

La regola proposta nell'equazione (19.19) è un buon compromesso, ma non deve essere interpretata troppo rigidamente. Valori di $P_{Acc} \in [0.3, 0.7]$ sono perfettamente accettabili nella maggior parte delle applicazioni pratiche. Nel Paragrafo 19.3 discutiamo una prima applicazione dell'algoritmo di Metropolis.

Laboratorio 19.3 *Convergenza dell'algoritmo di Metropolis*



Considerate una catena di Markov i cui stati sono tutti i numeri interi (in teoria questi formano un insieme infinito, ma svolgendo l'esercizio scoprirete che ciò non è un problema) e la cui distribuzione asintotica è proporzionale a una distribuzione gaussiana di media nulla: $\pi_j \propto \exp(-(j/\sigma)^2/2)$. Partite con il camminatore aleatorio nell'origine e fatelo muovere con un algoritmo di Metropolis in cui nella fase propositiva scegliete un nuovo stato in modo uniforme tra quelli entro una distanza R dallo stato attualmente occupato. Studiate la convergenza al valore asintotico di una quantità significativa, come la distanza quadratica media percorsa dal camminatore, $\langle r^2 \rangle - \langle r \rangle^2$, che dovrebbe convergere circa a σ^2 (piccoli scostamenti da questo valore sono dovuti alla discretizzazione dello spazio delle configurazioni). Verificate che il tempo di convergenza al valore asintotico vari con la scelta di R e che abbia un minimo per valori di R che hanno circa $P_{Acc} \simeq 0.5$. Usate un valore di σ che vi permetta di andare da un regime in cui $R \ll \sigma$ a un altro in cui $R \gg \sigma$; un possibile suggerimento è $\sigma = 100$, con R che varia logicamente tra 1 e 10^5 , cioè $R = 1, 10, 10^2, 10^3, 10^4, 10^5$.

19.2.5 Barriere e rottura dell'ergodicità

Concludiamo lo studio delle catene di Markov concentrandoci ancora una volta sul problema della convergenza all'equilibrio. Il problema che affrontiamo in questo paragrafo, quello della rottura dell'ergodicità, è spesso studiato in connessione con un modello specifico (soprattutto con quelli usati in meccanica statistica), ma in realtà il fenomeno è ben più generale ed è quindi utile studiarlo ancora al livello di generica catena di Markov.

Partiamo con un esempio concreto. Supponiamo di avere una catena di Markov definita sui numeri interi e di voler ottenere come distribuzione di probabilità asintotica la somma di due gaussiane di varianza unitaria, una centrata in $-L$ e l'altra in L . Per semplicità ammettiamo solo transizioni fra stati primi vicini, cioè tra numeri che differiscono di un'unità. Nella simulazione usiamo ovviamente l'algoritmo più efficiente tra quelli che conosciamo al momento, l'algoritmo di Metropolis.

Nei tre grafici in alto nella Figura 19.5 mostriamo la posizione del camminatore aleatorio, partito nell'origine al tempo iniziale, per i primi 10^4 passi e per tre valori diversi di L . Mentre per $L = 2$ il camminatore passa da un massimo della distribuzione all'altro con relativa facilità, per $L = 3$ il tempo tra un salto e l'altro è già dell'ordine del migliaio di passi (si vedono chiaramente 11 passaggi $L \leftrightarrow -L$) e infine, per $L = 4$, il camminatore rimane per tutti i 10^4 passi intorno al massimo della gaussiana con media

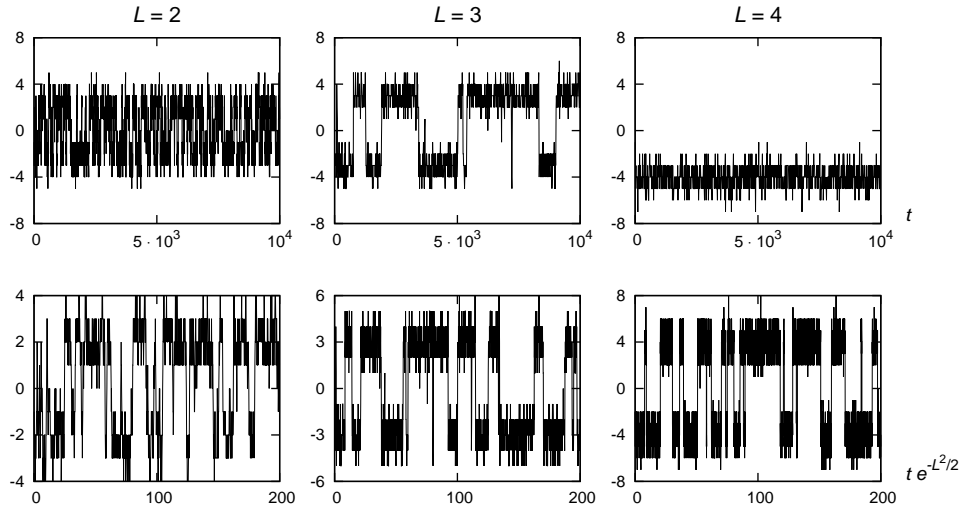


Figura 19.5 Posizione del camminatore aleatorio in una catena di Markov con probabilità asintotica concentrata in L e $-L$, per tre valori di L . Nei tre grafici in basso i tempi sono stati moltiplicati per un fattore $e^{-L^2/2}$.

negativa. In quest'ultimo caso è necessario fare una simulazione molto più lunga affinché il camminatore possa visitare entrambe le zone in cui la probabilità $\bar{\pi}$ è diversa da zero.

Il motivo alla base di questo comportamento è relativamente semplice: tutte le sequenze di spostamenti che portano il camminatore dalla zona circostante un massimo della $\bar{\pi}$ alla zona circostante l'altro massimo passano per una regione, quella intorno all'origine, in cui la probabilità asintotica è estremamente piccola. In questo caso particolare la nostra affermazione può sembrare ovvia e scontata (siamo in una dimensione e per andare da $-L$ a L o viceversa si deve passare per l'origine), ma il discorso che stiamo facendo si applica anche a casi più complicati, ad esempio a domini d'integrazione multidimensionali.

Nel semplice caso delle due distribuzioni gaussiane centrate in L e $-L$ la distribuzione asintotica è

$$\pi_j \propto e^{-(j+L)^2/2} + e^{-(j-L)^2/2},$$

e possiamo anche fare una stima approssimata del tempo necessario al camminatore per passare dagli stati negativi a quelli positivi. Il passaggio dai valori negativi a quelli positivi (e viceversa) avviene solo quando il camminatore si ritrova nell'origine e prosegue nella direzione opposta a quella di provenienza; il secondo di questi eventi avviene con probabilità $1/2$, mentre il primo evento avviene con probabilità $\pi_0 = \mathcal{O}(e^{-L^2/2})$ e quindi solo ogni $\tau = 1/\pi_0$ passi. Perciò è appropriato aspettarsi che le transizioni del camminatore da un massimo all'altro avvengano ogni $\mathcal{O}(e^{L^2/2})$ passi temporali.

Questo argomento è formalmente corretto solo nel limite $L \rightarrow \infty$ in cui la probabilità

π_0 tende a zero e il tempo di decorrelazione diverge, tuttavia anche per piccoli valori di L sembra dare un ottimo risultato. Infatti nei grafici in basso nella Figura 19.5, cambiando la scala dei tempi, riportati nelle ascisse, attraverso un fattore moltiplicativo $\exp(-L^2/2)$, abbiamo ottenuto tre curve che si somigliano molto (a parte le fluttuazioni maggiori nel caso $L = 2$, dove la separazione tra i massimi è un po' troppo piccola per evidenziare bene l'effetto).

Si usa spesso il termine *barriera* per indicare la regione in cui la $\bar{\pi}$ è molto piccola; il camminatore aleatorio deve tentare molte volte di scalarla prima di riuscire a superarla.

È importante osservare che se il numero di passi della nostra simulazione non è scelto accuratamente, potremmo non avere la certezza che tutte le barriere presenti nella catena di Markov che stiamo studiando siano state superate dal camminatore. Ad esempio, osservando il grafico in alto a destra nella Figura 19.5 potremmo essere convinti di aver raggiunto lo stato asintotico, ma in realtà il camminatore sta visitando solo metà degli stati rilevanti per la vera distribuzione asintotica. Talvolta l'unica maniera per convincerci che la catena di Markov è stata visitata completamente consiste nell'aumentare la durata della simulazione numerica. Nel Paragrafo 20.2.1 spieghiamo con maggiori dettagli come studiare la convergenza all'equilibrio di un metodo Monte Carlo basato su catene di Markov.

Laboratorio 19.4 *Tempi di attraversamento di una barriera*



Considerate la catena di Markov con distribuzione asintotica fatta di due distribuzioni gaussiane centrate in L e $-L$ e di una barriera intorno all'origine. Abbiamo appena visto che, per l'algoritmo di Metropolis, la transizione da un lato all'altro della barriera avviene ogni $\tau = \mathcal{O}(e^{L^2/2})$ passi, ossia è un processo dinamico molto raro per L grande. Inoltre, nel momento in cui tale transizione ha luogo, il processo di attraversamento della barriera sembra essere molto rapido, come nel grafico in basso a destra nella Figura 19.5.

Cercate di studiare numericamente il tempo di attraversamento di una barriera. A tal fine potete calcolare la lunghezza sia di quelle traiettorie del camminatore aleatorio che partono dalla posizione $-L$ e raggiungono la posizione L senza ripassare mai per la posizione $-L$, sia di quelle traiettorie che fanno il percorso speculare. Considerando che queste traiettorie devono avere una durata temporale di almeno $2L$ passi, dovrete concludere dalla vostra analisi che, quando il camminatore decide di attraversare la barriera, lo fa senza indugi, con una traiettoria quasi balistica!

Per un qualsiasi valore limitato di L la catena di Markov che abbiamo studiato in questo paragrafo è ergodica, ossia è possibile andare da un qualunque stato a qualsiasi altro stato (sebbene il tempo richiesto in alcuni casi sia molto lungo). Nel limite $L \rightarrow \infty$, invece, il tempo di decorrelazione τ diverge e la catena si decompone in due sottocatene, una formata dagli stati positivi e l'altra da quelli negativi: nel limite $L \rightarrow \infty$ avviene una *rottura dell'ergodicità*.

ESTRATTO DAL LIBRO **PROGRAMMAZIONE SCIENTIFICA** DI
L. M. BARONE, E. MARINARI, G. ORGANTINI E F. RICCI-TERSENGHI
I diritti di riproduzione e memorizzazione elettronica degli estratti appartengono a Pearson Education Italia S.r.l. e sono riservati per tutti i paesi. La stampa di essi è concessa gratuitamente solo a titolo personale. Ogni altro uso è vietato.

Quello illustrato qui è un sistema molto semplice tra quelli che presentano una rottura dell'ergodicità, tanto che non dovrebbe essere troppo difficile immaginare una diversa scelta delle probabilità u_{ij} , cioè una diversa connettività tra gli stati, che permetta di evitare il fenomeno della rottura dell'ergodicità. Ad esempio, aggiungendo alle transizioni già presenti nella catena quelle tra stati simmetrici rispetto all'origine, cioè $j \leftrightarrow -j$, tutti i fenomeni dinamici descritti in questo paragrafo (dal rallentamento dovuto alla barriera, alla rottura dell'ergodicità) svaniscono. La transizione $j \leftrightarrow -j$ rappresenta una sorta di scorciatoia che, aggirando la barriera, a tutti gli effetti la annulla. Notiamo che nell'algoritmo di Metropolis la transizione fra due stati simmetrici rispetto all'origine è sempre accettata, perché tali stati hanno la stessa probabilità asintotica, $\pi_j = \pi_{-j}$.

La ricerca di queste "scorciatoie" è uno degli aspetti più interessanti nello sviluppo di algoritmi Monte Carlo evoluti basati sulle catene di Markov. Purtroppo per sistemi più complicati di quello illustrato in questo paragrafo, le scorciatoie non sono affatto ovvie; ad esempio i cosiddetti algoritmi "a cluster" [40] sfruttano a fondo le conoscenze delle proprietà fisico-matematiche del modello di Ising, descritto nel Paragrafo 19.3.

19.3 Il modello di Ising

Il modello di Ising è caratterizzato da un insieme di variabili binarie disposte ai vertici di un reticolo regolare. Tali variabili sono dette *spin* perché in un'applicazione caratteristica del modello rappresentano momenti magnetici elementari. In questo paragrafo studiamo il modello definito su un reticolo quadrato bidimensionale, come quello della Figura 19.6. Gli spin su vertici primi vicini, cioè connessi da un arco, interagiscono tra di loro tramite un'interazione ferromagnetica che tende ad allinearli nella stessa direzione. L'energia di una configurazione $\mathbf{s} \equiv (s_1, \dots, s_N)$ degli N spin di cui è composto il sistema è data da

$$E(\mathbf{s}) = - \sum_{(i,j)} J s_i s_j . \quad (19.20)$$

Nell'espressione (19.20) le variabili di spin possono assumere solo due valori $s_i = \pm 1$, la somma è sulle coppie d'indici $i, j \in \{1, \dots, N\}$ corrispondenti a vertici adiacenti e la costante d'accoppiamento¹ J è positiva, perché l'interazione ferromagnetica favorisce l'allineamento degli spin. Il segno meno davanti alla somma è necessario perché in Natura le configurazioni preferite sono quelle che minimizzano l'energia: due spin allineati hanno energia $-J$, mentre due spin in direzioni opposte hanno energia $+J$. Non è difficile generalizzare questo modello ad altre topologie d'interazione: basta prendere un grafo qualsiasi, mettere uno spin su ogni vertice e far interagire con un accoppiamento di energia $-J s_i s_j$ ogni coppia di spin su vertici adiacenti, ossia legati da un arco.

Seguiamo i consigli del Paragrafo 10.4 sulle migliori unità di misura da usare in una simulazione. Essendo l'energia (19.20) un multiplo intero di J (che nelle unità di

¹Il simbolo J che usiamo per la costante d'accoppiamento non deve essere confuso con quello del Joule, l'unità di misura dell'energia (che tuttavia potrebbe essere proprio l'unità di misura dell'accoppiamento J visto che questo non è altro che una energia).