = t_n$$

È tutto. $\diamondsuit$

Prima di abbandonare l'argomento, mi corre l'obbligo di confessare che sono stato alquanto reticente sulle catene della forma (8.5); più in generale, quando le probabilità di transizione sono minorate da un multiplo positivo di una legge fissa[†]

$$\mathrm{P}(x, y) \geq c\,\lambda(y) \qquad (c > 0)$$

si può dimostrare che esiste unica la legge invariante (in generale non sarà uguale a λ), e inoltre si ha convergenza con distanze maggiorate da una progressione geometrica di ragione $1 - c$.

Rinvio il lettore al *Libro*, pagg. 27–28.

(8.9) ESERCIZIO. Nello spazio degli interi relativi, definiamo le probabilità di transizione

$$\mathrm{P}(x, x+1) = \mathrm{P}(x, x-1) = 1/2, \quad \mathrm{Q}(x, 0) = 1, \quad \overline{\mathrm{P}}(x, y) = \frac{4}{5}\mathrm{P}(x, y) + \frac{1}{5}\mathrm{Q}(x, y)$$

Verificare che le leggi di una catena retta dalle probabilità di transizione $\overline{\mathrm{P}}$ tendono geometricamente verso la legge invariante

$$\pi(x) = \frac{1}{3}\, 2^{-|x|}$$

[†] tecnicamente, è la *condizione di Doeblin*

Si noti che l'esempio non rientra tra quelli del tipo (8.5), dal momento che le P non ammettono alcuna legge invariante di probabilità.

9. oblio

Fermiamo l'attenzione su una conseguenza immediata della tendenza alla stazionarietà.

Se le leggi di una catena con date probabilità di transizione tendono alla legge invariante π, è manifesto che la distanza tra leggi di due catene che abbiano le stesse probabilità di transizione (ma condizioni iniziali diverse) tenderanno ad avvicinarsi.

Lo si evince dalla disuguaglianza ("triangolare")

$$\delta(\mu_n, \nu_n) \leq \delta(\mu_n, \pi) + \delta(\nu_n, \pi)$$

Intuitivamente, con il tempo le catene smarriscono il ricordo delle condizioni iniziali (con locuzione ad effetto, parlerò di *oblio*).

Può però accadere che l'effetto delle condizioni iniziali tenda a svanire senza tuttavia che le catene tendano a regime, cioè senza che si abbia convergenza a una legge limite.

Vediamo un esempio. Modifichiamo la passeggiata semplice sugli interi positivi ponendo

$$\mathrm{P}(x, x+1) = \mathrm{P}(x, x-1) = 1/4, \qquad \mathrm{P}(x, x) = 1/2 \tag{9.1}$$

e prendiamo due catene indipendenti X_n e Y_n, entrambe con probabilità di transizione date da (9.1) e stati iniziali diversi; si verifica senza difficoltà che $X_n - Y_n$ è una catena sugli interi relativi, omogenea con legge di spostamento data dalla (7.1), e connessa.

Nel capitolo 7 s'è visto che tale catena è ricorrente, pertanto visita qc l'origine, a un tempo aleatorio T; il ragionamento di pag. 9 del *Corso*, con piccole modifiche, permette di ottenere la maggiorazione

$$P(X_n \in A) - P(Y_n \in A) \leq P(T > n)$$

da cui, dette μ_n la legge di X_n e ν_n quella di Y_n, si deduce la conclusione voluta $\delta(\mu_n, \nu_n) \to 0$.

Ha dunque luogo l'oblio. Ma non può certo esservi convergenza in legge, poiché le probabilità di transizione (9.1), come tosto si constata, non ammettono leggi invarianti di probabilità.

Non si tratta di una coincidenza fortuita, giacché si dimostra (*Libro*, Teorema 8.20 pag. 115) si ha sempre oblio nel caso ultraconnesso ricorrente (teorema analogo a quello di tendenza alla stazionarietà, quando si sopprima l'ipotesi di ricorrenza veloce).

Tutto chiaro dunque ? Niente affatto, ché ora vedremo manifestarsi oblio in un caso *transitorio*.

Modifichiamo questa volta l'esempio della passeggiata asimmetrica, ponendo

$$\mathrm{P}(x, x+1) = 1/2, \qquad \mathrm{P}(x, x) = \mathrm{P}(x, x-1) = 1/4 \tag{9.2}$$

La natura transitoria è palese: la probabilità $u(x)$ di visitare l'origine partendo da x vale 1 per $x < 0$ e 2^{-x} per $x \geq 0$[†].

Come nell'esempio (9.1), introduciamo le catene X_n e Y_n; non serve sviluppare i dettagli: la catena $X_n - Y_n$ è connessa omogenea con legge di spostamento a supporto finito e *simmetrica*, dunque centrata.

La discussione del capitolo 7 basta e avanza per concludere che è ricorrente, e assodare come sopra che sussiste l'oblio[‡].

Tanto pervasivo appare l'oblio che vien fatto di cercare esempi nei quali non sussista, nei quali il marchio d'origine delle catene si riveli persistente (come l'accento dei sardi).

Scarteremo esempi banali: nella passeggiata semplice di pag. 3 del *Corso*, una catena che parte dall'origine e una che parte dal punto 1 conservano traccia permanente della loro diversità d'origine (quando l'una sta in uno stato pari l'altra sta in un dispari e viceversa).

È il problema discusso all'inizio del capitolo 5 del *Corso*, nell'esempio (5.1) pagg. 7 e 8. Lo eluderemo cercando catene ultraconnesse.

Un esempio facile facile si costruisce "simmetrizzando" l'esempio (9.2); poniamo

$$\mathrm{P}(x, x+1) = 1/2, \quad \mathrm{P}(x, x) = \mathrm{P}(x, x-1) = 1/4 \quad \text{per } \ x > 0$$

$$\mathrm{P}(x, y) = \mathrm{P}(-x, -y), \qquad \mathrm{P}(0,0) = \mathrm{P}(0,1) = \mathrm{P}(0,-1) = 1/3 \tag{9.3}$$

Si tratta di un "doppio saliscendi" (due saliscendi saldati nell'origine), chiaramente ultraconnesso, e altrettanto chiaramente transitorio come l'esempio (9.2) da cui deriva.

La probabilità di visitare l'origine partendo da x è data da

$$u_0(x) = 2^{-|x|} \tag{9.4}$$

† se siete pigri, verificate che la funzione descritta è armonica fuori dall'origine, vale 1 nell'origine, e usate la Proposizione (2.5) a pag. 4 del *Corso*

‡ un risultato più generale si ritrova nel *Libro*, a pag. 170

Una catena con le probabilità di transizione (9.3) non sa scavalcare nessuno stato; per passare da una semiretta all'altra deve visitare l'origine.

Poiché il numero di visite nell'origine è qc finito per la transitorietà, la catena deve prima o poi optare per una delle due semirette e rimanervi per sempre.

Per evidenti ragioni di simmetria, la probabilità di optare per la semiretta positiva partendo dall'origine sarà $1/2$ (e uguale per quella negativa); optare per una semiretta oppure farvi infinite visite è (in questo esempio) la stessa cosa.

Dunque la funzione $v_+(x)$, probabilità di infinite visite nella semiretta positiva, vale $1/2$ nell'origine.

Abbiamo un esempio del caso (c) dell'alternativa di pag. 4: a semiretta è un insieme polidromo, e la funzione v_+ è "divaricata" (ha estremo superiore 1 e estremo inferiore nullo).

Ma torniamo alla nostra discussione.

Intanto determiniamo compiutamente la funzione v_+.

Partendo da $x < 0$, per fare infinite visite nella semiretta positiva occorre intanto visitare l'origine (probabilità $u_0(x)$) e, ripartendo da lì, (assenza di memoria) compiere le infinite visite (probabilità $1/2$).

Quindi $v_+(x) = u_0(x)/2$ per $x < 0$.

Da qualsiasi punto si parta, una delle due semirette deve ricevere infinite visite e l'altra no; vale dunque 1 la somma di $v_+(x)$ e della probabilità di infinite visite nella semiretta negativa partendo da x; quest'ultima però coincide per simmetria con $v_+(-x)$.

Se ne ricava l'identità $v_+(x) = 1 - v_+(-x)$, e si ha in definitiva

$$v_+(x) = u_0(x)/2 \quad \text{per} \quad x \leq 0, \qquad v_+(x) = 1 - u_0(x)/2 \quad \text{per} \quad x \geq 0$$

Consideriamo due catene: la catena X_n che parte dal punto 1 e la catena Y_n che parte da -1; ne siano μ_n e ν_n le leggi rispettive.

La probabilità che X_n finisca per trovarsi nella semiretta positiva S è $v_+(1) = 3/4$; a tale valore tenderà dunque la successione $\mu_n(S)$; analogamente si vede che $\nu_n(S)$ tende a $1/4$.

Ciò comporta che da un certo indice n in poi dovrà essere

$$\mu_n(S) - \nu_n(S) \geq 1/3$$

e a fortiori

$$\delta(\mu_n, \nu_n) \geq 1/3 \tag{9.5}$$

(tenendo conto della (8.6), la minorazione (9.5) vale in realtà per ogni n).

Resta acclarato che non si ha oblio: le catene mentengono traccia permanente della loro origine. $\diamondsuit$

10. monodromia

Discuterò in questo capitolo la situazione che si determina quando le sole funzioni armoniche limitate sono le costanti.

Una prima immediata conseguenza di tale ipotesi riguarda la probabilità di compiere infinite visite in un insieme assegnato di stati; come funzione dello stato di partenza, questa è infatti armonica (e chiaramente limitata); nella casistica enumerata all'inizio di pag. 4, resta pertanto escluso il terzo caso.

Volendo usare la terminologia che adottammo allora, non esistono insiemi polidromi; parlerò pertanto di *monodromia*.

Ci sono due aspetti, legati ma distinti, nella monodromia.

Il primo è che la probabilità di infinite visite nell'insieme assegnato è costante, non dipende dallo stato di partenza della catena; il secondo che questa costante è ristretta ai valori estremi 0 e 1: se dunque l'evento "la catena visita l'insieme infinite volte" non è quasi certo, allora è quasi certo il complementare. Tornerò presto su questo punto.

Veniamo a un problema leggermente più complicato, quello delle *soste*.

Conveniamo che la catena X fa sosta (all'istante n) se X_n e X_{n+1} coincidono; che dire della probabilità di compiere infinite soste, nella nostra ipotesi di "costanza armonica" (le funzioni armoniche limitate siano costanti) ?

Il problema si riconduce a uno di infinite visite mediante un artificio abile quanto naturale.

La successione $Z_n = (X_n, X_{n+1})$ è una catena di Markov i cui stati sono coppie ordinate di stati della X, e le probabilità di transizione sono date da

$$\overline{\mathrm{P}}\big((x,y),(x',y')\big) = \mathrm{P}(y,y')\,\mathrm{I}(x',y)$$

dove $\mathrm{I}(x',y)$ vale 1 se $x' = y$ e zero altrimenti (se l'evidenza non basta, una verifica pedissequa fugherà i dubbi del lettore pignolo).

Sia $g(x,y)$ una funzione limitata, armonica per le probabilità $\overline{\mathrm{P}}$; vale dunque la relazione

$$g(x,y) = \sum_{x',y'} \overline{\mathrm{P}}\big((x,y),(x',y')\big)\, g(x',y') =$$

$$= \sum_{x',y'} \mathrm{P}(y,y')\,\mathrm{I}(x',y)\, g(x',y') = \sum_{y'} \mathrm{P}(y,y')\, g(y,y')$$

la quale rivela che g non dipende dal primo argomento, e mostra che (come funzione del secondo argomento) è armonica per le probabilità P, e pertanto costante.

Sussiste per $\overline{\mathrm{P}}$ l'ipotesi di costanza armonica e ne consegue in particolare che la "diagonale" $\{x = y\}$ non è un insieme polidromo. La probabilità che la catena X faccia infinite soste (cioè che Z faccia infinite visite nella diagonale) è costante e vale 0 oppure 1. $\diamondsuit$

Non sempre però si può fare affidamento su trucchi del genere, e viene spontaneo interrogarsi su eventi strutturalmente analoghi ai precedenti ma di tipo più generale.

Comune ai casi trattati è la natura "stazionaria" degli eventi, la loro "indifferenza allo slittamento" (infinite visite della catena $(X_0, X_1, X_2, \ldots)$ equivale a infinite visite della catena $(X_1, X_2, X_3, \ldots)$, e similmente per le soste).

Nell'ipotesi di fondo del capitolo (costanza delle funzioni armoniche limitate), la probabilità di un evento "stazionario" non dipende dallo stato di partenza della catena ed è ristretta ai valori 0 e 1.

Vediamo di darne una dimostrazione. Volendo essere più formali, fissiamo prima un sottoinsieme K dell'insieme di tutte le successioni di stati, la cui funzione indicatrice sia invariante per slittamento, ossia valga l'identità

$$I_K(x_0, x_1, x_2, \ldots) = I_K(x_1, x_2, x_3, \ldots)$$

da cui per iterazione

$$I_K(x_0, x_1, x_2, \ldots) = I_K(x_n, x_{n+1}, x_{n+2}, \ldots)$$

Data la catena di Markov X, l'evento stazionario corrispondente sarà poi

$$\{(X_0, X_1, \ldots) \in K\} = \{(X_n, X_{n+1}, \ldots) \in K\}$$

Lo scriverò succintamente $\{X \in K\}$, e chiamerò $h_K(x)$ la sua probabilità per una catena che parte da x.

L'assenza di memoria (*Corso*, pag. 2) per un evento A anteriore al tempo n garantisce che, condizionalmente all'evento $\{A, X_n = y\}$, la successione $(X_n, X_{n+1}, \ldots)$ ha la legge di una catena che parte da y.

Pertanto la probabilità dell'evento $\{X \in K\} = \{(X_n, X_{n+1}, \ldots) \in K\}$ condizionata all'evento $\{A, X_n = y\}$ vale $h_K(y)$. In simboli,

$$P\big(A, X_n = y, \{X \in K\}\big) = P(A, X_n = y)\, h_K(y) \tag{10.1}$$

La (10.1) è la relazione chiave. Scritta per X che parte da x, per $n = 1$ e prendendo per A l'evento totale, essa fornisce

$$P\big(X_1 = y, \{X \in K\}\big) = P(X_1 = y)\, h_K(y) = \mathrm{P}(x, y)\, h_K(y)$$

donde, sommando per tutti i valori di y, si ricava

$$h_K(x) = P\big(\{X \in K\}\big) = \sum_y \mathrm{P}(x,y)\, h_K(y)$$

cioè si apprende che la funzione (limitata) h_K è armonica, dunque costante, per le ipotesi in vigore in questo capitolo.

Tenendone conto nella (10.1) e sommando ancora per tutti gli y, si trova

$$P\big(A, \{X \in K\}\big) = P(A)\, h_K = P(A)\, P\big(\{X \in K\}\big)$$

Se l'evento $\{X \in K\}$ ha probabilità nulla, niente da dire; altrimenti, la relazione ottenuta garantisce che condizionare rispetto a quest'evento non altera la probabilità di A.

Ma A è un qualsiasi evento anteriore a un arbitrario istante n; dunque la probabilità condizionata $Q(\cdot) = P(\cdot \,|\{X \in K\})$ coincide con P su tutti gli eventi della forma

$$\{X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\}$$

Tanto basta† a concludere che sostituire P con Q non altera la legge di X, motivo per cui si deve avere $P\big(\{X \in K\}^c\big) = Q\big(\{X \in K\}^c\big) = 0$.

Se l'evento $\{X \in K\}$ non ha probabilità nulla, ha probabilità nulla il suo complementare. Come si voleva. $\diamondsuit$

Ci sono un paio di ipotesi facili che assicurano la costanza delle funzioni armoniche limitate.

(10.2) PROPOSIZIONE. *In una catena connessa ricorrente, le funzioni armoniche limitate sono costanti.*

DIM. Pur di sommarvi una conveniente costante, non è restrittivo limitarsi al caso di una funzione f armonica limitata e positiva.

Supponiamo per assurdo $f(a) < f(b)$, e notiamo che per ipotesi una catena X che parte da a visita quasi certamente lo stato b; sia T l'istante della sua prima visita in b.

Operiamo una chirurgia simile a quella di pagina 13, modificando la probabilità di transizione dallo stato b in modo che sia $\mathrm{P}(b,b) = 1$; verrà meno la connessità, ma f rimane armonica e niente cambia per quanto riguarda il tempo T.

† vi è una certa qual leggerezza nella conclusione; la stessa leggerezza, a pag. 2 del *Corso*, fu peraltro accolta con clemenza dalla critica

Dal lemma (7.2) segue

$$E\big(f(X_0)\big) = E\big(f(X_n)\big) \geq \int_{\{T\leq n\}} f(X_n) = \int_{\{T\leq n\}} f(X_T) = f(b)\,P(T\leq n)$$

Facendo tendere n all'infinito, l'ultimo termine tende a $f(b)$; ne segue l'assurda conclusione $f(a) = E\big(f(X_0)\big) \geq f(b) > f(a)$. ♢

(10.3) PROPOSIZIONE. *Se vale la proprietà di oblio, allora le funzioni armoniche limitate sono costanti.*

DIM. Sia f armonica limitata, che possiamo supporre senz'altro positiva, e sia M un numero reale che la maggiora.

La catena X parta da x e la Y da y; siano μ_n la legge di X_n e ν_n quella di Y_n.

Usando il lemma (7.2) e rammentando le prime righe di pag. 16, si ottiene

$$f(x) - f(y) = E\big(f(X_0)\big) - E\big(f(Y_0)\big) = E\big(f(X_n)\big) - E\big(f(Y_n)\big) =$$

$$= \sum_z f(z)\big(\mu_n(z) - \nu_n(z)\big) \leq M \sum_z \big(\mu_n(z) - \nu_n(z)\big)^+ \leq M\,\delta(\mu_n, \nu_n)$$

L'ultimo termine tende a zero per ipotesi, e d'altra parte è lecito scambiare x con y. La funzione f è pertanto costante. ♢

S'è visto nel corso del capitolo che la costanza delle funzioni armoniche limitate implica la degenerazione degli eventi stazionari, e dunque in particolare la monodromia. A volerla dir tutta, questa a sua volta comporta la costanza delle funzioni armoniche limitate, e il cerchio così si chiude.

Ne darò un cenno di dimostrazione†, che presuppone un teorema di convergenza di martingale.

Sia f armonica limitata, e supponiamo per assurdo $f(a) < f(b)$.

Scelti due numeri reali u e v con $f(a) < u < v < f(b)$, gli insiemi $\{f < u\}$ e $\{f > v\}$ sono non vuoti.

Per una catena X, la successione $f(X_n)$ è una *martingala* limitata, e (come sanno i dotti) converge quasi certamente; il che esclude infinite oscillazioni.

Per l'ipotesi di monodromia, la probabilità che la catena visiti un insieme infinite volte è necessariamente 0 oppure 1, e non dipende dallo stato di partenza.

Uno dei due insiemi riceve dunque un numero (qc) finito di visite. Per fissare le idee, supponiamo così sia per il primo, per cui sarà $f(X_n) \geq u$ quasi certamente a partire da un certo istante.

† non c'è nemmeno nel *Libro*

Se la catena parte da a, e dunque $f(X_0) = f(a) < u$, il lemma (7.2) porta facilmente a contraddizione. $\diamondsuit$

11. esodo

Il tema tocca un aspetto della transitorietà per noi nuovo, che meriterebbe una trattazione più vasta e dettagliata.

Un insieme è *transitorio* se rientra nel caso a) della casistica di pag. 4; una catena lo visita un numero quasi certamente finito di volte, nel frattempo magari ne visita il complementare, ma giunge inesorabile il momento in cui abbandona l'insieme in modo definitivo per entrare nel complementare e restarvi per sempre. È l'*esodo*.

Viene naturale studiare il tempo di esodo, raffrontato con quello totale di permanenza nell'insieme e con quello di primo ingresso nel complementare, come pure la distribuzione di probabilità dello stato dell'insieme nel quale si compie l'ultima visita, o dell'ingresso definitivo nel complementare.

Su quest'ampio panorama offrirò uno scorcio angusto, basato su un esempio fin troppo banale.

Si tratta della passeggiata asimmetrica sugli interi relativi, con probabilità di transizione

$$\mathrm{P}(x, x+1) = 2/3, \qquad \mathrm{P}(x, x-1) = 1/3 \tag{11.1}$$

Poiché la catena tende quasi certamente a $+\infty$, la semiretta negativa[†] è un insieme transitorio.

Supponiamo che lo stato di partenza sia l'origine, e ricordiamo che la catena non sa "scavalcare" gli stati (per andare da uno stato a un altro deve visitare tutti quelli intermedi).

Perciò il luogo dell'ultima visita nella semiretta negativa è per forza l'origine e lo stato dove avviene l'ingresso (definitivo o meno) nel complementare dev'essere il punto $+1$: in questo senso l'esempio scelto è un po' sempliciotto.

Tuttavia non mancano, come vedremo, aspetti interessanti.

Più che rifarmi a una teoria generale, userò ragionamenti intuitivi, anche a prezzo di qualche funambolismo.

Iniziamo con il tempo T del primo ingresso della catena X nello stato $+1$ e indichiamo con $t(0, +1)$ il valore atteso di T, con un simbolismo già impiegato in passato.

[†] qui intendo lo zero incluso

Guardiamo la prima transizione della catena a partire da 0; con probabilità 2/3 si sposta in $+1$, ed è fatta; con probabilità 1/3 si sposta in -1, da dove il tempo medio d'ingresso in $+1$ è $t(-1,+1)$.

Esso vale $t(-1,0)+t(0,+1)$, dato che la catena non sa fare salti; inoltre, per omogeneità, si ha $t(-1,0)=t(0,1)$, da cui $t(-1,+1)=2\,t(0,+1)$.

L'assenza di memoria al tempo 1 conduce all'equazione

$$t(0,1)=\frac{2}{3}\,1+\frac{1}{3}\big(1+2\,t(0,1)\big) \tag{11.2}$$

da cui è elementare trarre $E(T)=t(0,1)=3$.

Detto L l'istante (aleatorio) dell'ultima visita della catena nella semiretta negativa S, quello dell'ingresso definitivo nel complementare sarà $D=L+1$.

Facile da calcolare è la probabilità di esodo *immediato*; si ha quando la prima transizione porta (con probabilità 2/3) nello stato $+1$ e la catena non visita mai più S (cioè 0), il che accade con probabilità 1/2, come ormai abbiamo ripetuto alla noia.

Per la probabilità di esodo immediato si ottiene perciò $P(L=0)=1/3$.

Non meno facile il caso di esodo *brusco*. Parlo di esodo brusco quando il primo ingresso nel complementare di S è anche quello definitivo (quello immediato è ovviamente un caso particolare di esodo brusco).

Si tratta di guardare la catena all'istante T del primo ingresso in S^c, sfruttare la forma forte di assenza di memoria e imporre che, ripartendo ex novo da X_T $(=+1)$, la catena non visiti più S (cioè 0).

In definitiva: $P(\text{esodo brusco})=P(D=T)=1/2$.

Un ragionamento simile giova per il valor atteso del numero V di visite compiute nella semiretta S. Sino al tempo T la catena (contando l'istante iniziale) ha visitato S per un tempo T ininterrotto e si trova in $X_T=+1$; condizionalmente al visitare ancora S (probabilità 1/2), la catena si trova di nuovo nell'origine e la storia riparte daccapo.

Facendo uso della solita forma forte di assenza di memoria al tempo T, ne discende l'equazione

$$E(V)=E(T)+(1/2)\,E(V) \tag{11.3}$$

da cui segue agevolmente $E(V)=2\,E(T)=6$.

Il calcolo del valor atteso di D è un po' meno semplice.

Ancora si ragiona sul tempo T di primo ingresso, sulla falsariga delle argomentazioni che hanno condotto alla (11.3). Ma ora si deve tener conto del tempo medio che occorre per tornare eventualmente nell'origine *nell'ipotesi che la catena, ripartendo da $X_T=+1$, visiti S*[†].

[†] senza questa condizione in corsivo, il tempo per andare da $+1$ a 0 non è nemmeno qc finito, tantomeno ha valor medio finito

Va qui aperta una breve parentesi.

Sia $u(x)$ la probabilità di visitare l'origine partendo da x.

Partendo da $x > 0$, la probabilità di passare in y e poi visitare 0 è chiaramente $\mathrm{P}(x,y)\,u(y)$; le probabilità di transizione condizionate saranno allora

$$\mathrm{P}^*(x,y) = \mathrm{P}(x,y)\,u(y)/u(x)$$

Nel nostro caso $u(x) = 2^{-x}$ e si ottiene

$$\mathrm{P}^*(x,x+1) = \frac{2}{3}\frac{1}{2} = \frac{1}{3}, \qquad \mathrm{P}^*(x,x-1) = \frac{1}{3}\,2 = \frac{2}{3} \tag{11.4}$$

Sorpresa, è la copia speculare della (11.1) da cui siamo partiti.

Cercare il tempo medio per passare da $+1$ a 0 secondo queste probabilità (11.4) è la stessa cosa che calcolare il tempo medio di passaggio da 0 a $+1$ secondo le (11.1), cioè calcolare $E(T)$ (sappiamo che si ottiene 3).

Chiusa e archiviata la parentesi, giungiamo finalmente all'equazione

$$E(D) = E(T) + \frac{1}{2}\big(E(T) + E(D)\big) = 3 + \frac{1}{2}\big(3 + E(D)\big) \tag{11.5}$$

da cui la conclusione $E(D) = 9$.

Come si vede e ci s'aspettava, per l'ingresso definitivo nel complementare occorre (in media) assai più tempo che per il primo ingresso.

La tecnica ormai familiare suggerisce al lettore come calcolare il valor medio del numero di visite V_0 che la catena compie nell'origine: la probabilità di tornare a visitare l'origine dopo l'istante iniziale è $1/3 + (2/3)(1/2) = 2/3$.

Di qui l'equazione $E(V_0) = 1 + (2/3)E(V_0)$, e quindi $E(V_0) = 3$.

Le curiosità, si sa, non si esauriscono mai. Quando l'esodo non sia immediato, all'ultima visita nell'origine si giunge preferibilmente dall'insieme S oppure da fuori ?

Detto meglio: quanto vale la probabilità $P(L > 0, X_{L-1} = +1)$?

Anche questo è un quesito a suo modo impegnativo, ma non difficile.

Esaminando partitamente i possibili valori di L possiamo scrivere

$$P\big(L = k+1, X_{L-1} = +1\big) =$$

$$= P\big(X_k = +1, X_{k+1} = 0, (X_{k+1}, X_{k+2}, \ldots) \text{ non torna a visitare } 0\big) =$$

$$= P(X_k = +1)\,\mathrm{P}(+1,0)\,\frac{1}{3} = \frac{1}{9}P(X_k = +1) = \frac{1}{9}\,E\Big(I_{\{+1\}}(X_k)\Big)$$

Sommando per tutti i valori di k si ottiene

$$P(L > 0, X_{L-1} = +1) = \frac{1}{9}\,E\Big(\sum_k I_{\{+1\}}(X_k)\Big) = \frac{1}{9}\,E(V_1)$$

avendo riconosciuto, nella somma entro parentesi, il numero V_1 di visite che la catena compie nello stato $+1$.

Il valor medio di V_1 si determina al solito modo: al primo ingresso della catena in $+1$, per assenza di memoria e per l'omogeneità, dal nuovo stato di partenza $+1$ la catena compie in $+1$ un numero medio di visite uguale a quello che la catena originaria, partita dall'origine, compie in 0.

In altre parole, $E(V_1) = E(V_0) = 3$ e $P(L > 0, X_{L-1} = +1) = 1/3$.

Poiché $P(L > 0) = 1 - P(L = 0) = 2/3$, e d'altra parte i soli valori possibili di X_{L-1} sono $+1$ e -1, se ne deduce per differenza

$$P(L > 0, X_{L-1} = +1) = P(L > 0, X_{L-1} = -1) = 1/3$$

Sorprendentemente, alla visita di commiato in S si giunge da S oppure dal complementare con ugual probabilità.

Torniamo al tempo d'ingresso definitivo D in S^c per esaminare il seguito della vicenda.

Al pari della semiretta S, anche la semiretta $S_1 = S \cup \{1\}$ è un insieme transitorio; la catena ne esce prima o poi per entrare definitivamente nel complementare di S_1, ovviamente nel suo punto $+2$.

Bisognerà attendere un ulteriore tempo (aleatorio) D_1; quale ne sarà il valor medio ?

All'istante D, la catena si trova in $+1$ e fronteggia la stessa situazione iniziale, semplicemente slittata di un posto. In ragione dell'omogeneità, riesce spontaneo pensare (qualche lettore zelante lo avrà fatto) che D_1 abbia perciò la stessa legge e dunque lo stesso valor medio di D. Giusto ?

No, in un tale ragionamento si cela un sottile errore concettuale.

Ripartendo da $+1$ al tempo D, la catena non si muove con la stessa legge "come nulla fosse accaduto" (*Corso*, pag. 21), la sua dinamica è alterata per un condizionamento che proviene dal futuro: per costruzione, non può visitare mai più la semiretta S.

Qui non vige affatto l'assenza di memoria; del resto, le condizioni di pag. 21 del *Corso* non sussistono per D, che non è un *tempo d'arresto*.

La successione $(X_D, X_{D+1}, X_{D+2}, \ldots)$ obbedisce a probabilità di transizione $\overline{\mathrm{P}}$ condizionate a non visitare S: prova ne sia che $\overline{\mathrm{P}}(1,0)$ vale 0 invece che $1/3$.

Per $x > 0$ queste probabilità sono date da

$$\overline{\mathrm{P}}(x,y) = \frac{\mathrm{P}(x,y)\,(1-u(y))}{(1-u(x))} = \mathrm{P}(x,y)\,\frac{1-2^{-y}}{1-2^{-x}}$$

Se poniamo $p_x = \overline{\mathrm{P}}(x, x+1)$, esse definiscono un saliscendi con $r_x = 0$ e

$$p_1 = 7/9, \quad p_2 = 5/7, \quad p_3 = 31/45, \quad p_4 = 21/31, \quad \ldots \tag{11.6}$$

valori tutti maggiori del valore standard 2/3, al quale si avvicinano sempre più quanto più ci si allontana dall'origine e il condizionamento si fa sentire sempre meno. Ma l'omogeneità è perduta.

A partire dal tempo D la catena, obbligata a non tornare in S, "va più in fretta"; se pazientemente eseguiamo i calcoli per il saliscendi (11.6), troviamo che D_1 ha valor medio 3, decisamente minore di quello di D.

Ma perché penare tanto quando si può ragionare sul tempo totale $D+D_1$ e operare per differenza.

All'istante T la catena si trova in $+1$, dovrà passare un tempo D' per il suo ingresso definitivo in $+2$, che avverrà dunque all'istante $T+D' = D+D_1$.

Ma ora T è un tempo d'arresto e vige l'assenza di memoria; per omogeneità D' ha lo stesso valor atteso di D.

Si ritrova senza sforzo $E(D_1) = E(T) = 3$.

Un esercizio lasciato al lettore: riprendere le probabilità (11.1) assumendo un generico $p > 1/2$ al posto di $2/3$, e ponendo $q = 1-p$. Si ottiene

$$E(T) = \frac{1}{p-q}, \quad P(L=0) = p-q = 1-2q, \quad P(\text{esodo brusco}) = 1 - q/p$$

$$P(L>0, X_{L-1} = +1) = P(L>0, X_{L-1} = -1) = q, \quad E(V) = \frac{p}{(p-q)^2}$$

$$E(D) = \frac{1}{(p-q)^2}$$

12. appendice, la stella

L'esempio annunciato a pag. 4 consta di infiniti saliscendi $\Sigma_1, \Sigma_2, \ldots$ saldati in una comune origine (una variante infinita dell'esempio 9.3 di pag. 19).

Per nominare il punto che dista d dall'origine e sta nel saliscendi r uso la notazione "polare" (d, r), con l'abbreviazione $e_r = (1, r)$; le coppie $(0, r)$ denotano tutte l'origine o.

Indico semplicemente con x il punto $(x, 1)$ di Σ_1.

In Σ_1 le probabilità di transizione sono quelle solite della passeggiata asimmetrica

$$\mathrm{P}(x, x+1) = 2/3 \qquad \mathrm{P}(x, x-1) = 1/3 \quad (x > 0)$$

eccettuata l'origine, per la quale pongo $\mathrm{P}(o, e_r) = 2^{-r}$.

Come sappiamo, la probabilità di visitare l'origine partendo dallo stato $x \in \Sigma_1$ è 2^{-x}.

Entro il saliscendi r le probabilità di transizione sono scelte in modo che la probabilità di visitare l'origine partendo dal punto (d,r) sia 2^{-dr} (prop. 6.2 docet).

Abbiamo una catena connessa transitoria; la probabilità di infinite visite nell'origine è nulla.

Cerchiamo la probabilità v di visitare infinite volte Σ_1 in funzione dello stato di partenza.

Partendo entro Σ_r con $r \neq 1$, per visitare Σ_1 infinite volte occorre prima di tutto visitarne l'origine; da cui la relazione $v(d,r) = 2^{-dr}\, v(o)$.

Se si parte da Σ_1, va aggiunta la probabilità di muoversi sempre in Σ_1, cioè di *non* toccare mai l'origine: $v(d,1) = 2^{-d}\, v(o) + (1 - 2^{-d})$.

L'armonicità di v nell'origine fornisce l'equazione

$$v(o) = \sum_{j=1}^{\infty} 2^{-j}\, 2^{-j} v(o) + 2^{-1} 2^{-1} = \frac{1}{3} v(o) + \frac{1}{4}$$

da cui si ricava $v(o) = 3/8$ e finalmente

$$v(d,r) = \frac{3}{8}\, 2^{-dr} \quad (r > 1) \qquad\qquad v(d,1) = 1 - \frac{5}{8}\, 2^{-d} \tag{12.1}$$

Sia F l'insieme dei punti di Σ_1 la cui distanza dall'origine è un numero pari; non è possibile visitare Σ_1 infinite volte senza visitare infinite volte F; la probabilità di infinite visite in F resta dunque quella data dalla (12.1).

Nulla cambia se aggiungiamo a F l'insieme E dei punti della forma e_r; la probabilità di infinite visite nell'insieme $H = F \cup E$ è ancora la stessa; non si possono infatti fare infinite visite in E senza passare infinite volte dall'origine (il che ha probabilità nulla per la transitorietà).

Dalla (12.1), che dà dunque la probabilità $v = v_H$ di visitare H infinite volte, si legge la conclusione

$$\sup_{H^c} v_H \geq \lim_n v(2n+1, 1) = 1 \qquad\qquad \inf_H v_H \leq \lim_r v(e_r) = 0$$

Una curiosità: è possibile modificare leggermente la catena nei saliscendi con $r > 1$ in modo che la probabilità φ di visitare l'origine risulti

$$\varphi(1,r) = 2^{-r}, \quad \varphi(2,r) = 2^{-r}\big(1 - 2^{-r}\big), \quad \varphi(n+2,r) = \varphi(2,r)\, 2^{-nr}$$

In tal modo resta valido il calcolo che ha condotto alla (12.1); inoltre (vedi 6.1) la probabilità u_H di visitare H partendo dal punto $y_r = (2,r) \notin H$ diventa $\big(1 - 2^{-r}\big)$, mentre $v_H(2,r) = (3/8) 2^{-2r}$.

La differenza $(u_H - v_H)$ calcolata nei punti $y_r \notin H$ tende dunque verso 1, un'ulteriore stranezza. $\diamondsuit$

durante la stesura di queste note si è spenta la mia amatissima moglie
consumata lentamente anzitempo da un male crudele
ne serbi caro il ricordo chi la conobbe