

Appunti per le esercitazioni di Chimica Teorica

Giorgio F. Signorini

2007-2008

Indice

1	Metodi Monte Carlo	2
1.1	Importance Sampling	2
1.1.1	dimostrazione del valore della varianza della media del campione . .	4
1.2	Catene di Markov	5
1.3	Schema del metodo Metropolis Monte Carlo	6
1.4	Punti salienti di un programma MC	6
1.5	Calcolo di $g(r)$	7
1.5.1	in pratica	7
	Bibliografia	8

Capitolo 1

Metodi Monte Carlo

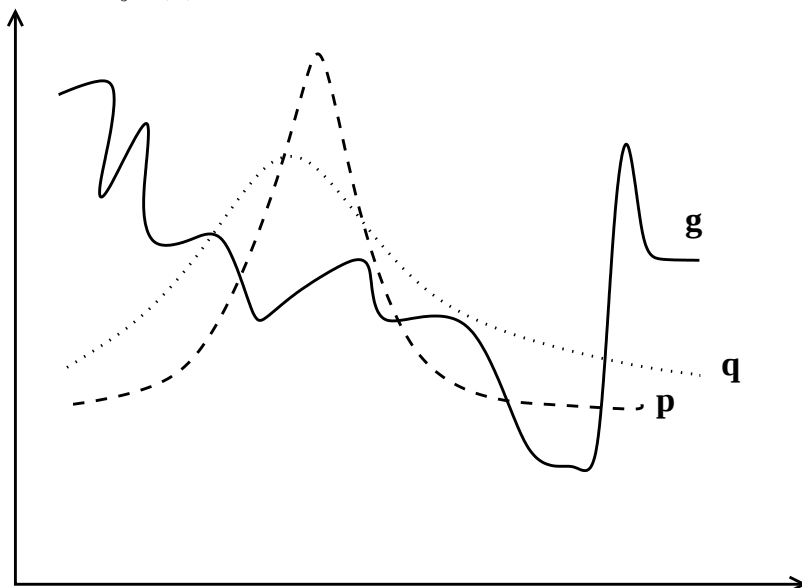
1.1 Importance Sampling

(basato su wik)

Cerchiamo il valore di aspettazione di una funzione $g(x)$ distribuita secondo una probabilità $p(x)$ sul dominio delle x :

$$E[g(x)|p] = \int g(x)p(x)dx$$

(assumiamo $\int p(x)dx = 1$).



Una stima empirica di $p(x)$, generando $x_1, x_2, \dots$ secondo p , è

$$P_n(x) = \frac{1}{n} \sum_P \delta(x - x_i)$$

La corrispondente stima empirica di g è

$$\begin{aligned} \hat{E}[g(x)|p] &= \int g(x)P_n(x)dx \\ &= \frac{1}{n} \sum_P g(x_i) \end{aligned}$$

e

$$\lim_{n \rightarrow \infty} \hat{E}[g(x)|p] \equiv \langle g \rangle_P$$

Se però non siamo in grado di generare punti secondo p , ma secondo un'altra distribuzione q (nel caso banale, una distribuzione uniforme), si può scrivere il valore di aspettazione come

$$\begin{aligned} E[g(x)|p] &= \int g(x) \frac{p(x)}{q(x)} q(x) dx \\ &= \int g(x) r(x) q(x) dx \\ &= E[g(x) \cdot r(x)|q] \end{aligned}$$

con

$$r(x) = \frac{p(x)}{q(x)}$$

La stima empirica su q del valore di aspettazione di $g(x)$ è

$$\hat{E}[g(x)r(x)|q] = \frac{1}{n} \sum_Q g(x_i)r(x_i)$$

e

$$\lim_{n \rightarrow \infty} \hat{E}[g(x)r(x)|q] \equiv \langle g \cdot r \rangle_Q$$

Quanto è accurata una certa stima di E fatta con n punti? Cioè, quanto si discosta $\hat{E}$ dal valore vero E ? Dipende da n ma anche dalle proprietà di $g(x)$ e di p (o q). Ad esempio, se $g(x) = G = \text{cost}$, il valore di aspettazione stimato con qualunque n è G , e l'errore è nullo per qualunque n . Negli altri casi, a parità di n , possiamo aspettarci che sarà tanto minore quanto minore è la variazione di $g(x)$ in un campionamento secondo p .

L'accuratezza della stima è misurata dalla *varianza della media* s^2 (da non confondere con la varianza *dei dati*, σ^2).

Si dimostra (vedi (1.1.1)) che

$$\begin{aligned} s_{P,n}^2 &\approx \frac{1}{n} \left(\langle g^2 \rangle_P - \langle g \rangle_P^2 \right) \\ &= \frac{\sigma_P^2}{n} \end{aligned} \tag{1.1}$$

L'equazione (1.1) conferma che a parità di n l'errore sulla stima è dato dalla variazione quadratica della g , e che se $g = \text{cost}$ allora $s^2 = 0$.

Se invece che su P , campioniamo su Q , l'errore sulla stima è

$$s_{Q,n}^2 = \frac{1}{n} \left(\langle (g \cdot r)^2 \rangle_Q - \langle g \cdot r \rangle_Q^2 \right) \tag{1.2}$$

cioè dipende dalla variazione quadratica di $g(x) \cdot r(x)$.

L'equazione (1.2) corrisponde esattamente alla (3.1.9) in Frenkel and Smit [2002], con $f = g \cdot p$ e $w = q$.

Consideriamo due casi limite:

1. Nel caso banale di una distribuzione U uniforme, $q = \frac{1}{x_{max}-x_{min}} = \frac{1}{L}$, e

$$s_{U,n}^2 = \frac{1}{n} \left(\left\langle (g \cdot p)^2 \right\rangle_U - \langle g \cdot p \rangle_U^2 \right) \quad (1.3)$$

Il confronto tra quest'ultima equazione e la (1.1) ci dice che la differenza tra fare una stima del valore di aspettazione di $g(x)$ campionando x in modo uniforme e campionandola secondo la sua distribuzione $p(x)$ è data dalla differenza nella variazione quadratica di $g \cdot p$ e di g *rispettivamente*. Se $p(x)$ è molto “piccata” (come è la distribuzione nell'insieme canonico) la differenza può essere enorme. Se al contrario il prodotto $g \cdot p$ varia meno di g , cioè g varia molto proprio dove p è bassa, l'errore nella stima può essere più basso se la stima viene calcolata con campionatura uniforme.

2. Se $q(x)$ è una buona approssimazione di $p(x)$, allora $r(x) \approx 1$, e

$$s^2 \approx \frac{1}{n} \left(\langle g^2 \rangle_Q - \langle g \rangle_Q^2 \right)$$

1.1.1 dimostrazione del valore della varianza della media del campione

Si immagina¹ di fare $N \rightarrow \infty$ stime ciascuna basata su n punti: $\hat{E}_1, \hat{E}_2, \dots$, e di fare la varianza di queste stime:

$$s^2 = \frac{1}{N} \sum_k \left(\hat{E}_k - \hat{E} \right)^2$$

con

$$\begin{aligned} \hat{E}_k &= \frac{1}{n} \sum_{i \in P} g(x_i^k) \\ \hat{E} &= \frac{1}{N} \sum_k \hat{E}_k = \frac{1}{Nn} \sum_k \sum_i g(x_i^k) \end{aligned}$$

Con notazione sintetica ($g_i^k = g(x_i^k)$ e $i = \{i \in P\}$) si ha:

$$\begin{aligned} s^2 &= \frac{1}{N} \sum_k \left(\hat{E} - \frac{1}{n} \sum_i g_i^k \right)^2 \\ &= \frac{1}{N} \sum_k \left(\frac{1}{n} \sum_i [g_i^k - \hat{E}] \right)^2 \\ &= \frac{1}{N} \sum_k \frac{1}{n^2} \left(\sum_i [g_i^k - \hat{E}] \right) \left(\sum_j [g_j^k - \hat{E}] \right) \end{aligned}$$

se i punti i, j sono estratti in maniera casuale sul dominio P , i termini misti si annullano:

$$\begin{aligned} s^2 &\approx \frac{1}{N} \sum_k \frac{1}{n^2} \sum_i [g_i^k - \hat{E}]^2 \\ &= \frac{1}{n} \frac{1}{N} \sum_k \frac{1}{n} \sum_i \left[(g_i^k)^2 - \hat{E}^2 \right] \\ &\approx \frac{1}{n} \left[\langle g^2 \rangle_P - \langle g \rangle_P^2 \right] \\ &= \frac{\sigma^2}{n} \end{aligned}$$

¹cfr. Haile [1997] e Chandler [1987]

dove il passaggio al valore di aspettazione vero $\langle \rangle$ si ha nel limite $N \rightarrow \infty$; l'ultima uguaglianza può essere presa come una definizione di σ^2

$$\sigma^2 = \left\langle g^2 - \langle g \rangle_P^2 \right\rangle_P$$

In sostanza, nel limite $N \rightarrow \infty$

$$\begin{aligned} s^2 &= \left\langle (g - \langle g \rangle_P)^2 \right\rangle_P \\ &= \frac{1}{n} \left\langle g^2 - \langle g \rangle_P^2 \right\rangle_P \\ &= \frac{1}{n} \sigma^2 \end{aligned}$$

1.2 Catene di Markov

(basato su Allen and Tildesley [1987])

Il problema è generare punti con una distribuzione di probabilità che approssimi la distribuzione canonica.

In una catena di Markov

- c'è un numero finito di stati
- ogni stato dipende solo dal precedente.

In qs condizioni si può costruire una matrice di transizione, π . Si ha

$$\begin{aligned} \rho^{(2)} &= \rho^{(1)} \pi \\ \rho^{(3)} &= \rho^{(2)} \pi = [\rho^{(1)} \pi] \pi \\ &\dots \\ \rho^{(n)} &= \rho^{(1)} \pi^n \end{aligned}$$

La catena ammette un limite

$$\rho = \lim_{\tau \rightarrow \infty} \rho^{(1)} \pi^\tau$$

tale che

$$\rho \pi = \rho$$

Dall'ultima equazione si capisce che lo stato limite ρ esiste se π ammette autovalore 1. Se il numero di stati è finito, lo è anche il numero di vettori $\rho^{(i)}$. (*approfondire: come si può essere sicuri che si arriva comunque al limite ρ ?*)

Nell'esempio del sistema a due stati (il calcolatore), si suppone che π sia nota.

Nelle simulazioni MC, si conosce ρ (ogni elemento $\rho_m = \rho_{NVT}(\Gamma_m)$), e si vuole trovare π .

Vi sono varie possibili soluzioni.

Si può porre la condizione (non necessaria) della *reversibilità microscopica*:

$$\rho_m \pi_{mn} = \rho_n \pi_{nm}$$

La soluzione *asimmetrica* di Metropolis prevede (α è una matrice simmetrica qualsiasi)

$$\begin{aligned}\pi_{mn} &= \alpha_{mn} & \rho_n > \rho_m & \quad m \neq n \\ \pi_{mn} &= \alpha_{mn} \frac{\rho_n}{\rho_m} & \rho_n \leq \rho_m & \quad m \neq n \\ \pi_{mm} &= 1 - \sum_{n \neq m} \pi_{mn}\end{aligned}$$

1.3 Schema del metodo Metropolis Monte Carlo

1. selezionare a caso una particella e calcolare la sua energia, $U(\mathbf{r}^N)$
2. spostarla a caso, e calcolare la sua nuova energia, $U(\mathbf{r}'^N)$
3. calcolare $p = e^{-\beta(U(\mathbf{r}') - U(\mathbf{r}))}$:
 - se $p \geq 1$, accettare la mossa
 - se $p < 1$, estrarre un numero ζ a caso nell'intervallo $[0, 1]$:
 - se $\zeta < p$, accettare la mossa
 - altrimenti, lasciare la particella nella vecchia posizione
4. accumulare le medie e ripartire dal punto 1.

1.4 Punti salienti di un programma MC

1. algoritmo base (v. punto precedente) Allen and Tildesley [1987]
2. a ogni mossa tentata non occorre calcolare tutta E , ma solo il ΔE dovuto allo spostamento della molecola spostata, i .
 N.B. che questa variazione è $\Delta E(i) = \sum_j \Delta E(i, j)$, e se la mossa viene accettata, va aggiornata sia la $E(i)$ che tutte le altre $E(j) = \sum_k E(j, k)$, di una quantità $\Delta E(j) = \Delta E(j, i)$. In sostanza si deve conservare in memoria ogni singola interazione $E(i, j)$.
3. Bilancio dettagliato: qualunque schema si adotti, la probabilità a priori di tornare indietro deve essere uguale a quella di andare avanti: $\alpha_{o \rightarrow n} = \alpha_{n \rightarrow o}$. Ad esempio: evitare di eseguire spostamenti casuali con un ordine fisso, tipo $T_x, T_y, T_z, R_x, R_y, R_z, T_x, \dots$, perché la mossa inversa è impossibile.
4. campionamento efficiente: v Frenkel, una buona misura di efficienza è la quantità di spazio delle fasi esplorato nell'unità di tempo, e di questo una buona misura è $\frac{\Delta r^2}{T_{CPU}}$. In particolare:
 - (a) conviene muovere una particella per volta
 - (b) non è detto che un rapporto di accettazione = 50% sia ottimale.
5. Muovere T_x, T_y, T_z insieme o solo uno per volta? Traslazioni e rotazioni insieme o separate?
6. Quanto deve essere grande Δ ?
7. PBC-MinimumImage: occhio a mantenere molecole intere

1.5 Calcolo di $g(r)$

1.5.1 in pratica

In pratica il $g(r)$ si calcola come:

$$g_{AB} \left(r + \frac{\Delta}{2} \right) = \frac{\langle n^{AB}(r, \Delta) \rangle}{n_{ideale}^{AB}(r, \Delta)}$$

cioè il rapporto tra il numero medio di particelle B a distanza da una particella A compresa tra r e $(r + \Delta)$

$$\langle n^{AB}(r, \Delta) \rangle = \frac{1}{TN_A} \sum_t^T \sum_i^{N_A} n_{it}^{AB}(r, \Delta)$$

e lo stesso in un sistema ideale

$$\begin{aligned} n_{ideale}^{AB}(r, \Delta) &= \rho_B \Delta V(r, \Delta) \\ n_{ideale}^{AA}(r, \Delta) &= \rho_A \frac{N_A - 1}{N_A} \Delta V(r, \Delta) \end{aligned}$$

Quindi:

$$g_{AB} \left(r + \frac{\Delta}{2} \right) = \frac{\sum_t^T \sum_i^{N_A} n_{it}^{AB}(r, \Delta)}{T(N_A - \delta_{AB}) \rho_B \Delta V(r, \Delta)}$$

Note pratiche:

1. Il calcolo di n_{it}^{AB} comprende comunque un loop sugli atomi di tipo B (per contarli)
2. Per le molecole i, j si contano solo le coppie $j \geq i$, ma all'interno di ciascuna delle due, tutti gli atomi A, B : $\{i_A j_A, i_A j_B, i_B j_A, i_B j_B\}$
3. Le interazioni di tipo $A - A$ e $B - B$ sono la metà di quelle $A - B$ (in cui vanno anche le $B - A$); ogni volta, vanno contate due volte tanto

Bibliografia

Importance Sampling. URL http://en.wikipedia.org/wiki/Importance_sampling.

M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Oxford University Press, New York, 1987. ISBN 0-19-855645-4.

D. Chandler. *Introduction to modern statistical mechanics*. Oxford University Press, New York, 1987. ISBN 0-19-504277-8.

D. Frenkel and B. Smit. *Understanding molecular simulation*. Academic Press, London and San Diego, California, 2002. ISBN 0-12-267351-4.

J. M. Haile. *Molecular dynamics simulation: elementary methods*. Wiley, 1997. ISBN 0-471-18439-X.