

Luca Leuzzi Enzo Marinari
Giorgio Parisi

Calcolo delle probabilità

Un trattatello per principianti volenterosi

Se vuoi accedere alle risorse online riservate

1. Vai su **my.zanichelli.it**
2. Clicca su *Registrati*.
3. Scegli *Studente*.
4. Segui i passaggi richiesti per la registrazione.
5. Riceverai un'email: clicca sul link per completare la registrazione.
6. Cerca il tuo codice di attivazione stampato in verticale sul bollino argentato in questa pagina.
7. Inseriscilo nella tua area personale su **my.zanichelli.it**

Se sei già registrato, per accedere ai contenuti riservati ti serve solo il codice di attivazione.

Indice generale

Prefazione	xiii
1 Introduzione alla probabilità	1
1.1 Definizione di probabilità	1
1.2 Proprietà di base delle probabilità	5
1.2.1 Probabilità di eventi caratterizzati da variabili discrete	5
1.2.2 Assiomi della Teoria della probabilità	6
1.2.3 Probabilità di eventi caratterizzati da variabili continue	7
1.3 Probabilità di intersezioni e unioni di insiemi	7
1.4 Probabilità condizionate	12
1.4.1 Probabilità congiunta di eventi indipendenti	14
1.5 La formula di Bayes	14
1.5.1 La probabilità marginale	15
1.5.2 Moltiplicazione della probabilità condizionata	16
1.5.3 La probabilità condizionata nei test diagnostici	18
2 Distribuzioni di probabilità	21
2.1 Proprietà delle distribuzioni	21
2.1.1 I momenti	21
2.1.2 I cumulanti	22
2.1.3 La mediana	23
2.2 Eventi di Bernoulli e distribuzione binomiale	26
2.2.1 Definizione e proprietà della distribuzione binomiale	26
2.2.2 Il limite di Poisson	29
2.3 La distribuzione di Poisson	29
2.4 La distribuzione gaussiana	31
2.4.1 Relazione tra le distribuzioni binomiale e gaussiana	32
2.4.2 Relazione tra le distribuzioni poissoniana e gaussiana	34
2.5 Distribuzione di Cauchy-Lorentz	35
2.5.1 Distribuzione di Cauchy generalizzata	36
Appendici matematiche	
2.A Integrali gaussiani	36
2.B Funzione Gamma di Eulero	37
2.B.1 La formula di Stirling	39
2.C Metodo di Laplace	39
3 Legge dei grandi numeri e Teorema del limite centrale	41
3.1 La Legge dei grandi numeri	42
3.1.1 Enunciato debole della Legge dei grandi numeri	42
3.1.2 Enunciato forte della Legge dei grandi numeri	44
3.1.3 Il Lemma di Čebyšëv	46
3.1.4 Dimostrazione dell'enunciato debole	47

3.2	Il Teorema del limite centrale	47
3.2.1	Dimostrazione costruttiva del Teorema del limite centrale	48
3.2.2	Dimostrazione dei momenti del Teorema del limite centrale	52
3.3	Teorema del limite centrale generalizzato	54
3.4	Distribuzioni stabili	57
Appendici matematiche		
3.A	Il limite di Markov	58
3.B	Cambiamento di variabili	59
3.C	Trasformata di Fourier, funzione generatrice	60
3.C.1	Funzione generatrice dei momenti di una distribuzione	61
3.C.2	Funzione generatrice dei cumulanti di una distribuzione	62
3.C.3	Generatrice dei cumulanti di una variabile gaussiana	62
3.C.4	Teorema di convoluzione	63
3.C.5	Alcuni dettagli sulla dimostrazione costruttiva del Teorema del limite centrale	64
3.C.6	Trasformata di Fourier di funzioni con decadimento a potenza	66
3.D	Trasformata di Laplace	69
3.E	Polinomi di Hermite	70
3.F	Distribuzioni esatte di somme di variabili aleatorie	70
3.F.1	Distribuzione della somma di variabili con distribuzione uniforme	70
3.F.2	Distribuzione della somma di variabili con distribuzione di Rademacher	72
3.F.3	Distribuzione della somma di variabili con distribuzione esponenziale	72
3.G	Analisi dimensionale	74
3.H	Miscellanea sulla distribuzione delta di Dirac	76
3.H.1	Integrale multiplo della delta di Dirac di una somma	77
3.I	Convergenza di funzioni	78
3.I.1	Convergenza in distribuzione	81
3.J	La generazione di numeri pseudocasuali	82
4	Grandi deviazioni	89
4.1	Teorema delle grandi deviazioni	89
4.2	Dimostrazione del Teorema delle grandi deviazioni	90
4.3	Esempi di applicazione del Teorema delle grandi deviazioni	93
4.3.1	Il caso gaussiano	93
4.3.2	La distribuzione bimodale (o di Rademacher)	93
4.3.3	La distribuzione semigaussiana	95
4.3.4	La distribuzione esponenziale	97
4.3.5	Un confronto	98
4.4	Formalismo termodinamico per le grandi deviazioni	99
4.5	La trasformata di Legendre	101
4.5.1	Proprietà della trasformata di Legendre	102
4.5.2	Trasformata di Legendre in fisica	103
4.5.3	La trasformata di Legendre per funzioni solo continue	104

4.6	Alcuni teoremi delle grandi deviazioni	105
4.6.1	Teorema di Cramer	106
4.6.2	Teorema di Gärtner-Ellis	106
4.6.3	Teorema di Varadhan	106
Appendici matematiche		
4.A	Il metodo del punto di sella	107
4.A.1	Alcune proprietà delle funzioni armoniche e delle funzioni analitiche	107
4.A.2	Proprietà dell'intorno di un punto di sella	109
4.A.3	Dimostrazione del Teorema del punto di sella	110
4.A.4	Esempio di calcolo di un integrale con il metodo del punto di sella	112
5	Inferenza statistica e analisi dei dati sperimentali	117
5.1	Analisi dei dati sperimentali nel caso semplice	117
5.2	L'uso del principio di Bayes	119
5.2.1	L'inferenza delle urne	121
5.2.2	L'inferenza bayesiana per eventi continui	122
5.3	La scelta della distribuzione a priori	124
5.3.1	Dati distribuiti secondo la distribuzione binomiale e un paradosso	125
5.3.2	Dati distribuiti in modo poissoniano	126
5.3.3	La probabilità a priori per la massima ignoranza: la soluzione del paradosso	127
5.3.4	Il caso di diverse categorie poissoniane	130
5.3.5	Dati distribuiti in modo gaussiano	133
5.3.6	Inferenza da dati con distribuzione di Cauchy	135
5.4	Il caso generale di una distribuzione di probabilità ignota	136
5.4.1	Il metodo del ripescaggio	136
5.4.2	Una prima verifica del ripescaggio: la distribuzione di Poisson	142
5.4.3	Una seconda verifica: il caso gaussiano	143
5.5	Metodi di ricampionamento	146
5.5.1	Il ricampionamento binario: come scovare un dato anomalo	146
5.5.2	Il metodo del raggruppamento e le correlazioni temporali	149
5.5.3	Ricampionamento jackknife	152
5.5.4	Ricampionamento bootstrap	155
5.5.5	Un confronto tra ricampionamenti	159
Appendice matematica		
5.A	La distribuzione a posteriori di eventi di Bernoulli	159
6	Analisi di dati sperimentali vettoriali e correlati	161
6.1	Il caso gaussiano	161
6.1.1	Propagazione degli errori	164
6.2	Il caso di sottocampionamento	165
6.3	Il metodo del ripescaggio nel caso vettoriale	166
6.4	I metodi di ricampionamento nel caso vettoriale	169

6.4.1	Ricampionamento jackknife per variabili correlate	169
6.4.2	Ricampionamento bootstrap per variabili correlate	171
6.5	Il metodo dei minimi quadrati	173
6.5.1	Il caso di dati scorrelati	173
6.5.2	La distribuzione del χ^2	179
6.5.3	Il caso di dati correlati	180
6.5.4	Il metodo dei minimi quadrati con il ripesaggio	184
6.5.5	La scelta della matrice della funzione dei minimi quadrati	186
7	Cammini aleatori	189
7.1	Cammini aleatori in un mezzo omogeneo	189
7.1.1	Il caso reticolare	189
7.1.2	Il caso continuo	193
7.2	Cammini aleatori in mezzi non omogenei	194
7.3	Il limite al continuo dei processi di diffusione	195
7.3.1	L'equazione di Fokker-Planck	197
7.4	Cammini aleatori con trappole	198
7.5	Soluzione stazionaria dell'equazione di Fokker-Planck del moto browniano	200
7.5.1	La relazione di Stokes-Einstein	202
7.5.2	Unicità della soluzione di Fokker-Planck stazionaria	203
7.5.3	Equazione di Fokker-Planck ed equazione di Schrödinger	205
7.6	L'equazione di Langevin	206
7.6.1	I processi di Wiener	206
7.6.2	L'autocorrelazione del processo di Wiener	208
7.6.3	La delicata questione dell'integrazione stocastica	209
	Appendici matematiche	
7.A	La trasformata di Fourier su reticolo	210
7.B	Derivazione e proprietà dell'equazione di Fokker-Planck	213
7.B.1	Limite al continuo del camminatore aleatorio	213
7.B.2	Spettro e autofunzioni dell'operatore di Fokker-Planck	215
8	Funzioni generatrici e reazioni a catena	219
8.1	Funzioni generatrici	219
8.1.1	Funzioni generatrici e trasformate di Fourier e Laplace	220
8.1.2	Funzioni generatrici inverse	221
8.1.3	Funzioni generatrici delle probabilità delle code e cumulativa	222
8.1.4	Funzione generatrice della convoluzione	223
8.1.5	Teorema delle probabilità composte	225
8.2	Reazioni a catena	226
8.2.1	Teorema fondamentale delle reazioni a catena	228
8.2.2	Dimostrazione del Teorema fondamentale delle reazioni a catena	229

9	Eventi ricorrenti	233
9.1	Definizioni	233
9.1.1	Esempi di eventi ricorrenti	233
9.2	Classificazione degli eventi ricorrenti	234
9.2.1	Eventi periodici e aperiodici	234
9.2.2	Eventi certi, nulli e transienti	235
9.3	Relazioni fondamentali	236
9.3.1	Probabilità di primo ritorno	236
9.3.2	Funzioni generatrici delle probabilità degli eventi ricorrenti	238
9.4	Teorema della probabilità limite degli eventi ricorrenti	240
9.4.1	Generalizzazione del Teorema al caso periodico	241
9.4.2	L'approccio alla distribuzione limite per tempi lunghi	243
Appendici matematiche		
9.A	Due teoremi per la somma di serie divergenti	244
9.A.1	Dimostrazione del Teorema di divergenza della funzione generatrice nel limite di una serie	245
9.A.2	Dimostrazione del Teorema del limite di una serie data la divergenza della sua funzione generatrice	247
10	Catene di Markov	253
10.1	Esempi di catene di Markov	253
10.1.1	Il camminatore aleatorio	253
10.1.2	Le reazioni a catena	254
10.1.3	Gli eventi ricorrenti	254
10.2	Proprietà generali delle catene di Markov	255
10.2.1	Probabilità di transizione e probabilità condizionata	255
10.2.2	Matrice stocastica	256
10.2.3	Probabilità congiunta	256
10.2.4	Catene di Markov stazionarie	257
10.2.5	Equazione di Chapman-Kolmogorov per le catene	258
10.3	Altri esempi di catene di Markov	258
10.3.1	Ancora gli eventi ricorrenti	258
10.3.2	La catena a due stati	259
10.3.3	La rovina del giocatore, o del perché c'è lo zero alla roulette	260
10.4	Classificazione delle catene di Markov	262
10.4.1	Tipi di catene di Markov	262
10.4.2	Teorema di riducibilità della matrice stocastica	264
10.4.3	La chiusura di uno stato	265
10.4.4	Classificazione degli stati	265
10.5	Probabilità limite di transizione a uno stato persistente	268
10.6	Teoremi fondamentali delle catene di Markov irriducibili	269
10.6.1	Tutti gli stati sono dello stesso tipo	269
10.6.2	Tutti gli stati di una catena finita sono persistenti non-nulli	271
10.6.3	Teorema della chiusura di uno stato persistente	272

10.6.4	Teorema dell'indipendenza dalle condizioni iniziali	273
10.6.5	Teorema della ripartizione degli stati	274
10.7	La relazione di bilancio per catene irriducibili ergodiche	275
10.7.1	Teorema dell'equazione di bilancio	276
10.7.2	Teorema di unicità della soluzione dell'equazione di bilancio	277
10.8	Catene finite e decomposizione spettrale	277
10.8.1	L'ordinamento delle pagine	277
10.8.2	Decomposizione spettrale delle matrici non-hermitiane	279
10.9	Esempi di catene non di Markov	283
10.9.1	Processi non-markoviani continui	284
Appendici matematiche		
10.A	Scambio di limite e somma nell'ipotesi di assoluta convergenza	285
10.B	Decomposizione spettrale	286
10.B.1	Decomposizione delle matrici hermitiane	286
10.B.2	Decomposizione delle matrici non-hermitiane	287
10.C	Teorema di Perron-Frobenius	288
11	Simulazioni numeriche	291
11.1	Catene di Markov inverse	291
11.1.1	Catene di Markov reversibili	292
11.2	Bilancio dettagliato per catene reversibili	292
11.3	Metodi Monte Carlo	296
11.3.1	Integrali in poche e molte dimensioni	296
11.3.2	Campionamento per importanza	299
11.3.3	L'algoritmo di Metropolis	300
11.3.4	Il modello di Ising	302
12	Eventi correlati	307
12.1	Il limite centrale per catene di Markov	307
12.2	Il limite centrale per eventi ricorrenti	312
12.3	Funzioni di correlazione	320
12.3.1	Definizioni	320
12.3.2	Costruzione di van Kampen	322
12.3.3	Funzioni generatrici delle funzioni di correlazione	324
12.3.4	Funzione generatrice delle correlazioni di variabili gaussiane multivariate	325
12.4	Limite centrale e grandi deviazioni per eventi correlati	327
12.4.1	Grandi deviazioni per variabili correlate	331
12.4.2	Il limite centrale per catene di Markov, rivisto	332
12.5	Eventi fortemente correlati e transizioni di fase	333
12.5.1	Limite centrale gaussiano e assenza di transizione di fase	335
12.5.2	Limite centrale generalizzato e comportamento critico	335
12.5.3	Probabilità asintotica nel caso correlato generale	336
Appendice matematica		
12.A	Comportamenti asintotici delle serie e singolarità complesse	337

13	Processi dipendenti dal tempo	341
13.1	Processi di Poisson	341
13.2	Processi di nascita pura	343
13.2.1	Processi divergenti di nascita e il Teorema di Feller	347
13.2.2	Un'analisi approssimata	348
13.2.3	La dimostrazione di Feller	349
13.2.4	Una dimostrazione costruttiva del Teorema di Feller	351
13.2.5	Un calcolo esplicito del tempo di divergenza	354
13.3	Processi di nascita e morte	356
13.4	I processi di Markov	358
14	Entropia, probabilità e meccanica statistica	363
14.1	L'entropia microscopica e la Teoria dell'informazione	364
14.1.1	Definizione dell'entropia di Shannon	364
14.1.2	Codifica di un messaggio	368
14.1.3	Primo teorema di Shannon, o Teorema della codifica della sorgente	370
14.1.4	Dimostrazione del Primo teorema di Shannon	371
14.1.5	Trasmissione di informazione in presenza di rumore	374
14.1.6	Secondo teorema di Shannon, o Teorema della codifica del canale con rumore	376
14.1.7	Dimostrazione del Secondo teorema di Shannon	377
14.2	L'entropia nei sistemi dinamici	379
14.2.1	L'esponente principale di Lyapunov	379
14.2.2	La produzione di entropia	382
14.2.3	L'entropia di Kolmogorov-Sinai	383
14.3	Intermezzo: richiami di meccanica statistica	384
14.3.1	La distribuzione di Boltzmann	385
14.3.2	Il Teorema di risposta lineare	385
14.3.3	Potenziali termodinamici	386
14.3.4	Principi variazionali	387
14.4	Il principio di massima entropia	387
14.4.1	Una derivazione probabilistica della meccanica statistica	389
14.5	Grandi deviazioni e termodinamica	390
14.5.1	Energie libere di Helmholtz e di Gibbs	392
14.5.2	Entropia e funzione di decrescita delle grandi deviazioni	394
14.6	L'entropia configurazionale dei sistemi vetrosi	395
Appendici matematiche		
14.A	Derivazione delle espressioni di informazione ed entropia di Shannon	397
14.A.1	Costruzione della funzione informazione	398
14.A.2	Dimostrazione della formula dell'entropia di Shannon	399
14.B	Messaggi autodelimitanti	401
Bibliografia		403
Indice analitico		409
Esercizi risolti		online.universita.zanichelli.it/leuzzi-probabilita

Prefazione

Questo libro, che per molti anni è stato per noi affettuosamente “Il Trattatello” (e ringraziamo Giovanni Gallavotti per averci donato questa colta ispirazione), nasce e cresce con il corso di Calcolo delle Probabilità che abbiamo insegnato a staffetta agli studenti di Fisica della Sapienza, attraverso diversi ordinamenti. Una staffetta così lunga che uno degli studenti degli inizi (L.L.) ha fatto in tempo a diventare portatore del testimone.

Partendo da un’asciutta presentazione dei fondamenti della moderna teoria della probabilità, cui dedichiamo il Capitolo 1, e da una panoramica, nel Capitolo 2, sulle distribuzioni di probabilità e le loro proprietà, la nostra intenzione è di fornire ai lettori concetti fondamentali, metodi di calcolo e loro applicazioni allo studio di problemi probabilistici. Non sempre in quest’ordine.

Dopo aver approfondito lo studio della Legge dei grandi numeri, del Teorema del limite centrale e della Teoria delle grandi deviazioni (nei Capitoli 3 e 4) abbiamo sperimentato che, alle volte, risulta didatticamente più incisivo partire dalle applicazioni, con approcci mirati, per poi astrarre a teorie più generali.

Nei Capitoli 5 e 6 diamo numerosi esempi di analisi dei dati sperimentali, scalari e vettoriali, alternando l’applicazione di metodi specifici all’approfondimento della teoria su cui si fondano. Nel Capitolo 7 studiamo il primo esempio di fenomeno correlato privo di memoria: il camminatore aleatorio (il famoso passeggiatore ubriaco). Questo, naturalmente, senza resistere alla tentazione di una incursione in campo continuo con l’equazione della diffusione di Fokker-Planck, che scopriamo essere un’equazione di Schrödinger nel tempo immaginario, e la sua parente stretta, l’equazione differenziale stocastica di Langevin.

Nei Capitoli 8 e 9 vediamo altri processi discreti smemorati: le reazioni a catena e gli eventi ricorrenti. Scopriamo, quindi, nel Capitolo 10, che sono tutte catene di Markov e deriviamo una teoria generale per poi continuare, anche nel Capitolo 11, con le applicazioni, che analizzano per esempio l’ordinamento delle pagine web da parte di motori di ricerca o la simulazione al computer della dinamica di modelli di fisica statistica.

Il Capitolo 12 è dedicato alle correlazioni. Riprendiamo il Teorema del limite centrale, prima in contumacia, con un paio di esempi specifici risolti con una notevole – ma istruttiva – fatica e poi con il Teorema del limite centrale per eventi correlati, che semplifica di gran lunga le cose. Ritorniamo a Markov nel Capitolo 13, questa volta con i processi a tempi continui e le equazioni maestre. Finiamo in bellezza con un Capitolo, il 14, dedicato all’entropia, in

Teoria dell'informazione, nei sistemi dinamici, in fisica statistica, riallacciando alcuni nodi lasciati in sospeso nei capitoli precedenti.

In questo viaggio appassionato nello scrivere, nell'insegnare, nel confrontarci tra di noi e con colleghi, amici e studenti, abbiamo imparato e curiosato molto e, soprattutto, ci siamo molto divertiti. Il nostro augurio è ora che almeno un po' di questo divertimento possa arrivare ai nostri lettori.

Roma, 31 marzo 2023

Luca Leuzzi, Enzo Marinari, Giorgio Parisi

■ RINGRAZIAMENTI

Troppi colleghi e amici ci hanno insegnato cose importanti e hanno discusso con noi durante la stesura: sarebbe impossibile elencarli tutti. Alcuni però balzano alla mente: Andrea Baldassarri, Sebastiano Carpi, Chiara Cammarota, Andrea Crisanti, Persi Diaconis, Andrea Gabrielli, Irene Giardina, Andrea Puglisi, Federico Ricci-Tersenghi, Angelo Vulpiani.

Un incredibile contributo, di stimoli, di domande curiose e interessate e di oculatissimo controllo è venuto, anno dopo anno, dagli studenti del corso. Senza di loro questo libro non sarebbe mai stato scritto.

Luca ringrazia Donatella e Pino, per aver pazientemente assecondato e sostenuto una passione così diversa dalle loro, Alessandra, per il suo sostegno centrale, come e oltre il limite, e grande come le deviazioni (che la vita ci offre), e Aurora, che trasforma la prospettiva e proietta ogni esperienza, persino scrivere un manuale di probabilità, in uno spazio mai banale e molto luminoso.

Enzo è grato a Margherita e a Chiara, per le quali, come per tutti noi, saper valutare la probabilità di quel che succederà è sempre molto importante.

Giorgio ringrazia Daniella, Lorenza e Leonardo, che mentre lui rimuginava questo libro non se ne sono neanche accorti, dato che lui con loro parlava di tutt'altro: senza saperlo lo hanno aiutato come sempre, perché quando si sta bene insieme si pensa meglio.

10 Catene di Markov

■ 10.1 ESEMPI DI CATENE DI MARKOV

Per *catene di Markov* si intendono processi dinamici aleatori di variabili discrete a passi temporali discreti che hanno la particolarità di essere descritti semplicemente conoscendo la probabilità di transire da un valore all'altro della variabile in un passo mediante la semplice concatenazione di queste probabilità nel tempo. Chiameremo spesso *stati* i valori delle variabili. Conoscendo la distribuzione degli stati a un dato passo, serve soltanto conoscere la probabilità di transizione in un passo temporale a un altro stato per determinare lo stato attuale del sistema. Non importa come si è arrivati allo stato di partenza: la storia precedente non condiziona il passo futuro, dunque per conoscere (stocasticamente) l'evoluzione del sistema è completamente irrilevante ricordarsi che cosa è successo prima. La distribuzione dei valori delle variabili al passo precedente sarà a sua volta data dalla distribuzione al passo ancora precedente e dalla probabilità di transizione. E così a ritroso, formando una catena. Le catene di Markov sono una struttura matematica molto generale che contiene come caso particolare molti dei sistemi che abbiamo studiato fino ad adesso. Prima di entrare nei dettagli formali della teoria delle catene di Markov, quindi, vediamo di caratterizzare meglio questo tipo di fenomeni a noi già noti.

10.1.1 Il camminatore aleatorio

Nel Capitolo 7 abbiamo studiato il camminatore aleatorio su di un reticolo e abbiamo visto che la probabilità $P(\vec{m}, N)$ che il camminatore si trovi al passo temporale N nel sito \vec{m} soddisfa l'equazione (7.25), che riportiamo qui per comodità

$$P(\vec{m}, N) = \sum_{\vec{\ell}} J(\vec{m} - \vec{\ell}, \vec{\ell}) P(\vec{m} - \vec{\ell}, N - 1), \quad (10.1)$$

dove $P(\vec{m} - \vec{\ell}, N - 1)$ è la probabilità di trovarsi nel sito $\vec{m} - \vec{\ell}$ al passo precedente e $J(\vec{m} - \vec{\ell}, \vec{\ell})$ è la probabilità di transire da $\vec{m} - \vec{\ell}$ a \vec{m} in un passo. Notiamo che abbiamo sommato soltanto sulle probabilità al passo precedente $N - 1$. La probabilità di trovarsi nel sito \vec{m} dipende solo dalle probabilità di trovarsi da

Iterazione dinamica della probabilità assoluta

Data la semplice regola di evoluzione (10.8) della matrice di transizione da uno a n passi, la probabilità assoluta di stare in un dato stato k al passo n si itera come

$$p_k(n) = \sum_j p_j(n-1) P_{jk} = \dots = \sum_j p_j(0) P_{jk}^{(n)}. \quad (10.10)$$

10.2.5 Equazione di Chapman-Kolmogorov per le catene

Dalle proprietà di iterazione nei passi della matrice stocastica della catena di Markov possiamo scrivere l'equazione per la probabilità di transire da uno stato j a uno stato k in $n+m$ passi

$$P_{jk}^{(n+m)} = (P^{n+m})_{jk} = \sum_l (P^n)_{jl} (P^m)_{lk} = \sum_l P_{jl}^{(n)} P_{lk}^{(m)}, \quad (10.11)$$

che è la forma a stati e passi temporali discreti della cosiddetta equazione di Chapman-Kolmogorov per processi stocastici.

■ 10.3 ALTRI ESEMPI DI CATENE DI MARKOV

10.3.1 Ancora gli eventi ricorrenti

È vero, ne abbiamo già parlato, ma dopo aver introdotto la matrice stocastica vogliamo provare a dare una rappresentazione formalmente più esplicita della catena di eventi ricorrenti come una catena di Markov. Consideriamo quindi una catena di Markov con stati S_k , $k = 0, 1, \dots$ e prendiamo, per esempio, come evento ricorrente $\mathcal{E} = \{\text{realizzare lo stato } S_0 \text{ della catena}\}$. Possiamo descrivere la dinamica dell'evento ricorrente \mathcal{E} mediante una catena di Markov con la seguente matrice stocastica:

$$P = \begin{pmatrix} S_0 & S_1 & S_2 & S_3 & \dots & \dots \\ f_1 & f_2 & f_3 & f_4 & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & \dots \\ 0 & 1 & 0 & 0 & \dots & \dots \\ 0 & 0 & 1 & 0 & \dots & \dots \\ 0 & 0 & 0 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (10.12)$$

Gli elementi della riga 0 sono le probabilità di primo ritorno, $P_{0k} = f_{k+1}$, di un evento certo (o ritornante, o persistente). Gli elementi della matrice delle righe inferiori sono tutti nulli tranne

$$P_{k-1, k-2} = P_{k-2, k-3} = \dots = P_{2, 1} = P_{1, 0} = 1.$$

La proprietà di chiusura della matrice stocastica P_{jk} è banalmente soddisfatta per $j > 0$ (un solo elemento non nullo pari a 1). Per $j = 0$ si ha

$$\sum_{k=0}^{\infty} P_{0k} = \sum_{k=1}^{\infty} f_k = 1,$$

perché stiamo considerando un evento ricorrente certo. Se il primo passo porta da S_0 a S_{k-1} , nei $k-1$ passi successivi la catena è vincolata a fare con probabilità 1 le transizioni $k-1 \rightarrow k-2$, $k-2 \rightarrow k-3$, ..., $1 \rightarrow 0$. L'evento S_0 è perciò un evento ricorrente (persistente) e i tempi di ritorno k hanno distribuzione $\{f_k\}$. Va da sé che la stessa dinamica di ritorno si può rappresentare per qualsiasi stato S_k .

10.3.2 La catena a due stati

Prendiamo una catena di Markov con soli due stati, 0 e 1, come quella riprodotta in Figura 10.1. Indichiamo $P_{01} = \gamma$ e $P_{10} = \alpha$. Per la condizione di chiusura si ha $P_{00} = 1 - \gamma$ e $P_{11} = 1 - \alpha$. Denotiamo con $p_k(n)$, $k = 0, 1$ la probabilità di trovarsi nello stato k al passo n .

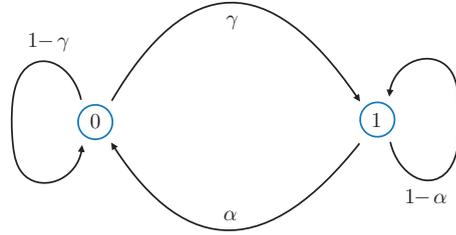


Figura 10.1. Catena di Markov a due stati.

A ogni passo n varrà la chiusura $p_0(n) + p_1(n) = 1$. Possiamo scrivere le equazioni della dinamica della catena a due stati in modo ricorsivo partendo dalla formula (10.10)

$$p_k(n) = \sum_{j=0,1} p_j(n-1)P_{jk} = p_0(n-1)P_{0k} + p_1(n-1)P_{1k}, \quad k = 0, 1. \quad (10.13)$$

Nel caso non banale, $\alpha, \gamma \in (0, 1)$, iteriamo $p_0(n)$ a partire da n passi

$$\begin{aligned} p_0(n) &= p_0(n-1)P_{00} + p_1(n-1)P_{10} \\ &= p_0(n-1)(1-\gamma) + [1-p_0(n-1)]\alpha \\ &= \alpha + (1-\gamma-\alpha)p_0(n-1) \\ &= \alpha + (1-\gamma-\alpha)[\alpha + (1-\gamma-\alpha)p_0(n-2)] \\ &= \alpha + (1-\gamma-\alpha)[\alpha + (1-\gamma-\alpha)[\alpha + (1-\gamma-\alpha)p_0(n-3)]] \\ &= \alpha + \alpha(1-\gamma-\alpha) + \alpha(1-\gamma-\alpha)^2 + (1-\gamma-\alpha)^3 p_0(n-3) \\ &= \alpha + \alpha(1-\gamma-\alpha) + \alpha(1-\gamma-\alpha)^2 + \dots + \alpha(1-\gamma-\alpha)^{n-1} + \\ &\quad + (1-\gamma-\alpha)^n p_0(0) \\ &= \alpha \sum_{k=0}^{n-1} (1-\gamma-\alpha)^k + (1-\gamma-\alpha)^n p_0(0). \end{aligned} \quad (10.15)$$

Ci sono due casi limite. Per $\alpha = \gamma = 1$ si ha un ciclo continuo tra 0 e 1, se parte da 0 ai passi dispari il sistema si troverà certamente in 1 e ai passi pari certamente in 0 (e viceversa se si parte da 1). Se $\alpha = \gamma = 0$, è $P_{00} = P_{11} = 1$ e dai due stati non è mai possibile uscire. Questi stati si dicono *assorbenti*.

Per tutti gli altri casi, $1 - \alpha - \gamma \in (-1, 1)$, si ottiene la soluzione

$$p_0(n) = \frac{\alpha}{\gamma + \alpha} + (1 - \alpha - \gamma)^n \left[p_0(0) - \frac{\alpha}{\gamma + \alpha} \right] \quad (10.16)$$

partendo dalla condizione iniziale $p_0(0)$. La distribuzione limite ($n \rightarrow \infty$) è, quindi, $p_0 = \alpha/(\alpha + \gamma)$, $p_1 = \gamma/(\alpha + \gamma)$ e perde la dipendenza dalla condizione iniziale. Inoltre, se scegliamo simmetrica la matrice stocastica di transizione, $\gamma = \alpha$, l'equazione (10.16) si riduce a

$$p_0(n) = \frac{1}{2} + (1 - 2\alpha)^n \left[p_0(0) - \frac{1}{2} \right] \tag{10.17}$$

e, per $n \rightarrow \infty$, $p_0 = p_1 = 1/2$ indipendentemente da α .

10.3.3 La rovina del giocatore, o del perché c'è lo zero alla roulette

Prendiamo una catena di Markov composta da un camminatore aleatorio che si può muovere in una dimensione lungo $T + 1$ stati spostandosi solo verso stati limitrofi, tra due barriere in 0 e in T che diremo assorbenti. Lo stato 0 e lo stato T hanno la peculiarità che una volta raggiunti non si possono più lasciare. Questo assorbimento è rappresentato dagli elementi della matrice di transizione $P_{00} = P_{TT} = 1$.

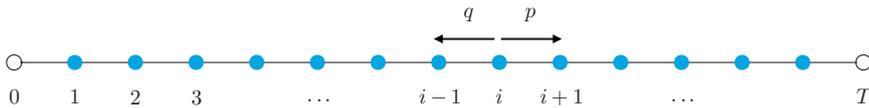


Figura 10.2. Catena di Markov con transizioni a “primi vicini” e barriere assorbenti agli estremi 0 e T .

Denotiamo la probabilità di muoversi in avanti di un passo con $P_{i\ i+1} = p$ e quella di muoversi all'indietro con $P_{i\ i-1} = q$, con $p + q = 1$. Alla fine la matrice stocastica di questa catena di Markov avrà la forma

$$P \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & q & 0 & p & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{pmatrix} \tag{10.18}$$

Usando questa catena di Markov possiamo calcolare qual è la probabilità di giungere nello stato *assorbente* 0 partendo da uno stato $i > 0$. *Mutatis mutandis*, possiamo riformulare il tutto in termini di gioco d'azzardo e chiederci qual è la probabilità che un giocatore vada in rovina se parte da un gruzzolo di i € e a ogni puntata può vincere o perdere 1 €. Inoltre, la sala da gioco escluderà il giocatore dal fare nuove puntate una volta raggiunto un certo tetto T €.

Chiamiamo \mathcal{E} l'evento di finire i soldi e indichiamo con \mathcal{R}_i la probabilità che si verifichi \mathcal{E} iniziando a giocare con i €. Diremo che il gruzzolo del giocatore al passo zero (la posizione iniziale del camminatore aleatorio) è $x(0) = i$, dunque $\mathcal{R}_i \equiv \text{Prob}(\mathcal{E} | x(0) = i)$.

L'origine dei tempi non ha in realtà importanza, in quanto \mathcal{E} è definito per qualsiasi intervallo temporale. Non è l'evento di fallire in n passi, ma solo di fallire, prima o poi. Di conseguenza la probabilità \mathcal{R}_i è la probabilità di fallire partendo da $i \in$ in un tempo qualsiasi, o, equivalentemente, di fallire avendo a un certo momento un gruzzolo di $i \in$.

Due casi sono già dati. Il primo è $\mathcal{R}_0 = 1$: se si parte senza soldi, con certezza si rimane senza. L'altro caso è $\mathcal{R}_T = 0$: non si entra in gioco con una quota pari al tetto, per cui non si può andare in rovina.

Per le probabilità dei casi intermedi possiamo scrivere una formula iterativa. Chiamiamo $p(x(n+1) = k | x(n) = j)$ la probabilità condizionata di avere $k \in$ al passo $n+1$ se si hanno $j \in$ al passo n , che è pari alla probabilità di transizione P_{jk} dallo stato j allo stato nella matrice (10.18). Scriviamo, quindi, la seguente uguaglianza

$$\begin{aligned} \mathcal{R}_i &= \mathcal{R}_{i+1} p(x(1) = i+1 | x(0) = i) + \mathcal{R}_{i-1} p(x(1) = i-1 | x(0) = i) \\ &= \mathcal{R}_{i+1} p + \mathcal{R}_{i-1} q, \end{aligned} \quad (10.19)$$

che esprime il fatto che la probabilità di andare in bancarotta avendo $i \in$ sia pari alla probabilità di vincere $1 \in$ [$p(x(1) = i+1 | x(0) = i) = P_{i, i+1} = p$] e fallire partendo da $i+1 \in$ sommata alla probabilità dell'evento disgiunto di perdere $1 \in$ [$p(x(1) = i-1 | x(0) = i) = P_{i, i-1} = q$] e andare in rovina partendo da $i-1 \in$. Ripetiamo nuovamente che \mathcal{R}_i non è la probabilità di avere $i \in$ (cioè di trovarsi nello stato i), ma è la probabilità di andare in bancarotta partendo da un gruzzolo di $i \in$.

Una soluzione elementare dell'equazione (10.19) avrà la forma $\mathcal{R}_k = A^k$. Il valore della base A si può ricavare verificando la (10.19). Si trovano due possibili valori $A = 1, q/p$. Nel caso in cui $q \neq p$ la soluzione generale si trova come combinazione lineare delle due soluzioni elementari

$$\mathcal{R}_k = c_1 + c_2 \left(\frac{q}{p}\right)^k. \quad (10.20)$$

I valori delle costanti si trovano imponendo le condizioni agli estremi, $\mathcal{R}_0 = 1$ e $\mathcal{R}_T = 0$

$$c_1 = \frac{1}{1 - (q/p)^T}; \quad c_2 = -\frac{(q/p)^T}{1 - (q/p)^T}, \quad (10.21)$$

da cui

$$\mathcal{R}_k = \frac{(p/q)^{T-k} - 1}{(p/q)^T - 1}. \quad (10.22)$$

Se, invece, $q = p = 1/2$, la soluzione elementare del tipo A^k fornisce una unica soluzione ($A = 1$). Per formulare la soluzione generale utilizziamo allora un'altra soluzione elementare, di forma $\mathcal{R}_k = k$, da cui $\mathcal{R}_k = c_1 k + c_2$. Imponendo le condizioni ai bordi si ottiene, quindi,

$$\mathcal{R}_k = 1 - \frac{k}{T}. \quad (10.23)$$

Possiamo notare che, se il gioco è equo ($p = q = 1/2$), come testa o croce, oppure pari e dispari, la probabilità di andare in rovina cresce linearmente all'alleggerirsi del gruzzolo del giocatore. Il che è molto ragionevole. Se un giocatore parte da 500 €, la probabilità che vada in rovina se il tetto massimo è 1000 € è $\mathcal{R}_{500} = 1/2$.

Vediamo che cosa succede se il gioco non è equo: per esempio, se il giocatore punta sul rosso alla roulette. Ci sono 18 caselle rosse, 18 nere e una verde (lo zero). Dunque $p = 18/37$, $q = 19/37$ e $p/q = 18/19$. Se prendiamo il caso di prima, una somma di 500 € da puntare e un tetto di 1000 € imposto dal croupier, la formula (10.22) ci fornisce la probabilità $\mathcal{R}_{500} = 1 - 2 \cdot 10^{-12}$. Una differenza di probabilità tra vincita e perdita di $1/37$ fa passare la probabilità di rovina da $1/2$ a 1 ! Per avere una probabilità pari a quella di testa o croce, nella puntata su rosso o nero nella roulette bisogna partire da una disponibilità di almeno 987 € da puntare (su 1000 € di tetto massimo).

Che cosa succede se leviamo il tetto massimo? Se il gioco è equo, vediamo dalla (10.23) con $T \rightarrow \infty$ che $\mathcal{R}_k = 1, \forall k$. A maggior ragione, la rovina è certa se $p < q$, come conferma l'equazione (10.22) quando si manda $T \rightarrow \infty$. Se invece $p > q$, sempre dalla (10.22) abbiamo $\mathcal{R}_k = (q/p)^k$: se non esiste lo stato assorbente superiore, ci può essere in questo caso una probabilità finita pari a $1 - (q/p)^k$ di poter giocare in eterno partendo da k €.

In generale, in tutti gli esempi sopra esposti è evidente la proprietà di dipendere solo dal passo precedente che *identifica* le catene di Markov: sono *processi correlati* in cui la probabilità di finire in uno stato dipende dalla probabilità dello stato precedente, ma sono *senza memoria*, in quanto tutto ciò che riguarda la storia precedente al verificarsi dell'evento precedente non conta, è dimenticata.

■ 10.4 CLASSIFICAZIONE DELLE CATENE DI MARKOV

Allo scopo di semplificare la trattazione delle catene di Markov è conveniente introdurre una prima classificazione. A una catena di Markov possiamo associare un grafico, si vedano le Figure 10.3 e 10.4 a pagina seguente, in maniera tale che gli stati (o eventi) della catena siano i nodi del grafico e le transizioni permesse siano indicate con linee orientate da un nodo all'altro.

10.4.1 Tipi di catene di Markov

Catene decomponibili e indecomponibili

Una catena è detta *decomponibile* se si possono dividere gli stati in due sottoinsiemi \mathcal{A} e \mathcal{B} tali che non sia possibile fare transizioni da \mathcal{A} a \mathcal{B} e viceversa. In questo caso la matrice P può essere pensata come il prodotto tensoriale di due matrici che agiscono su due spazi indipendenti. Una catena di Markov è detta *indecomponibile* se non è decomponibile. Lo studio di una catena decomponibile si può separare nello studio delle due sottocatene di Markov \mathcal{A} e \mathcal{B} . Molto spesso verranno quindi considerate catene indecomponibili. Ogni catena di Markov può essere decomposta nella somma numerabile di catene indecomponibili (si veda il Paragrafo 10.6.5).

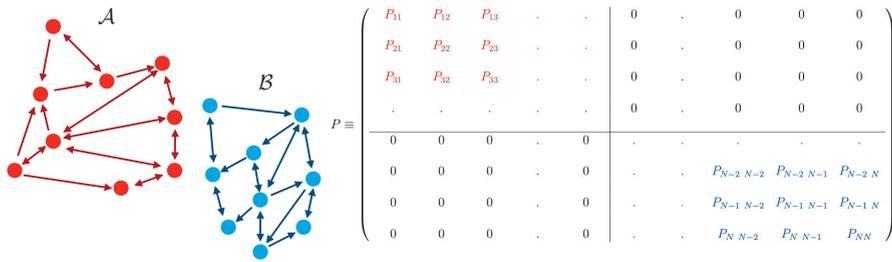


Figura 10.3. Catena di Markov di N stati decomponibile in due sottocatene \mathcal{A} e \mathcal{B} e relativa matrice stocastica.

Catene riducibili e irriducibili

Una catena è detta *riducibile* se si possono dividere gli stati in due sottoinsiemi \mathcal{C} e \mathcal{D} tali che non sia possibile fare transizioni da \mathcal{C} a \mathcal{D} , mentre le transizioni da \mathcal{D} a \mathcal{C} possono avvenire. In questo caso possiamo definire una sottocatena di Markov limitata all'insieme \mathcal{C} . Naturalmente una catena è *irriducibile* se non è riducibile. In generale una catena di Markov può contenere una o più sottocatene irriducibili.

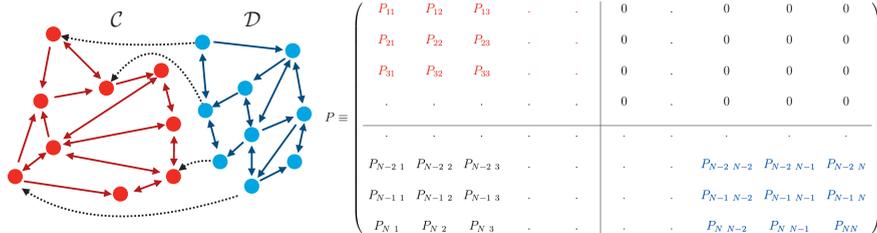


Figura 10.4. Catena di Markov composta da due sottoinsiemi di stati \mathcal{C} e \mathcal{D} riducibile alla sottocatena \mathcal{C} e relativa matrice stocastica.

Stato accessibile, stato assorbente, insieme chiuso

Uno stato k si dice *accessibile* dallo stato j se esiste un numero di passi n tale che la probabilità di transizione $P_{jk}^{(n)} > 0$: prima o poi $j \rightarrow k$ è una transizione possibile.

Un sottoinsieme di stati \mathcal{C} verso i quali è possibile accedere da stati fuori di \mathcal{C} ma dai quali non è possibile accedere a nessuno stato fuori di \mathcal{C} è detto *insieme chiuso*. Se in una catena non esiste nessun insieme chiuso, vuol dire che l'intera catena è l'insieme chiuso e questa è irriducibile. *In una catena di Markov irriducibile ogni stato è accessibile da ogni altro.*

Può esistere un insieme chiuso composto da un singolo stato, che è detto *stato assorbente* (si veda il Paragrafo 7.4, il cammino aleatorio con trappole, o il Paragrafo 10.3.3, la rovina del giocatore), per il quale $P_{kk} = 1$: una volta raggiunto lo stato k , si rimane in k con certezza.

10.4.2 Teorema di riducibilità della matrice stocastica

Un insieme chiuso \mathcal{C} di stati di una catena può essere studiato indipendentemente da tutti gli altri stati. Se guardiamo gli elementi della matrice stocastica dell'intera catena, avremo che, se $j \in \mathcal{C}$ e $k \notin \mathcal{C}$, allora $P_{jk}^{(n)} = 0$ per ogni n . In dettaglio avremo che

$$0 = P_{jk}^{(n)} = \sum_{i_1, \dots, i_{n-1}} P_{ji_1} P_{i_1 i_2} \dots P_{i_{n-1} k}, \quad (10.24)$$

dove le varie $P_{i_{r-1} i_r}$ sono elementi della matrice stocastica di transizione. L'espressione sopra sarà nulla, perché negli n passi prima o poi si proporrà la transizione tra uno stato $i_{r-1} \in \mathcal{C}$ e uno i_r fuori dall'insieme chiuso e, dunque, sarà nullo almeno l'elemento $P_{i_{r-1} i_r}$. Non è vero, in generale, l'opposto, si veda la Figura 10.4: da uno stato esterno all'insieme chiuso si può transire a uno stato interno, $P_{kj}^{(n)} \geq 0$.

A questo punto si può enunciare il Teorema di riducibilità per la matrice stocastica di una catena di Markov riducibile composta da $M > N$ stati (eventualmente anche con $M \rightarrow \infty$). Una catena di Markov riducibile di matrice stocastica P contiene almeno un insieme chiuso \mathcal{C}_N di N stati. Gli altri stati appartengono al complemento $\bar{\mathcal{C}}_N$. Eliminando da P tutte le righe contenenti le probabilità di transizione verso stati appartenenti a \mathcal{C}_N da stati in $\bar{\mathcal{C}}_N$ e tutte le colonne contenenti le probabilità di transizione verso stati del complemento, si ottiene una matrice ridotta $N \times N$, $P(\mathcal{C}_N)$, che è ancora stocastica e che caratterizza la catena di Markov composta esclusivamente dall'insieme chiuso \mathcal{C}_N .

Prendiamo per esempio la catena di Markov riducibile riprodotta in Figura 10.5. La matrice stocastica P delle probabilità di transizione tra stati è

$$P \equiv \begin{pmatrix} 0 & P_{12} & P_{13} & 0 & P_{15} & 0 \\ P_{21} & 0 & P_{23} & 0 & P_{25} & 0 \\ P_{31} & P_{32} & 0 & 0 & P_{35} & 0 \\ P_{41} & P_{42} & P_{43} & 0 & P_{45} & 0 \\ P_{51} & P_{52} & P_{53} & 0 & 0 & 0 \\ P_{61} & P_{62} & P_{63} & P_{64} & P_{65} & 0 \end{pmatrix}, \quad (10.25)$$

dove abbiamo esplicitamente messo a zero le probabilità di transizioni assenti in figura. La matrice ridotta 4×4 delle transizioni tra stati della sottocatena

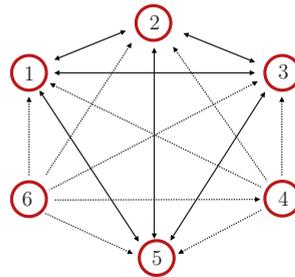


Figura 10.5. Esempio di catena di Markov riducibile. Gli stati $\{1, 2, 3, 5\} \equiv \mathcal{C}_4$ formano una sottocatena chiusa irriducibile: ognuno di essi è accessibile a ciascuno degli altri e viceversa. Gli stati 4 e 6, invece, non fanno parte di un insieme chiuso: da questi due stati si possono raggiungere gli stati dell'insieme \mathcal{C}_4 , ma essi non sono raggiungibili da nessuno degli stati in \mathcal{C}_4 .

chiusa irriducibile \mathcal{C}_4 vale, quindi,

$$P(\mathcal{C}_4) = \begin{pmatrix} 0 & P_{12} & P_{13} & P_{15} \\ P_{21} & 0 & P_{23} & P_{25} \\ P_{31} & P_{32} & 0 & P_{35} \\ P_{51} & P_{52} & P_{53} & 0 \end{pmatrix}. \quad (10.26)$$

Se togliamo le righe e le colonne delle transizioni da e verso stati fuori dall'insieme chiuso $\mathcal{C}_4 \equiv \{1, 2, 3, 5\}$, le condizioni di chiusura di ciascuna riga non cambiano, perché dalle righe delle transizioni da stati in \mathcal{C}_4 eliminiamo solo elementi nulli. Conseguentemente la $P(\mathcal{C}_4)$ è ancora stocastica e vale l'equazione di Chapman-Kolomogorov (10.11). Rivedremo e dimostreremo questa proprietà nel Paragrafo 10.6.5, come parte di un teorema più generale.

10.4.3 La chiusura di uno stato

Si definisce *chiusura di uno stato* l'insieme di stati chiuso più piccolo che contiene lo stato in questione. In altri termini, la chiusura di uno stato è costituita da tutti gli stati raggiungibili da quello stato. Per alcuni tipi di stati, che definiremo tra poco, la loro chiusura è un insieme irriducibile, per cui si avrà la stessa chiusura per tutti gli stati dell'insieme irriducibile. Altri tipi di stati, invece, hanno chiusure più complicate, che possono contenere le chiusure di altri stati (anche in numero infinito) che, a loro volta, *non* le contengono. Prendiamo, per esempio, la catena definita dalla matrice stocastica

$$\begin{aligned} P_{jk} &= 0 && \text{se } k \leq j \\ P_{jk} &> 0 && \text{se } k = j + 1 \\ P_{jk} &\geq 0 && \text{se } k \geq j + 1. \end{aligned} \quad (10.27)$$

Si possono avere soltanto transizioni in avanti: per questo motivo la chiusura di j sarà composta da j e da tutti gli stati $k > j$ della catena di Markov. Questo insieme contiene anche la chiusura di $j + 1$, che non contiene j ma contiene la chiusura di $j + 2$ e così via dicendo (si veda la Figura 10.6).

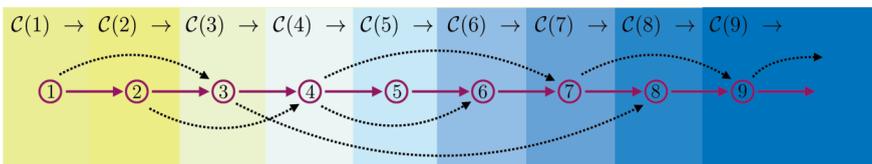


Figura 10.6. Esempio di chiusure di stati della catena di Markov definita in (10.27). Con il simbolo $\mathcal{C}(k) \rightarrow$ indichiamo la chiusura dello stato k (aperta verso destra).

10.4.4 Classificazione degli stati

Abbiamo già visto che, dato uno stato k arbitrario e definito l'evento «la catena di Markov si trova nello stato k al tempo n », questo soddisfa tutte le condizioni per essere un evento ricorrente al variare di n . In particolare, la condizione di

l'insieme $\mathcal{A} = \cup_m \mathcal{A}(m)$ di tutti gli stati raggiungibili da a (prima o poi). È immediato vedere che non ci possono essere transizioni tra un elemento di \mathcal{A} e un elemento che non appartiene ad \mathcal{A} . Prendiamo, per esempio, uno stato d generico che sia accessibile in un passo partendo da un elemento $c \in \mathcal{A}$. Per definizione di \mathcal{A} c è accessibile a partire da a dopo un qualche m : $c \in \mathcal{A}(m)$. Quindi d risulterà raggiungibile in $m + 1$ passi a partire da a : $d \in \mathcal{A}(m + 1) \in \mathcal{A}$.

La condizione di irriducibilità implica che l'insieme \mathcal{A} coincida con la catena stessa, dunque qualunque stato b appartiene ad \mathcal{A} e, di conseguenza, esiste un numero m tale che $P_{ab}^{(m)}$ sia diversa da zero. Nello stesso modo esiste un numero n tale che $P_{ba}^{(n)}$ sia diversa da zero. Per brevità diremo $P_{ab}^{(m)} = \alpha > 0$ e $P_{ba}^{(n)} = \beta > 0$.

In generale, la probabilità di partire da a e ritornare in a in $\ell + m + n$ passi sarà maggiore o uguale alla probabilità che il sistema vada da a a b in m passi, ritorni in b in ℓ passi e successivamente vada da b a a in n passi, si veda la Figura 10.7. Infatti la possibile esistenza di altri percorsi alternativi da a a a non può che aumentare la probabilità di ritorno in a . Abbiamo quindi la disuguaglianza

$$P_{aa}^{(\ell+m+n)} \geq P_{ab}^{(m)} P_{bb}^{(\ell)} P_{ba}^{(n)} = \alpha\beta P_{bb}^{(\ell)}. \tag{10.42}$$

Nello stesso modo, scambiando $a \leftrightarrow b$, otteniamo

$$P_{bb}^{(\ell+m+n)} \geq \alpha\beta P_{aa}^{(\ell)}, \tag{10.43}$$

dove ricordiamo che sia α sia β sono probabilità strettamente non nulle e, dunque, $\alpha\beta \leq 1$. Le due relazioni precedenti implicano che il comportamento asintotico per grandi ℓ di $P_{aa}^{(\ell)}$ e $P_{bb}^{(\ell)}$ sia lo stesso. Di conseguenza si avrà

$$\sum_{\ell=0}^{\infty} P_{aa}^{(\ell)} \sim \sum_{\ell=0}^{\infty} P_{bb}^{(\ell)}, \tag{10.44}$$

dove \sim indica l'andamento a meno di un numero finito di termini. Quindi le serie saranno o entrambe serie convergenti, e a e b saranno entrambi transienti, oppure entrambe divergenti, e allora gli stati saranno entrambi persistenti. In quest'ultimo caso, visto che anche

$$\lim_{\ell \rightarrow \infty} P_{aa}^{(\ell)} = \lim_{\ell \rightarrow \infty} P_{bb}^{(\ell)}, \tag{10.45}$$

a e b saranno entrambi persistenti nulli o entrambi persistenti non-nulli.

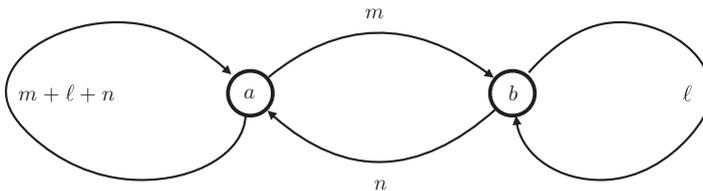


Figura 10.7. Schema per la dimostrazione del teorema 1, si veda il Paragrafo 10.6.1. Transizione $a \rightarrow b \rightarrow a$ composta da passaggio da a a b in m passi, seguita da ritorno in b in ℓ passi e da $b \rightarrow a$ in n passi. Nel testo, nella disuguaglianza (10.42) viene comparato con un più generico ritorno in a ($a \rightarrow a$) in $m + l + n$ passi.

Inoltre, è vero anche che due stati in una catena irriducibile devono essere entrambi periodici, con la stessa periodicità. Dunque anche entrambi aperiodici (di periodicità 1). Vediamo di dimostrarlo.

Se a è periodico di periodicità M , per definizione esiste un ℓ multiplo di M tale che $P_{aa}^{(\ell)} > 0$. Dalla (10.42) con $\ell = 0$ abbiamo, inoltre,

$$P_{aa}^{(m+n)} \geq \alpha\beta P_{bb}^{(0)} = \alpha\beta > 0. \quad (10.46)$$

Siccome a è periodico di periodicità M , dovrà essere $m+n \pmod{M} = 0$: $m+n$ è un multiplo del periodo. Sempre per la periodicità di a , usando la (10.43) abbiamo che

$$P_{bb}^{(\ell+m+n)} \geq \alpha\beta P_{aa}^{(\ell)} > 0, \quad \text{per } \ell \pmod{M} = 0. \quad (10.47)$$

Si evince che, se a è periodico, allora $(\ell + m + n) \pmod{M} = 0$, perché $m + n$ è multiplo di M , e anche b sarà periodico di periodicità M .

Infine, possiamo anche facilmente dimostrare la proprietà ulteriore che

$$P_{ab}^{(\ell+m)} \geq P_{ab}^{(m)} P_{bb}^{(\ell)} = \alpha P_{bb}^{(\ell)}. \quad (10.48)$$

La probabilità di transizione tra stati diversi ha lo stesso comportamento delle probabilità di ritorno nello stesso stato.

QED.

In conclusione, dato che in una catena irriducibile tutti gli eventi hanno la stessa natura, la classificazione degli eventi si riporta alla catena intera, che sarà quindi detta periodica o aperiodica, persistente (nulla o non-nulla) e transiente. In particolare, si dice catena di Markov *ergodica* una catena *persistente non-nulla aperiodica*.

10.6.2 Tutti gli stati di una catena finita sono persistenti non-nulli

Un corollario del teorema precedente implica che in una catena irriducibile *finita* non è possibile che tutti gli stati siano transienti o persistenti nulli. Infatti, la condizione di normalizzazione della probabilità implica che

$$\sum_{b=1}^N P_{ab}^{(n)} = 1, \quad (10.49)$$

dove la somma su b si estende su un numero finito N di stati. Possiamo quindi scambiare il limite $n \rightarrow \infty$ con la sommatoria e ottenere

$$1 = \lim_{n \rightarrow \infty} \sum_{b=1}^N P_{ab}^{(n)} = \sum_{b=1}^N \lim_{n \rightarrow \infty} P_{ab}^{(n)}. \quad (10.50)$$

Come abbiamo visto, per il Teorema (10.41) della probabilità limite delle catene di Markov per stati persistenti non-nulli si ha

$$\lim_{n \rightarrow \infty} P_{ab}^{(n)} = \lim_{n \rightarrow \infty} P_{bb}^{(n)} = \frac{1}{\mu_b} > 0,$$

mentre per stati persistenti nulli, si veda la (10.38), o per stati transienti, si veda la (10.36), ogni termine della somma del lato destro dell'equazione (10.50) è nullo. Siccome devono sommare a 1, tutti gli stati saranno persistenti non-nulli.

Come sottocorollario notiamo che in questo caso le frequenze medie di primo ritorno $\{\mu^{-1}\}$ compongono una distribuzione di probabilità degli stati ben normalizzata. Torneremo a questa proprietà nel Paragrafo 10.7.

10.6.3 Teorema della chiusura di uno stato persistente

Vediamo e dimostriamo un altro teorema, che afferma che per un dato stato persistente p

- esiste un unico insieme chiuso irriducibile \mathcal{C}_p che lo contiene;
- per ogni coppia di stati appartenenti a \mathcal{C}_p il *primo passaggio* in entrambi i versi è un evento certo, $f_{ab} = f_{ba} = 1$.

Se uno stato ignoto i appartiene alla chiusura di p (si veda il Paragrafo 10.4.1), vuol dire che i è accessibile da p . Per definizione esiste almeno un tempo n per il quale la probabilità di primo passaggio da p a i è strettamente positiva: $f_{pi}^{(n)} = \alpha > 0$.

La probabilità di andare da p a i per la prima volta in un numero qualsiasi di passi è $f_{pi} = \sum_{n=1}^{\infty} f_{pi}^{(n)}$. Analogamente, indicheremo con f_{ip} la probabilità di primo passaggio da i a p . La probabilità dell'evento complementare, ovvero di non andare mai da i a p , sarà, quindi, $1 - f_{ip}$.

Con questa terminologia possiamo scrivere la probabilità di lasciare p per non ritornarci mai più come

$$R_{\setminus p} \equiv \sum_i f_{pi}(1 - f_{ip}) = \sum_{n=1}^{\infty} \sum_i f_{pi}^{(n)}(1 - f_{ip}) > \alpha(1 - f_{ip}). \quad (10.51)$$

Però, per ipotesi, p è persistente, ovvero è certo il ritorno in p , e dunque per definizione deve essere $R_{\setminus p} = 0$, il che implica

$$\alpha(1 - f_{ip}) = 0 \rightarrow f_{ip} = 1.$$

Riassumendo, se p è persistente, allora p è accessibile da qualsiasi stato i (di cui ignoriamo il tipo) della sua chiusura. Rammentiamo che nella classificazione del Paragrafo 10.4.4 avevamo visto che uno stato persistente p è definito dalla proprietà $f_{pp} = 1$. Ora vediamo che la natura persistente di p implica anche che $f_{ip} = 1, \forall i$.

In Figura 10.8a, mostriamo un esempio dell'insieme \mathcal{C}_p degli stati accessibili da p . Prendiamo due stati $a, b \in \mathcal{C}_p$. Abbiamo appena visto, e lo disegniamo nella Figura 10.8b, che il solo fatto che p sia persistente implica

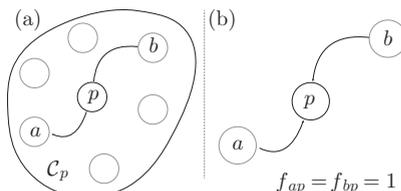


Figura 10.8. (a) Insieme \mathcal{C} degli stati accessibili dallo stato persistente p . (b) Da qualsiasi stato raggiungibile da p è certo il ritorno in p , perché questo è persistente.

■ 10.9 ESEMPI DI CATENE NON DI MARKOV

In generale, la maggior parte dei processi non è descrivibile semplicemente in termini di probabilità di transizione da uno stato all'altro e di probabilità assoluta che il sistema si trovi in uno stato. La probabilità congiunta $p(k(0), k(1), \dots, k(n))$ di una data catena di eventi k nel tempo discreto n non è decomponibile come nella (10.5) nel prodotto di probabilità $P_{k(i)k(i+1)}$ di passare da uno stato all'altro in un passo. Tutti i processi per cui questa semplificazione speciale non è valida vengono denominati genericamente *non-markoviani*.

In diversi casi è possibile "markovianizzare" una catena ampliando lo spazio delle variabili. Una catena non di Markov in una variabile può diventare una catena di Markov in due variabili sotto determinate condizioni. Invertendo il discorso, una catena non di Markov può essere vista come la proiezione su di un sottospazio di una dinamica markoviana in uno spazio più grande.

Prendiamo l'esempio di un cammino aleatorio simmetrico in una dimensione. Come sappiamo è un processo aleatorio di Markov X che prende valori $k = -L, \dots, -1, 0, 1, \dots, L$ sui numeri interi (L anche infinito), la cui matrice stocastica è

$$P_{ik} = \frac{1}{2}\delta_{k,i+1} + \frac{1}{2}\delta_{k,i-1} \tag{10.81}$$

Modifichiamo la dinamica come proposto da van Kampen [68] favorendo la tendenza a proseguire lungo la stessa direzione. Se alla mossa precedente il camminatore è andato in una data direzione, allora avrà una probabilità $p > 1/2$ di continuare in quella direzione e una probabilità $q < 1/2$ di tornare indietro. Dunque la probabilità di una transizione al passo $n + 1$ non dipende più soltanto dalla posizione X al passo n ma anche da quale posizione X proveniva al passo $n - 1$. La probabilità congiunta dell'intera dinamica può decomporsi al più in prodotti di termini del tipo $p(k(n - 1), k(n), k(n + 1))$ e $X(n)$ non è più una catena di Markov.

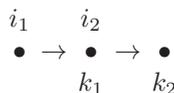
Possiamo, però, rappresentare la stessa dinamica mediante una variabile vettoriale \vec{X} a due componenti

$$\vec{X}(n) \equiv \{X_1(n) = X(n - 1), X_2(n) = X(n)\}.$$

Se guardiamo la transizione di questa variabile bidimensionale, due sono gli eventi che si possono verificare in un passo:

- il sistema si sposta (per il camminatore originale compie due mosse nella stessa direzione), con probabilità p ;
- il sistema rimane fermo (nel sistema originale il camminatore si sposta e torna indietro), con probabilità q .

Il passo della dinamica nella variabile \vec{X} che assume valori $\vec{k} = \{k_1, k_2\}$ può essere illustrato così



e può scriversi mediante la semplice probabilità di transizione in un passo dallo stato \vec{i} allo stato \vec{k}

$$P_{\vec{i}, \vec{k}} = \delta_{k_1, i_2} (p\delta_{k_2 - k_1, i_2 - i_1} + q\delta_{k_2, i_1}), \quad (10.82)$$

che è la matrice stocastica di una catena di Markov per \vec{X} .

Possiamo fare un altro esempio sempre partendo dal camminatore aleatorio unidimensionale X che prende valori sui siti i . Raggruppiamo i siti in scatole di lunghezza L e consideriamo la dinamica tra eventi che consistono nello stare dentro una scatola. In pratica, arrozziamo il processo di cammino aleatorio tra siti in un processo tra scatole (ma sempre con spostamento di un sito alla volta), come in Figura 10.10.

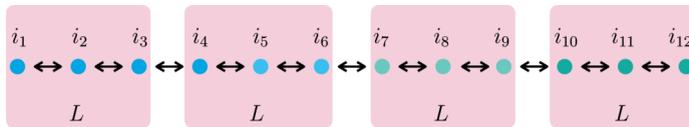


Figura 10.10. Esempio di catena non di Markov. Gli stati i_k sono raggruppati in scatole di L stati l'una e le variabili dipendono dalla scatola e non dallo stato. Eventuali cambi di valore delle variabili non seguono una dinamica di tipo markoviano.

Definiamo una variabile aleatoria “di scatola” Y_L che al passo n prende il valore $k(n) = [i(n)/L]$, dove $[\dots]$ indica la parte intera. Se $L > 1$, non sappiamo esattamente in che sito si trovi il camminatore dentro la scatola, per cui non possiamo valutare la probabilità che rimanga nella scatola o che vada in un'altra scatola soltanto come una semplice probabilità di transizione da una scatola all'altra. Abbiamo bisogno di conoscere quando e come il camminatore è entrato nella scatola di partenza e tutte le possibili dinamiche interne alla scatola (a noi invisibili). La dinamica tra scatole richiede, perciò, la conoscenza di Y_L ai passi precedenti e la catena di eventi non è più di Markov. In questo esempio si mostra esplicitamente come il meccanismo di *coarse graining* che mette insieme posizioni diverse nella scatola porti da una catena di Markov a una non-Markov.

Lo stesso si può dire dell'estrazione da un'urna di palline di tipo diverso, caratterizzato da un dato peso. Se la variabile aleatoria di cui osserviamo la dinamica è il peso totale di tutte le palline estratte, il processo non è markoviano. Lo diventa se consideriamo come variabile aleatoria il vettore i cui elementi rappresentano il peso dei diversi tipi di palline.

In generale, se partiamo da una catena di Markov e riduciamo lo spazio degli eventi, perdendo alcune informazioni, la catena risultante può non essere una catena di Markov; al contrario, partendo da una catena non di Markov questa può diventare equivalente a una catena di Markov se ampliamo lo spazio degli eventi.

10.9.1 Processi non-markoviani continui

Passando al continuo, sia nelle variabili sia nel tempo, come abbiamo fatto per il caso del camminatore aleatorio nei Paragrafi 7.3-7.6, un processo markoviano

soddisfa un'equazione differenziale stocastica del primo ordine, detta *equazione di Langevin*, come per esempio l'equazione (7.48). Nel Paragrafo 7.5 abbiamo visto che questa è una approssimazione per grande viscosità dell'equazione (7.49). Quest'ultima è una equazione del secondo ordine, che riportiamo qui per comodità in una dimensione

$$\ddot{x} + \eta \dot{x} = F(x) + \sqrt{v} \xi(t),$$

con

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t) \xi(t') \rangle = \delta(t - t'),$$

la cui soluzione è un processo non-markoviano. È però possibile anche in questo caso ampliare lo spazio delle variabili definendo una variabile vettoriale (x, v) per cui l'equazione diventa un sistema

$$\begin{cases} v = \dot{x} \\ \dot{v} = -\eta v + F(x) + \sqrt{v} \xi(t) \end{cases}$$

di equazioni differenziali stocastiche (o almeno la seconda) del primo ordine, la cui soluzione $(x(t), v(t))$ è, invece, un processo di Markov. La densità di probabilità $p(x, v, t | x_0, v_0, t_0)$ del processo bidimensionale $(x(t), v(t))$ condizionata alle condizioni iniziali (x_0, v_0) al tempo t_0 del sistema di equazioni è la soluzione dell'equazione di Fokker-Planck del moto browniano in assenza di sovrasmorzamento (si vedano i Paragrafi 7.3 e 7.5).

APPENDICI MATEMATICHE

■ 10.A SCAMBIO DI LIMITE E SOMMA NELL'IPOTESI DI ASSOLUTA CONVERGENZA

Vogliamo ora approfondire l'ultimo passaggio della dimostrazione del Teorema della probabilità limite del Paragrafo 10.5 e mostrare esplicitamente perché questo sia lecito e come entri in gioco l'ipotesi di assoluta convergenza. La prima cosa da fare è giustificare il limite

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n f_{ab}^{(m)} P_{bb}^{(n-m)} = \lim_{n \rightarrow \infty} \sum_{m=1}^{\infty} f_{ab}^{(m)} P_{bb}^{(n-m)}. \quad (10.83)$$

Banalmente si potrebbe mettere a zero ogni $P^{(n < 0)}$ prima di prendere il limite, ma si può vedere che è sufficiente che le $P^{(n < 0)}$ siano probabilità (ovvero ≤ 1) per garantire questo passaggio. Infatti, se aggiungiamo e togliamo gli elementi della serie fino a ∞ , usando la (10.31) possiamo scrivere

$$\left| \lim_{n \rightarrow \infty} P_{ab}^{(n)} - \lim_{n \rightarrow \infty} \sum_{m=1}^{\infty} f_{ab}^{(m)} P_{bb}^{(n-m)} \right| = \lim_{n \rightarrow \infty} \sum_{m=n+1}^{\infty} f_{ab}^{(m)} P_{bb}^{(n-m)} \leq \lim_{n \rightarrow \infty} \Delta(n),$$

dove abbiamo definito

$$\Delta(n) \equiv \sum_{m=n+1}^{\infty} f_{ab}^{(m)}. \quad (10.84)$$

Siccome $\sum_{\ell=1}^{\infty} f_{ab}^{(\ell)} = 1$, $\lim_{n \rightarrow \infty} \Delta(n) = 0$ e vale la (10.83).

Luca Leuzzi, Enzo Marinari, Giorgio Parisi

Calcolo delle probabilità

Un trattatello per principianti volenterosi

Calcolo delle probabilità presenta la probabilità e la statistica in modo originale e profondo, pur restando intuitivo e facilmente accessibile. Anche quando elabora, e lo fa di frequente, argomenti intellettualmente sofisticati che sono spesso alla base di queste discipline, sceglie di appoggiarsi ai molti esempi svolti, a figure significative e ben scelte, e di esporre la teoria in un linguaggio di elegante essenzialità. Il testo copre tutti gli argomenti fondamentali, sempre con approfondimenti mirati a inserire sviluppi delicati ma cruciali in una comprensione di base. Gli strumenti matematici necessari a impadronirsi della teoria sono discussi in modo chiaro e originale, in nove appendici dedicate a questo scopo. Esempi di stime usano questioni e dati attuali e aggiornati: alcune argomentazioni si riferiscono per esempio all'analisi della recente pandemia di Covid-19.

Il capitolo finale *Entropia, probabilità e meccanica statistica* introduce a una dimensione che richiama esplicitamente il mondo della fisica caro agli autori.

«Questo libro è un affascinante distillato degli aspetti più utili della probabilità e della statistica. Ed è doppiamente affascinante. Innanzitutto, perché ciò che gli autori hanno scelto di includere (e di tralasciare) è un ottimo equilibrio tra l'utile e il bello. In secondo luogo, perché il testo è intessuto di intuizioni e commenti brillanti. Folgoranti, poi, le connessioni tra fisica statistica, equazione di Fokker-Planck ed equazione di Schrödinger».

Persi Diaconis
professore di Matematica e Statistica
Stanford University

Luca Leuzzi è primo ricercatore presso l'Istituto di Nanotecnologia del CNR. Dal 2002 è ricercatore del centro Statistical Mechanics and Complexity dell'Istituto Nazionale di Fisica della Materia e presso l'Istituto dei Processi Chimico-Fisici del CNR.

Enzo Marinari è professore ordinario di Fisica Teorica alla Sapienza Università di Roma. Ha svolto la sua ricerca a Parigi, a Roma Tor Vergata, a Cagliari, alla Syracuse University (NY, USA). È stato anche direttore per la fisica di NPAC a Syracuse, USA, e di BIFI a Saragozza, in Spagna.

Giorgio Parisi, nato a Roma nel 1948, si è laureato in Fisica nel 1970; è professore emerito di Fisica Teorica alla Sapienza Università di Roma. Nel 2021 ha vinto il premio Wolf e il premio Nobel per la Fisica.

Le risorse digitali



online.universita.zanichelli.it/leuzzi-probabilita

A questo indirizzo sono disponibili le risorse digitali di complemento al libro.

Per accedere alle risorse protette è necessario registrarsi su **my.zanichelli.it** inserendo il codice di attivazione personale contenuto nel libro.

Libro con ebook



Chi acquista il libro nuovo può accedere gratuitamente all'**ebook**, seguendo le istruzioni presenti nel sito. L'ebook si legge con l'applicazione *Booktab*, che si scarica gratis da App Store (sistemi operativi Apple) o da Google Play (sistemi operativi Android).

L'accesso all'ebook e alle risorse digitali protette è personale, non condivisibile e non cedibile, né autonomamente né con la cessione del libro cartaceo.

LEUZZI*CALCOLO DELLE PROBABIL LUM

ISBN 978-88-08-49966-0



9 788808 499660

4 5 6 7 8 9 0 1 2 (60B)