

Computational Physics Laboratory

20 ottobre 2023

Indice

1	Introduzione	1
I	Integrazione numerica elementare	2
2	Formule di Newton–Cotes	2
2.1	Integrazione secondo quadrature numeriche	3
2.1.1	Regola trapezoidale	4
2.1.2	Regola di Simpson	5
3	Quadrature gaussiane	6
3.1	Polinomi ortogonali	7
3.2	Quadrature gaussiane	7
4	Integrazione numerica composta	8
5	Integrali multidimensionali	9
II	Metodi Monte Carlo	10
6	Prerequisiti teorici	10
6.1	Polinomi ortogonali di Hermite	10
6.2	Trasformata di Fourier	10
6.3	Trasformata di una convoluzione	11
6.4	Variabile aleatoria e distribuzione di probabilità	11
6.5	Funzione generatrice	12
6.6	Variabile standardizzata	13
6.7	Proprietà dei cumulanti	13
7	Teorema del limite centrale	13

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 regole 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 il 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

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

ven 13 set
2023 14:30

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 [r]

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

□

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 Prerequisiti teorici

Si studiano le fondamenta su cui si basa il metodo Monte Carlo.

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

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

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

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

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 $\frac{x-x_0}{\sigma}$ corrisponde al valore di cui si vuole conoscere la densità. Per alcune distribuzioni la corrispondenza è biunivoca, come per la gaussiana:

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

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.

7 Teorema del limite centrale

Teorema 7.1 (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

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

Dimostrazione. Si dimostra il caso di distribuzioni identiche

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

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

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

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

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

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

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

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 7.2. 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

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