Computational Physics Laboratory

Maso*

4 novembre 2023

Indice

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

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

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

Definizione 2.3. Dati dei punti $x_0, \ldots, x_n$ 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

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

Proposizione 2.13. Si vuole calcolare l'integrale mendiate 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):

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

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

$$w_i \equiv \int_a^b l_i^n(x)\,\mathrm{d}x$$

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

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

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

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

La forma dei pesi è data da

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

[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 $\int_a^b \mathrm{d}x$

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

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

Pertanto

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

dove l'errore è dato da

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

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

$$f(c) = \frac{\int_a^b f(x)g(x)\,\mathrm{d}x}{\int_a^b g(x)\,\mathrm{d}x}$$

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

Applicando il teorema, si ottiene un errore pari a

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

Usando il precedente cambio di variabili

$$z = \frac{x-a}{h}\,, \quad \mathrm{d}x = h\,\mathrm{d}z$$

si ha

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

Quindi il risultato delle formule per la regole trapezoidale è

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

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

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

Pertanto

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

L'errore è dato da

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

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

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

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

$$\int_a^b Q_2(x)\,\mathrm{d}x = 0$$

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

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

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

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

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

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

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 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

ven 13 set 2023 14:30

3 Quadrature gaussiane

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

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

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	$He_n(x)$	$-\infty$	$+\infty$	$\mathrm{e}^{-\frac{x^2}{2}}$
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

$$\int_a^b w(x) f(x) \, \mathrm{d}x$$

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 $x_0, \ldots, x_{n-1}$ 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

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

dove

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

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

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

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

$$\int_a^b w(x)p(x)\,\mathrm{d}x = \int_a^b w(x)Q(x)p_n(x)\,\mathrm{d}x + \int_a^b w(x)R(x)\,\mathrm{d}x$$

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

$$\int_a^b w(x)Q(x)p_n(x)\,\mathrm{d}x = 0 \implies \int_a^b w(x)p(x)\,\mathrm{d}x = \int_a^b w(x)R(x)\,\mathrm{d}x$$

Nei punti $x_0, \ldots, x_{n-1}$, 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

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

□

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

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

con

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

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

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

□

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

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

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

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

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

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

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

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

Come sopra, l'errore può essere stimato come

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

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,

$$I = \int \mathrm{d}x_1 \cdots \mathrm{d}x_d\, f(x_1, \ldots, x_d)$$

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

$$I = \sum_{j_1=0}^{n} \cdots \sum_{j_d=0}^{n} w_{j_1} \cdots w_{j_d} f(\tfrac{j_1}{n}, \ldots, \tfrac{j_d}{n}) + O(n^{-2})$$

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 $w(x) = \mathrm{e}^{-\frac{x^2}{2}}$. L'intervallo di integrazione è l'asse reale. I polinomi sono ortogonali rispetto al peso

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

I polinomi si ottengono dalla formula di Rodrigues:

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

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

$$f(x) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} \widetilde{f}(k) \mathrm{e}^{-\mathrm{i}kx} \,, \quad \widetilde{f}(k) = \int_{\mathbb{R}} \mathrm{d}x f(x) \mathrm{e}^{\mathrm{i}kx}$$

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

$$\delta(x) = \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{-\mathrm{i}kx} \,, \quad \int \delta(x) \, \mathrm{d}x = 1$$

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

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

L'integrale ed il secondo momento sono

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

La sua trasformata è data da

$$\int_{\mathbb{R}} P(x) \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}x = \mathrm{e}^{-\frac{k^2\sigma^2}{2}}$$

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

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

Dimostrazione. Segue

$$\begin{aligned}
\widetilde{He}_n(k) &= \int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\frac{x^2}{2}} He_n(x) \mathrm{e}^{\mathrm{i}kx} = (-1)^n \int_{\mathbb{R}} \left[\mathrm{d}_x^n \mathrm{e}^{-\frac{x^2}{2}} \right] \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}x \\
&= \int_{\mathbb{R}} \mathrm{e}^{-\frac{x^2}{2}} \left[\mathrm{d}_x^n \mathrm{e}^{\mathrm{i}kx} \right] \, \mathrm{d}x = \int_{\mathbb{R}} \mathrm{e}^{-\frac{x^2}{2}} (\mathrm{i}k)^n \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}x \\
&= \sqrt{2\pi} (\mathrm{i}k)^n \mathrm{e}^{-\frac{k^2}{2}}
\end{aligned}$$

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 $\sigma = 1$. □

6.3 Trasformata di una convoluzione

Definizione 6.4. La convoluzione di due funzioni è definita come

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

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

$$\begin{aligned}
\widetilde{g}(k) &= \int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{\mathrm{i}kx} \int_{\mathbb{R}} f_1(y) f_2(x-y) \, \mathrm{d}y = \int_{\mathbb{R}^2} f_1(y) \mathrm{e}^{\mathrm{i}ky} f_2(x-y) \mathrm{e}^{\mathrm{i}k(x-y)} \, \mathrm{d}x \, \mathrm{d}y \,, \quad z \equiv x - y \\
&= \int_{\mathbb{R}^2} f_1(y) \mathrm{e}^{\mathrm{i}ky} f_2(z) \mathrm{e}^{\mathrm{i}kz} \, \mathrm{d}y \, \mathrm{d}z = \widetilde{f}_1(k) \widetilde{f}_2(k)
\end{aligned}$$

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

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

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

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

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

$$\mathrm{e}^{-F(k)} = z(k) \equiv \int_{\mathbb{R}} P(x) \mathrm{e}^{\mathrm{i}kx} \,\mathrm{d}x$$

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

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

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

$$\mathrm{e}^{-F(k)} = \int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{(x-x_0)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x = \mathrm{e}^{\mathrm{i}kx_0} \int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{(x-x_0)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{\mathrm{i}k(x-x_0)}\,\mathrm{d}x = \mathrm{e}^{\mathrm{i}kx_0}\mathrm{e}^{-\frac{k^2\sigma^2}{2}}$$

da cui si ottiene

$$F(k) = -\mathrm{i}kx_0 + \frac{k^2\sigma^2}{2}$$

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

$$P_s(z) = \int_{\mathbb{R}} P(x)\delta(\tfrac{x-x_0}{\sigma} - z)\,\mathrm{d}x$$

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

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

ma altre distribuzioni possono avere più contributi.

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

$$\mathrm{e}^{-\Phi(k)} = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \mathrm{e}^{\mathrm{i}kz}\,\mathrm{d}z = \mathrm{e}^{-\frac{k^2}{2}} \implies \Phi(k) = \frac{1}{2}k^2$$

6.7 Proprietà dei cumulanti

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

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

si vuole ricavare la distribuzione della variabile somma

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

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

$$P(y) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{-\mathrm{i}ky} \int_{\mathbb{R}} \mathrm{d}x_1\,\mathrm{d}x_2\, P_1(x_1)P_2(x_2)\mathrm{e}^{\mathrm{i}kx_1}\mathrm{e}^{\mathrm{i}kx_2} = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{-\mathrm{i}ky}\widetilde{P}_1(k)\widetilde{P}_2(k)$$

Per l'unicità della trasformata di Fourier, segue

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

dove $\widetilde{P}(k)$ 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

ven 27 ott 2023 14:30

6.8 Teorema del limite centrale

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

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

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

$$P(y) = \int \frac{\mathrm{d}k}{2\pi} [p(k)]^n \mathrm{e}^{-\mathrm{i}ky}$$

la distribuzione della variabile standardizzata è data da

$$\pi(z) = \int \mathrm{d}y \, P(y) \delta(z - \tfrac{y - \langle y \rangle}{\sigma_y}) = \int \frac{\mathrm{d}k}{2\pi} [p(k)]^n \int \mathrm{d}y \, \mathrm{e}^{-\mathrm{i}ky} \delta(z - \tfrac{y - \langle y \rangle}{\sigma_y})$$

Notando

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

si ottiene

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

Ricordando

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

segue

$$\pi(z) = \sigma_y \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{-\mathrm{i}kz\sigma_y} \exp\left[n \sum_{m=2}^{\infty} \frac{(\mathrm{i}k)^m}{m!} c_{x,m} \right] = \int \frac{\mathrm{d}k'}{2\pi} \mathrm{e}^{-\mathrm{i}k'z} \exp\left[n \sum_{m=2}^{\infty} \left(\frac{\mathrm{i}k'}{\sigma_y} \right)^m \frac{c_{x,m}}{m!} \right]$$

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

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

Il secondo addendo nel generatore dei cumulanti si annulla nel limite $n \to \infty$. 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

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

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

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

si ha

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

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

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

Ricordando che la trasformata di Fourier dei polinomi di Hermite è

$$\widetilde{He}_n = \int_{\mathbb{R}}\mathrm{d}x\,\mathrm{e}^{-\frac{x^2}{2}}He_n(x)\mathrm{e}^{\mathrm{i}kx} = \sqrt{2\pi}(\mathrm{i}k)^n\mathrm{e}^{-\frac{k^2}{2}}$$

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

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

Osservazione 7.2. Risulta chiaro che nel limite $n \to \infty$, 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)**.** Estraendo n variabili aleatorie con stessa distribuzione, la media è

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

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 $n \to \infty$. In questo modo si può calcolare un integrale.

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

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

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

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

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

$$y = f(x) \,, \quad P(y) = \int P(x) \delta(y - f(x)) \,\mathrm{d}x$$

Si ha

$$\int y P(y) \,\mathrm{d}y = \int \mathrm{d}x \,\mathrm{d}y \, y P(x) \delta(y - f(x)) = \int \mathrm{d}x \, P(x) f(x) = I$$

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

$$I = \int_0^1 f(x)\,\mathrm{d}x$$

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

$$y = f(x)\,, \quad P(y) = \int \mathrm{d}x_i\, P(x_i)\delta(y - f(x_i))$$

il cui valor medio è

$$\langle y \rangle = \int y P(y)\,\mathrm{d}y = \int \mathrm{d}y\,\mathrm{d}x_i\, P(x_i)\delta(y - f(x_i))y = \int \mathrm{d}x_i\, P(x_i) f(x_i)$$

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

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

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

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

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

$$I = \int f(x) \, \mathrm{d}x$$

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

$$g(x) \geq 0 \,, \quad \int_0^1 g(x) \, \mathrm{d}x = 1$$

Allora vale

$$I = \int \frac{f(x)}{g(x)} g(x) \, \mathrm{d}x$$

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

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

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

$$\langle O \rangle = \frac{\int \pi \, \mathrm{d}x_i \, \mathrm{e}^{-S} v(x) \cdots}{\int \pi \, \mathrm{d}x_i \, \mathrm{e}^{-S}}$$

in cui si pone

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

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

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

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

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

da cui

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

La jacobiana ed il jacobiano sono

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

Pertanto, le variabili y sono distribuite secondo

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

[r]

Osservazione 8.7. Estraendo 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

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

Si definisce

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

La jacobiana ed il jacobiano sono

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

Quindi

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

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):

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

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

Osservazione 8.10. L'azione dell'oscillatore armonico è

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

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.

Lezione 4

ven 03 nov 2023 14:30

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

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

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

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

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

così come

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

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

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

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

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

I due formalismi sono equivalenti

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

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

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

e si definisce il propagatore ritardato:

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

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

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

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 un 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

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

La trasformata di Fourier nel tempo è [r]

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

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

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

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

$$G(x, t, x_0, t_0) = \langle x| \hat{G}(t, t_0) |x_0\rangle = \theta(t - t_0) \sum_n \mathrm{e}^{-\mathrm{i}E_n(t-t_0)} u_n(x) u_n^*(x_0)$$
$$\widetilde{G}(x, x_0, z) = \mathrm{i} \sum_n \frac{1}{z - E_n} u_n(x) u_n^*(x_0)$$

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 $\operatorname{Re} z < 0$, $\operatorname{Im} z = 0$. Per autovalori continui, cioè $\operatorname{Re} z > 0$, $\operatorname{Im} z = 0$, 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

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

Inserendo l'identità [r]

$$\hat{I} = \int \mathrm{d}x_1 \, |x_1\rangle\langle x_1|$$

si ottiene

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

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 posizione 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

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

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

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

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

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

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 N^{-2}. Infatti, sebbene in questo caso non succeda, si hanno N termini [r] è $O(N^{-1})$ e non $O(N^{-2})$.

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

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

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

Utilizzando la formula di Trotter si ha

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

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

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

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 che viene scritto come

$$1 = \mathrm{e}^{\frac{\mathrm{i}}{2}aV(x)}\mathrm{e}^{-\frac{\mathrm{i}}{2}aV(x)}$$

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

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

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

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

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

$$\hat{\tau}_a = \mathrm{e}^{-\mathrm{i}a\hat{\hat{H}}}\,, \quad \hat{\hat{H}} = \hat{\hat{H}}^\dagger$$

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

Osservazione 10.2. L'operatore di traferimento $\hat{\tau}_a$ è l'operatore che propaga uno stato generico nel tempo $\Delta t = a$ a meno di un errore di ordine $O(a^2)$.

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

$$\hat{G}_{\mathrm{E}}(t,t_0) = \hat{G}(-\mathrm{i}t,-\mathrm{i}t_0) = \theta(t-t_0)\mathrm{e}^{-\hat{H}(t-t_0)}$$

che corrisponde a

$$\hat{G}_{\mathrm{E}}(t_{\mathrm{E}},t_{0,\mathrm{E}}) = \hat{G}(-\mathrm{i}t_{\mathrm{E}},-\mathrm{i}t_{0,\mathrm{E}})\,, \quad t_{\mathrm{E}} = \mathrm{i}t$$

[r] definizione tempo euclideo
In questo modo, per autovalori grandi, l'esponenziale tende a zero. In teoria dei campi, la metric 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 $G_{\mathrm{E}}(t_{\mathrm{E}})$ 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 le formule.

Operatore di trasferimento euclideo. Il passo reticolare diventa $a_{\mathrm{M}} \to \mathrm{i} a_{\mathrm{E}}$ da cui

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

Osservazione 10.6. Per costruzione si ha

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

A questo punto si può inserire l'identità $N-1$ volte tra 0 e T.

Elementi di matrice di trasferimento nella base delle coordinate. Inserire l'identità implica dover calcolare

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

Per definizione si ha

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

poiché il potenziale è un operatore diagonale nella base delle coordinate. Mentre il termine cinetico è

$$\begin{aligned}\left\langle x_i\right| \mathrm{e}^{-\frac{\hat{p}^2}{2m}}\left|x_j\right\rangle &= \int \mathrm{d}q_i\,\mathrm{d}q_j\,\left\langle x_i|q_i\right\rangle\left\langle q_i\right|\mathrm{e}^{-\frac{\hat{p}^2}{2m}}\left|q_j\right\rangle\left\langle q_j|x_j\right\rangle = \cdots\\ &= \left(\frac{m}{2\pi a}\right)^{\frac{1}{2}}\mathrm{e}^{-\frac{(x_i-x_j)^2 m}{2a}}\end{aligned}$$

[r] Il tempo euclideo si utilizza perché così si ha una gaussiana che si può integrare, si può trovare il termine cinetico. Unendo quanto trovato si ha

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

Propagatore euclideo come integrale sui cammini. [r] Il propagatore ritardato è ora scritto come un integrale mulidimensionale. Inoltre, si nota che

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

cioè si ottiene la velocità classica. Inoltre, si ha segno positivo al potenziale a causa del tempo euclideo. La lagrangiana euclidea e l'azione sono definite come

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

pertanto

$$G_{\mathrm{E}}(x_N, T, x_0, 0) = \left(\frac{m}{2\pi a}\right)^{\frac{N}{2}}\int\prod_{i=1}^{N-1}\mathrm{d}x_i\,\mathrm{e}^{-S_{\mathrm{E}}}$$

con x_0, x_N fissati. Una volta discretizzato il tempo, il cammino di Feynman corrisponde ad una n-upla 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 scritto come integrale sui cammini. [r] L'integrale scritto in questo modo non ha senso, ma è solo un'operazione formale.

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

[r] In meccanica quantistica, una particella può seguire qualsiasi traiettoria il cui peso è dato dall'azione (non solo dalla sua forma).

Limite classico. Una deviazione di un sistema macroscopico $\delta S \approx 10^{20}\hbar$ implica una soppressione enorme, mentre per un sistema microscopico si ha $\delta S \approx \hbar$ dunque molte traiettorie contribuiscono.

Fenditura e doppia fenditura si può fare bene con il path integral (Grasiorowich).

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

$$\int \mathrm{d}x_0 \, G(x_N, T, x_0, 0) \, , \quad x_N = x_0$$

cioè la traccia del propagatore ritardato

$$Z_a(T) = \mathrm{Tr}\Big[\hat{G}(T,0)\Big] = \mathrm{Tr}\Big[\hat{G}(T,0) \int \mathrm{d}x_0 \, |x_0\rangle\langle x_0|\Big] = \int \mathrm{d}x_0 \, G(x_0, T, x_0, 0) \, , \quad t > t_0 \, , \quad T = t - t_0$$

[r]

La traccia è pari a

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

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.