Monte Carlo - Simulationen

Quellen: Thijssen, Kap. 10

⋮

Def. Berechnung der Eigenschaften physikalischer Modelle mit stochastischen Methoden

Physikalische Fragestellung (Beispiel)

Klassische statistische Physik: Erwartungswerte im kanonischen Ensemble

N Teilchen mit Positionen $\vec{r}_i \in V$ und Impulsen $\vec{p}_i$

Energie $E = E(\{\vec{r}_i\}, \{\vec{p}_i\})$

Observable $A = A(\{\vec{r}_i\}, \{\vec{p}_i\})$

Erwartungswert

$$\langle A \rangle = \frac{\int_{V^N} d^{3N}_r \int_{\mathbb{R}^{3N}} d^{3N}_p \, A \, e^{-\beta E}}{\int_{V^N} d^{3N}_r \int_{\mathbb{R}^{3N}} d^{3N}_p \, e^{-\beta E}}$$

Für ein typisches Vielteilchensystem (s. Molekuldynamik)

$$E = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m_i} + \underbrace{\sum_{i=1}^{N} V_i^{(a)}(\vec{r}_i) + \sum_{i<j=1}^{N} V_{ij}^{(w)}(|\vec{r}_i - \vec{r}_j|)}_{= V = V(\{\vec{r}_i\})}$$

$\Rightarrow$ Wir können die Integrale über die Impulse durchführen, falls A nur von den Koordinaten abhängt

$$\langle A \rangle = \frac{\int_{V^N} d^{3N}r \; A(\{\vec{r}_i\}) \; e^{-\beta V}}{\int_{V^N} d^{3N}r \; e^{-\beta V}}$$

Beispiele:

- Potentielle Energie $\quad A = V$
- Interne Energie $\quad U = \frac{3}{2} N k_B T + \langle V \rangle$
- Druck (Virialsatz)

$$PV = N k_B T - \frac{1}{3V} \left\langle \frac{1}{2} \sum_{i=1}^{N} \vec{r}_i \, \vec{\nabla}_i \sum_{j=1}^{N} V^{(w)}_{ij}(|\vec{r}_i - \vec{r}_j|) \right\rangle$$

Anmerkung

Mit Hilfe verschiedener Relationen der Thermodynamik und statistischer Physik, können wir thermodynamische Grössen berechnen, die sich nicht als Erwartungswert ausdrücken lassen.

Beispiel: Entropie und freie Energie

$$S(T) = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \approx \frac{1}{kT^2} \frac{F(T+\Delta T) - F(T-\Delta T)}{2\Delta T}$$

Aber $F(T+\Delta T) - F(T-\Delta T) \approx -k_B T \ln\left(\frac{Z(T+\Delta T)}{Z(T-\Delta T)}\right)$

Für die Zustandssumme

$$Z = \frac{1}{N!}\left(\frac{2m\pi}{\beta h^2}\right)^{3/2} \int_{V^N} dr^{3N}\, e^{-\beta V}$$

erhalten wir

$$\frac{Z(T+\Delta T)}{Z(T-\Delta T)} \approx \frac{\int_{V^N} dr^{3N}\, e^{-\Delta\beta V}\, e^{-\beta V}}{\int_{V^N} dr^{3N}\, e^{-\beta V}}$$

mit $\Delta\beta = \frac{1}{T+\Delta T} - \frac{1}{T-\Delta T} \approx \frac{-2\Delta T}{T^2}$

Math. Aufgabe   Integration in hohen Dimensionen

$$I = \int_{V \subset \mathbb{R}^D} dX^D \; f(\vec{X})$$

Problem mit Quadratur-Formeln

Volumen V wird in M Zellen mit Kantenlängen h und Volumen $\Delta V = h^D = \frac{V}{M}$ zerlegt

$$I_M \cong \frac{V}{M} \sum_{i=1}^{M} w(\vec{X}_i)\, f(\vec{X}_i)$$

wobei $\vec{X}_i$ ist ein Punkt in der Zelle i und $w(\vec{X}_i) > 0$

ist ein Gewicht, das von der Quadratur-Formel abhängt.

Der Fehler ist $|I_M - I| \sim h^k \sim M^{-k/D}$
wobei

$k = 2$ für die erweiterte Trapezoid-Formel

$k = 4$ " " " Simpson-Formel

...

$\Rightarrow$ { Sehr langsame Konvergenz mit zunehmenden M
Schwierigkeiten mit den Integrationsgrenzen

Monte Carlo-Integration (Importance Sampling)

Grundidee: • Wahrscheinlichkeitsdichte $P(\vec{x}) > 0$, $\vec{x} \in V$ auswählen

• $I = \int_V dx^D \, g(\vec{x}) \, P(\vec{x})$

mit $g(\vec{x}) = \frac{f(\vec{x})}{P(\vec{x})}$

• M $\boxed{\text{unabhängige}}$ Zufallspunkte $\vec{x}_i \in V$ mit Wahrscheinlichkeitsdichte $P(\vec{x})$ erzeugen

$$I_M^{MC} = \frac{1}{M} \sum_{i=1}^{M} g(\vec{x}_i)$$

Zentraler Grenzwertsatz (Gesetz der grossen Zahlen):

I_M^{MC} ist eine Zufallsvariable mit Wahrscheinlichkeitsdichte

$$P(I_M^{MC}) \approx \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(I_M^{MC}-I)^2}{2\sigma^2}\right)$$

für $M \gg 1$.

Die Breite der Verteilung ist

$$\sigma^2 = \frac{\sigma_g^2}{M}$$

mit dem Schwankungsquadrat der Funktion $g(\vec{x})$

$$\sigma_g^2 = \frac{1}{V}\int_V dx^D\, g^2(\vec{x}) - \left|\frac{1}{V}\int_V dx^D\, g(\vec{x})\right|^2$$

Interpretation: I_M^{MC} "konvergiert" gegen I für $M \to \infty$

Der Fehler ist $|I_M^{MC} - I| < n\frac{\sigma_g^2}{\sqrt{M}}$

mit Wahrscheinlichkeit $\approx$ 68 % für $n=1$

$\approx$ 95 % für $n=2$

$\approx$ 99,7 % für $n=3$

...

Schlussfolgerung: MC-Integration ist effizienter als Quadraturformeln für $D > 2k$

Vorsicht: Die Genauigkeit hängt sehr stark von den Schwankungen der Funktion $g(\vec{x})$

Anmerkung: σ_q^2 kann aus der MC-Berechnung abgeschätzt werden

$$\sigma_q^2 \approx \frac{1}{M} \sum_{i=1}^{M} q^2(\vec{x}_i) - \left| \frac{1}{M} \sum_{i=1}^{M} q(\vec{x}_i) \right|^2$$

Praktische Probleme:

1) Wahl der Verteilung $\rho(\vec{x})$. Für unser Beispiel

$$\rho(\vec{x}) = \frac{e^{-\beta V}}{\int_V dx^D e^{-\beta V}}$$

2) Wie erzeugt man die Zufallspunkte $\vec{x}_i$ mit Wahrscheinlichkeitsdichte $\rho(\vec{x})$?

→ Künstliche diskrete "Moleküldynamik" mit einer sogenannten Markov-Kette

Markov-Ketten

Konfigurationen $\vec{X}_t$ mit $t = 0, 1, 2, .., S$

Für unser Problem $\vec{X}_t = \{\vec{r}_i\}$

Künstliche diskrete Dynamik mit "Zeit" t

Grundidee:
- Anfangskonfiguration $\vec{X}_0$ mit Verteilung $\rho_0(\vec{X})$
- Übergänge $\vec{X}_t \to \vec{X}_{t+1}$ mit Wahrscheinlichkeit $T(\vec{X} \to \vec{X}')$
- M unabhängige Konfigurationen $\vec{X}_{\tau_R + m\tau_c}$ $\quad m = 1, \ldots, M$ mit Verteilung $\approx \rho(\vec{X})$

<u>Beispiel</u>: Metropolis-Algorithmus (für einen Übergang)

1. Vorgegeben $\vec{X}_t$
2. Neue Konfiguration $\vec{Y}$ aus $\vec{X}_t$ erzeugen
3. Zufallszahl z mit gleichmässiger Verteilung in $[0,1]$ erzeugen
4. Falls $\rho(\vec{Y}) / \rho(\vec{X}_t) > z$ wird die Konfiguration angenommen, d.h. $\vec{X}_{t+1} = \vec{Y}$
5. Sonst wird $\vec{Y}$ weggeworfen und zurück zu 2.

Anmerkungen:

1) Die Verteilung P_t der Zufallvariable X_t konvergiert gegen $P(\vec{x})$ für $t \gg 1$:

$$P_t(\vec{x}) \approx P(\vec{x}) + \mathcal{O}\left(e^{-t/\tau_R}\right)$$

τ_R ist die Relaxationszeit. Für $t \gg \tau_R$ können wir annehmen, dass $P_t \approx P$

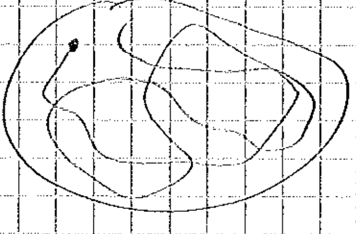
2) Die Korrelation zwischen 2 Konfigurationen $\vec{X}_t$ und $\vec{X}_{t'}$ verschwinden für $|t-t'| \gg 1$:

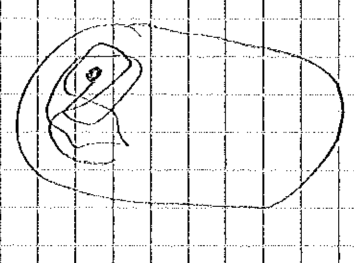
$$\mathcal{C}(t,t') = \langle \vec{X}_t \vec{X}_{t'} \rangle - \langle \vec{X}_t \rangle \langle \vec{X}_{t'} \rangle$$

$$\approx \mathcal{O}\left(e^{-|t-t'|/\tau_c}\right)$$

τ_c ist die (Auto-) Korrelationszeit. Für $|t-t'| \gg \tau_c$ können wir annehmen, dass die Zufallskonfigurationen $\vec{X}_t$ und $\vec{X}_{t'}$ unabhängig voneinander sind.

3) Die Erzeugung der Konfiguration $\vec{Y}$ hängt stark vom System ab. Eine wichtige Bedingung ist dass jede Konfiguration erreichbar sein muss.





4) Im Metropolis - Algorithmus müssen wir nie die Warscheinlichkeit $P(\vec{X})$ berechnen aber nur das Verhältnis $P(\vec{Y})/P(\vec{X}_t)$ für 2 Konfigurationen

In unserem Beispiel, es ist

$$\frac{P(\{\vec{r}_i'\})}{P(\{\vec{r}_i\})} = \frac{e^{-\beta V(\{\vec{r}_i'\})}}{e^{-\beta V(\{\vec{r}_i\})}} = e^{-\beta\left[V(\{\vec{r}_i'\}) - V(\{\vec{r}_i\})\right]}$$

und kann ohne Schwierigkeit berechnet werden.

5) Bei Phasenübergängen gibt es oft ein "critical slowing down" der MC - Dynamick, d.h.

$$\tau_R, \tau_C \to \infty$$

τ_R hängt sehr stark von der Anfangskonfiguration $\vec{X}_0$ ab und kann mit einer passenden Wahl von $\vec{X}_0$ verkürzt werden

τ_C hängt sehr stark vom Verfahren für den Übergang $\vec{X}_t \to \vec{X}_{t+1}$. Sie kann oft mit Änderungen des Erzeugungsverfahrens für $\vec{Y}$ oder sogar des Algorithmus (Heat bath statt Metropolis, ...) beeinflusst werden.

Zusammenfassung: MC-Simulation

- Wir erzeugen eine Markov-Kette der Länge S
- Daraus erhalten wir M näherungsweise unabhängige und gemäss $\rho(\vec{x})$ verteilte Zufallskonfigurationen

$$X_{\tau_R + j\tau_c} \qquad j = 1, \ldots, M$$

$$(S \cong \tau_R + M\tau_c)$$

- Das Integral wird approximiert durch

$$I_M^{MC} = \frac{1}{M} \sum_{j=1}^{M} g(\vec{X}_{\tau_R + j\tau_c})$$

- Der Fehler $|I - I_M^{MC}|$ verschwindet wie $\frac{1}{\sqrt{M}}$