

Computational Physics Laboratory

Maso*

6 gennaio 2024

Indice

1	Introduzione	2
I	Integrazione numerica elementare	2
2	Formule di Newton–Cotes	3
2.1	Integrazione secondo quadrature numeriche	4
2.1.1	Regola trapezoidale	5
2.1.2	Regola di Simpson	6
3	Quadrature gaussiane	7
3.1	Polinomi ortogonali	7
3.2	Quadrature gaussiane	8
4	Integrazione numerica composta	9
5	Integrali multidimensionali	10
II	Metodi Monte Carlo	10
6	Teorema del limite centrale	10
6.1	Polinomi ortogonali di Hermite	11
6.2	Trasformata di Fourier	11
6.3	Trasformata di una convoluzione	12
6.4	Variabile aleatoria e distribuzione di probabilità	12
6.5	Funzione generatrice	12
6.6	Variabile standardizzata	13
6.7	Proprietà dei cumulanti	14
6.8	Teorema del limite centrale	14
7	Metodi Monte Carlo	16
7.1	Serie di Edgeworth	16
7.2	Metodo Monte Carlo	17
7.2.1	Integrali multidimensionali	18
8	Campionamento di importanza	19
8.1	Metodo del cambio di variabili	19
8.1.1	Cambio multidimensionale	20
9	Conclusioni	22

*<https://github.com/M-a-s-o/notes>

III	Integrale sui cammini in meccanica quantistica	22
10	Integrale sui cammini di Feynman	22
10.1	Propagatore ritardato	23
10.2	Operatore di trasferimento	25
10.2.1	Operatore di trasferimento euclideo	26
11	Funzioni di correlazione	29
11.1	Prodotto ordinato temporale di operatori	30
11.2	Funzioni di correlazione	31
12	Oscillatore armonico	33
12.1	Integrale sui cammini	34
IV	Catene di Markov e teorema ergodico	37
13	Catene di Markov	38
13.1	Matrice di probabilità di transizione	39
13.2	Probabilità di transizione dopo un numero arbitrario di passi	39
13.2.1	Esempio — parte prima	40
13.3	Classificazione degli insiemi di stati	41
13.4	Classificazione degli stati	41
13.4.1	Esempio — parte seconda	45
13.5	Distribuzioni all'equilibrio	46
13.6	Distribuzione invariante per un insieme di stati ergodici	46
13.7	Riassunto sulle catene	48
13.8	Bilancio dettagliato	49
14	Algoritmo di Metropolis	49
15	Riassunto delle formule teoriche	51
15.1	Monte Carlo standard per l'oscillatore armonico	51
15.2	Autocorrelazione	52

Lezione 1

ven 06 ott
2023 14:30

Esame. Orale di teoria, discussione relazione (circa 10, 15 ore per scrivere cosa fatto durante tutto l'anno) e correttezza programmi. Orale prevede interrogazione su teoremi e dimostrazioni. Simulazione QCD si analizzano dati (si può fare in python), ma il programma è dato dal prof.

1 Introduzione

L'utilizzo della computazione per la fisica teorica permette di risolvere problemi prima inattaccabili. Ad esempio, la cromodinamica quantistica (QCD) è una teoria non perturbativa alle basse energie: calcolare le masse di varie particelle non si può fare analiticamente, ma bisogna affidarsi ad un calcolatore. Inoltre, la maggioranza dei problemi interessanti della fisica vengono risolti numericamente perché le equazioni di tali problemi sono così complesse da rendere difficile trovare una soluzione analitica. Il corso tratta i metodi per calcolare integrali sui cammini.

I metodi numerici permettono di formulare teoremi proprio come con i metodi analitici. L'analisi numerica studia i metodi numerici sviluppandoli in modo da sapere l'errore associato ad un metodo ed il significato di tale errore. Una differenza con i metodi analitici è data dai parametri: nei metodi numerici non si ottengono funzioni di parametri, ma si ottengono numeri perché bisogna specificare il valore dei parametri.

Parte I

Integrazione numerica elementare

Si studiano i metodi di integrazione deterministici.

2 Formule di Newton–Cotes

Si consideri una funzione di una variabile

$$y = f(x), \quad x, y \in \mathbb{R}$$

Definizione 2.1. Sia P_n l'insieme di tutti i polinomi di grado minore o uguale a n .

Definizione 2.2. Dati dei punti x_0, \dots, x_n distinti in un intervallo $[a, b]$, un polinomio $p(x) \in P_n$ interpola $f(x)$ in ognuno dei punti se

$$p(x_i) = f(x_i), \quad i = 0, \dots, n$$

Definizione 2.3. Dati dei punti x_0, \dots, x_n distinti, il j -esimo polinomio di Lagrange (detto anche funzione cardinale) di grado n è

$$l_j(x) = \prod_{i \neq j}^n \frac{x - x_i}{x_j - x_i}$$

[r] mettere $i = 0$ sotto o sopra $i \neq j$.

Proposizione 2.4. Questi polinomi hanno la seguente proprietà:

$$l_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

I polinomi di Lagrange formano una base dell'insieme P_n : i polinomi sono non nulli solamente nei punti in cui $i = j$.

Teorema 2.5. Dati dei punti x_0, \dots, x_n distinti in un intervallo $[a, b]$ e dato un insieme y_0, \dots, y_n di numeri reali, esiste ed è unico il polinomio $p(x) \in P_n$ tale che

$$p(x_j) = y_j, \quad j = 0, \dots, n$$

Dimostrazione. L'unicità è data dal teorema fondamentale dell'algebra. Si dimostra l'esistenza. Si consideri

$$p(x) = \sum_{i=0}^n y_i l_i(x)$$

Il suo valore in x_j risulta essere

$$p(x_j) = \sum_{i=0}^n y_i l_i(x_j) = \sum_{i=0}^n y_i \delta_{ij} = y_j$$

□

Osservazione 2.6. L'unicità deriva dal fatto che i polinomi di Lagrange l_i sono una base dello spazio vettoriale P_n .

Osservazione 2.7. Se una funzione $f(x)$ è tale per cui $f(x_i) = y_i$, allora $p(x) \in P_n$ è il polinomio interpolante di tale funzione $f(x)$ nei punti x_i . Pertanto

$$p(x) = \sum_{i=0}^n f(x_i) l_i(x)$$

Teorema 2.8. Data una funzione $f \in C^{n+1}[a, b]$ e dato il polinomio $p \in P_n$ che interpola la funzione $f(x)$ in $x_0, \dots, x_n \in [a, b]$, in ogni punto dell'intervallo vale

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi(x)) \prod_{i=0}^n (x - x_i)$$

Ad ogni punto $x \in [a, b]$ corrisponde un punto $\xi \in (a, b)$ tale che la formula è versa. La scrittura $f^{(r)}$ indica la derivata r -esima.

Osservazione 2.9. Il polinomio $p(x)$ è generico (non è detto nella base di Lagrange) e x_0, \dots, x_i non devono essere distinti come nell'altro caso. La dimostrazione seguente non assume mai che i punti siano distinti. [r]

Dimostrazione. Si consideri una funzione

$$g(z) = f(z) - p(z) - Q(z) \frac{f(x) - p(x)}{Q(x)}, \quad Q(x) \equiv \prod_{i=0}^n (x - x_i)$$

dove x è da intendersi come parametro. Essa ha $n+2$ zeri nell'intervallo $[a, b]$: i primi $n+1$ dati quando $z = x_i$ e l'ultimo per $z = x$. Per il teorema di Rolle, la derivata $g'(z)$ ha $n+1$ zeri. Similmente, la derivata $g^{(n+1)}(z)$ ha almeno uno zero in un punto $\xi \in (a, b)$. Dunque

$$\begin{aligned} g^{(n+1)}(\xi) = 0 &\implies 0 = f^{(n+1)}(\xi(x)) - (n+1)! \frac{f(x) - p(x)}{Q(x)} \\ &\implies f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi(x)) \prod_{i=0}^n (x - x_i) \end{aligned}$$

La funzione g introdotta contiene la funzione resto cioè la frazione. □

Osservazione 2.10. Si noti che la dimostrazione non fa uso dell'ipotesi di punti x_j distinti.

Osservazione 2.11. In questo modo si può porre un limite superiore all'errore derivante dall'interpolazione polinomiale: il limite è dato dal massimo della derivata moltiplicata per i coefficienti della formula sopra.

2.1 Integrazione secondo quadrature numeriche

Una volta imparato ad approssimare le funzioni, si passa ad approssimare gli integrali.

Definizione 2.12 (quadratura numerica). Si consideri una funzione reale, l'integrale in un intervallo si può stimare come

$$\int_a^b f(x) dx \approx \sum_{i=0}^n w_i f(x_i)$$

Proposizione 2.13. Si vuole calcolare l'integrale mediante il calcolo di un numero finito di volte della funzione f . [r]

Proposizione 2.14. Si possono scegliere i pesi w_i ed i punti x_i in modo da avere la migliore approssimazione possibile.

Osservazione 2.15. Un metodo naturale è utilizzare la formula di Lagrange di interpolazione polinomiale (Theorem 2.8):

$$\int_a^b f(x) dx = \int_a^b p(x) dx + \frac{1}{(n+1)!} \int_a^b f^{(n+1)}(\xi(x)) \prod_{i=0}^n (x - x_i) dx$$

Il primo integrale è l'approssimazione tramite quadrature numeriche, il secondo addendo è l'errore associato. Poiché non si conosce ξ per ogni punto x , bisogna utilizzare i metodi numerici per il calcolo dell'integrale.

Definizione 2.16. Si definiscono i pesi della quadratura come

$$w_i \equiv \int_a^b l_i^n(x) dx$$

L'indice n specifica il grado del polinomio l_i di Lagrange e non è da intendersi come potenza. Segue

$$\int_a^b f(x) dx = \sum_{i=0}^n w_i f(x_i) + \frac{1}{(n+1)!} \int_a^b f^{(n+1)}(\xi(x)) \prod_{i=0}^n (x - x_i) dx$$

Definizione 2.17. Rimane ancora un grado di libertà dato dalla scelta dei punti x_i . Esistono vari metodi più o meno complessi per la scelta di tali punti (ad esempio i nodi di Chebyshev). Il metodo più semplice è considerare dei sotto-intervalli: [r]

$$x_i = a + ih, \quad h \equiv \frac{b-a}{n}, \quad i = 0, \dots, n$$

Proposizione 2.18. Utilizzare questo metodo permette di calcolare i pesi w_i indipendentemente dall'intervallo. Infatti

$$z \equiv \frac{x-a}{h} \implies w_j = h \int_0^n l_j(a+zh) dz = h \int_0^n \prod_{i \neq j} \frac{z-i}{j-i} dz$$

La forma dei pesi è data da

$$w_j = h \gamma_j, \quad \gamma_j = \int_0^n \prod_{i \neq j} \frac{z-i}{j-i} dz$$

[r] specificare limiti della produttoria, da $i = 0$ a n ?

Si noti che la somma dei γ_i è n , per vederlo si può calcolare $\int_a^b dx$

2.1.1 Regola trapezoidale

Si definisce un metodo a intervalli costanti con $n = 1$ che viene detto trapezoidale perché si interpola la funzione con una retta: l'area di integrazione è un trapezio. Si consideri un intervallo $[a, b]$ e siano gli estremi i punti in cui la funzione $f(x)$ coincide con il polinomio $p(x)$:

$$x_0 = a, \quad x_1 = b, \quad h = b - a$$

I pesi sono proporzionali a

$$\gamma_0 = \int_0^1 \frac{z-1}{0-1} dz = \frac{1}{2}, \quad \gamma_1 = \int_0^1 \frac{z}{1-0} dz = \frac{1}{2}$$

Pertanto

$$\int_a^b f(x) dx = \frac{h}{2} [f(a) + f(b)] + E_1(f)$$

dove l'errore è dato da

$$E_1(f) = \frac{1}{2} \int_a^b f''(\xi(x))(x-a)(x-b) dx$$

Il prodotto degli ultimi due fattori è sempre non positivo.

Teorema 2.19 (media pesata). Si consideri una funzione $f \in C^2[a, b]$. Se g è integrabile su (a, b) [r] chiuso o aperto?, e ivi non negativa (o non positiva), allora esiste un valore $c \in (a, b)$ tale per cui

$$f(c) = \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx}$$

Dimostrazione. Vedere appunti di Analisi I. [r] Mettere dim?

□

Applicando il teorema, si ottiene un errore pari a

$$E_1(f) = \frac{1}{2} f''(\xi) \int_a^b (x-a)(x-b) dx, \quad \xi \in (a, b)$$

Usando il precedente cambio di variabili

$$z = \frac{x-a}{h}, \quad dx = h dz$$

si ha

$$E_1(f) = \frac{h^3}{2} f''(\xi) \int_0^1 z(z-1) dz = -\frac{h^3}{12} f''(\xi)$$

Quindi il risultato delle formule per la regola trapezoidale è

$$\int_a^b f(x) dx = \frac{h}{2} [f(a) + f(b)] - \frac{h^3}{12} f''(\xi), \quad h = b-a, \quad \xi \in (a, b)$$

Il secondo addendo non si sa calcolare, tuttavia sapendo che ξ fa parte dell'intervallo, si stima l'errore come il valore massimo di tale addendo.

Osservazione 2.20. La regola trapezoidale è esatta per polinomi di grado minore o uguale di 1.

Definizione 2.21. Si dice che la formula è esatta all'ordine 1.

Nelle integrazioni numeriche si riformula il problema in modo che sia più facilmente risolvibile dal calcolatore.

2.1.2 Regola di Simpson

Si consideri l'intervallo $[a, b]$ e tre punti equispaziati:

$$n = 2, \quad h = \frac{b-a}{2}, \quad x_0 = a, \quad x_1 = \frac{a+b}{2}, \quad x_2 = b$$

I pesi sono proporzionali a

$$\begin{aligned} \gamma_0 &= \int_0^2 \frac{(z-1)(z-2)}{(0-1)(0-2)} dz = \frac{1}{3}, \\ \gamma_1 &= \int_0^2 \frac{(z-0)(z-2)}{(1-0)(1-2)} dz = \frac{4}{3}, \\ \gamma_2 &= \int_0^2 \frac{(z-0)(z-1)}{(2-0)(2-1)} dz = \frac{1}{3} \end{aligned}$$

Pertanto

$$\int_a^b f(x) dx = \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] + E_2(f)$$

L'errore è dato da

$$E_2(f) = \frac{1}{6} \int_a^b f^{(3)}(\xi(x))(x-a)(x-\frac{a+b}{2})(x-b) dx$$

Non si può più applicare il teorema della media pesata poiché un fattore cambia segno al centro dell'intervallo.

Proposizione 2.22. Il prodotto degli ultimi tre fattori

$$Q_2(x) \equiv (x-a)(x-\frac{a+b}{2})(x-b)$$

risulta essere una funzione antisimmetrica rispetto al centro dell'intervallo $[a, b]$:

$$\int_a^b Q_2(x) dx = 0$$

Aggiungendo il polinomio $Q_2(x)$ alla base $l_i(x)$ interpolante (cioè al polinomio $p(x)$, la base non è ortogonale), il suo contributo all'integrale della funzione $f(x)$ è nullo (cfr. Definizione 2.16). In questo modo si considerano anche i polinomi di terzo grado. Questo equivale ad aggiungere un punto interpolante arbitrario. Tale punto non cambia la formula.

Aggiungendo il punto

$$x_3 = x - \frac{a+b}{2} = x_1$$

e ripetendo il procedimento sopra per quattro punti, si ha

$$E_2(f) = E_3(f) = \frac{1}{24} f^{(4)}(\xi) \int_a^b (x-a)(x-\frac{a+b}{2})^2(x-b) dx$$

dove si è applicato il teorema della media pesata poiché il prodotto nell'integrale è negativo. Tramite il cambio di variabile

$$z = \frac{x-a}{h} \implies \int_a^b (x-a)(x-\frac{a+b}{2})^2(x-b) dx = h^5 \int_0^2 z(z-1)^2(z-2) dz = -\frac{4}{15} h^5$$

si ottiene

$$E_2(f) = E_3(f) = -\frac{h^5}{90} f^{(4)}(\xi)$$

Nonostante la regola di Simpson si calcola per tre punti, l'errore è uguale a utilizzare l'ordine successivo con quattro punti. Questo fenomeno si ripete spesso. A questo punto, l'integrale della funzione è dato da

$$\int_a^b f(x) dx = \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{90} f^{(4)}(\xi)$$

Questa formula è esatta fino al terzo ordine perché la derivata quarta è nulla. Come per il metodo precedente, l'errore tra il valore vero e quello calcolato è al più pari al massimo del secondo addendo della formula sopra.

Osservazione 2.23. Si vuole una grandezza del passo h piccola: di solito, utilizzare ordini maggiori permette di ridurre l'errore. Questo non vale in tutti i casi, ma dipende dalla funzione $f(x)$ e dalle sue derivate. Se si ha un intervallo $[a, b]$ grande, basta dividerlo in sotto-intervalli di modo che h diventi piccolo.

Osservazione 2.24. Rispetto al metodo dei trapezi, questo metodo utilizza più risorse perché valuta la funzione tre volte.

Lezione 2

3 Quadrature gaussiane

ven 13 set
2023 14:30

Oltre a scegliere i pesi w_i per diminuire l'errore, si possono scegliere anche i punti x_i a tal fine.

3.1 Polinomi ortogonali

Vedere Abramowitz.

Definizione 3.1. Dato un intervallo $[a, b]$ e ivi una funzione peso $w(x) \geq 0$, un insieme di polinomi p_n è ortogonale rispetto alla funzione peso (detta anche funzione di misura) se

$$\int_a^b w(x) p_n(x) p_m(x) dx = c_m \delta_{mn}$$

Risulta possibile normalizzare ogni polinomio in modo da ottenere una base ortonormale.

Esempio 3.2. Esempi di sistemi di polinomi ortogonali sono:

Nome	$p_n(x)$	a	b	$w(x)$
Legendre	$P_n(x)$	-1	1	1
Laguerre	$L_n(x)$	0	$+\infty$	e^{-x}
Hermite	$He_n(x)$	$-\infty$	$+\infty$	$e^{-\frac{x^2}{2}}$
Chebyshev	$T_n(x)$	-1	1	$(1-x^2)^{-\frac{1}{2}}$
Jacobi	$P_n^{(\alpha, \beta)}$	-1	1	$(1-x)^\alpha (1+x)^\beta$

con $\alpha, \beta > -1$. I polinomi di Hermite si utilizzano nei metodi Monte Carlo.

Teorema 3.3. Un polinomio ortogonale di grado n ha n zeri (radici) reali nell'intervallo $[a, b]$.

3.2 Quadrature gaussiane

Si scelgono i punti x_i pari alle radici di polinomi ortogonali in modo da ridurre l'errore. Si vuole approssimare numericamente l'integrale

$$\int_a^b w(x)f(x) dx$$

dove $w(x) \geq 0$ in $[a, b]$.

Osservazione 3.4. L'uso dei pesi $w(x)$ altrove permette di rimuovere eventuali complessità.

Teorema 3.5. Date le radici x_0, \dots, x_{n-1} del polinomio ortogonale di ordine n associato alla misura $w(x)$ in $[a, b]$, se $p(x)$ è un polinomio di grado minore di $2n$ allora

$$\int_a^b w(x)p(x) dx = \sum_{i=0}^{n-1} w_i p(x_i)$$

dove

$$w_i \equiv \int_a^b w(x) l_i^{n-1}(x) dx$$

con l_i^{n-1} polinomio i -esimo di Lagrange di grado $n-1$.

Dimostrazione. Si dimostra prima che l'integrazione è esatta per polinomi $p(x)$ di grado minore di n . Tale polinomio coincide con il proprio polinomio interpolante e per la Osservazione 2.7 con $f(x) \equiv p(x)$ si ottiene

$$p(x) = \sum_{i=0}^{n-1} p(x_i) l_i^{n-1}(x)$$

Moltiplicando per $w(x)$ e integrando, segue

$$\int_a^b w(x)p(x) dx = \sum_{i=0}^{n-1} w_i p(x_i), \quad w_i = \int_a^b w(x) l_i^{n-1}(x) dx$$

Per la Definizione 2.16, l'integrale è esatto. Si dimostra che vale anche per un polinomio $p(x)$ di grado minore di $2n$. Tale polinomio si può sempre scrivere come

$$p(x) = Q(x)p_n(x) + R(x)$$

dove $Q(x)$ e $R(x)$ sono polinomi di grado minore di n . Insieme, questi due polinomi hanno al più $[r] 2n$ gradi di libertà, cioè i gradi di libertà di un polinomio di grado massimo $2n-1$. Pertanto

$$\int_a^b w(x)p(x) dx = \int_a^b w(x)Q(x)p_n(x) dx + \int_a^b w(x)R(x) dx$$

Poiché $Q(x)$ ha grado minore di n , si può riscrivere in termini dei polinomi ortogonali $p_n(x)$ fino al grado $n-1$. Per ortogonalità con $p_n(x)$ si ha

$$\int_a^b w(x)Q(x)p_n(x) dx = 0 \implies \int_a^b w(x)p(x) dx = \int_a^b w(x)R(x) dx$$

Nei punti x_0, \dots, x_{n-1} , cioè gli zeri del polinomio $p_n(x)$, si ha $p(x_i) = R(x_i)$. Utilizzando la relazione precedente e applicando quanto dimostrato sopra per polinomi di grado massimo $n-1$, si ha

$$\int_a^b w(x)p(x) dx = \int_a^b w(x)R(x) dx = \sum_{i=0}^{n-1} w_i R(x_i) = \sum_{i=0}^{n-1} w_i p(x_i)$$

□

Osservazione 3.6. Rispetto ai metodi precedenti (cfr. Definizione 2.16), la formula sopra è esatta, senza errore fino a polinomi con grado $2n-1$, sebbene si utilizzino solo n punti interpolanti. Tuttavia, bisogna calcolare gli zeri dei polinomi ortogonali $p_n(x)$ di grado n .

Osservazione 3.7. Per una funzione generica si ha

$$\int_a^b w(x)f(x) dx = \sum_{i=0}^{n-1} w_i f(x_i) + E_n(f)$$

con

$$E_n(f) = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b p_n^2(x) dx, \quad p_n(x) = \prod_{i=0}^{n-1} (x - x_i)$$

Per integrare una funzione generica si sceglie $w(x) = 1$ e si utilizzano i polinomi di Legendre.

Dimostrazione. Generalizzando quanto visto per la regola di Simpson, ogni volta che l'approssimazione vale per l'ordine successivo, si può scegliere un altro punto e renderlo coincidente con un punto già presente. In questo modo, si passa dal grado massimo $n-1$ del polinomio integrabile a $2n-1$ duplicando tutti gli n punti. Pertanto si ha

$$E_n(f) = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \prod_{i=0}^{n-1} (x - x_i)^2 dx$$

□

4 Integrazione numerica composta

L'integrazione numerica composta viene anche detta integrazione tramite formule estese. Per ottenere un passo h piccolo, si può dividere un grande intervallo in m sotto-intervalli e applicare a ciascuno un metodo visto in precedenza. Si pone

$$a_i = a + i \frac{b-a}{m}$$

e si riscrive

$$\int_a^b f(x) dx = \sum_{i=0}^{m-1} \int_{a_i}^{a_{i+1}} f(x) dx$$

Formula estesa trapezoidale. Si valuta la funzione $m+1$ volte

$$\int_a^b f(x) dx = h \left[\frac{1}{2} f(a_0) + \sum_{i=1}^{m-1} f(a_i) + \frac{1}{2} f(a_m) \right] - \frac{h^3}{12} \sum_{i=1}^m f''(\xi_i), \quad h = \frac{b-a}{m}$$

L'assenza del fattore $\frac{1}{2}$ dagli estremi dei sotto-intervalli interni deriva dal fatto che ogni estremo appartiene a due sotto-intervalli. Per calcolare l'errore si può prendere il massimo della derivata seconda in ogni intervallo, oppure, poiché interessa l'ordine di grandezza dell'errore, si può utilizzare la seguente formula

$$E = -\frac{h^3}{12} m f''(\xi) = -\frac{(b-a)^3}{12m^2} f''(\xi) \sim m^{-2}$$

Da questa si nota che l'errore si riduce secondo m^2 .

Formula estesa di Simpson. Si hanno m intervalli e si calcola la funzione $2m+1$ volte

$$\int_a^b f(x) dx = \frac{h}{3} \left[f(x_0) + 4 \sum_{i=1}^m f(x_{2i-1}) + 2 \sum_{i=1}^{m-1} f(x_{2i}) + f(x_{2m}) \right] - \frac{h^5}{90} \sum_{i=1}^{2m} f^{(4)}(\xi_i), \quad h = \frac{b-a}{2m}$$

Come sopra, l'errore può essere stimato come

$$E = -\frac{h^5}{90} m f^{(4)}(\xi) = -\frac{1}{2880} \frac{(b-a)^5}{m^4} f^{(4)}(\xi) \sim m^{-4}$$

5 Integrali multidimensionali

I metodi visti finora valgono per una variabile. Per simulare una teoria quantistica relativistica, in generale una teoria di campo con integrali sui cammini, si ha bisogno di gestire miliardi di variabili reali, cioè le componenti di un campo in vari punti. La dimensione dell'integrale in una teoria dei campi tende all'infinito. Per definire una teoria dei campi in modo non perturbativo si discretizza lo spazio-tempo quadridimensionale su di un reticolo con passo a . Per simulare una teoria di campo, ad esempio $\lambda\phi^4$, si considera un volume finito ed un passo finito, da cui si ottiene un reticolo finito; ponendo un campo scalare in ogni punto, si ha un numero finito di gradi di libertà. Utilizzando reticoli da 128 punti in ogni dimensione spazio-temporale e considerando ogni componente del campo, il numero di gradi di libertà, e quindi di variabili da utilizzare, risulta essere importante. Della simulazione si vuole studiare il limite per volume infinito e passo reticolare infinitesimo: il numero di punti cresce.

Si consideri un integrale in d variabili, cioè in d gradi di libertà,

$$I = \int dx_1 \cdots dx_d f(x_1, \dots, x_d)$$

Si calcola il suo valore applicando ripetutamente una delle regole precedenti. Ad esempio, per la regola trapezoidale estesa si ha

$$I = \sum_{j_1=0}^n \cdots \sum_{j_d=0}^n w_{j_1} \cdots w_{j_d} f\left(\frac{j_1}{n}, \dots, \frac{j_d}{n}\right) + O(n^{-2})$$

La funzione viene valutata $N = (n+1)^d$ volte. L'errore è dato da

$$E \sim \frac{1}{n^2} \sim \frac{1}{N^{\frac{2}{d}}}$$

ma per un numero enorme di dimensioni d dell'integrale, l'errore è confrontabile con il valore numerico dell'integrale stesso. I metodi visti non sono adatti ad integrare molte variabili. L'errore nel metodo di Simpson va come

$$E \sim \frac{1}{N^{\frac{4}{d}}}$$

Parte II

Metodi Monte Carlo

Il primo metodo Monte Carlo fu l'ago di Buffon. Successivamente ci fu l'algoritmo di Metropolis. Per una dimensionalità d molto grande, i metodi sopra sono inapplicabili. Il metodo Monte Carlo permette di avere un errore che va come

$$E \sim \frac{1}{N^{\frac{1}{2}}}$$

indipendente dalla dimensionalità dell'integrale. Tuttavia, si vede poi che anch'esso fallisce con il path integral e bisogna aggiungere le catene di Markov.

6 Teorema del limite centrale

Si studiano le fondamenta su cui si basa il metodo Monte Carlo. Si vedono alcuni prerequisiti teorici.

6.1 Polinomi ortogonali di Hermite

Un polinomio di Hermite di grado n è indicato come $He_n(x)$ ed ha peso $w(x) = e^{-\frac{x^2}{2}}$. L'intervallo di integrazione è l'asse reale. I polinomi sono ortogonali rispetto al peso

$$\int_{\mathbb{R}} dx e^{-\frac{x^2}{2}} He_n(x) He_m(x) = \sqrt{2\pi} n! \delta_{nm} \quad n, m \in \mathbb{N}_0$$

I polinomi si ottengono dalla formula di Rodrigues:

$$d_x^n e^{-\frac{x^2}{2}} = (-1)^n He_n(x) e^{-\frac{x^2}{2}}, \quad n \in \mathbb{N}_0$$

I primi polinomi sono

$$\begin{aligned} He_0 &= 1 \\ He_1 &= x \\ He_2 &= x^2 - 1 \\ He_3 &= x^3 - 3x \\ He_4 &= x^4 - 6x^2 + 3 \\ He_5 &= x^5 - 10x^3 + 15x \\ He_6 &= x^6 - 15x^4 + 45x^2 - 15 \end{aligned}$$

I polinomi di Hermite sono utili poiché la loro trasformata di Fourier, insieme alla misura, è un monomio.

6.2 Trasformata di Fourier

Definizione 6.1. La trasformata di Fourier è definita come

$$f(x) = \int_{\mathbb{R}} \frac{dk}{2\pi} \tilde{f}(k) e^{-ikx}, \quad \tilde{f}(k) = \int_{\mathbb{R}} dx f(x) e^{ikx}$$

Osservazione 6.2. La trasformata di Fourier è in corrispondenza biunivoca con la funzione che viene trasformata (sempre che le funzioni siano regolari). Ad ogni funzione corrisponde una sola trasformata e ad ogni trasformata corrisponde una sola funzione.

Definizione 6.3. La funzione delta di Dirac è definita come

$$\delta(x) = \int \frac{dk}{2\pi} e^{-ikx}, \quad \int \delta(x) dx = 1$$

Trasformata di una gaussiana. Si utilizza ampiamente la trasformata di una gaussiana. In un certo senso, essa corrisponde al principio di indeterminazione in meccanica quantistica non relativistica. Si consideri la funzione gaussiana normalizzata

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

L'integrale ed il secondo momento sono

$$\int_{\mathbb{R}} P(x) dx = 1, \quad \int_{\mathbb{R}} x^2 P(x) dx = \sigma^2$$

La sua trasformata è data da

$$\int_{\mathbb{R}} P(x) e^{ikx} dx = e^{-\frac{k^2\sigma^2}{2}}$$

Una gaussiana localizzata nello spazio delle posizioni, corrisponde ad una gaussiana estesa nello spazio dei momenti e viceversa. La varianza passa da denominatore a numeratore dell'esponente.

Trasformata di Fourier dei polinomi di Hermite. La trasformata di Fourier considerando anche la misura è

$$\widetilde{He}_n(k) = \sqrt{2\pi}(ik)^n e^{-\frac{k^2}{2}}$$

Dimostrazione. Segue

$$\begin{aligned}\widetilde{He}_n(k) &= \int_{\mathbb{R}} dx e^{-\frac{x^2}{2}} He_n(x) e^{ikx} = (-1)^n \int_{\mathbb{R}} \left[d_x^n e^{-\frac{x^2}{2}} \right] e^{ikx} dx \\ &= \int_{\mathbb{R}} e^{-\frac{x^2}{2}} [d_x^n e^{ikx}] dx = \int_{\mathbb{R}} e^{-\frac{x^2}{2}} (ik)^n e^{ikx} dx \\ &= \sqrt{2\pi}(ik)^n e^{-\frac{k^2}{2}}\end{aligned}$$

Alla prima riga si utilizza la formula di Rodrigues prima moltiplicando per $(-1)^{2n}$. Alla seconda riga si è integrato per parti n volte. All'ultima riga si è utilizzata la trasformata di Fourier della gaussiana con $\sigma = 1$. \square

6.3 Trasformata di una convoluzione

Definizione 6.4. La convoluzione di due funzioni è definita come

$$g(x) \equiv [f_1 * f_2](x) = \int_{\mathbb{R}} f_1(y) f_2(x - y) dy$$

Proposizione 6.5. La trasformata di Fourier di una convoluzione è data da

$$\begin{aligned}\widetilde{g}(k) &= \int_{\mathbb{R}} dx e^{ikx} \int_{\mathbb{R}} f_1(y) f_2(x - y) dy = \int_{\mathbb{R}^2} f_1(y) e^{iky} f_2(x - y) e^{ik(x-y)} dx dy, \quad z \equiv x - y \\ &= \int_{\mathbb{R}^2} f_1(y) e^{iky} f_2(z) e^{ikz} dy dz = \widetilde{f}_1(k) \widetilde{f}_2(k)\end{aligned}$$

La trasformata di Fourier di un prodotto di convoluzione è il prodotto delle trasformate di Fourier.

6.4 Variabile aleatoria e distribuzione di probabilità

Si utilizzano ampiamente le variabili aleatorie. Si indica il valore x della variabile aleatoria \hat{x} rimuovendo il cappuccio. Una variabile aleatoria \hat{x} è una variabile reale che assume i valori x a cui è associato un valore della distribuzione di probabilità $P(x)$. Da queste definizioni segue

$$-\infty < x < +\infty, \quad \int_{\mathbb{R}} P(x) dx = 1, \quad P(x) \geq 0$$

La funzione $P(x)$ si può interpretare come la densità di probabilità che la variabile aleatoria \hat{x} abbia valore x .

Osservazione 6.6. Di seguito si presume che i momenti della distribuzione esistano

$$\langle x^n \rangle \equiv \int_{\mathbb{R}} P(x) x^n dx$$

6.5 Funzione generatrice

Si consideri una variabile aleatoria \hat{x} ed una distribuzione di probabilità $P(x)$. Si vuole trasformata di Fourier della distribuzione $P(x)$ ed il logaritmo di tale trasformazione.

La funzione generatrice $F(k)$ dei cumulanti di una distribuzione è data da

$$e^{-F(k)} = z(k) \equiv \int_{\mathbb{R}} P(x) e^{ikx} dx$$

In teoria dei campi $F(k)$ è la free-energy, mentre $z(k)$ è la funzione di partizione. Sviluppando la funzione generatrice intorno a $k = 0$, si ottiene

$$F(k) = - \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} c_n, \quad c_n = -\frac{1}{i^n} d_k^n F(0)$$

dove c_n sono i cumulanti della distribuzione. Si noti che $c_0 = 0$ poiché $P(x)$ è normalizzata, per questo si parte da $n = 1$.

Notando $F(k) = -\ln z(k)$, le espressioni dei primi cumulanti sono

$$\begin{aligned} c_1 &= \langle x \rangle \\ c_2 &= \langle x^2 \rangle - \langle x \rangle^2 \\ c_3 &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 \\ c_4 &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle - 3\langle x^2 \rangle^2 + 12\langle x^2 \rangle \langle x \rangle^2 - 6\langle x \rangle^4 \end{aligned}$$

Proposizione 6.7. Sotto alcune ipotesi di regolarità della distribuzione di probabilità, i cumulanti sono unicamente definiti dalla funzione di distribuzione di probabilità. Vale anche il viceversa: dati tutti i cumulanti, la distribuzione di probabilità è univocamente determinata.

Come funzioni con la stessa trasformata di Fourier sono la stessa funzione, così distribuzioni con gli stessi cumulanti sono la stessa distribuzione. I cumulanti caratterizzano le distribuzioni.

Esempio 6.8. Si vede l'esempio di una gaussiana centrata in x_0 :

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right]$$

La sua trasformata è

$$e^{-F(k)} = \int_{\mathbb{R}} \frac{e^{-\frac{(x-x_0)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} e^{ikx} dx = e^{ikx_0} \int_{\mathbb{R}} \frac{e^{-\frac{(x-x_0)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} e^{ik(x-x_0)} dx = e^{ikx_0} e^{-\frac{k^2\sigma^2}{2}}$$

da cui si ottiene

$$F(k) = -ikx_0 + \frac{k^2\sigma^2}{2}$$

I cumulanti della gaussiana sono tutti nulli tranne i primi due pari al valor medio e la varianza:

$$c_1 = x_0, \quad c_2 = \sigma^2$$

Osservazione 6.9. Per altre distribuzioni si definiscono la skewness (indice di asimmetria, obliquità) e la kurtosis (curtosi, gobba) come

$$\frac{c_3}{c_2^{3/2}}, \quad \frac{c_4}{c_2^2}$$

La skewness descrive il grado di asimmetria di una distribuzione rispetto al proprio valor medio. La kurtosis descrive quanto la distribuzione sia piccata rispetto ad una gaussiana con stessa varianza.

6.6 Variabile standardizzata

Definizione 6.10. La variabile standardizzata di una distribuzione è una variabile che non dipende dalle proprietà principali della distribuzione. Per una distribuzione di valor medio x_0 e varianza σ^2 , la variabile standardizzata è definita come

$$\hat{z} = \frac{\hat{x} - x_0}{\sigma}, \quad z = \frac{x - x_0}{\sigma}, \quad \sigma = \sqrt{\sigma^2} = \sqrt{c_2}$$

La variabile standardizzata ha una propria distribuzione

$$P_s(z) = \int_{\mathbb{R}} P(x) \delta\left(\frac{x-x_0}{\sigma} - z\right) dx$$

La densità di probabilità corrispondente ad un particolare z è data dalla somma di densità di probabilità quando il valore di $\frac{z-x_0}{\sigma}$ corrisponde al valore di cui si vuole conoscere la densità. Per alcune distribuzioni la corrispondenza è biunivoca, come per la gaussiana:

$$P_s(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

ma altre distribuzioni possono avere più contributi.

Funzionale generatore. Il funzionale generatore per la variabile standardizzata di una gaussiana è una parabola:

$$e^{-\Phi(k)} = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} e^{ikz} dz = e^{-\frac{k^2}{2}} \implies \Phi(k) = \frac{1}{2}k^2$$

6.7 Proprietà dei cumulanti

Si vede la proprietà di cumulo dei cumulanti. Date due variabili aleatorie

$$\hat{x}_1, x_1, P_1(x_1); \quad \hat{x}_2, x_2, P_2(x_2)$$

si vuole ricavare la distribuzione della variabile somma

$$\hat{y} = \hat{x}_1 + \hat{x}_2, \quad y = x_1 + x_2, \quad P(y) = \int_{\mathbb{R}^2} dx_1 dx_2 P_1(x_1) P_2(x_2) \delta(y - x_1 - x_2)$$

Si considera la somma perché essa compare quando si calcola il valor medio di tante misure in un esperimento. Pertanto, riscrivendo la funzione delta come trasformata di Fourier, si ha

$$P(y) = \int_{\mathbb{R}} \frac{dk}{2\pi} e^{-iky} \int_{\mathbb{R}} dx_1 dx_2 P_1(x_1) P_2(x_2) e^{ikx_1} e^{ikx_2} = \int_{\mathbb{R}} \frac{dk}{2\pi} e^{-iky} \tilde{P}_1(k) \tilde{P}_2(k)$$

Per l'unicità della trasformata di Fourier, segue

$$\tilde{P}(k) = \tilde{P}_1(k) \tilde{P}_2(k) \implies F(k) = F_1(k) + F_2(k) \implies c_n = c_n^1 + c_n^2$$

dove $\tilde{P}(k)$ è la trasformata di Fourier di $P(y)$. I cumulanti della distribuzione di probabilità della variabile somma sono uguali alla somma dei cumulanti delle singole distribuzioni: i cumulanti cumulano. I cumulanti caratterizzano in modo più semplice la distribuzione di probabilità di quanto faccia la distribuzione stessa.

Lezione 3

6.8 Teorema del limite centrale

ven 27 ott
2023 14:30

Teorema 6.11 (Lindeberg-Feller CLT). Date n variabili aleatorie $\hat{x}_1, \dots, \hat{x}_n$ indipendenti con distribuzioni di probabilità $P_1(x_1), \dots, P_n(x_n)$ arbitrarie (che però soddisfano condizioni piuttosto generali), per un numero di variabili n arbitrariamente grande, la variabile somma

$$\hat{y} = \sum_{i=1}^n \hat{x}_i$$

ha una distribuzione di probabilità gaussiana definita da

$$\langle \hat{y} \rangle = \sum_{i=1}^n \langle \hat{x}_i \rangle, \quad \sigma_y^2 = \sum_{i=1}^n \sigma_i^2$$

In termini di misure, bisogna interpretare il teorema come una misura di n variabili aleatorie e non n misure di una variabile aleatoria.

Dimostrazione. Si dimostra il caso di distribuzioni identiche

$$\hat{x}_1, x_1, p(x_1); \quad \cdots; \quad \hat{x}_n, x_n, p(x_n)$$

La variabile somma è definita come

$$\hat{y} = \sum_{i=1}^n \hat{x}_i, \quad y = \sum_{i=1}^n x_i, \quad P(y) = \int dx_1 \cdots dx_n p(x_1) \cdots p(x_n) \delta(y - \sum_{i=1}^n x_i)$$

La proprietà di cumulo dei cumulanti implica

$$\langle y \rangle = n \langle x \rangle, \quad \sigma_y^2 = n \sigma_x^2, \quad c_{y,m} = n c_{x,m}, \quad m \geq 3$$

Per capire com'è fatta la distribuzione si utilizza la variabile standardizzata

$$\hat{z} = \frac{\hat{y} - \langle y \rangle}{\sigma_y}$$

Considerato

$$P(y) = \int \frac{dk}{2\pi} [p(k)]^n e^{-iky}$$

la distribuzione della variabile standardizzata è data da

$$\pi(z) = \int dy P(y) \delta(z - \frac{y - \langle y \rangle}{\sigma_y}) = \int \frac{dk}{2\pi} [p(k)]^n \int dy e^{-iky} \delta(z - \frac{y - \langle y \rangle}{\sigma_y})$$

Notando

$$\delta(z - \frac{y - \langle y \rangle}{\sigma_y}) = \sigma_y \delta(\sigma_y z - y + \langle y \rangle)$$

si ottiene

$$\pi(z) = \int \frac{dk}{2\pi} [p(k)]^n e^{-ik(\sigma_y z + \langle y \rangle)} \sigma_y$$

Ricordando

$$p(k) = e^{-F_x(k)} = \exp \left[\sum_{m=1}^{\infty} \frac{(ik)^m}{m!} c_{x,m} \right]$$

segue

$$\pi(z) = \sigma_y \int \frac{dk}{2\pi} e^{-ikz\sigma_y} \exp \left[n \sum_{m=2}^{\infty} \frac{(ik)^m}{m!} c_{x,m} \right] = \int \frac{dk'}{2\pi} e^{-ik'z} \exp \left[n \sum_{m=2}^{\infty} \left(\frac{ik'}{\sigma_y} \right)^m \frac{c_{x,m}}{m!} \right]$$

dove $k' = k\sigma_y$. Ricordando $\sigma_y^2 = n\sigma_x^2 = nc_{x,2}$ si ha

$$\pi(z) = \int \frac{dk'}{2\pi} e^{-ik'z} e^{-\Phi(k')}, \quad \Phi(k') = \frac{(k')^2}{2} - \frac{1}{\sqrt{n}} \sum_{m=3}^{\infty} \frac{c_{x,m}}{m!} \frac{1}{n^{\frac{m-3}{2}}} \left(\frac{ik'}{\sigma_x} \right)^m$$

Il secondo addendo nel generatore dei cumulanti si annulla nel limite $n \rightarrow \infty$. I cumulanti di ordine maggiore di due si annullano in tale limite e quindi la distribuzione tende ad una gaussiana. \square

Osservazione 6.12. Nel caso di una media per variabili identicamente distribuite

$$\langle \hat{y} \rangle = \frac{1}{n} \sum_{i=1}^n \langle \hat{x}_i \rangle$$

quando si calcola la varianza si ottiene

$$\sigma_y^2 = \frac{1}{n^2} \sum_{i=1}^n \sigma_x^2 = \frac{\sigma_x^2}{n}$$

7 Metodi Monte Carlo

7.1 Serie di Edgeworth

Si studia come la distribuzione $\pi(z)$ della somma tende ad una gaussiana.

Definizione 7.1. La serie di Edgeworth è l'espansione in potenze di $n^{-\frac{1}{2}}$ della funzione

$$e^{\frac{1}{2}k^2} e^{-\Phi(k)}$$

ossia l'espansione della correzione rispetto alla gaussiana della trasformata di Fourier della distribuzione. L'espansione è diversa da quella precedente poiché essa è in k .

Ricordando che l'espansione in serie di potenze dell'esponenziale è

$$e^x = \sum_{l=0}^{\infty} \frac{x^l}{l!}$$

si ha

$$e^{\frac{1}{2}k^2} e^{-\Phi(k)} = \sum_{l=0}^{\infty} \frac{1}{l!} \left[\frac{1}{n^{\frac{1}{2}}} \sum_{m=3}^{\infty} \frac{c_{x,m}}{m!} \frac{1}{n^{\frac{m-3}{2}}} \left(\frac{ik}{\sigma_x} \right)^m \right]^l = \sum_{l=0}^{\infty} \frac{(ik)^{2l}}{l!} \left[\sum_{m=1}^{\infty} \frac{(ik)^m}{(m+2)!} \frac{c_{x,m+2}}{n^{\frac{m}{2}} \sigma_x^{m+2}} \right]^l$$

Da questo si capisce che la distribuzione gaussiana è la distribuzione attorno alla quale si studiano tutte le altre: il termine gaussiano è trattato esattamente, mentre il resto viene visto come correzioni. I primi due termini sono

- primo termine: $l = 1, m = 1, O(n^{-\frac{1}{2}})$,
- secondo termine: $l = 1, m = 2, O(n^{-1})$ e $l = 2, m = 1, O(n^{-1})$.

Quindi

$$\begin{aligned} e^{\frac{1}{2}k^2 - \Phi(k)} &= 1 + (ik)^2 \left[\frac{ik}{6} \frac{c_{x,3}}{n^{\frac{1}{2}} \sigma_x^3} + \frac{(ik)^2}{24} \frac{c_{x,4}}{n \sigma_x^4} \right] + \frac{(ik)^4}{2} \left[\frac{ik}{6} \frac{c_{x,3}}{n^{\frac{1}{2}} \sigma_x^3} \right]^2 + o(n^{-1}) \\ &= 1 + \frac{1}{n^{\frac{1}{2}}} \frac{c_{x,3}}{6 \sigma_x^3} (ik)^3 + \frac{1}{n} \left[\frac{c_{x,4}}{24 \sigma_x^4} (ik)^4 + \frac{(ik)^6}{72} \frac{c_{x,3}^2}{\sigma_x^6} \right] + o(n^{-1}) \end{aligned}$$

Ricordando che la trasformata di Fourier dei polinomi di Hermite è

$$\widetilde{He}_n = \int_{\mathbb{R}} dx e^{-\frac{x^2}{2}} He_n(x) e^{ikx} = \sqrt{2\pi} (ik)^n e^{-\frac{k^2}{2}}$$

la distribuzione della variabile standardizzata è data dalla trasformata di Fourier inversa

$$\begin{aligned} \pi(z) &= \int \frac{dk}{2\pi} e^{-ikz} e^{-\Phi(k)} \\ &= \int \frac{dk}{2\pi} e^{-ikz} e^{-\frac{1}{2}k^2} \left[1 + \frac{1}{n^{\frac{1}{2}}} \frac{c_{x,3}}{3! \sigma_x^3} (ik)^3 + \frac{1}{n} \frac{c_{x,4}}{4! \sigma_x^4} (ik)^4 + \frac{1}{n} \frac{10c_{x,3}^2}{6! \sigma_x^6} (ik)^6 + o(n^{-1}) \right] \\ &= \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} \left[1 + \frac{1}{n^{\frac{1}{2}}} \frac{c_{x,3}}{3! \sigma_x^3} He_3(z) + \frac{1}{n} \frac{c_{x,4}}{4! \sigma_x^4} He_4(z) + \frac{1}{n} \frac{10c_{x,3}^2}{6! \sigma_x^6} He_6(z) + o(n^{-1}) \right] \end{aligned}$$

Osservazione 7.2. Risulta chiaro che nel limite $n \rightarrow \infty$, la distribuzione $\pi(z)$ tende ad una gaussiana.

Corollario 7.3. La media aritmetica

$$\hat{y} = \frac{1}{n} \sum_{i=1}^n \hat{x}_i$$

è distribuita, per valori asintotici di n , in modo gaussiano con

$$\langle y \rangle = \langle x \rangle, \quad \sigma_y^2 = \frac{\sigma_x^2}{n}$$

Questa è la conclusione più importante per quanto riguarda i metodi Monte Carlo.

Esercizio. Rifare la dimostrazione del teorema del limite centrale per la variabile media al posto della variabile standardizzata. Basti notare che si può porre

$$z = \frac{\sum_i x_i - n\langle x \rangle}{\sqrt{n}\sigma_x} = \frac{\frac{1}{n} \sum_i x_i - \langle x \rangle}{\sigma_x / \sqrt{n}}$$

nella variabile standardizzata.

7.2 Metodo Monte Carlo

Si vuole trovare un metodo per calcolare degli integrali. Si vede un esempio.

Esempio 7.4 (Integrale Monte Carlo). Estruendo n variabili aleatorie con stessa distribuzione, la media è

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{x} = \langle x \rangle = \int_a^b P(x) x \, dx$$

dove $\langle x \rangle$ indica il valore di aspettazione di una variabile aleatoria. L'uguaglianza $\bar{x} = \langle x \rangle$ vale per il teorema del limite centrale $n \rightarrow \infty$. In questo modo si può calcolare un integrale.

Si consideri la variabile media e l'integrale [r]

$$\hat{y} = \frac{1}{n} \sum_{i=1}^n \hat{x}_i, \quad I = \langle x \rangle = \int_{\mathbb{R}} x P(x) \, dx = \int_0^1 x \, dx, \quad P(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{altrimenti} \end{cases}$$

con variabili \hat{x}_i distribuite uniformemente in $[0, 1]$. La stima Monte Carlo dell'integrale è data da

$$\bar{I} = \frac{1}{n} \sum_{i=1}^n x_i$$

Per una funzione generica si ha

$$I = \langle f(x) \rangle = \int_{\mathbb{R}} f(x) P(x) \, dx, \quad \bar{I} = \frac{1}{n} \sum_{i=1}^n f(x_i)$$

che si può calcolare dal caso elementare precedente. Infatti, si può definire una nuova variabile aleatoria

$$y = f(x), \quad P(y) = \int P(x) \delta(y - f(x)) \, dx$$

Si ha

$$\int y P(y) \, dy = \int dx dy y P(x) \delta(y - f(x)) = \int dx P(x) f(x) = I$$

da cui

$$\bar{I} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n f(x_i)$$

cioè la stima sopra.

Si vuole calcolare l'integrale

$$I = \int_0^1 f(x) \, dx$$

Si può ricondurre l'integrale in un intervallo arbitrario a quello sopra tramite il cambio di variabile

$$z = \frac{x - a}{b - a}$$

Definizione 7.5. Si supponga di avere un generatore di numeri casuali distribuiti in maniera uniforme nell'intervallo $[0, 1]$.

Teorema 7.6. Se $\hat{x}_1, \dots, \hat{x}_n$ sono n variabili aleatorie indipendenti, allora le variabili

$$\hat{f}_i = f(\hat{x}_i)$$

sono n variabili aleatorie indipendenti tutte con valore di aspettazione I .

Si definisce la variabile aleatoria

$$y = f(x), \quad P(y) = \int dx_i P(x_i) \delta(y - f(x_i))$$

il cui valor medio è

$$\langle y \rangle = \int y P(y) dy = \int dy dx_i P(x_i) \delta(y - f(x_i)) y = \int dx_i P(x_i) f(x_i)$$

Corollario 7.7. Il teorema precedente insieme al teorema del limite centrale implica che la variabile aleatoria

$$\hat{f} = \frac{1}{n} \sum_{i=1}^n f(\hat{x}_i)$$

ha come valor medio

$$\langle \hat{f} \rangle = I$$

e come varianza

$$\frac{1}{n} \int_0^1 [f(x) - I]^2 dx = \frac{1}{n} \langle f^2(x) \rangle - \langle f(x) \rangle^2 = \frac{\sigma_x^2}{n} = \sigma^2$$

Estraendo n numeri casuali x_1, \dots, x_n distribuiti uniformemente in $[0, 1]$, si approssima l'integrale con

$$I \approx \frac{1}{n} \sum_{i=1}^n f(x_i) + O(n^{-\frac{1}{2}})$$

Osservazione 7.8. L'errore nella procedura Monte Carlo ha un significato statistico e non deterministico. La stima è distribuita in modo gaussiano con varianza $\frac{\sigma^2}{n}$ e pertanto si può fornire la probabilità con cui il risultato differisca dal valore vero, ma non è possibile dare delle stime o limiti deterministici dell'errore.

Il limite centrale fornisce una interpretazione dell'errore. Poiché la variabile somma o media è distribuita come una gaussiana, i valori all'interno di 1σ sono il 68% dei casi. L'errore nel metodo Monte Carlo ha un significato diverso rispetto ai metodi deterministici.

Osservazione 7.9. L'errore Monte Carlo decresce come $n^{-\frac{1}{2}}$ con n numero di volte che si estraggono numeri casuali e si valuta la funzione.

7.2.1 Integrali multidimensionali

Risulta ovvio che nel caso monodimensionale, il metodo Monte Carlo non è il più efficace. Si passa a d variabili di integrazione

$$dx \rightarrow d^d x, \quad x \rightarrow \mathbf{x} = \{x_1, \dots, x_d\}, \quad I = \int d^d x f(\mathbf{x})$$

La stima Monte Carlo è data da

$$I \approx \frac{1}{n} \sum_{i=1}^n f(x_1^i, \dots, x_d^i) + O(n^{-\frac{1}{2}})$$

Ogni componente va estratta n volte.

Osservazione 7.10. L'errore decresce come $n^{-\frac{1}{2}}$ indipendentemente dal numero di dimensioni d . Il metodo Monte Carlo è il metodo da usare per integrali con un grande numero di dimensioni.

Osservazione 7.11. Dato un integrale, l'obiettivo è descrivere una procedura Monte Carlo che con il minimo tempo di calcolo si parte dell'errore desiderato. Nella maggior parte dei casi, questo significa ridurre la varianza $\frac{\sigma^2}{n}$ ottimizzando la procedura Monte Carlo al problema dato.

La varianza si può ridurre inserendo informazioni riguardo il sistema fisico. Ad esempio, in QCD agli algoritmi si può insegnare la presenza di rottura spontanea e la libertà asintotica.

Osservazione 7.12. Estrarre i numeri non deve diventare molti più costoso, altrimenti si calcola meno volte la funzione [r]

8 Campionamento di importanza

Il metodo visto è elementare, semplice ma inefficiente. Per funzioni piccate, il metodo estrae con probabilità uniforme in tutto l'intervallo: molti valori sono lontani dal picco. Questo si manifesta in una grande varianza. Per risolvere questo problema si utilizza il campionamento di importanza: si fornisce al calcolatore una informazione (più o meno approssimata) di com'è fatta la funzione. L'informazione è costituita da una funzione che si sa integrare analiticamente vicina alla funzione da integrare.

Dato l'integrale

$$I = \int f(x) dx$$

si supponga di conoscere una funzione $g(x)$ tale che

$$g(x) \geq 0, \quad \int_0^1 g(x) dx = 1$$

Allora vale

$$I = \int \frac{f(x)}{g(x)} g(x) dx$$

Se x è estratto secondo la distribuzione $g(x)$, allora

$$\bar{I} = \frac{1}{n} \sum_{i=1}^n \frac{f(x_i)}{g(x_i)}, \quad \sigma^2 = \frac{1}{n} \int_0^1 \left[\frac{f(x)}{g(x)} - I \right]^2 g(x) dx$$

Per vedere ciò basta ricondursi al caso precedente con

$$\bar{f}(x) = \frac{f(x)}{g(x)}, \quad P(x) = g(x)$$

Nel limite in cui $f(x) = g(x)$, cioè si sa integrare la funzione, allora la varianza è nulla.

Osservazione 8.1. Risulta chiaro che più $f(x)$ è vicina a $kg(x)$ più la varianza è piccola.

Osservazione 8.2. In una teoria dei campi, la distribuzione delle variabili di campo va come il peso di Boltzmann e^{-S} dove l'azione S è una variabile estensiva ed è proporzionale al volume V : la probabilità va come e^{-V} . L'espressione per l'integrale sui cammini è

$$\langle O \rangle = \frac{\int \pi dx_i e^{-S} v(x) \dots}{\int \pi dx_i e^{-S}}$$

in cui si pone

$$g(x) = \frac{e^{-S}}{\int \pi dx_i e^{-S}}, \quad f(x) = v(x) \dots$$

Poiché il volume è dell'ordine dei miliardi di punti, le funzioni sono molto piccate. Senza il campionamento d'importanza, il metodo Monte Carlo non funzionerebbe.

Osservazione 8.3. Questo [r]

Osservazione 8.4. Per applicare il campionamento di importanza bisogna essere in grado di estrarre numeri casuali con la distribuzione $g(x)$.

8.1 Metodo del cambio di variabili

Si vede il caso in basso numero di dimensioni. Si supponga di avere un generatore di numeri casuali che generi numeri x secondo una distribuzione $p(x)$

$$\int_{\mathbb{R}} p(x) dx = 1$$

Si supponga di calcolare $g(x)$ [r] Sia $y = f(x)$ dove f è una funzione che lega le due variabili casuali x ed y . Dunque

$$|p(y) dy| = |p(x) dx| \implies p(y) = p(x) |d_y x|$$

[r]

Esempio 8.5 (Esponenziale). Si consideri

$$p(y) = e^{-y}, \quad p(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{altrimenti} \end{cases}, \quad \int_0^1 p(x) dx = 1$$

Si utilizza il cambio di variabile

$$y = -\ln(1-x) \implies p(y) = e^{-y}, \quad \int_0^\infty p(y) dy = 1$$

La procedura da seguire è

- estrarre x in modo uniforme in $[0, 1)$,
- calcolare $y(x)$ che è distribuita secondo $p(y) = e^{-y}$ in $[0, \infty)$.

8.1.1 Cambio multidimensionale

Si supponga di estrarre d variabili casuali x_1, \dots, x_d con probabilità

$$p(x_1, \dots, x_d) dx_1 \cdots dx_d$$

allora le variabili

$$y_1 = f_1(x_1, \dots, x_d), \quad \dots, \quad y_d = f_d(x_1, \dots, x_d)$$

sono distribuite come

$$p_y(y_1, \dots, y_d) dy_1 \cdots dy_d = p_x(x_1, \dots, x_d) \left| \frac{\partial(x_1, \dots, x_d)}{\partial(y_1, \dots, y_d)} \right| dy_1 \cdots dy_d$$

dove compare il jacobiano del cambio di variabili.

Esempio 8.6 (Guassiana, Box–Muller). Il metodo di Box–Muller permette di estrarre numeri distribuiti in modo gaussiano. Dati due numeri casuali x_1, x_2 uniformi in $[0, 1)$, si definisce

$$y_1^2 = -\ln(1-x_2) \sin^2\left(\frac{\pi}{2}x_1\right), \quad y_2^2 = -\ln(1-x_2) \cos^2\left(\frac{\pi}{2}x_1\right)$$

Pertanto vale

$$e^{-(y_1^2+y_2^2)} = 1-x_2, \quad \left[\frac{y_1}{y_2}\right]^2 = \tan^2\left(\frac{\pi}{2}x_1\right)$$

da cui

$$x_1 = \frac{2}{\pi} \arctan \frac{y_1}{y_2}, \quad x_2 = 1 - e^{-(y_1^2+y_2^2)}$$

La jacobiana ed il jacobiano sono

$$J = \begin{bmatrix} \partial_{y_1} x_1 & \partial_{y_2} x_1 \\ \partial_{y_1} x_2 & \partial_{y_2} x_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\pi} \frac{y_2}{y_1^2+y_2^2} & -\frac{2}{\pi} \frac{y_1}{y_1^2+y_2^2} \\ 2y_1 e^{-(y_1^2+y_2^2)} & 2y_2 e^{-(y_1^2+y_2^2)} \end{bmatrix}, \quad \det J = \frac{4}{\pi} e^{-(y_1^2+y_2^2)}$$

Pertanto, le variabili y sono distribuite secondo

$$P(y_1, y_2) = \frac{2}{\sqrt{\pi}} e^{-y_1^2} \frac{2}{\sqrt{\pi}} e^{-y_2^2}$$

[r]

Osservazione 8.7. Estrahendo due numeri in modo uniforme, il cambio di variabili porta a due numeri distribuiti in modo gaussiano.

Osservazione 8.8. Si noti che il metodo più veloce per ottenere tanti numeri casuali distribuiti in modo gaussiano è utilizzare l'algoritmo ziggurat.

Esempio 8.9. Si vuole generare una variabile distribuita secondo

$$p(y) = \frac{2}{\sqrt{\pi}} \sqrt{y} e^{-y}$$

Unendo gli esempi precedenti, si estraggono due variabili distribuite come

$$p(x_1) = \frac{2}{\sqrt{\pi}} e^{-x_1^2}, \quad p(x_2) = e^{-x_2}$$

Si definisce

$$y = x_1^2 + x_2, \quad y_1 = y, \quad y_2 = x_2$$

La jacobiana ed il jacobiano sono

$$J^{-1} = \begin{bmatrix} \partial_{x_1} y_1 & \partial_{x_2} y_1 \\ \partial_{x_1} y_2 & \partial_{x_2} y_2 \end{bmatrix} = \begin{bmatrix} 2x_1 & 1 \\ 0 & 1 \end{bmatrix} \Rightarrow \det J = \frac{1}{\det J^{-1}} = \frac{1}{2x_1} = \frac{1}{2\sqrt{y_1 - y_2}}$$

Quindi

$$p(y_1, y_2) = \frac{2}{\sqrt{\pi}} \frac{e^{-y_1}}{2\sqrt{y_1 - y_2}}$$

I domini delle variabili trasformate sono

$$y_2 : [0, \infty), \quad y_1 : [y_2, \infty)$$

La regione del piano di tali due variabili si può anche riscrivere come

$$y_1 : [0, \infty), \quad y_2 : [0, y_1)$$

La variabile y_2 non è di interesse, cioè importa il valore di y_1 a prescindere da y_2 , pertanto si integra (e si rinomina y_1 a y):

$$\begin{aligned} p(y) dy &= \int_0^y dy_2 \frac{2}{\sqrt{\pi}} \frac{e^{-y}}{2\sqrt{y - y_2}} = \frac{1}{\sqrt{2\pi}} e^{-y} \int_0^y dy_2 \frac{1}{\sqrt{y - y_2}} \\ &= \frac{2}{\sqrt{\pi}} e^{-y} (-\sqrt{y - y_2})_0^y = \frac{2}{\sqrt{\pi}} \sqrt{y} e^{-y} \end{aligned}$$

Questo esempio, insieme al primo, è utile per le teorie di gauge.

Osservazione 8.10. L'azione dell'oscillatore armonico è

$$S = \int dt \left(\frac{1}{2} m \dot{x}^2 - \frac{1}{2} m \omega^2 x^2 \right)$$

dove non si utilizza il tempo euclideo. Si ha un numero infinito di variabili $x(t)$, ma non si sa fare un integrale di una produttoria, cioè l'esponenziale e^{-S} . Invece, si discretizza la derivata temporale e l'integrale. Si integra su tutti i tempi e su tutte le variabili x :

$$S = a \sum_i \frac{m}{2} \frac{(x_{i+1} - x_i)^2}{a} - \frac{1}{2} m \omega^2 x_i^2$$

Osservazione 8.11. Il metodo del cambio di variabili funziona per un basso numero di dimensioni. Quanto visto è utile avere variabili distribuite secondo una funzione arbitraria. Se si vuole simulare una teoria in un grande numero di dimensioni e $g(x)$ non fattorizzabile, il metodo del cambio di variabili non si sa fare. Bisogna trovare una strategia che faccia in modo automatico il cambio di variabili: le catene di Markov. Una catena di Markov è un processo iterativo che permette di arrivare ad un generatore che distribuisce i numeri casuali con un cambio di variabili arbitrario, senza sapere la formula analitica che connette le due variabili.

9 Conclusioni

Se la varianza di della distribuzione di $\hat{f}(x_i)$ è finita, la stima Monte Carlo $[r]$ al valore vero dell'integrale per $[r]$ il numero di configurazioni $N_{\text{conf.}}$. In seguito, il numero di estrazioni è indicato come numero di configurazioni.

La stima Monte Carlo è unbiased per tutti i valori del numero di configurazione, il valore di aspettazione della stima Monte Carlo è il valore vero dell'integrale.

La stima Monte Carlo è distribuita in modo gaussiano per valori del numero di configurazione asintotici.

La deviazione standard della stima Monte Carlo è data da

$$\sigma = \sqrt{\frac{\text{Var } f}{N_{\text{conf}}}}$$

dove $\text{Var } f$ è la varianza teorica della funzione f . $[r]$

Lezione 4

ven 03 nov
2023 14:30

Parte III

Integrale sui cammini in meccanica quantistica

10 Integrale sui cammini di Feynman

L'integrale sui cammini (path integral) fornisce un'altra interpretazione della meccanica quantistica (non relativistica) e della sua relazione con la meccanica classica.

La formulazione di Dirac-Feynman delle teorie quantistiche permette di definire non perturbativamente le teorie di campo utili in fisica (QED, QCD, etc.) che altrimenti si definirebbero solo come sviluppi perturbativi attorno alla teorie libere.

Si deriva l'integrale sui cammini di Feynman per il propagatore ritardato per un sistema quantistico monodimensionale. L'estensione a più dimensioni è immediata e non pone particolari problemi concettuali.

Per teorie di campo scalari, la formulazione è stata trovata da Kohrut? mentre per campi fermionici è stata trovata da Kusher. $[r]$

Notazione. Si utilizza

$$\hat{x} |x\rangle = x |x\rangle, \quad \hat{p} |p\rangle = p |p\rangle, \quad \hbar = 1, \quad \hat{p} = -i \partial_x, \quad [\hat{x}, \hat{p}] = i$$

così come

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \quad \langle x|p\rangle = \frac{1}{\sqrt{2\pi}} e^{ipx}, \quad \langle x|x'\rangle = \delta(x-x'), \quad \langle p|q\rangle = \delta(p-q)$$

Descrizione di Schrödinger e di Heisenberg. Nella descrizione di Schrödinger della meccanica quantistica, gli operatori, le osservabili non dipendono dal tempo, mentre lo stato del sistema si evolve secondo

$$i \partial_t |\alpha, t\rangle_S = \hat{H} |\alpha, t\rangle_S \implies |\alpha, t\rangle_S = e^{-i\hat{H}t} |\alpha, 0\rangle, \quad t \geq 0$$

Mentre nella descrizione di Heisenberg, lo stato è costante e gli operatori si evolvono secondo

$$i \partial_t \hat{O}_H = [\hat{O}_H, \hat{H}] \implies \hat{O}_H(t) = e^{i\hat{H}t} \hat{O}_H(0) e^{-i\hat{H}t}, \quad t \geq 0$$

I due formalismi sono equivalenti

$$|\alpha\rangle_{\text{H}} = |\alpha, 0\rangle_{\text{S}}, \quad \hat{O}_{\text{H}}(0) = \hat{O}_{\text{S}} \implies {}_{\text{S}}\langle\alpha', t|\hat{O}_{\text{S}}|\alpha, t\rangle_{\text{S}} = {}_{\text{H}}\langle\alpha'|\hat{O}_{\text{H}}|\alpha\rangle_{\text{H}}$$

In teoria dei campi si utilizza la descrizione di Heisenberg, mentre la descrizione di Schrödinger è utilizzata nella meccanica quantistica non relativistica.

Da questo punto bisogna costruire l'integrale sui cammini e verificare che sia equivalente a quanto già noto dall'equazione di Schrödinger, in particolare per quanto riguarda la teoria delle perturbazioni. In teoria dei campi si ha solamente l'integrale sui cammini come formulazione della meccanica quantistica.

10.1 Propagatore ritardato

Indipendentemente dal formalismo utilizzato, la dinamica è codificata nel propagatore ritardato. Si utilizzano gli stati nella descrizione di Schrödinger, ma per gli operatori nella descrizione di Heisenberg è analogo. Si parte dalla soluzione formale di uno stato ad un tempo t

$$|\alpha, t\rangle_{\text{S}} = e^{-i\hat{H}t} |\alpha, 0\rangle$$

e si definisce il propagatore ritardato:

$$\hat{G}(t, t_0) \equiv \theta(t - t_0) e^{-i\hat{H}(t - t_0)}$$

dove si ha la funzione θ di Heaviside. Tale operatore soddisfa un'equazione differenziale al prim'ordine con una condizione al contorno

$$i\partial_t \hat{G}(t, t_0) = \hat{H} \hat{G}(t, t_0) + i\delta(t - t_0) \hat{I}, \quad \hat{G}(t_0, t_0) = \hat{I}$$

Poiché tale operatore risolve la dinamica del sistema, allora deve contenere tutte le informazioni della dinamica: gli autostati e gli autovalori dell'hamiltoniana. Essi si possono trovare in due modi, uno dei quali consiste nel diagonalizzare l'hamiltoniana. Dati gli autovalori e autostati dell'hamiltoniana

$$\hat{H} |E_n\rangle = E_n |E_n\rangle, \quad u_n(x) = \langle x | E_n \rangle$$

dove $u_n(x)$ sono gli autostati nella base della posizione. [r] Ci si pone nella base di autostati dell'hamiltoniana

$$\hat{G}(t, t_0) = \theta(t - t_0) \sum_n e^{-iE_n(t - t_0)} |E_n\rangle \langle E_n|$$

La trasformata di Fourier nel tempo è [r]

$$\hat{\tilde{G}}(z) = \int_{\mathbb{R}} e^{izt} \hat{G}(t) dt = \int_0^\infty e^{izt} e^{-i\hat{H}t} dt, \quad t_0 = 0$$

L'integrale non è ben definito per z reale. Si considera una continuazione analitica di tale parametro in modo da calcolare l'integrale e poi applicarlo a z reale. Per $\text{Im } z > 0$, l'integrale è convergente

$$\hat{\tilde{G}}(z) = \frac{i}{z - \hat{H}} = i \sum_n \frac{1}{z - E_n} |E_n\rangle \langle E_n|$$

Nel resto del piano complesso, si continua analiticamente la funzione tramite l'espressione sopra. [r]

Elementi di matrice. Risulta interessante calcolare gli elementi di matrice del propagatore e della sua trasformata

$$G(x, t, x_0, t_0) = \langle x | \hat{G}(t, t_0) | x_0 \rangle = \theta(t - t_0) \sum_n e^{-iE_n(t - t_0)} u_n(x) u_n^*(x_0)$$

$$\tilde{G}(x, x_0, z) = i \sum_n \frac{1}{z - E_n} u_n(x) u_n^*(x_0)$$

Gli elementi di matrice sono l'ampiezza di probabilità di una particella di passare dall'evento (t_0, x_0) all'evento (t, x) . Gli elementi di matrice della trasformata sono funzione della variabile z coniugata del tempo, mentre la posizione di partenza x_0 e arrivo x sono parametri. Dalle relazioni sopra è chiaro che il propagatore \hat{G} contiene tutte le informazioni della dinamica: l'hamiltoniana ha autovalori reali e dunque il propagatore \hat{G} ha singolarità sull'asse reale.

Per gli stati legati, cioè stati ad energia negativa, la funzione gli elementi di matrice della trasformata ha dei poli in corrispondenza degli autovalori, cioè per $\text{Re } z < 0$, $\text{Im } z = 0$. Per autovalori continui, cioè $\text{Re } z > 0$, $\text{Im } z = 0$, la funzione ha un taglio. Per i poli, i residui sono proporzionali alla funzione d'onda degli stati legati: il residuo al polo, come funzione di x (quindi x_0 è fisso) fornisce l'autofunzione associata all'autovalore del polo; per il taglio è il coefficiente al variare di z .

Le proprietà di analiticità della trasformata di Fourier del propagatore ritardato sono molto generali. La posizione dei poli e dei tagli permette di ricavare gli autovalori e autofunzioni dell'hamiltoniana del sistema in esame. Conoscere il propagatore è equivalente ad aver risolto la dinamica.

Quando si generalizza ad una teoria dei campi non perturbativa, non si sa definire bene un'hamiltoniana né risolvere l'equazione di Schrödinger, ma si possono trovare i poli ed i tagli della trasformata del propagatore. Tuttavia, tali poli e tagli non si sanno ricavare analiticamente, pertanto bisogna sviluppare delle tecniche numeriche.

Interpretazione. Gli elementi di matrice del propagatore sono l'ampiezza di probabilità che il sistema si trovi al punto x nell'istante t (quando venga misurato) se nell'istante t_0 si trovava nel punto x_0 . Ossia, all'istante t_0 la coordinata viene misurata e vale x_0 cioè è un autostato $|x_0\rangle$.

Proprietà di convoluzione. Per $t_0 < t_1 < t$ si ha

$$G(x, t, x_0, t_0) = \langle x | \hat{G}(t, t_0) | x_0 \rangle = \langle x | e^{-i\hat{H}(t-t_0)} | x_0 \rangle = \langle x | e^{-i\hat{H}(t-t_1)} e^{-i\hat{H}(t_1-t_0)} | x_0 \rangle$$

Inserendo l'identità [r]

$$\hat{I} = \int dx_1 |x_1\rangle\langle x_1|$$

si ottiene

$$\begin{aligned} G(x, t, x_0, t_0) &= \int dx_1 \langle x | e^{-i\hat{H}(t-t_1)} | x_1 \rangle \langle x_1 | e^{-i\hat{H}(t_1-t_0)} | 0 \rangle \\ &= \int dx_1 G(x, t, x_1, t_1) G(x_1, t_1, x_0, t_0) \end{aligned}$$

cioè andare da (x_0, t_0) a (x, t) equivale ad andare prima da (x_0, t_0) a (x_1, t_1) e poi in (x, t) . Questa formula di convoluzione è detta formula di Chapman–Kolmogorov. L'ampiezza di probabilità è la somma di tutti i possibili prodotti di ampiezze di probabilità (cioè tutte le posizioni x_1). Si noti che questa discussione non è relativistica. Vale anche per [r]

$$\theta(t - t_0) = \theta(t - t_1) \theta(t_1 - t_0), \quad t_0 < t_1 < t$$

Al fine di costruire l'integrale sui cammini si ripete questo passaggio un numero N di volte arbitrariamente grande con intervallo temporale $\delta t \equiv t_{i+1} - t_i$ infinitesimo. Bisogna prima dimostrare alcune formule.

Teorema 10.1 (Formula di Trotter). Dati due operatori A e B ragionevoli (ossia $[A, B]$ non diverge), vale

$$e^{A+B} = \lim_{N \rightarrow \infty} (e^{\frac{A}{N}} e^{\frac{B}{N}})^N$$

Questo teorema viene successivamente applicato all'hamiltoniana poiché somma dell'operatore cinetico e dell'operatore potenziale che non commutano tra loro.

Dimostrazione. Considerando l'espansione in serie di potenze dell'esponenziale di una matrice (oppure usando la formula di Zassenhaus derivata da quella di Baker–Campbell–Hausdorff), vale

$$e^{\frac{A+B}{N}} = e^{\frac{A}{N}} e^{\frac{B}{N}} + o(N^{-1})$$

dove nel secondo addendo sono presenti i termini con i commutatori. Dunque

$$\begin{aligned}
 e^{A+B} - (e^{\frac{A}{N}} e^{\frac{B}{N}})^N &= (e^{\frac{A+B}{N}})^N - (e^{\frac{A}{N}} e^{\frac{B}{N}})^N \\
 &= e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-1} - e^{\frac{A}{N}} e^{\frac{B}{N}} e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-2} \\
 &\quad + e^{\frac{A}{N}} e^{\frac{B}{N}} e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-3} - (e^{\frac{A}{N}} e^{\frac{B}{N}})^2 e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-4} \\
 &\quad + (e^{\frac{A}{N}} e^{\frac{B}{N}})^2 e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-5} + \dots - (e^{\frac{A}{N}} e^{\frac{B}{N}})^{N-1} e^{\frac{A+B}{N}} \\
 &\quad + (e^{\frac{A}{N}} e^{\frac{B}{N}})^{N-1} e^{\frac{A+B}{N}} - (e^{\frac{A}{N}} e^{\frac{B}{N}})^N \\
 &= [e^{\frac{A+B}{N}} - e^{\frac{A}{N}} e^{\frac{B}{N}}] e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-1} \\
 &\quad + e^{\frac{A}{N}} e^{\frac{B}{N}} [e^{\frac{A+B}{N}} - e^{\frac{A}{N}} e^{\frac{B}{N}}] e^{\frac{A+B}{N}} (e^{\frac{A+B}{N}})^{N-2} + \dots \\
 &\quad + (e^{\frac{A}{N}} e^{\frac{B}{N}})^{N-1} [e^{\frac{A+B}{N}} - (e^{\frac{A}{N}} e^{\frac{B}{N}})] \\
 &= o(N^{-1})
 \end{aligned}$$

Alla prima riga, si noti che $A+B$ si può considerare come un singolo operatore, dunque

$$\frac{A+B}{N} N = A+B$$

Alla terza uguaglianza, le parentesi quadre sono tutte $o(N^{-1})$ e i termini non si sommano ad annullarsi, perciò rimane almeno un termine di ordine $O(N^{-2})$. Infatti, sebbene in questo caso non succeda, gli N termini potrebbero dare un termine di ordine $O(N^{-1})$ e non $O(N^{-2})$.

In teorie fermioniche tali termini si accumulano e si ha un termine di ordine N^{-1} . \square

10.2 Operatore di trasferimento

Si introducono delle modifiche alla definizione degli elementi di matrice del propagatore ritardato per ottenere una definizione rigorosa dell'integrale dei cammini. Si ipotizzi di discretizzare il tempo in modo che un intervallo venga diviso in N sotto-intervalli

$$T = t - t_0 = aN$$

con a passo reticolare (temporale). Si può scrivere il propagatore ritardato come

$$\hat{G}(t, t_0) = e^{-i\hat{H}(t-t_0)} = e^{-iaN\hat{H}} = \exp\left[-iaN\left(\frac{\hat{p}^2}{2m} + V(\hat{x})\right)\right]$$

Utilizzando la formula di Trotter si ha

$$\begin{aligned}
 \hat{G}(t, t_0) &= \exp\left[-iaN\left(\frac{\hat{p}^2}{2m} + V(\hat{x})\right)\right] \approx \left[e^{-ia\frac{\hat{p}^2}{2m}} e^{-iaV(\hat{x})}\right]^N \\
 &= e^{i\frac{a}{2}V(\hat{x})} \left[e^{-i\frac{a}{2}V(\hat{x})} e^{-ia\frac{\hat{p}^2}{2m}} e^{-i\frac{a}{2}V(\hat{x})}\right]^N e^{-i\frac{a}{2}V(\hat{x})}
 \end{aligned}$$

Dalla prima riga si nota che l'operatore che propaga una particella da t_0 a t_1 è dato da $e^{-ia\hat{H}}$ a meno di un termine $o(a)$ (applicando l'identità utilizzata nella dimostrazione). Nel limite continuo

$$a \rightarrow 0, \quad N \rightarrow \infty, \quad aN = \text{cost.}$$

gli errori spariscono. Per vedere che la seconda riga è vera, l'esponente N equivale a ripetere la base N volte e si moltiplica all'inizio per

$$1 = e^{i\frac{a}{2}V(\hat{x})} e^{-i\frac{a}{2}V(\hat{x})}$$

mentre tra gli N termini si divide l'esponentiale del potenziale. Si definisce l'operatore di trasferimento come

$$\hat{\tau}_a \equiv e^{-i\frac{a}{2}V(\hat{x})} e^{-i\frac{\hat{p}^2}{2m}a} e^{-i\frac{a}{2}V(\hat{x})}$$

Nel limite continuo, l'operatore sopra evolve il sistema da un istante a quello successivo (senza errori). Quindi

$$\hat{G}(t, t_0) = e^{i\frac{a}{2}V(\hat{x})} \hat{\tau}_a^N e^{-i\frac{a}{2}V(\hat{x})}$$

Per costruzione, vale

$$\hat{\tau}_a \hat{\tau}_a^\dagger = \hat{\tau}_a^\dagger \hat{\tau}_a = \hat{I}$$

dunque si può definire un operatore hermitiano $\hat{\hat{H}}$ tale che

$$\hat{\tau}_a = e^{-ia\hat{\hat{H}}}, \quad \hat{\hat{H}} = \hat{\hat{H}}^\dagger$$

dove $\hat{\hat{H}} \neq \hat{H}$ per N finito.

Osservazione 10.2. L'operatore di trasferimento $\hat{\tau}_a$ è l'operatore che propaga uno stato generico nel tempo $\Delta t = a$ a meno di un errore $o(a)$.

10.2.1 Operatore di trasferimento euclideo

Tempo euclideo. Si prolunga analiticamente il propagatore ritardato nella variabile t e si definisce il propagatore euclideo come

$$\hat{G}_E(t, t_0) = \hat{G}(-it, -it_0) = \theta(t - t_0) e^{-\hat{H}(t-t_0)}, \quad t \rightarrow -it$$

che corrisponde esplicitamente a

$$\hat{G}_E(t_E, t_{0,E}) = \hat{G}(-it_E, -it_{0,E}), \quad t_E = it$$

In questo modo, per autovalori grandi, l'esponenziale tende a zero. In teoria dei campi, la metrica di Minkowski diventa la metrica euclidea quanto si considera il tempo euclideo.

Osservazione 10.3. Il tempo euclideo t_E è una variabile reale che corrisponde alla parte immaginaria del tempo t prolungato analiticamente e preso sull'asse complesso.

Osservazione 10.4. Successivamente si potrebbe omettere la distinzione tra tempo euclideo e tempo di Minkowski, ma si dovrebbe capire dal contesto.

Osservazione 10.5. Risulta ovvio che nel tempo euclideo, il propagatore $G_E(t_E)$ contiene tutta l'informazione sulla dinamica del sistema. Infatti, si può fare il prolungamento analitico, tornare al tempo di Minkowski e utilizzare le formule precedenti. Oppure si modificano direttamente le formule.

Operatore di trasferimento euclideo. Il passo reticolare diventa $a_M \rightarrow -ia_E$ da cui

$$\hat{T}_a \equiv e^{-\frac{a}{2}V(\hat{x})} e^{-\frac{p^2}{2m}a} e^{-\frac{a}{2}V(\hat{x})}$$

A questo punto si può inserire l'identità

$$\hat{I} = \int dx |x\rangle\langle x|$$

tra 0 e T per $N - 1$ volte.

Osservazione 10.6. Per costruzione si ha

$$\hat{T}_a = e^{-a\hat{\hat{H}}} = e^{-a\hat{H}} + o(a), \quad \hat{T}_a |\mathcal{E}_n\rangle = e^{-a\mathcal{E}_n} |\mathcal{E}_n\rangle, \quad \mathcal{E}_n = E_n + o(a)$$

Elementi di matrice di trasferimento nella base delle coordinate. Si utilizza il tempo euclideo. Si calcola

$$\langle x_i | \hat{T}_a | x_j \rangle = \langle x_i | e^{-\frac{a}{2}V(\hat{x})} e^{-\frac{\hat{p}^2}{2m}} e^{-\frac{a}{2}V(\hat{x})} | x_j \rangle$$

Per definizione si ha

$$e^{-\frac{a}{2}V(\hat{x})} | x_i \rangle = e^{-\frac{a}{2}V(x_i)} | x_i \rangle$$

poiché il potenziale è un operatore diagonale nella base delle coordinate. Inserendo l'identità di autostati del momento nel termine cinetico si ha

$$\begin{aligned} \langle x_i | e^{-\frac{\hat{p}^2}{2m}a} | x_j \rangle &= \int dq_i dq_j \langle x_i | q_i \rangle \langle q_i | e^{-\frac{\hat{p}^2}{2m}a} | q_j \rangle \langle q_j | x_j \rangle \\ &= \frac{1}{2\pi} \int dq_i dq_j e^{iq_i x_i} e^{-\frac{q_i^2}{2m}a} \delta(q_i - q_j) e^{-iq_j x_j} \\ &= \frac{1}{2\pi} \int dq e^{-\frac{q^2}{2m}a} e^{iq(x_i - x_j)} \\ &= \left(\frac{m}{2\pi a} \right)^{\frac{1}{2}} e^{-\frac{(x_i - x_j)^2 m}{2a}} \end{aligned}$$

alla quarta riga si è applicato

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{q^2}{2\sigma^2}} e^{iqx} dq = e^{-\frac{x^2\sigma^2}{2}}$$

Si utilizza il tempo euclideo perché così si può svolgere l'integrale di una gaussiana e si può trovare il termine cinetico. Unendo quanto trovato si ha

$$\langle x_i | \hat{T}_a | x_j \rangle = \left(\frac{m}{2\pi a} \right)^{\frac{1}{2}} \exp \left[-a \left(\frac{m}{2} \left(\frac{x_i - x_j}{a} \right)^2 + \frac{1}{2}V(x_i) + \frac{1}{2}V(x_j) \right) \right]$$

Propagatore euclideo come integrale sui cammini. Dato il propagatore euclideo nella base delle coordinate

$$G_E(x_N, T, x_0, 0) = \langle x_N | \hat{G}_E(T, 0) | x_0 \rangle$$

si inserisce l'identità $N - 1$ volte

$$\begin{aligned} G_E(x_N, T, x_0, 0) &= \int \prod_{i=1}^{N-1} dx_i e^{\frac{a}{2}V(x_N)} \left[\prod_{i=0}^{N-1} \langle x_{i+1} | \hat{T}_a | x_i \rangle \right] e^{-\frac{a}{2}V(x_0)} \\ &= \left(\frac{m}{2\pi a} \right)^{\frac{N}{2}} \int \prod_{i=1}^{N-1} dx_i \prod_{i=0}^{N-1} \exp \left[-a \left(\frac{m}{2} \left(\frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i) \right) \right] \end{aligned}$$

Il propagatore ritardato è ora scritto come un integrale multidimensionale. Inoltre, si nota che

$$\frac{x_{i+1} - x_i}{a} \rightarrow \dot{x}, \quad a \rightarrow 0$$

cioè si ottiene la velocità classica. Si noti il segno positivo al potenziale a causa del tempo euclideo. La densità di lagrangiana euclidea e l'azione sono definite come

$$\mathcal{L}_E(x_{i+1}, x_i) \equiv \frac{1}{2}m \left(\frac{x_{i+1} - x_i}{a} \right)^2 + V(x_i), \quad S_E = a \sum_{i=0}^{N-1} \mathcal{L}_E(x_{i+1}, x_i)$$

pertanto

$$G_E(x_N, T, x_0, 0) = \left(\frac{m}{2\pi a} \right)^{\frac{N}{2}} \int \prod_{i=1}^{N-1} dx_i e^{-S_E}$$

con x_0 e x_N fissati. Questa è la forma standard per il propagatore ritardato espresso come integrale sui cammini di un sistema quantistico nel tempo euclideo. Una volta discretizzato il tempo, un cammino di Feynman corrisponde ad una n -pla di coordinate. Nel limite continuo, l'ampiezza di probabilità è data dalla somma su tutti i cammini pesati con l'esponenziale dell'azione classica.

Propagatore Minkowskiano come integrale sui cammini. Per tornare al tempo fisico di Minkowski si sostituisce $a_E \rightarrow ia_M$. La lagrangiana diventa

$$\mathcal{L} = -\mathcal{L}_E(x_{i+1}, x_i) = \frac{m}{2} \left(\frac{x_{i+1} - x_i}{a_M} \right)^2 - V(x_i), \quad S = a \sum_{i=0}^{N-1} \mathcal{L}(x_{i+1}, x_i)$$

Dunque

$$G(x_N, t, x_0, 0) = \left(\frac{m}{2\pi ia} \right)^{\frac{N}{2}} \int \prod_{i=1}^{N-1} dx_i e^{iS}$$

con x_0 e x_N fissati. L'integrale scritto in questo modo non ha senso, è solo un'operazione formale. Esso ha senso solo se inteso come un modo conciso per indicare la procedura usata sopra: il prolungamento analitico del propagatore ritardato al tempo euclideo.

Limite al continuo. Si fa prima il limite al continuo e poi l'integrale. Visto che la simbologia è formale, si presume che il limite possa passare all'interno dell'integrale. Si definisce

$$\mathcal{L}(t) = \frac{1}{2} m \dot{x}^2 - V(x), \quad S = \int dt \mathcal{L}(t)$$

così come

$$\lim_{a \rightarrow 0} \left(\frac{2\pi ia}{m} \right)^{\frac{N}{2}} G(x_N, t, x_0, 0) = \int \prod_{i=1}^{N-1} dx_i e^{iS}$$

con x_0 e x_N fissati. L'azione S è l'azione classica del sistema. Il fattore davanti al propagatore è una costante di rinormalizzazione irrilevante che viene ignorata.

L'equazione sopra è un'equazione formale. Scritta in tale modo non è definita matematicamente. Per darle un senso bisogna

- considerare $a \neq 0$,
- calcolare il prolungamento analitico del propagatore G mediante l'integrale sui cammini,
- ritornare al tempo fisico,
- fare il limite per $a \rightarrow 0$.

Questo è il significato di tale espressione formale.

Interpretazione fisica dell'integrale sui cammini. Nella meccanica classica, un punto materiale si muove lungo una traiettoria $x(t)$. Essa estremizza l'azione

$$S = \int_0^t \mathcal{L} dt, \quad \mathcal{L} = \frac{1}{2} m \dot{x}^2 - V(x)$$

Le equazioni del moto di Eulero-Lagrange si ottengono variando ed estremizzando l'azione rispetto $\delta x(t)$ mantenendo fissi gli estremi

$$\delta x(0) = \delta x(t) = 0$$

da cui

$$d_t \partial_{\dot{x}} \mathcal{L} - \partial_x \mathcal{L} = 0$$

Questo implica che, in meccanica classica, il moto è determinato dalla forma dell'azione attorno al proprio estremo.

In meccanica quantistica, tutte le possibili traiettorie del punto materiale contribuiscono. L'ampiezza di probabilità è data da

$$\int \prod_i dx_i e^{i\frac{S}{\hbar}}$$

A causa della presenza della costante ridotta di Planck \hbar , non solo la forma dell'azione S è rilevante, ma anche il suo valore. Esso definisce il peso delle vere traiettorie.

Limite classico. Per un sistema classico si ha $|S| \gg \hbar$ e quindi, in generale, le fasi sono molto diverse per due traiettorie vicine: i contributi di diverse traiettorie, come funzioni di x_i , si cancellano tra loro e non contribuiscono. Equivalentemente, nel formalismo euclideo, le traiettorie che sono lontane dal minimo non contribuiscono poiché l'esponenziale è piccolo.

Scegliendo delle traiettorie vicino a quella classica si ha $\delta S = 0$ e quindi tali traiettorie hanno la stessa fase: la cancellazione non avviene. Nel formalismo euclideo, l'azione di tali cammini sono vicini al minimo e contribuiscono tutte le traiettorie? [r] con $\delta S = \hbar$ vicino al minimo.

Per un sistema classico contribuiscono solo i termini vicini all'estremo, vicini nel senso $\delta S \sim \hbar$. Con strumenti di misura macroscopici non ci si accorge di tali termini e dunque si misura una traiettoria deterministica.

Misurare la traiettoria con precisioni di \hbar su δS allora ci si accorge che la traiettoria non è più unica.

Per sistemi atomici si ha $|S| \sim \hbar$ e quindi tutti gli integrali sui cammini contribuiscono, non solo quelli vicini alla traiettoria classica. In questo caso, nessuna traiettoria è significativamente più importante di ogni altra.

La fenditura e la doppia fenditura si può fare bene con il path integral (v. Gasiorowicz).

Funzione di partizione di un sistema quantistico. Tipicamente non si considera l'ampiezza di probabilità $G(x, t, x_0, t_0)$, ma si calcola

$$\int dx_0 G(x_N, T, x_0, 0), \quad x_N = x_0$$

cioè la traccia del propagatore ritardato con condizioni periodiche

$$Z_a(T) = \text{Tr} [\hat{T}_a^N] = \int dx_0 G(x_N, T, x_0, 0), \quad t > t_0, \quad T = t - t_0, \quad x_N = x_0$$

La sua espressione integrale è

$$Z_a(T) = \left(\frac{m}{2\pi a} \right)^{\frac{N}{2}} \int \prod_{i=0}^{N-1} dx_i e^{-S_E}, \quad x_N = x_0$$

Il fattore di normalizzazione è spesso ignorato. Per i sistemi con azione S reale, l'integrale è identico alla funzione di partizione con fattore di Boltzmann e^{-S} . Inoltre risulta chiaro che si possono usare le tecniche di risoluzione sviluppate in meccanica statistica per calcolare l'integrale sopra. In particolare, i metodi Monte Carlo possono essere utilizzati per calcolare le sue funzioni di correlazione.

Inoltre, vale

$$Z_a(T) = \sum_n e^{-\mathcal{E}_n T}, \quad T > 0$$

Come per il propagatore ritardato, si può studiare la dipendenza di T_0 facendo la trasformata di Fourier ed è possibile calcolare tutti gli autovalori \mathcal{E}_n .

Se si potesse calcolare, si va a T grande, si estrae l'esponenziale dominante, cioè quello con autovalore minore, lo si sottrae alla funzione, poi si va a T grande, si estrae l'esponenziale dominante, cioè il secondo autovalore minore, lo si sottrae e così via.

Si fa l'oscillatore armonico per imparare le tecniche (e visto che c'è la soluzione analitica si confronta). Poi si può usare un potenziale anarmonico.

Lezione 5

11 Funzioni di correlazione

Gli integrali sopra non si fanno risolvere tramite il metodo Monte Carlo. Inoltre si calcolano solamente gli autovalori. Si è interessati agli elementi di matrice di osservabili su stati fisici.

ven 10 nov
2023 14:30

11.1 Prodotto ordinato temporale di operatori

In generale, in meccanica quantistica si è interessati al calcolo di elementi di matrice di operatori. In particolare di elementi tra autostati dell'hamiltoniana, cioè particelle. Per esempio dagli elementi di matrice di \hat{x} , \hat{x}^2 , \hat{x}^3 , etc, si può ricostruire la funzione d'onda

$$u_n(x) = \langle x | E_n \rangle$$

Si vede come estrarre gli elementi di matrice nel formalismo dell'integrale sui cammini. Si definisce il prodotto ordinato temporale (T-prodotto o tempo-ordinato) degli operatori $\hat{O}_1(t_1)$ e $\hat{O}_2(t_2)$ (in rappresentazione di Heisenberg) come

$$\langle E_0 | \mathcal{T}\{\hat{O}_1(t_1)\hat{O}_2(t_2)\} | E_0 \rangle \equiv \theta(t_1 - t_2) \langle E_0 | \hat{O}_1(t_1)\hat{O}_2(t_2) | E_0 \rangle + \theta(t_2 - t_1) \langle E_0 | \hat{O}_2(t_2)\hat{O}_1(t_1) | E_0 \rangle$$

dove $|E_0\rangle$ è lo stato di vuoto (la cui energia è tipicamente posta a zero o ad una fase banale). Se \hat{O}_i è lo stesso campo, allora essa viene detta funzione di correlazione a due punti. Da essa, segue

$$\begin{aligned} \langle E_0 | \mathcal{T}\{\hat{O}_1(t_1)\hat{O}_2(t_2)\} | E_0 \rangle &= \theta(t_1 - t_2) \langle E_0 | e^{i\hat{H}t_1} \hat{O}_1(0) e^{-i\hat{H}(t_1-t_2)} \hat{O}_2(0) e^{-i\hat{H}t_2} | 0 \rangle \\ &+ \theta(t_2 - t_1) \langle E_0 | e^{i\hat{H}t_2} \hat{O}_2(0) e^{-i\hat{H}(t_2-t_1)} \hat{O}_1(0) e^{-i\hat{H}t_1} | E_0 \rangle \end{aligned}$$

L'esponenziale nel mezzo è il propagatore ritardato \hat{G} che determina l'evoluzione temporale anche dei prodotti ordinati temporali.

Si inserisce l'identità

$$I = \sum_n |E_n\rangle\langle E_n|$$

tra i due operatori per ottenere

$$\begin{aligned} \langle E_0 | \mathcal{T}\{\hat{O}_1(t_1)\hat{O}_2(t_2)\} | E_0 \rangle &= \theta(t_1 - t_2) \sum_n \langle E_0 | \hat{O}_1(0) | E_n \rangle \langle E_n | \hat{O}_2(0) | E_0 \rangle e^{-i(E_n - E_0)(t_1 - t_2)} \\ &+ \theta(t_2 - t_1) \sum_n \langle E_0 | \hat{O}_2(0) | E_n \rangle \langle E_n | \hat{O}_1(0) | E_0 \rangle e^{-i(E_n - E_0)(t_2 - t_1)} \end{aligned}$$

Risulta chiaro che dalla dipendenza temporale della funzione di correlazione a due punti si possono estrarre gli autovalori e gli elementi di matrice.

Esempio. Per esempio, come il propagatore ritardato, si consideri la trasformazione di Fourier con $\text{Im } z > 0$:

$$\int_0^\infty dt e^{izt} \langle E_0 | \mathcal{T}\{\hat{O}_1(t)\hat{O}_2(0)\} | E_0 \rangle = i \sum_n \frac{1}{z - (E_n - E_0)} \langle E_0 | \hat{O}_1 | E_n \rangle \langle E_n | \hat{O}_2 | E_0 \rangle$$

dove $t = t_1 - t_2$. Scegliendo $\hat{O}_1 = \hat{O}_2$, dai poli o dai tagli della funzione si possono estrarre gli autovalori E_n e la densità di probabilità

$$|\langle E_0 | \hat{O}_1(0) | E_n \rangle|^2$$

Risulta ovvio come generalizzare il prodotto ordinato temporale per una funzione a tre operatori (dunque tre punti). Dalla funzione a tre punti si può estrarre l'elemento di matrice tra due stati generici, e così via.

Conclusione. Per determinare gli elementi di matrice (equivalenti alla funzione d'onda) si può studiare l'andamento temporale (o la loro trasformata di Laplace) del prodotto ordinato temporale di operatori. Dall'andamento temporale (o trasformata di Laplace) si possono estrarre anche gli autovalori dell'energia.

11.2 Funzioni di correlazione

La funzione di partizione è un integrale di funzioni numeriche. I prodotti ordinati temporali sono lo strumento per studiare non perturbativamente una teoria. I valori di aspettative dei prodotti ordinati temporali sono in relazione biunivoca con le funzioni di correlazione del sistema statistico associato.

Si utilizza il formalismo euclideo con le condizioni al bordo periodiche $x_N = x_0$. Si definisce

$$t_1 = i_1 a, \quad t_2 = i_2 a, \quad T = N a, \quad i_j \in \mathbb{N}$$

dove i è l'indice che identifica il passo temporale discretizzato. Successivamente, si potrebbe confondere con t [r]. Si aggiunge un'ipotesi semplificativa. Senza perdita di generalità, si considerano solo operatori diagonali nella base delle coordiante

$$\hat{O}_1 |x\rangle = O_1(x) |x\rangle, \quad \hat{O}_2 |x\rangle = O_2(x) |x\rangle$$

In meccanica statistica, la funzione di correlazione tra $O_1(x)$ e $O_2(x)$ è definita come

$$C(t_1, t_2) = \frac{\int \prod_{i=0}^{N-1} dx_i O_1(x_{i_1}) O_2(x_{i_2}) e^{-S}}{\int \prod_{i=0}^{N-1} dx_i e^{-S}}$$

In teoria dei campi, gli autovalori degli operatori sono i campi. Nel formalismo operatoriale, essa corrisponde a

$$C(t_1, t_2) = \frac{1}{Z_a(T)} \left[\theta(t_2 - t_1) \text{Tr} \left(\hat{T}_a^{N-|i_2-i_1|} \hat{O}_2 \hat{T}_a^{|i_1-i_2|} \hat{O}_1 \right) + \theta(t_1 - t_2) \text{Tr} \left(\hat{T}_a^{N-|i_1-i_2|} \hat{O}_1 \hat{T}_a^{|i_1-i_2|} \hat{O}_2 \right) \right]$$

il sistema si evolve poi subisce l'operatore \hat{O}_1 poi si evolve con \hat{T}_a , poi agisce l'operatore \hat{O}_2 e poi si evolve. Si noti che per la proprietà di ciclicità, si portano i primi operatori di evoluzione alla fine.

Utilizzare l'identità

$$\hat{I} = \int dx_i |x_i\rangle \langle x_i|$$

inserita a destra degli operatori \hat{O}_i , come fatto per il propagatore ritardato, permette di ottenere la definizione della meccanica statistica (perché gli operatori sono diagonali, dunque si possono sostituire con i propri autovalori, cioè numeri).

Relazione con prodotti ordinati temporali. Inserendo una volta l'identità

$$\hat{I} = \sum_n |\mathcal{E}_n\rangle \langle \mathcal{E}_n|$$

il denominatore diventa

$$Z_a(T) = \sum_n e^{-\mathcal{E}_n T}$$

mentre al numeratore, l'identità si pone prima del primo operatore \hat{O}_i a sinistra

$$Z_a(T) C(t_1, t_2) = \theta(t_2 - t_1) \sum_n e^{-\mathcal{E}_n (T-|t_1-t_2|)} \langle \mathcal{E}_n | \hat{O}_2 \hat{T}_a^{|t_1-t_2|} \hat{O}_1 | \mathcal{E}_n \rangle + \theta(t_1 - t_2) \sum_n e^{-\mathcal{E}_n (T-|t_1-t_2|)} \langle \mathcal{E}_n | \hat{O}_1 \hat{T}_a^{|t_1-t_2|} \hat{O}_2 | \mathcal{E}_n \rangle$$

Ipotizzando per convenienza che $\mathcal{E}_n > 0$, nel limite $T \rightarrow \infty$ (cioè il reticolo tende all'infinito) e a costante si ha

$$N \rightarrow \infty \implies Z_a(T) = e^{-\mathcal{E}_0 T}$$

cioè si ha un contributo solamente dall'autovalore più piccolo. Similmente a numeratore rimane solo $n = 0$ della somma: si ha il solo contributo del vuoto. Ricordando che nel formalismo euclideo gli operatori in descrizione di Heisenberg si evolvono come

$$\hat{O}(t) = \hat{T}^{-\frac{t}{a}} \hat{O} \hat{T}^{\frac{t}{a}}$$

allora, semplificando $Z_a(T)$, riportando l'esponenziale nell'elemento di matrice e riscrivendolo in termini di \hat{T}_a , si ha

$$\lim_{a \rightarrow 0} \lim_{\substack{T \rightarrow \infty \\ a = \text{cost}}} C(t_1, t_2) = \theta(t_2 - t_1) \langle E_0 | \hat{O}_2(t_2) \hat{O}_1(t_1) | E_0 \rangle \\ + \theta(t_1 - t_2) \langle E_0 | \hat{O}_1(t_1) \hat{O}_2(t_2) | E_0 \rangle$$

Nel formalismo euclideo, il contributo del vuoto nel numeratore e denominatore richiede soltanto il limite per $T \rightarrow \infty$.

Elementi di matrice dalla funzione a due punti. Si consideri la funzione di correlazione nel formalismo operatoriale nel limite $T \rightarrow \infty$. Si inserisce un'altra identità per ottenere

$$C(t_1, t_2) = \sum_m e^{-\mathcal{E}_m |t_2 - t_1|} [\theta(t_2 - t_1) \langle \mathcal{E}_0 | \hat{O}_2 | \mathcal{E}_m \rangle \langle \mathcal{E}_m | \hat{O}_1 | \mathcal{E}_0 \rangle \\ + \theta(t_1 - t_2) \theta(t_2 - t_1) \langle \mathcal{E}_0 | \hat{O}_1 | \mathcal{E}_m \rangle \langle \mathcal{E}_m | \hat{O}_2 | \mathcal{E}_0 \rangle]$$

vera perché T è grande. Facendo anche il limite della distanza temporale tra i due operatori $|t_2 - t_1| \rightarrow \infty$ si ha

$$\lim_{\substack{|t_1 - t_2| \rightarrow \infty \\ T \rightarrow \infty}} C(t_1, t_2) = e^{-\bar{\mathcal{E}} |t_2 - t_1|} [\theta(t_2 - t_1) \langle \mathcal{E}_0 | \hat{O}_2 | \bar{\mathcal{E}} \rangle \langle \bar{\mathcal{E}} | \hat{O}_1 | \mathcal{E}_0 \rangle \\ + \theta(t_1 - t_2) \langle \mathcal{E}_0 | \hat{O}_1 | \bar{\mathcal{E}} \rangle \langle \bar{\mathcal{E}} | \hat{O}_2 | \mathcal{E}_0 \rangle]$$

dove $|\bar{\mathcal{E}}\rangle$ è il primo stato (il più libero) per il quale la combinazione di elementi di matrice è diversa da zero (il motivo per cui rimane è analogo a sopra). Questo è valido per ogni coppia di operatori \hat{O}_1 e \hat{O}_2 e quindi si ha accesso a tutte le informazioni dinamica del sistema.

Non si usa direttamente il vuoto perché il valore di aspettazione potrebbe essere nullo.

Questa espressione è quella che si usa per calcolare la massa di un nucleo a partire dall'azione della cromodinamica quantistica.

Esempio 11.1. Si consideri un sistema invariante per trasformazione di parità. Gli stati si possono classificare in base alla parità

$$\mathcal{P} |\mathcal{E}_0\rangle = |\mathcal{E}_0\rangle, \quad \mathcal{P} |\mathcal{E}_1\rangle = |\mathcal{E}_1\rangle$$

ad esempio un oscillatore armonico. Ricordando che

$$\mathcal{P} \hat{x} \mathcal{P} |x\rangle = \mathcal{P} \hat{x} |-x\rangle = -x \mathcal{P} |-x\rangle = -\hat{x} |x\rangle \implies \mathcal{P} \hat{x} \mathcal{P} = -\hat{x}$$

da cui

$$\langle \mathcal{E}_0 | \hat{x} | \mathcal{E}_0 \rangle = \langle \mathcal{E}_0 | \mathcal{P} \hat{x} \mathcal{P} | \mathcal{E}_0 \rangle = -\langle \mathcal{E}_0 | \hat{x} | \mathcal{E}_0 \rangle \implies \langle \mathcal{E}_0 | \hat{x} | \mathcal{E}_0 \rangle = 0$$

Il vuoto non contribuisce, dunque il correlatore $\hat{x}\hat{x}$ è

$$\lim_{\substack{|t_1 - t_2| \rightarrow \infty \\ T \rightarrow \infty}} C_{xx}(t_1, t_2) = e^{-\mathcal{E}_1 |t_2 - t_1|} |\langle \mathcal{E}_0 | \hat{x} | \mathcal{E}_1 \rangle|^2$$

Studiando il correlatore a grandi distanze si estrae l'autovalore e l'elemento di matrice. Mentre il correlatore $\hat{x}^2 \hat{x}^2$ è

$$\lim_{\substack{|t_1 - t_2| \rightarrow \infty \\ T \rightarrow \infty}} C_{x^2 x^2}(t_1, t_2) = e^{-\mathcal{E}_0 |t_2 - t_1|} |\langle \mathcal{E}_0 | \hat{x}^2 | \mathcal{E}_1 \rangle|^2$$

I polinomi di Hermite si accoppiano con stati $|\mathcal{E}_i\rangle$ con indice i maggiore. Per altri stati si può studiare la funzione a tre punti oppure estrarre l'esponenziale sub-leading (cioè dell'ordine successivo a quello dominante) in quello a due punti.

Studiando l'andamento temporale dei correlatori si possono estrarre gli autovalori e gli elementi di matrice.

Per calcolare i correlatori si utilizza il campionamento di importanza in cui le variabili x sono estratte con probabilità e^{-S} e dove le osservabili sono $O_1(x_{i_1})O_2(x_{i_2})$.

Osservazione 11.2. Le teorie quantistiche relativistiche interessanti per la fisica, ossia le teorie quantistiche dei campi, si sanno definire in modo non perturbativo soltanto mediante l'integrale sui cammini. Per risolvere non perturbativamente teorie interessanti come la cromodinamica quantistica, si usano metodi Monte Carlo. La massa del protone, o del neutrone, si calcola da principi primi utilizzando tali metodi.

Osservazione 11.3. In meccanica quantistica, per quantizzare una teoria si scrive l'hamiltoniana ed i commutatori tra gli operatori. Lo stesso avviene in teoria dei campi. In questo caso, la quantizzazione è più semplice e data da

$$Z = \int \prod_i dx_i e^{-S}$$

dove S è l'azione classica. Ad esempio, per la teoria scalare $\lambda\varphi^4$

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\varphi)^2 + \frac{1}{2}m\varphi^2 + \frac{\lambda}{4!}\varphi^4$$

Allora la quantizzazione della teoria di campo è la formula sopra integrando nelle coordinate fondamentali (cioè il campo).

12 Oscillatore armonico

La soluzione dell'oscillatore armonico è analitica e si trova la versione discreta, diversa da quella analitica. La soluzione discreta va confrontata con il calcolo numerico.

Integrale gaussiano multidimensionale. L'integrale gaussiano multidimensionale è dato da

$$I(A, b) = \int \prod_{i=1}^n dx_i \exp \left[-\frac{1}{2} \sum_{ij} x_i A_{ij} x_j + \sum_{i=1}^n b_i x_i \right]$$

dove A è una matrice simmetrica reale con autovalori $\lambda_i > 0$ positivi. Definendo

$$x_i = \sum_j (A^{-1})_{ij} b_j + y_i \implies \prod_i dx_i = \prod_i dy_i, \quad x = A^{-1}b + y$$

e, notando che l'inversa di una matrice simmetrica è ancora una matrice simmetrica, si ottiene

$$\sum_{ij} x_i A_{ij} x_j = x^\top A x = (y^\top + b^\top A^{-1}) A (A^{-1}b + y) = y^\top A y + b^\top A^{-1}b + y^\top b + b^\top y$$

così come

$$\sum_i b_i x_i = b^\top x = b^\top A^{-1}b + b^\top y$$

Pertanto

$$\begin{aligned} I(A, b) &= \int \prod_i dy_i \exp \left[-\frac{1}{2} (y^\top A y + b^\top A^{-1}b + y^\top b + b^\top y) + b^\top A^{-1}b + b^\top y \right] \\ &= e^{\frac{1}{2} b^\top A^{-1}b} \int \prod_i dy_i e^{-\frac{1}{2} y^\top A y} \end{aligned}$$

Operando un cambio di variabile che diagonalizza la matrice (che si può fare poiché la matrice è simmetrica)

$$z = Ry, \quad R^\top R = I, \quad \det R = 1 = \det J \implies z^\top R^\top A R z = \sum_i \lambda_i z_i^2$$

si ha

$$I(A, b) = e^{\frac{1}{2} b^\top A^{-1}b} \int \prod_i dz_i e^{-\frac{1}{2} \sum_i z_i^2 \lambda_i} = e^{\frac{1}{2} b^\top A^{-1}b} \prod_i \sqrt{\frac{2\pi}{\lambda_i}} = (2\pi)^{\frac{n}{2}} (\det A)^{-\frac{1}{2}} e^{\frac{1}{2} b^\top A^{-1}b}$$

12.1 Integrale sui cammini

Considerando condizioni al bordo periodiche, sia

$$T = aN$$

La funzione di partizione è

$$Z_a(T) = \left(\frac{m}{2\pi a}\right)^{\frac{N}{2}} \int \prod_{i=0}^{N-1} dx_i e^{-S}, \quad x_N = x_0$$

L'azione è data da

$$S = a \sum_{i=0}^{N-1} \left[\frac{1}{2} m \left(\frac{x_{i+1} - x_i}{a} \right)^2 + \frac{1}{2} m \omega^2 x_i^2 \right]$$

Visto che l'integrale è quadratico allora si può svolgere. Si potrebbe fare l'integrale quadratico calcolando autovalori ed autovettori della matrice. Invece si segue una procedura ad hoc¹. Si definisce

$$\xi = 1 + \frac{1}{2} a^2 \omega^2, \quad y_i = \sqrt{\frac{m}{a}} x_i$$

da cui l'azione è

$$S = \sum_{i=0}^{N-1} \xi y_i^2 - \sum_{i=0}^{N-1} y_{i+1} y_i$$

[r] manca 1/2? Si considerano come variabili di integrazione tutte tranne y_0 e y_N (che sono anche identiche). Definendo la matrice $(N-1) \times (N-1)$

$$Q = \begin{bmatrix} 2\xi & -1 & 0 & & & \\ -1 & 2\xi & -1 & 0 & & \\ 0 & -1 & 2\xi & -1 & 0 & \\ & 0 & -1 & 2\xi & -1 & 0 \\ & & & \ddots & \ddots & \ddots \end{bmatrix}, \quad i = 1, \dots, N-1$$

si ha

$$y^\top \equiv (y_1, \dots, y_{N-1}), \quad b^\top = (y_0, 0, 0, \dots, 0, 0, y_0)$$

per cui l'azione risulta essere

$$S = \frac{1}{2} y^\top Q y + \xi y_0^2 - y_0(y_1 + y_{N-1}) = \frac{1}{2} y^\top Q y + \frac{1}{2} \xi b^\top b - y^\top b$$

Quindi l'integrale sui cammini si scrive come

$$Z_a(T) = \left(\frac{m}{2\pi a}\right)^{\frac{N}{2}} \left(\frac{a}{m}\right)^{\frac{N}{a}} \int dy_0 \int \prod_{i=1}^{N-1} dy_i \exp \left[- \left(\frac{1}{2} y^\top Q y + \frac{1}{2} \xi b^\top b - y^\top b \right) \right]$$

Sia l'ultimo integrale nominato I . Utilizzando il risultato dell'integrale gaussiano si ha

$$I(Q, b) = \int \prod_{i=1}^{N-1} dy_i \exp \left[- \frac{1}{2} y^\top Q y + y^\top b \right] = (2\pi)^{\frac{N-1}{2}} (\det Q)^{-\frac{1}{2}} e^{\frac{1}{2} b^\top Q^{-1} b}$$

pertanto

$$Z_a(T) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\det Q}} \int dy_0 \exp \left[- \frac{1}{2} b^\top (\xi - Q^{-1}) b \right]$$

Bisogna calcolare il determinante ed il termine $Q^{-1}b$.

¹Si veda Rothe, "Lattice Gauge Theories", §Test of the Energy Sum Rule. The Harmonic Oscillator.

Determinante. Si calcola il determinante tramite ricorsione. Si definisce

$$D_n \equiv \det Q$$

dove Q è la matrice sopra con dimensione $n \times n$. Dunque

$$D_0 = 1, \quad D_1 = 2\xi, \quad D_2 = (2\xi)^2 - 1 = 2\xi D_1 - D_0, \quad D_3 = (2\xi)^3 - 4\xi = 2\xi D_2 - D_1$$

La regola di ricorsione è

$$D_{n+1} = 2\xi D_n - D_{n-1}$$

cioè un'equazione differenziale discretizzata al secondo ordine nella variabile n . Definendo

$$\xi \equiv \cosh(a\tilde{\omega}) \implies a\tilde{\omega} = \ln \left(1 + a\bar{\omega} + \frac{1}{2}a^2\omega^2 \right), \quad \bar{\omega}^2 \equiv \omega^2 \left(1 + \frac{1}{4}a^2\omega^2 \right)$$

Le omega sono tutte uguali nel limite del continuo, le differenze sono dovute alla formula di Trotter cioè all'errore di discretizzazione. Si dimostra la relazione sopra cioè la soluzione all'equazione differenziale

$$2 \left(1 + \frac{1}{2}a^2\omega^2 \right) = e^{a\tilde{\omega}} + e^{-a\tilde{\omega}} \iff 2 \left(1 + \frac{1}{2}a^2\omega^2 \right) = x + \frac{1}{x}, \quad x = e^{a\tilde{\omega}}$$

da cui si ottiene

$$x = 1 + \frac{1}{2}a^2\omega^2 \pm a\bar{\omega} \implies a\tilde{\omega} = \ln \left(1 + a\bar{\omega} + \frac{1}{2}a^2\omega^2 \right)$$

dove si è presa la soluzione maggiore.

Quindi l'equazione ricorsiva si può scrivere come

$$D_{n+1} = [e^{a\tilde{\omega}} + e^{-a\tilde{\omega}}]D_n - D_{n-1}$$

La cui soluzione è

$$D_n = ce^{na\tilde{\omega}} + de^{-na\tilde{\omega}}, \quad c = \frac{e^{a\tilde{\omega}}}{2 \sinh(a\tilde{\omega})}, \quad d = -\frac{e^{a\tilde{\omega}}}{2 \sinh(a\tilde{\omega})}$$

Pertanto

$$D_n = \frac{\sinh[(n+1)a\tilde{\omega}]}{\sinh(a\tilde{\omega})}$$

Si verifica la correttezza

$$D_0 \equiv 1, \quad D_1 = \frac{\sinh(2a\tilde{\omega})}{\sinh(a\tilde{\omega})} = 2 \cosh(a\tilde{\omega}) = 2\xi$$

Eliminando il denominatore, per ricorsione si ha

$$\sinh[(n+2)a\tilde{\omega}] = 2 \cosh(a\tilde{\omega}) \sinh[(n+1)a\tilde{\omega}] - \sinh(na\tilde{\omega})$$

che è vera utilizzando le formule di somma delle funzioni iperboliche.

Quindi il determinante è

$$\det Q = \frac{\sinh[Na\tilde{\omega}]}{\sinh(a\tilde{\omega})}$$

Osservazione 12.1. Successivamente si vede che gli autovalori dell'oscillatore armonico nel continuo e nel discreto sono

$$E_n = \omega \left(n + \frac{1}{2} \right), \quad \mathcal{E}_n = \tilde{\omega} \left(n + \frac{1}{2} \right)$$

Secondo termine. Definendo

$$a = Q^{-1}b \implies Qa = b$$

si può risolvere per a il sistema di equazioni lineari. La matrice Q è sparsa (sparse matrix). Dato

$$a = (a_1, \dots, a_{N-1})$$

si ha

$$Qa = b \implies 2\xi a_1 - a_2 = y_0, \quad -a_{n+2} + 2\xi a_{n+1} - a_n = 0, \quad -a_{N-2} + 2\xi a_{N-1} = y_0$$

con $n = 1, \dots, N-3$. L'equazione di mezzo è uguale a quella per il determinante. Le altre fissano le costanti c e d . Allora

$$a_n = ce^{na\tilde{\omega}} + de^{-na\tilde{\omega}}, \quad c = \frac{y_0}{2 \sinh(Na\tilde{\omega})} [1 - e^{-Na\tilde{\omega}}], \quad d = -\frac{y_0}{2 \sinh(Na\tilde{\omega})} [1 - e^{Na\tilde{\omega}}]$$

da cui

$$a_n = \frac{y_0}{\sinh(Na\tilde{\omega})} [\sinh(na\tilde{\omega}) + \sinh((N-n)a\tilde{\omega})]$$

Pertanto, ricordando la forma di b^\top , si ha

$$b^\top Q^{-1}b = b^\top a = \frac{2y_0^2}{\sinh(Na\tilde{\omega})} [\sinh(a\tilde{\omega}) + \sinh((N-1)a\tilde{\omega})]$$

da cui

$$\begin{aligned} b^\top (\xi - Q^{-1})b &= \xi b^\top b - b^\top a \\ &= \frac{2y_0^2}{\sinh(Na\tilde{\omega})} [-\sinh(a\tilde{\omega}) - \sinh((N-1)a\tilde{\omega}) + \cosh(a\tilde{\omega}) \sinh(Na\tilde{\omega})] \\ &= 2y_0^2 \frac{\sinh(a\tilde{\omega})}{\sinh(Na\tilde{\omega})} [\cosh(Na\tilde{\omega}) - 1] \end{aligned}$$

Alla terza riga si sono manipolati gli ultimi due termini.

Funzione di partizione. Quindi l'integrale cercato per la funzione di partizione è un integrale gaussiano con una complicata espressione per la varianza

$$\int dy_0 \exp \left[-\frac{1}{2} b^\top (\xi - Q^{-1})b \right] = \left[\frac{2\pi \sinh(Na\tilde{\omega})}{2 \sinh(a\tilde{\omega}) [\cosh(Na\tilde{\omega}) - 1]} \right]^{\frac{1}{2}}$$

dove la varianza è

$$\sigma^{-2} = 2 \frac{\sinh(a\tilde{\omega})}{\sinh(Na\tilde{\omega})} [\cosh(Na\tilde{\omega}) - 1]$$

Quindi la funzione di partizione è data da

$$\begin{aligned} Z_a(t) &= \frac{1}{\sqrt{2\pi}} \left[\frac{\sinh(a\tilde{\omega})}{\sinh(Na\tilde{\omega})} \right]^{\frac{1}{2}} \left[\frac{2\pi \sinh(Na\tilde{\omega})}{2 \sinh(a\tilde{\omega}) [\cosh(Na\tilde{\omega}) - 1]} \right]^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{2 [\cosh(Na\tilde{\omega}) - 1]}} = [e^{Na\tilde{\omega}} + e^{-Na\tilde{\omega}} - 2]^{-\frac{1}{2}} \\ &= [e^{\frac{a}{2}N\tilde{\omega}} - e^{-\frac{a}{2}N\tilde{\omega}}]^{-1} = \frac{e^{-\frac{a}{2}N\tilde{\omega}}}{1 - e^{-aN\tilde{\omega}}} \end{aligned}$$

Osservazione 12.2. Nel limite del passo reticolare infinitesimo $a \rightarrow 0$, e fissando m e ω , si ha

$$\lim_{a \rightarrow 0} \tilde{\omega} = \lim_{a \rightarrow 0} \bar{\omega} = \omega$$

Gli autostati dell'energia e gli elementi di matrice hanno come limite le espressioni del continuo.

Osservazione 12.3. A parte l'energia del vuoto, la funzione di partizione sopra corrisponde alla funzione di partizione per un fotone. Essa è legata alla radiazione di corpo nero con lunghezza temporale $T = (k_B \Theta)^{-1}$ dove Θ è la temperatura. Per $a \rightarrow 0$

$$\lim_{a \rightarrow 0} Z_a(T) = \frac{e^{-\frac{\omega T}{2}}}{1 - e^{-\omega T}}, \quad T = \frac{1}{k_B \Theta}$$

Per i sistemi quantistici, la funzione di partizione corrisponde alla funzione dei sistemi statistici con temperatura data dall'espressione sopra. Il vuoto termico di una teoria è

$$\text{Tr} [e^{-\frac{\hat{H}}{k_B \beta}}]$$

[r]

Osservazione 12.4. Gli errori di discretizzazione nei livelli energetici sono determinati da $\tilde{\omega}$, mentre quelli negli elementi di matrice degli operatori da $\bar{\omega}$. Essi sono solitamente diversi.

Osservazione 12.5. Dallo stesso calcolo per la funzione di correlazione, si nota che le posizioni dei poli sono gli autovalori E_n , mentre i residui sono i prodotti degli elementi di matrice degli operatori.

Osservazione 12.6. Partendo dall'espressione della funzione di partizione

$$Z_a(T) = \sum_n e^{-\mathcal{E}_n T}$$

si ottiene

$$Z_a(T) = e^{-\frac{1}{2} \tilde{\omega} T} \sum_n (e^{-\tilde{\omega} T})^n = \sum_n e^{-T \tilde{\omega}^2 (n + \frac{1}{2})}$$

Si estrapolano (non interpolano) vari valori per ottenere ω al continuo in un grafico (a^2, Z) [r]
In QCD definita opportunamente, il termine leading va come a^2 . [r]

Formalismo Minkowskiano. Una volta risolto l'integrale sui cammini, si può tornare al tempo di Minkowski sviluppando in serie

$$Z_a^M(T) = \sum_{n=0}^{\infty} e^{-i a N \tilde{\omega} (n + \frac{1}{2})}$$

Nel limite al continuo con fissando $T = aN$, m e ω si ha

$$a \rightarrow 0, \quad \tilde{\omega} \rightarrow \omega, \quad Z_a^M(T) = \sum_{n=0}^{\infty} e^{-i T \omega (n + \frac{1}{2})}$$

Si applica la trasformata di Laplace

$$\int_0^{\infty} e^{i z t} Z_a^M(t) dt$$

la cui posizione è

$$E_n = \omega \left(n + \frac{1}{2} \right)$$

Lezione 6

ven 17 nov
2023 14:30

Parte IV

Catene di Markov e teorema ergodico

Si veda Feller e Parisi. Al fine di calcolare le funzioni di correlazione

$$\langle O_1(t_1) O_2(t_2) \rangle = \frac{\int \prod_{i=0}^{N-1} dx_i e^{-S} O_1(t_1) O_2(t_2)}{\int \prod_{i=0}^{N-1} dx_i e^{-S}}$$

bisogna trovare un metodo per campionare secondo una distribuzione

$$P(x) = \frac{e^{-S}}{\int \prod_{i=0}^{N-1} dx_i} e^{-S}$$

in modo automatico.

In precedenza ci si è occupate dell'estrazione di valori di \hat{x} indipendenti tra loro, cioè la probabilità di ottenere la sequenza

$$x_1, \dots, x_n, \quad P(x_1 \cdots x_n) = P_1 P_2 \cdots P_n$$

Ogni estrazione dipende dalla probabilità dello stato finale del sistema, ma mai dallo stato iniziale. Nelle catene di Markov, si considera la più semplice tra le generalizzazioni, ossia la probabilità in ogni estrazione dipende dallo stato in cui si trova il sistema e dallo stato finale $P_{1,2}$

$$P\{(E_{j_0}, E_{j_1}, \dots, E_{j_n})\} = a_{j_0} P_{j_0 j_1} P_{j_1 j_2} \cdots P_{j_{n-1} j_n}$$

dove P_{ij} è la probabilità condizionata. La distribuzione di probabilità cercata si ottiene dopo n passaggi.

13 Catene di Markov

Notazione. Il simbolo E_k indica i risultati definiti ed i possibili stati del sistema (l'indice k indica quante estrazioni sono state fatte). Ad esempio, per l'oscillatore armonico, ogni E_k è una ennupla (x_1, \dots, x_N) (ogni ennupla può avere elementi diversi). Esso può indicare molte variabili, mai solo una come in precedenza. Il simbolo P_{jk} è la probabilità di ottenere una transizione dallo stato E_j allo stato E_k . Infine, il simbolo a_k è la probabilità di ottenere lo stato E_k nell'estrazione iniziale.

Definizione 13.1. Una sequenza di estrazioni E_1, E_2, \dots si chiama catena di Markov se la probabilità di estrazione della sequenza è data da

$$P\{(E_{j_0}, E_{j_1}, \dots, E_{j_n})\} = a_{j_0} P_{j_0 j_1} P_{j_1 j_2} \cdots P_{j_{n-1} j_n}$$

dove a_{j_0} è la distribuzione di probabilità della prima estrazione e P_{jk} è la probabilità condizionata (fissa nel tempo di Minkowski) di ottenere E_k partendo da E_j .

Proposizione 13.2. Poiché a_k e P_{jk} sono probabilità, vale

$$\begin{aligned} a_k &\geq 0, \quad \forall k; \quad \sum_k a_k = 1 \\ P_{jk} &\geq 0, \quad \forall k, j; \quad \sum_k P_{jk} = 1, \quad \forall j \end{aligned}$$

Teorema 13.3. La distribuzione $P\{(E_{j_0}, E_{j_1}, \dots, E_{j_n})\} \geq 0$ è una distribuzione di probabilità nello spazio degli stati di $n+1$ estrazioni (ossia spazio di $n+1$ combinazioni). [r]

Dimostrazione. Applicando la proposizione sopra si ha

$$\sum_{j_0 \cdots j_n} a_{j_0} P_{j_0 j_1} P_{j_1 j_2} \cdots P_{j_{n-1} j_n} = \sum_{j_0 \cdots j_{n-1}} a_{j_0} P_{j_0 j_1} \cdots P_{j_{n-2} j_{n-1}} = \cdots = 1$$

□

Teorema 13.4. La probabilità $P\{(E_{j_0}, E_{j_1}, \dots, E_{j_m})\}$ è indipendente dal numero $n \geq m$ di estrazioni.

13.1 Matrici di probabilità di transizione

Definizione 13.5. Le probabilità di transizione possono essere rappresentate da una matrice di probabilità di transizione

$$P \equiv \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots \\ P_{21} & P_{22} & P_{23} & \cdots \\ P_{31} & P_{32} & P_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Essa è una matrice quadrata, finita o infinita (in base alla catena di Markov studiata).

Proposizione 13.6. Per i sistemi con un numero finito di stati, la matrice P è una matrice stocastica perché è quadrata e vale

$$[P]_{ij} \geq 0, \quad \forall ij, \quad \sum_j [P]_{ij} = 1$$

La matrice stocastica P e la distribuzione $[r]$ definiscono completamente la catena di Markov con gli stati E_0, E_1 , etc.

Proposizione 13.7 (non fatta). Dalla proprietà $[r]$

$$\sum_j [P]_{ij} = 1 \implies P v_I = v_I \implies v_I = (1, 1, 1, \dots)^T$$

cioè le matrici stocastiche hanno un autovalore 1 ed l'autovettore associato è v_I .

Proposizione 13.8. Definendo la norma di un vettore di stati? $[r]$ come

$$\|v\| \equiv \sum_j |v_j|$$

si ha

$$\|vP\| = \sum_j \left| \sum_i v_i P_{ij} \right| \leq \sum_{ij} |v_i P_{ij}| = \sum_{ij} |v_i| P_{ij} = \|v\|, \quad \forall v$$

dove si applica la disuguaglianza triangolare. Quindi

$$\|vP\| \leq \|v\|, \quad \forall v$$

Corollario 13.9. Questo implica che tutti gli autovalori sinistri sono $\lambda \leq 1$. Infatti

$$vP = \lambda v \implies \|vP\| = |\lambda| \|v\|$$

quindi la relazione precedente implica

$$|\lambda| \|v\| \leq \|v\| \implies |\lambda| \leq 1$$

13.2 Probabilità di transizione dopo un numero arbitrario di passi

Definizione 13.10. Si definisce la probabilità $P_{jk}^{(n)}$ di transizione dallo stato E_j allo stato E_k in esattamente n passi.

Proposizione 13.11 (Relazione di Chapman–Kolmogorov). Vale

$$P_{jk}^{(1)} = P_{jk}, \quad P_{jk}^{(2)} = \sum_{\nu} P_{j\nu} P_{\nu k}, \quad \dots, \quad P_{jk}^{(n+1)} = \sum_{\nu} P_{j\nu} P_{\nu k}^{(n)}$$

così come

$$P_{jk}^{(n+m)} = \sum_{\nu} P_{j\nu}^{(m)} P_{\nu k}^{(n)}$$

cioè il prodotto righe per colonne. Questa relazione è analoga a quella per il propagatore ritardato dell'oscillatore armonico.

Definizione 13.12. Matrice stocastica della probabilità composta è

$$[P^{(n)}]_{jk} \equiv P_{jk}^{(n)}$$

Proposizione 13.13. Vale

$$P^{(n)} = P^n, \quad P^{(n+1)} = P P^n, \quad P^{(n+m)} = P^{(m)} P^{(n)}$$

13.2.1 Esempio — parte prima

Definizione 13.14. Si definisce una catena di Markov per un sistema con due stati: 0 e 1, falso e vero. La matrice di probabilità di transizione è

$$P \equiv \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}, \quad \alpha, \beta > 0$$

Rappresentazione spettrale. Si cercano gli autovalori e gli autovettori della matrice. Il polinomio caratteristico è

$$\begin{aligned} \det \begin{bmatrix} 1-\alpha-\lambda & \alpha \\ \beta & 1-\beta-\lambda \end{bmatrix} &= (1-\alpha)(1-\beta) - \lambda(2-\alpha-\beta) + \lambda^2 - \alpha\beta \\ &= \lambda^2 - \lambda(2-\alpha-\beta) + (1-\alpha-\beta) = 0 \end{aligned}$$

da cui gli autovalori sono

$$\lambda = \frac{(2-\alpha-\beta) \pm (\alpha+\beta)}{2} \implies \lambda_1 = 1, \quad \lambda_2 = 1-\alpha-\beta$$

Autovettori. Gli autovettori associati sono

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix}$$

Dati gli autovettori, si può costruire la matrice Q del cambio di base per diagonalizzare la matrice di transizione P :

$$Q = \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix}, \quad Q^{-1} = \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ 1 & -1 \end{bmatrix}$$

La matrice di transizione diventa

$$P = Q \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{bmatrix} Q^{-1}$$

Matrice stocastica della probabilità composta. La matrice stocastica della probabilità composta è

$$P^n = Q \begin{bmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{bmatrix} Q^{-1}$$

Se il modulo di un autovalore è minore di 1, allora il suo peso viene soppresso per $n \rightarrow \infty$. La matrice si può riscrivere come

$$\begin{aligned} P^n &= \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ 1 & -1 \end{bmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta} \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta & \alpha \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix} \end{aligned}$$

Proposizione 13.15. Nel limite $n \rightarrow \infty$ si ha

$$\lim_{n \rightarrow \infty} P^n = \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} \equiv \begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_0 & \Pi_1 \end{bmatrix}$$

Proposizione 13.16 (I). Per costruzione, la matrice sopra è una matrice stocastica $\Pi_0 + \Pi_1 = 1$.

Proposizione 13.17 (II). Vale

$$\begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_0 & \Pi_1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Proposizione 13.18 (III). Per costruzione vale

$$\begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_0 & \Pi_1 \end{bmatrix} = \begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_0 & \Pi_1 \end{bmatrix} P \implies (\Pi_0, \Pi_1) = (\Pi_0, \Pi_1) P$$

cioè è autovettore sinistro della matrice i cui elementi sono una distribuzione di probabilità [r] detta anche distribuzione di probabilità asintotica di P ed è pari a

$$\Pi = \frac{e^{-S}}{Z}$$

Bisogna definire la matrice P di modo che la sua distribuzione asintotica sia quella cercata.

Proposizione 13.19 (IV). Data una distribuzione iniziale

$$a = [a_0, a_1] \mid a_0 + a_1 = 1$$

segue

$$(a_0, a_1) \begin{bmatrix} \Pi_0 & \Pi_1 \\ \Pi_0 & \Pi_1 \end{bmatrix} = (\Pi_0, \Pi_1)$$

La distribuzione di probabilità dello stato finale (la distribuzione asintotica) è indipendente dalla distribuzione di probabilità dello stato iniziale.

Osservazione 13.20. Risulta chiaro che per estrarre eventi con probabilità Π_i , le catene di Markov sono un metodo adatto.

13.3 Classificazione degli insiemi di stati

Definizione 13.21. Un insieme di stati C è chiuso se nessuno stato fuori di C può essere raggiunto da qualunque stato in C tramite una catena di Markov.

Definizione 13.22. Una catena di Markov è irriducibile se l'unico insieme chiuso è quello che contiene ogni stato.

Proposizione 13.23. Una catena è irriducibile se e solo se ogni stato può essere raggiunto da ogni altro stato.

Esempio 13.24. Si considerino degli stati E_1, \dots, E_ρ tali che E_1, \dots, E_r con $r < \rho$ formino un insieme chiuso. Allora la matrice di transizione è

$$P = \begin{bmatrix} Q & 0 \\ U & V \end{bmatrix} \implies P^n = \begin{bmatrix} Q^n & 0 \\ U^n & V^n \end{bmatrix}$$

dove Q è la sottomatrice di transizione $r \times r$. Le probabilità di transizione tra stati nell'insieme chiuso si può calcolare solo come somma sugli stati dell'insieme chiuso. Lo stesso vale per V .

Osservazione 13.25. La richiesta che una catena sia irriducibile serve per eliminare la possibilità che la distribuzione asintotica $\lim_{n \rightarrow \infty} P^n$ presenti degli zero o [r] due sistemi allo stesso tempo.

13.4 Classificazione degli stati

Definizione 13.26. Sia $f_{jk}^{(n)}$ la probabilità che sia presente un processo che parta dallo stato E_j e arrivi per la prima volta allo stato E_k in n passi. Per convenzione si pone $f_{jk}^{(0)} = 0$.

Osservazione 13.27. Tale probabilità non è la probabilità che di arrivare in E_k al passo n poiché tale probabilità è descritta da $P_{jk}^{(n)}$, ma è la probabilità di arrivarci per la prima volta al passo n .

Definizione 13.28. La probabilità di partire da E_j e arrivare a E_k è

$$f_{jk} \equiv \sum_{n=1}^{\infty} f_{jk}^{(n)} \leq 1$$

Definizione 13.29. Quando $f_{jk} = 1$, la probabilità $f_{jk}^{(n)}$ è una distribuzione di probabilità nella variabile n , detta distribuzione di primo passaggio. Si può definire il tempo medio di ricorrenza

$$\mu_j \equiv \sum_{n=1}^{\infty} n f_{jj}^{(n)}$$

Relazione tra le due probabilità ad un passo arbitrario. Utilizzando le definizioni precedenti con le convenzioni

$$P_{kk}^{(0)} = 1 \quad f_{jk}^{(0)} = 0$$

segue

$$\begin{aligned} P_{jk}^{(1)} &= f_{jk}^{(1)} P_{kk}^{(0)} \\ P_{jk}^{(2)} &= f_{jk}^{(2)} P_{kk}^{(0)} + f_{jk}^{(1)} P_{jk}^{(1)} \\ P_{jk}^{(3)} &= f_{jk}^{(3)} P_{kk}^{(0)} + f_{jk}^{(2)} P_{kk}^{(1)} + f_{jk}^{(1)} P_{jk}^{(2)} \\ &\vdots \\ P_{jk}^{(n)} &= \sum_{\nu=1}^n f_{jk}^{(\nu)} P_{kk}^{(n-\nu)}, \quad n > 0 \end{aligned}$$

Questa relazione definisce $f_{jk}^{(\nu)}$ in modo implicito.

Classificazione. Gli stati si possono classificare come

- uno stato E_j è persistente se $f_{jj} = 1$,
- uno stato E_j è transiente se $f_{jj} < 1$,
- uno stato persistente E_j è nullo se il tempo medio di ricorrenza diverge $\mu_j \rightarrow \infty$,
- uno stato E_j è periodico con periodo $t > 1$ intero se $P_{jj}^{(n)} = 0$ tranne per $n = \nu t$ e t è l'intero più grande con questa probabilità,
- uno stato E_j è aperiodico se non è periodico,
- uno stato E_j è ergodico se è persistente con $\mu_j < \infty$ e aperiodico.

Teorema 13.30 (Ergodico). Si veda Parisi, p. 110?.

I. Uno stato E_j è transiente se e solo se

$$\sum_{n=0}^{\infty} P_{jj}^{(n)} < \infty$$

In tal caso

$$\sum_{n=1}^{\infty} P_{ij}^{(n)} < \infty, \quad \forall i$$

II. Uno stato E_j è uno stato persistente nullo se e solo se

$$\sum_{n=0}^{\infty} P_{jj}^{(n)} \rightarrow \infty, \quad \lim_{n \rightarrow \infty} P_{jj}^{(n)} = 0$$

In tal caso

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = 0, \quad \forall i$$

III. Uno stato persistente aperiodico è ergodico se e solo se il tempo medio di ricorrenza è finito $\mu_j < \infty$. In tal caso

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = f_{ij} \mu_j^{-1}$$

Si noti che come il tempo medio abbia un solo indice.

Dimostrazione. Si definiscono le funzioni generatrici per P_{jk} ed f_{jk}

$$P_{jk}(s) \equiv \sum_{n=0}^{\infty} P_{jk}^{(n)} s^n, \quad F_{jk}(s) \equiv \sum_{n=0}^{\infty} f_{jk}^{(n)} s^n$$

La relazione di ricorrenza si può trovare come

$$P_{jj}^{(n)} = \sum_{\nu=0}^n f_{jj}^{(\nu)} P_{jj}^{(n-\nu)} + \delta_{n,0}$$

Risulta chiaro che questa è la convoluzione discreta di ν ed n . Dunque

$$\begin{aligned} P_{jj}(s) &= \sum_{n=0}^{\infty} \sum_{\nu=0}^n f_{jj}^{(\nu)} P_{jj}^{(n-\nu)} s^n + 1 = \sum_{\nu=0}^{\infty} \sum_{n=\nu}^{\infty} f_{jj}^{(\nu)} P_{jj}^{(n-\nu)} s^n + 1 \\ &= \sum_{\nu=0}^{\infty} f_{jj}^{(\nu)} s^{\nu} + \sum_{n=\nu}^{\infty} P_{jj}^{(n-\nu)} s^{n-\nu} + 1, \quad m \equiv n - \nu \\ &= \sum_{\nu=0}^{\infty} f_{jj}^{(\nu)} s^{\nu} + \sum_{m=0}^{\infty} P_{jj}^{(m)} s^m + 1 = F_{jj}(s) P_{jj}(s) + 1 \\ &= \frac{1}{1 - F_{jj}(s)} \end{aligned}$$

alla seconda uguaglianza della prima riga si cambiano i limiti di integrazione poiché si mantiene il dominio di somma lo stesso. Ripetendo il calcolo per $j \neq k$ e ricordando che $P_{jk}^{(0)} = 0$ in tal caso, si ha

$$P_{jk}^{(n)} = \sum_{\nu=0}^n f_{jk}^{(\nu)} P_{kk}^{(n-\nu)}$$

da cui

$$P_{jk}(s) = F_{jk}(s) P_{kk}(s) = \frac{F_{jk}(s)}{1 - F_{kk}(s)}$$

I. Per uno stato transiente si ha

$$f_{jj} = \sum_{n=0}^{\infty} f_{jj}^{(n)} < 1 \iff F_{jj}(1) < 1 \iff F_{jj}(1) < \infty \iff \sum_{n=0}^{\infty} P_{jj}^{(n)} < \infty$$

Inoltre

$$P_{jk}(1) = \frac{F_{jk}(1)}{1 - F_{kk}(1)} \implies \sum_{n=0}^{\infty} P_{jk}^{(n)} < \infty, \quad \forall j$$

II. Per uno stato persistente nullo si ha

$$f_{jj} = 1 \iff F_{jj}(1) = 1 \iff P_{jj}(1) \rightarrow \infty \iff \sum_{n=0}^{\infty} P_{jj}^{(n)} \rightarrow \infty$$

così come

$$\sum_{n=0}^{\infty} P_{jk}^{(n)} \rightarrow \infty$$

Utilizzando il teorema successivo si ha

$$\lim_{n \rightarrow \infty} P_{jj}^{(n)} = \lim_{s \rightarrow 1^-} (1 - s)[P_{jj}(s) - 1]$$

mentre dai risultati precedenti segue

$$\begin{aligned} P_{jj}(s)A &= [1 - F_{jj}(1) - F'_{jj}(1)(s - 1) - F''_{jj}(1)(s - 1)^2 + \dots]^{-1} \\ &= -[F'_{jj}(1)(s - 1) + F''_{jj}(1)(s - 1)^2 + \dots]^{-1} \end{aligned}$$

sapendo che $F_{jj}(1) = 1$. Quindi

$$\lim_{n \rightarrow \infty} P_{jj}^{(n)} = \frac{1}{F'_{jj}(1)} = \frac{1}{\mu_j}$$

dove si ha

$$F'_{jj}(s) = \sum_{n=0}^{\infty} n f_{jj}^{(n)} s^{n-1} \implies F'_{jj}(1) = \sum_{n=0}^{\infty} n f_{jj}^{(n)} = \mu_j$$

Ripetendo i calcoli per $P_{jk}^{(n)}$ e ricordando

$$P_{jk}(s) = \frac{F_{jk}(s)}{1 - F_{kk}(s)}$$

segue

$$\lim_{n \rightarrow \infty} P_{jk}^{(n)} = \frac{f_{jk}}{\mu_k}$$

Se lo stato è persistente nullo, allora $\mu_j \rightarrow \infty$ e quindi

$$\lim_{n \rightarrow \infty} P_{jj}^{(n)} = 0, \quad \lim_{n \rightarrow \infty} P_{jk}^{(n)} = 0$$

III. Uno stato è ergodico $\mu_j < \infty$ se e solo se

$$\lim_{n \rightarrow \infty} P_{jk}^{(n)} = \frac{f_{jk}}{\mu_k}$$

Si noti che [r]. □

Teorema 13.31. Vale

$$\lim_{n \rightarrow \infty} P^{(n)} = \lim_{s \rightarrow 1^-} (1-s) \sum_{k=1}^{\infty} P^{(k)} s^k$$

dove $P^{(r)}$ è una successione limitata. Più precisamente, vale se esiste il limite da uno dei due lati, allora l'altro limite esiste e sono uguali.

Dimostrazione. Si dimostra il teorema in un senso. Si ipotizza

$$\lim_{n \rightarrow \infty} P^{(n)} = A$$

e si verifica che l'altro limite esista e sia pari ad A . Per la dimostrazione del viceversa, vedere Feller. Un teorema generale di analisi afferma che se esiste il limite sopra per $n > n(\varepsilon)$, allora

$$|P(n) - A| < \varepsilon$$

Ricordando la serie geometrica

$$\sum_{k=0}^{\infty} s^k = \frac{1}{1-s}$$

segue

$$\begin{aligned} \Delta &= \left| \sum_{k=0}^{\infty} P^{(k)} s^k - \frac{A}{1-s} \right| = \left| \sum_{k=0}^{\infty} P^{(k)} s^k - \sum_{k=0}^{\infty} A s^k \right| = \left| \sum_{k=0}^{\infty} (P^{(k)} - A) s^k \right| \\ &\leq \left| \sum_{k=0}^{n(\varepsilon)} (P^{(k)} - A) s^k \right| + \left| \sum_{k=n(\varepsilon)}^{\infty} (P^{(k)} - A) s^k \right| \\ &\leq \sum_{k=0}^{n(\varepsilon)} |P^{(k)} - A| s^k + \sum_{k=n(\varepsilon)}^{\infty} |P^{(k)} - A| s^k \\ &\leq n(\varepsilon) + \varepsilon \sum_{k=n(\varepsilon)}^{\infty} s^k \leq n(\varepsilon) + \frac{\varepsilon}{1-s} \end{aligned}$$

alla quarta riga si è applicato il fatto che $|P^{(k)} - A| \leq 1$ così come $A \leq 1$. Questo implica che

$$\left| (1-s) \sum_{k=0}^{\infty} P^{(k)} s^k - A \right| \leq n(\varepsilon)(1-s) + \varepsilon$$

Nel limite $s \rightarrow 1^-$ si ha

$$\lim_{s \rightarrow 1^-} \left| (1-s) \sum_{k=0}^{\infty} P^{(k)} s^k - A \right| \leq \varepsilon$$

cioè la tesi. □

13.4.1 Esempio — parte seconda

Ricordando che

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

segue

$$f_{00}^{(n)} = \alpha(1-\beta)^{n-2}\beta, \quad f_{11}^{(n)} = \beta(1-\alpha)^{n-2}\alpha, \quad n = 2, 3, 4, \dots$$

così come

$$f_{01}^{(n)} = (1-\alpha)^{n-1}\alpha, \quad f_{10}^{(n)} = (1-\beta)^{n-1}\beta, \quad n = 1, 2, 3, \dots$$

Quindi il tempo medio di ricorrenza è dato da

$$\mu_0 = 1 - \alpha - \sum_{n=2}^{\infty} n f_{00}^{(n)} = 1 - \alpha + \alpha\beta \sum_{n=2}^{\infty} n(1-\beta)^{n-2}$$

Utilizzando la serie geometrica

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k, \quad \frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$$

e sostituendo $m = n - 1$ si ha

$$\sum_{n=2}^{\infty} n(1-\beta)^{n-2} = \sum_{m=1}^{\infty} (m+1)(1-\beta)^{m-1} = \sum_{m=1}^{\infty} m(1-\beta)^{m-1} + \sum_{m=0}^{\infty} (1-\beta)^m = \frac{1}{\beta^2} + \frac{1}{\beta}$$

Pertanto

$$\mu_0 = 1 - \alpha + \alpha\beta \left[\frac{1}{\beta} + \frac{1}{\beta^2} \right] = 1 - \alpha + \alpha \left[1 + \frac{1}{\beta} \right] = 1 + \frac{\alpha}{\beta} = \frac{\alpha + \beta}{\beta}$$

In modo analogo si ha

$$\mu_1 = 1 - \beta + \sum_{n=2}^{\infty} n f_{11}^{(n)} = 1 - \beta + \alpha\beta \sum_{n=2}^{\infty} n(1-\alpha)^{n-2} = \frac{\alpha + \beta}{\alpha}$$

Pertanto, la probabilità di partire da E_0 ed arrivare in E_0 è

$$f_{00} = \sum_{n=0}^{\infty} f_{00}^{(n)} = 0 + 1 - \alpha + \alpha\beta \sum_{n=2}^{\infty} (1-\beta)^{n-2} = 1 - \alpha + \alpha\beta \frac{1}{\beta} = 1$$

Similmente per scambio $\alpha \leftrightarrow \beta$ si ha

$$f_{11} = 1$$

I termini misti sono

$$f_{01} = \alpha \sum_{n=1}^{\infty} (1-\alpha)^{n-1} = \frac{\alpha}{\alpha} = 1, \quad f_{10} = 1$$

Unendo quanto trovato si ha

$$\lim_{n \rightarrow \infty} P^{(n)} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix}$$

così come

$$f_{00} = f_{11} = f_{10} = f_{01} = 1$$

pure

$$\mu_0 = \frac{\alpha + \beta}{\beta}, \quad \mu_1 = \frac{\alpha + \beta}{\alpha}$$

Pertanto

$$\lim_{n \rightarrow \infty} P_{jk}^{(n)} = \frac{f_{jk}}{\mu_k}$$

a conferma della tesi del teorema ergodico.

Osservazione 13.32. Gli stati di questa catena di Markov sono tutti ergodici cioè persistenti, aperiodici e con tempo medio finito $\mu_j < \infty$. Inoltre, $f_{ij} = 1$ per ogni ij .

Lezione 7

13.5 Distribuzioni all'equilibrio

ven 24 nov
2023 14:30

Le catene a cui si è interessati trattano solo stati ergodici.

Definizione 13.33. Due stati di una catena sono dello stesso tipo se hanno le stesse caratteristiche definite in precedenza.

Teorema 13.34. Tutti gli stati di una catena sono dello stesso tipo.

Dimostrazione. Si veda Feller, vol. 1, p. 391. □

Teorema 13.35. Per uno stato persistente E_j esiste un solo insieme chiuso C irriducibile che contiene tale stato E_j e tale che

$$\forall E_i, E_k \in C, \quad f_{ik} = f_{ki} = 1, \quad \forall i, k$$

Dimostrazione. Si supponga che E_k sia uno stato raggiungibile da E_j . Allora la probabilità di passare da E_j a E_k senza ripassare da E_j è diversa da zero e vale α . Una volta in E_k , la probabilità di non tornare più in E_j è $1 - f_{kj}$. Quindi, la probabilità che, partendo da E_j , non si torni mai in E_j è almeno

$$\alpha(1 - f_{kj})$$

Per definizione, uno stato persistente ha probabilità di non ritorno pari a zero e quindi $\alpha \neq 0$. Dunque $f_{kj} = 1$.

Sia C l'insieme di tutti gli stati che possono essere raggiunti da E_j . Due stati E_i ed E_k in C possono essere raggiunti l'uno dall'altro tramite E_j e quindi ogni stato in C può essere raggiunto da ogni altro stato in C . Pertanto C è irriducibile per quanto definito in precedenza. Tutti gli stati sono persistenti e si può applicare la prima parte della dimostrazione ad ogni stato E_i per cui si ottiene

$$E_{ik} = 1$$

□

13.6 Distribuzione invariante per un insieme di stati ergodici

Teorema 13.36. I. In una catena irriducibile con soltanto stati ergodici, il limite

$$\Pi_k = \lim_{n \rightarrow \infty} P_{jk}^{(n)}$$

esiste ed è indipendente dallo stato iniziale j . Inoltre

$$\sum_k \Pi_k = 1, \quad \Pi_k = \sum_i \Pi_i P_{ik}$$

II. Viceversa, se la catena è irriducibile e aperiodica, ed esiste Π_k tale che

$$\sum_k \Pi_k = 1 \quad \Pi_k = \sum_i \Pi_i P_{ik}$$

allora tutti gli stati sono ergodici e vale

$$\lim_{n \rightarrow \infty} P_{jk}^{(n)} = \Pi_k, \quad \Pi_k = \frac{1}{\mu_k}$$

con μ_k il tempo medio di ricorrenza.

Il viceversa è più utile perché permette di costruire la matrice di transizione a partire dalla distribuzione asintotica (che nel caso desiderato è e^{-S}).

Dimostrazione. I. Per i sistemi ergodici vale

$$\lim_{n \rightarrow \infty} P_{jk}^{(n)} = \frac{f_{jk}}{\mu_k}$$

e si ricorda la relazione di Chapman–Kolmogorov

$$P_{jk}^{(n+1)} = \sum_{\nu} P_{j\nu}^{(n)} P_{\nu k}$$

Per un numero finito di stati si avrebbe

$$\lim_{n \rightarrow \infty} P_{jk}^{(n+1)} = \Pi_k \iff \lim_{n \rightarrow \infty} \sum_{\nu} P_{j\nu}^{(n)} P_{\nu k} = \sum_{\nu} \Pi_{\nu} P_{\nu k}$$

Per un numero infinito di stati si utilizza il lemma seguente per avere

$$\lim_{n \rightarrow \infty} \sum_j P_{ij}^{(n)} P_{jk} \geq \sum_j \lim_{n \rightarrow \infty} P_{ij}^{(n)} P_{jk} = \sum_j \Pi_j P_{jk}$$

Dunque

$$\Pi_k - \sum_j \Pi_j P_{jk} \geq 0$$

Per ogni k il membro di sinistra è un numero non negativo. Vale anche

$$\sum_k \Pi_k - \sum_j \Pi_j \sum_k P_{jk} = \sum_k \Pi_k - \sum_j \Pi_j = 0$$

Dato che sono tutti numeri non negativi, allora ogni addendo dev'essere nullo, quindi

$$\Pi_k = \sum_j \Pi_j P_{jk}$$

Inoltre, per ogni n ed i si ha

$$\sum_k P_{ik}^{(k+1)} = 1 \implies \lim_{k \rightarrow \infty} 1 = 1 \implies \sum_k \Pi_k = 1$$

II. Se

$$\sum_k \Pi_k = 1, \quad \Pi_k = \sum_i \Pi_i P_{ik}$$

allora vale

$$\Pi_k = \sum_j \Pi_j P_{jk}^{(n)}, \quad \forall n > 1$$

Se gli stati sono transienti o nulli, allora $P_{jk}^{(n)} \rightarrow 0$ e dunque

$$\Pi_k = 0, \quad \forall k$$

che contraddice $\sum_k \Pi_k = 1$. Quindi gli stati sono ergodici. Analogamente a prima, si può fare il limite [r] per avere

$$\lim_{k \rightarrow \infty} \Pi_k = \Pi_k, \quad \lim_{n \rightarrow \infty} \sum_j \Pi_j P_{jk}^{(n)} = \sum_j \Pi_j \frac{1}{\mu_k}$$

poiché i Π_j si sommano ad 1 si ha

$$\Pi_k = \frac{1}{\mu_k}$$

Pertanto, si conclude che

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \Pi_k, \quad \sum_k \Pi_k = 1, \quad \Pi_k = \sum_j \Pi_j P_{jk}, \quad \Pi_k = \frac{1}{\mu_k}$$

□

Lemma 13.37. Date delle quantità $P^{(n)}$ non negative, vale

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P_i^{(n)} \geq \sum_{i=1}^{\infty} \lim_{n \rightarrow \infty} P_i^{(n)}$$

Dimostrazione. Poiché le probabilità $P_i^{(n)}$ sono definite positive, vale

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P_i^{(n)} &\geq \sum_{i=1}^k P_i^{(n)}, \quad \forall n \\ \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P_i^{(n)} &\geq \lim_{n \rightarrow \infty} \sum_{i=1}^k P_i^{(n)} \\ \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P_i^{(n)} &\geq \sum_{i=1}^k \lim_{n \rightarrow \infty} P_i^{(n)} \end{aligned}$$

Nel limite $k \rightarrow \infty$ si ha

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P_i^{(n)} \geq \sum_{i=1}^{\infty} \lim_{n \rightarrow \infty} P_i^{(n)}$$

□

Osservazione 13.38. Con le catene di Markov si possono generare distribuzioni Π_k indipendentemente dal dato iniziale.

13.7 Riassunto sulle catene

Definizione 13.39. Uno stato è ergodico se e solo se è persistente, aperiodico e con un tempo medio di ricorrenza finito, ossia

$$f_{jj} = 1, \quad \mu_j = \sum_n n f_{jj}^{(n)} < \infty$$

Un insieme chiuso di stati ergodici di una catena irriducibile è detto ergodico.

Definizione 13.40. Una distribuzione di probabilità Π_k che soddisfa

$$\sum_k \Pi_k = 1, \quad \Pi_k = \sum_j \Pi_j P_{jk}, \quad \forall k$$

è detta invariante o stazionaria per una data catena di Markov.

Teorema 13.41. Una catena aperiodica e irriducibile ha una distribuzione invariante Π_k se e solo se è ergodica. In questo caso

$$\Pi_k > 0, \quad \forall k$$

La probabilità di transizione tende a Π_k indipendentemente dalla distribuzione iniziale.

13.8 Bilancio dettagliato

L'obiettivo è costruire una catena di Markov che sia ergodica e che abbia la distribuzione asintotica Π_k desiderata.

Teorema 13.42. Data una distribuzione di probabilità

$$\mathcal{P}_i = \frac{R_i}{\sum_i R_i}, \quad R_i \geq 0, \quad \forall i$$

e data una catena di Markov definita da P_{ij} , condizione sufficiente affinché la catena sia ergodica

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \mathcal{P}_i$$

è che P_{ij} soddisfi il bilancio dettagliato

$$R_i P_{ij} = R_j P_{ji}$$

La probabilità di andare in una direzione e tornare indietro è la stessa.

Dimostrazione. Per definizione

$$\sum_j P_{ij} = 1, \quad \forall i$$

dunque

$$\sum_j R_i P_{ij} = \sum_j R_j P_{ji} \implies R_i = \sum_j R_j P_{ji}$$

dividendo per la somma $\sum_i R_i$ si ha

$$\mathcal{P}_i = \sum_j \mathcal{P}_j P_{ji}$$

cioè è un autovettore sinistro con autovalore 1. Per una catena di Markov ergodica, Π_j è unico e dunque $\mathcal{P}_j = \Pi_j$ (si veda definizione e teorema precedenti?) [r]. \square

Osservazione 13.43. La condizione per il bilancio dettagliato non determina la probabilità P_{ij} di transizione in modo univoco, quindi si ha la libertà di [r] algoritmi efficienti per il problema particolare che si studia.

14 Algoritmo di Metropolis

Dato uno stato E_j [r] uno stato E_k con probabilità Q_{jk} dove Q soddisfa la condizione di micro-reversibilità [r]

$$Q_{jk} = Q_{kj}$$

Per generare eventi con distribuzione

$$\mathcal{P}_i = \frac{R_i}{\sum_i R_i}$$

si accetta la configurazione generata al passo precedente con probabilità diverse. Per

$$\frac{R_k}{R_j} \geq 1$$

si accetta con probabilità 1, mentre per

$$\frac{R_k}{R_j} < 1$$

si accetta con probabilità $\frac{R_k}{R_j}$. [r]

Dimostrazione bilancio dettagliato [r] Nel primo caso

$$P_{jk} = Q_{jk}, \quad P_{kj} = Q_{kj} \frac{R_j}{R_k} = Q_{jk} \frac{R_j}{R_k}$$

quindi

$$P_{jk} = Q_{jk} = P_{kj} \frac{R_k}{R_j} \implies R_j P_{jk} = R_k P_{kj}$$

Nel secondo caso

$$P_{jk} = Q_{jk} \frac{R_k}{R_j}, \quad R_{kj} = Q_{kj}$$

[r]

Metropolis per oscillatore armonico. Per passare da uno stato E al tempo Markoviano k allo stato E' al tempo $k + 1$, si propone

$$x_i - \Delta \leq x'_i < x_i + \Delta$$

con distribuzione uniforme. Pertanto

$$x'_i = x_i + 2\Delta \left(r_1 - \frac{1}{2} \right)$$

dove r_1 è un numero casuale con distribuzione uniforme $P(r_1)$. Questa implementa

$$Q_{x_i, x'_i} = \begin{cases} 1, & x_i - \Delta \leq x'_i < x_i + \Delta \\ 0, & \text{altrove} \end{cases}$$

fatto per una componente alla volta. Il secondo passo dell'algoritmo metropolis considera

$$\frac{R_{x'}}{R_x} = \frac{e^{-S(x')}}{e^{-S(x)}} > 1 \implies \Delta S \leq 0$$

che viene accettato dove x' è la ennupla con dentro x'_i . Se tale configurazione è minore di 1, allora si accetta con probabilità data da tale rapporto. Se questo è il caso, si estrae un altro numero casuale uniforme, e se è minore (o uguale) di tale rapporto, allora si accetta, altrimenti si rifiuta la configurazione e si rimette la x_i iniziale. Le due condizioni si possono mettere insieme e considerare solo l'estrazione del secondo numero casuale (quindi non controllare se > 1).

Si cambia una coordinata alla volta e il parametro Δ dipendono da una scelta non dettata dal teorema. Si cambia una coordinata alla volta perché cambiandole tutte, allora $\Delta S \sim N$ cioè grande in modulo, allora e^{-cN} la probabilità di accettare il passo è quasi nullo: la catena di Markov ci mette tanto a progredire. Una cosa simile succede scegliendo Δ grande poiché $\Delta S \propto \Delta$. Bisogna cambiare la configurazione abbastanza perché la catena si muova, ma non troppo perché si rifiuta sempre la proposta.

L'azione è locale e si risolve il problema di ΔS piccolo cambiando una coordinate una volta. Nelle teorie fermioniche, l'azione non è locale e cambiare una coordinata alla volta può implicare ΔS grande. Quello che si fa è cambiare tutte le coordinate in modo correlato (seguendo le equazioni classiche del moto) così che l'azione cambi di poco.

Prima di considerare effettivamente la catena, si termalizza la catena, buttando via i primi tanti n stati.

Todo: Fare routine Metropolis, loop ogni variabile, cambiare una variabile alla volta, proposta viene vagliata e iterare tante volte.

Estrarre in modo random al configurazione x iniziale (la partenza da freddo è tutti zero), poi si fa loop sulle coordinate dove ne si cambia una x'_i con probabilità piatta compresa $\pm \Delta$, poi ne si estrae un altro r , se $e^{-\Delta S} \geq r$ allora si accetta la configurazione. Dopo aver fatto ogni sweep si calcola l'azione. Plottare in funzione delle sweep il valore dell'azione. Partendo da caldo si vede un'esponenziale che convergente dall'alto a $\frac{N}{2}$ per $M = W = 1$, cioè il valore medio di una configurazione termalizzata: quando si vede tale fenomeno la catena è termalizzata e si può

usare per calcolare le configurazioni. Dopo 100 sweep le enupple x sono distribuite come e^{-S} . In questo modo non si fanno i cambi di variabile.

Fare routine che fa una sweep. Fare main program che calcola routine e fare plot dopo ogni sweep. Generare subito $2N$ numeri random all'inizio di ogni sweep.

Si definisce sweep passare sul reticolo almeno una volta, cioè cambiare per la prima volta le N variabili.

Lezione 8

ven 01 dic
2023 14:30

15 Riassunto delle formule teoriche

[r] dove l'energia di vuoto è

$$\tilde{E}_0 = \frac{1}{2} \tilde{\omega}$$

[r] p.2 fine Il denominatore è $Z(T, a)$, mentre il numeratore è

p.3 metà, nel primo esponenziale si fa il limite $N \rightarrow \infty$ e l'unico termine che rimane è \tilde{E}_0 ; per il secondo integrale si fa $|l - k| \rightarrow \infty$ l'esponenziale che rimane ha \tilde{E}_1 perché $\langle \tilde{E}_0 | \hat{x} | \tilde{E}_0 \rangle = 0$. Poiché il calcolatore non può fare N infinito, allora a causa delle condizioni periodiche si ha $|l - k| > N - |l - k|$, i due ruoli si scambiano e dunque si aggiunge la seconda riga di esponenziali.

nella formula dopo, la tilde di a va su E

deve valere l'invarianza per trasformazione di parità

fare plot termalizzazione catena di Markov e plot in scala logaritmica plottare la funzione di correlazione in funzione di $|l - k|$: il grafico ha la forma di una parabola x^2 , ma in modo esponenziale con stessa pendenza, ai lati più esterni (cioè poco dopo l'inizio e poco prima di N) non si sa perché possono contribuire molti esponenziali, per l'oscillatore armonico questo non avviene. Se il reticolo fosse infinito, ci sarebbe solo l'esponenziale di sinistra.

bisogna determinare l'elemento di matrice e la differenza tra i due stati $\tilde{E}_0 - \tilde{E}_1$

Funzione di correlazione a due punti del quadrato. p.4 [r] il quadrato è un operatore pari, quindi il valore di aspettazione del vuoto non è nullo. Si usa la convenzione che l'energia del vuoto sia nulla. Il grafico della funzione di correlazione è una costante, mentre ai lati esterni non si sa, come sopra. Creutz per calcolare l'auto-funzione dello stato di vuoto.

15.1 Monte Carlo standard per l'oscillatore armonico

Si estraggono cammini distribuiti come $\frac{e^{-S}}{Z}$

[r] appunti quaderno, retro

Errore su correlazione Bisogna scegliere uno stimatore affinché nel limite $N_{\text{sweep}} \rightarrow \infty$, si ottenga quanto desiderato. Quindi l'errore è

$$\sigma^2 = \frac{1}{N_{\text{conf}}} [\langle (x_l x_k)^2 \rangle - \langle x_l x_k \rangle^2]$$

Il rapporto tra l'errore ed il segnale è

$$\frac{\sigma^2}{\langle x_l x_k \rangle^2} = \frac{1}{N_{\text{conf}}} \left[\frac{\langle (x_l x_k)^2 \rangle}{\langle x_l x_k \rangle^2} - 1 \right]$$

Il numeratore va costante perché dominato dal vuoto. Al denominatore contribuisce solo il vuoto ed il primo stato, e scende in modo esponenziale. A grandi distanze vale

$$\frac{\sigma^2}{\langle x_l x_k \rangle^2} \approx \frac{1}{N_{\text{conf}}} \frac{\left| \langle \tilde{E}_0 | \hat{x}^2 | \tilde{E}_0 \rangle \right|^2}{\left| \langle \tilde{E}_0 | \hat{x} | \tilde{E}_0 \rangle \right|^4} e^{aN(\tilde{E}_1 - \tilde{E}_0)} \cosh^{-2} \dots$$

Nel limite $N \rightarrow \infty$

$$\frac{\sigma^2}{\langle x_l x_k \rangle^2} \approx \frac{1}{N_{\text{conf}}} \frac{\left| \langle \tilde{E}_0 | \hat{x}^2 | \tilde{E}_0 \rangle \right|^2}{\left| \langle \tilde{E}_0 | \hat{x} | \tilde{E}_0 \rangle \right|^4} e^{2a|l-k|(\tilde{E}_1 - \tilde{E}_0)}$$

L'errore relativo aumenta esponenzialmente con la distanza tra gli operatori. L'unico modo per vincere l'esponenziale bisogna usare $\frac{1}{N_{\text{conf}}}$, per questo si fanno milioni di sweep. Per l'oscillatore armonico, questo stimatore è semplice, ma cattivo. Questo problema non è ancora risolto in QCD per cui molte cose che si potrebbero calcolare in teoria, in verità non si può fare perché l'errore è troppo grande. Bisogna calcolare la varianza.

Calcolo da fare. L'integrando di

$$C(l, k) = \langle x_l x_k \rangle$$

dipende separatamente da l e k . Ma visto che l'hamiltoniana è invariante per traslazione del tempo fisico, il correlatore dipende solo dalla distanza $|l - k|$. In questo modo si può cambiare stimatore riducendo l'errore

$$C(k - l) = C(j) = \frac{1}{N} \sum_{m=0}^{N-1} \langle x_m x_{m+j} \rangle = \left\langle \frac{1}{N} \sum_m x_m x_{m+j} \right\rangle, \quad j = k - l$$

Nel calcolatore usare, la seconda espressione dentro la media. Poi mediarla. Fare almeno 50-100 sweep termalizzazione, per questo fare 10k-20k sweep. Il plot in scala logaritmica in funzione di j è simile ad una parabola.

Prossima volta. Una volta calcolata l'osservabile e la funzione di correlazione allo sweep k -esimo. Esso non è indipendente dal correlazione $k - 1$ -esimo. La formula della varianza presume che

$$\langle c^k(j) C^{k+1}(j) \rangle = \langle C^k(j) \rangle \langle C^{k+1}(j) \rangle$$

cioè le misure sono indipendenti.

La varianza è

$$(\Delta C)^2 = \frac{1}{N_{\text{sweep}}} [\langle C^2 \rangle - \langle C \rangle^2]$$

Bisogna modificare la formula, secondo la funzione di auto-correlazione

$$\Gamma(t^M) = \frac{\langle C^k C^{k+t_M} \rangle - \langle C^k \rangle^2}{\langle C^k \rangle^2 - \langle C^k \rangle^2}$$

L'auto-correlazione va a zero ad un certo tempo t^M oltre cui si possono considerare le variabili senza correlazione. [r] Si calcola una funzione di correlazione nel tempo markoviano delle correlazioni nel tempo fisico. Per l'oscillatore armonico, con le estrazioni correlate si possono calcolare le osservabili e mediare, così con quelle successive non correlate. Si dimostra che anche le medie sono decorrelate. Questo non si fa sempre, tipo in QCD misurare le osservabili è costoso, più del Monte Carlo.

Lezione 9

15.2 Autocorrelazione

Bisogna capire come calcolare l'errore statistico quando gli eventi sono correlati. Si indica con y una generica osservabile. Per la correlazione è

$$y_i = C^i(j), \quad C^i(j) = \frac{1}{N} \sum_{l=0}^{N-1} x_l^i x_{l+j}^i$$

ven 15 dic
2023 14:30

Per ogni j bisogna calcolare la funzione di auto-correlazione. Si usa un indice i degli sweep. Il valor medio è

$$\bar{y} = \frac{1}{N} \sum_i y_i, \quad \langle \bar{y} \rangle = \langle y \rangle, \quad \sigma_{\bar{y}}^2 = \langle (\bar{y} - \langle y \rangle)^2 \rangle$$

Per variabili y_i e y_j indipendenti, si ha

$$\langle \bar{y} \bar{y} \rangle = \dots = \langle y \rangle^2 + \frac{1}{N} (\langle y^2 \rangle - \langle y \rangle^2)$$

infatti vale

$$\langle y_i y_j \rangle = \langle y_i \rangle \langle y_j \rangle = \langle y \rangle^2$$

Quindi

$$\sigma_{\bar{y}}^2 = \frac{1}{N} \sigma_y^2$$

p.7 Si studia la varianza per due variabili correlate. Infatti, nelle catene di Markov, la probabilità dipende da y_i e y_{i-1} . La varianza è

$$\begin{aligned} \text{var}(\bar{y}) &= \dots = \frac{1}{N} (\langle y^2 \rangle - \langle y \rangle^2) + \frac{1}{N} (\langle y^2 \rangle - \langle y \rangle^2) \frac{1}{N} \sum_{i \neq j} \frac{\langle y_i y_j \rangle - \langle y \rangle^2}{\langle y^2 \rangle - \langle y \rangle^2} \\ &= \frac{\sigma_y^2}{N} \left[1 + \frac{1}{N} \sum_{i \neq j} \frac{\langle y_i y_j \rangle - \langle y \rangle^2}{\sigma_y^2} \right] = \frac{\sigma_y^2}{N} \left[1 + 2 \frac{1}{N} \sum_{i=1}^N \sum_{j=i+1}^N \frac{\langle y_i y_j \rangle - \langle y \rangle^2}{\sigma_y^2} \right] \\ &= \frac{\sigma_y^2}{N} \left[1 + 2 \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{N-i} \frac{\langle y_i y_{i+t} \rangle - \langle y \rangle^2}{\sigma_y^2} \right] \end{aligned}$$

alla terza riga si pone $t = j - i$: la catena di Markov è invariante per tempo Markoviano, pertanto la varianza? $[r]$ dipende solo dalla distanza temporale, non dal valore assoluto. Con N numero di sweep.

Si studia il limite $N \rightarrow \infty$:

$$\lim_{N \rightarrow \infty} \frac{2}{N} \sum_{i=1}^N \sum_{t=1}^{N-i} \frac{\langle y_i y_{i+t} \rangle - \langle y \rangle^2}{\sigma_y^2} = \lim_{N \rightarrow \infty} \frac{2}{N} \sum_{i=1}^{\infty} \sum_{t=1}^{\infty} \frac{\langle y_i y_{i+t} \rangle - \langle y \rangle^2}{\sigma_y^2}$$

Visto che la dipendenza è solamente dalla distanza temporale, la sommatoria su i fornisce N :

$$\text{var}(\bar{y}) = \frac{\sigma_y^2}{N} \left[1 + 2 \sum_{t=1}^{\infty} \frac{\langle y_i y_{i+t} \rangle - \langle y \rangle^2}{\sigma_y^2} \right]$$

dove i è arbitrario. Si definisce la funzione di auto-correlazione

$$\Gamma(t) = \langle y_i y_{i+t} \rangle - \langle y \rangle^2$$

dunque

$$\text{var}(\bar{y}) = \frac{\sigma_y^2}{N} \left[1 + 2 \sum_{t=1}^{\infty} \frac{\Gamma(t)}{\Gamma(0)} \right]$$

Si definisce

$$\tau_{\text{int}} = \frac{1}{2} \left[1 + 2 \sum_{t=1}^{\infty} \frac{\Gamma(t)}{\Gamma(0)} \right] \implies \text{var}(\bar{y}) = \frac{2\sigma_y^2}{N} \tau_{\text{int}}$$

Catene di Markov. L'auto-correlazione per le catene di Markov decresce come un esponenziale ed oscilla intorno a zero quando l'errore diventa più grande del valore di aspettazione.

$$\frac{\Gamma(t)}{\Gamma(0)} \propto e^{-\frac{t}{\tau}}$$

dove τ è il tempo di auto-correlazione.

Si può usare un altro metodo: negli sweep si misura la variabile (calcolare il correlatore) ogni 10τ ; calcolando il tempo di auto-correlazione si ottiene zero e dunque si può usare la formula dell'errore standard. Nel caso in cui l'osservabile costa poco, questo non conviene. Prima bisogna stimare τ , poi conviene calcolare l'osservabile per ogni sweep e poi definire un'osservabile

$$\tilde{y} = \frac{1}{d_{\text{bin}}} \sum_{j=1}^{d_{\text{bin}}} y_j$$

come media delle misure dell'osservabile all'interno di un bin $d_{\text{bin}} = 10\tau$. Pertanto, il numero di bin è $\text{NSweep}/d_{\text{bin}}$. Si ricalcola la funzione di autocorrelazione e si ottiene subito zero (dopo 1). Per le medie binnate si può usare la formula semplice per l'errore.

Main program termalizzazione, salvare plot, per relazione
Main program calcola auto-correlazione per i primi 10 j . Il τ dipende dal passo reticolare (e altri) dunque nel limite $a \rightarrow 0$ diventa grande, salvare plot per relazione (con dati dei parametri)
Misurato il τ , esso diventa un parametro mettendolo dentro il global e la dimensione del bin d_{bin} .
Fare main program che calcola le osservabili: calcola correlatore, fare media su ogni bin, e di tali osservabili medie fare l'analisi statistica.

L'osservabile $y^i = C^i(j)$, quando si fa il bin bisogna fare

$$\frac{1}{d_{\text{bin}}} \sum_{i=1}^{d_{\text{bin}}} C^i(s)$$

prima calcolare $C^i(j)$ poi fare la media sul bin. Il bin serve per facilitare l'analisi statistica.

Prossima volta. Si fa il valor medio e l'errore sui correlatori. Si usa la formula banale per l'incertezza

$$\sigma^2 = \frac{\langle y^2 \rangle - \langle y \rangle^2}{N}$$

In una catena di Markov, i dati sono correlati nel tempo Markoviano e sono correlati nel tempo fisico. Una volta da $C(j)$ si calcola il primo autovalore come

$$a\Delta E(j) = \text{arccosh} \frac{C(j+1) + C(j-1)}{2C(j)}, \quad j \neq 0, 1$$

fino dove ΔE ha errore di 10% perché cresce esponenzialmente. Il problema è che $C(j)$ e $C(j+1)$ sono correlati perché si sono usati gli stessi cammini di Feynman per calcolare. Una funzione di due variabili correlate, la sua varianza è

$$f(a, b) \implies \sigma_f^2 = \partial_a f(\bar{a}, \bar{b})^2 \sigma_a^2 + \partial_b f(\bar{a}, \bar{b})^2 \sigma_b^2 + 2 \partial_a f \partial_b f \text{ cov}(a, b)$$

Esiste un meccanismo che manipola i risultati binnati e calcola le derivate automaticamente. automatic derivation?. I dati sono manipolati di modo che la formula sopra equivale a calcolare

$$\langle y^2 \rangle - \langle y \rangle^2$$