Kapitel 11

Monte Carlo Simulationen

12.1.2006

Erzeugung von Konfigurationen im Computer gemäß vorgegebenen Verteilung (z.B. kanonisch). Einführende Bücher:

- M.E.J. Newman and G.T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon Press, Oxford, 1999)
- D.P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, (Cambridge University Press, Cambridge 2000)

11.1 Markov Ketten

Gegeben: System mit (endlich) vielen Zuständen $\underline{y} = \underline{y}_1, \underline{y}_2, \ldots, \underline{y}_K$ und Wahrscheinlichkeiten $\rho(\underline{y})$.
Typisch: $\rho(\underline{y}) \in O(1)$ nur für wenige Zustände, aber sehr klein für die meisten Zustände.
Ziel: Messung von Mittelwerten von Observablen $A(\underline{y})$

$$\langle A \rangle = \mathrm{Sp}\,(A\rho) = \sum_{\underline{y}} A(\underline{y})\rho(\underline{y}) \tag{11.1}$$

Annahme: K sehr groß (z.B. $K = 2^N$ für N Ising Spins), $\Rightarrow$nicht alle Zustände enummerierbar
Weiterhin $\rho(A) \in O(1)$ (üblicherweise) für exponentiell kleinem Bereich (berücksichte Zustandsdichte $D(E)$!)

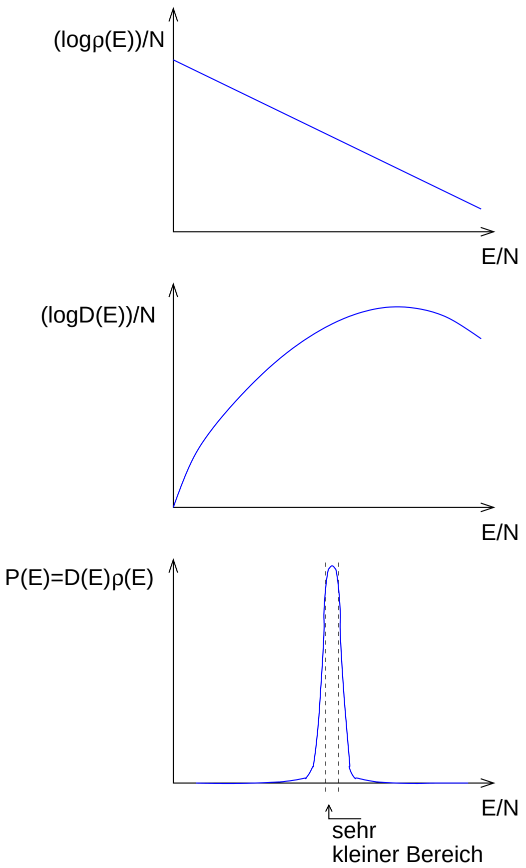
Simple Sampling:

Generiere M Zustände $\{\underline{y}^i\}$ $(i = 1, \ldots, M)$ zufällig, alle Zustände gleich wahrscheinlich⇒

$$\langle A \rangle \approx \overline{A}^{(1)} := \sum_{\underline{y}^i} A(\underline{y}^i)\rho(\underline{y}^i) / \sum_{\underline{y}_i} \rho(\underline{y}^i)$$

Nachteil: Für fast alle Werte A ist Beitrag $\sum_{\underline{y}} \delta_{A(\underline{y}),A} \rho(\underline{y})$ sehr klein
⇒man zählt meistens “0” in $\overline{A}^{(1)}$
⇒nur wenige $\underline{y}^{(i)}$ tragen bei
$\overline{A}^{(1)}$ sehr ungenau.

Beispiel Energie $\longrightarrow$
Energie scharf
⇒Anteil der beitragenden Zustände exponentiell klein in N



Importance Sampling:

Besser: erzeuge M Konfigurationen $\underline{y}^i$ so, dass sie gemäß $\rho(\underline{y}^i)$ verteilt sind (also wichtigen Zustände häufiger) ⇒

$$\langle A \rangle \approx \overline{A}^{(2)} := \sum_{\underline{y}^i} A(\underline{y}^i)/M \tag{11.2}$$

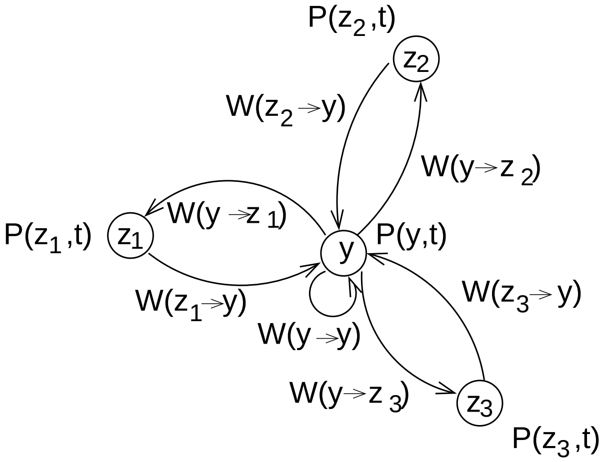
Aber: meistens unmöglich $\underline{y}^i$ direkt gemäß $\rho(\underline{y}^i)$ zu erzeugen (Verteilungsfunktion nicht ausrechenbar bzw. nicht invertierbar).

$\rightarrow$
Grundidee: Zustände $\underline{y}^i$ nicht unabhängig erzeugen, sondern durch probabilistische Dynamik zwischen Zuständen $\underline{y}(t)$ an diskreten Zeiten $t = 0, 1, 2, \ldots$: $\underline{y}(0) \rightarrow \underline{y}(1) \rightarrow \underline{y}(2) \rightarrow \ldots$.

Annahme: Zustände $\underline{y}(t+1)$ hängen nur von Zufallszahlen und von $\underline{y}(t)$ ab. $\{\underline{y}(t)|t = 0, 1, 2, \ldots\}$ heiß dann Markov Kette.
Beschreibung Übergänge $\underline{y}(t) \rightarrow \underline{y}(t+1)$ durch $W_{\underline{y}\underline{z}} = W(\underline{y} \rightarrow \underline{z}) =$ Wahrscheinlichkeit vom Zustand $\underline{y}$ (zur Zeit t) in den Zustand $\underline{z}$ (zur Zeit $t+1$) zu gehen, mit $W_{\underline{y}\underline{z}} \neq W_{\underline{y}\underline{z}}(t)$ (zeitunabhängig).
Eigenschaften:

$$\begin{aligned} W_{\underline{y}\underline{z}} &\geq 0 \quad \forall \underline{y}, \underline{z} \quad (\textit{Positivität}) \\ \sum_{\underline{z}} W_{\underline{y}\underline{z}} &= 1 \quad \forall \underline{y} \quad (\textit{Erhaltung}) \end{aligned}$$

Def.: Zustandsraum + Übergangsraten = Markov Prozess.
$P(\underline{y}, t) :=$ Wahrscheinlichkeit, dass Markov Prozess zur Zeit t im Zustand $\underline{y}(t) = \underline{y}$ ist. Bilanz für den Zustand $\underline{y}$:



⇒(Master Gleichung)

$$\Delta P(\underline{y},t) := P(\underline{y},t+1) - P(\underline{y},t) = \sum_{\underline{z}\neq\underline{y}} W_{\underline{z}\underline{y}} P(\underline{z},t) - \sum_{\underline{z}\neq\underline{y}} W_{\underline{y}\underline{z}} P(\underline{y},t) \quad \forall \underline{y} \tag{11.3}$$

$$\Rightarrow P(\underline{y},t+1) = \sum_{\underline{z}} Q_{\underline{z}\underline{y}} P(\underline{z},t)$$

mit

$$Q_{\underline{z}\underline{y}} = \begin{cases} W_{\underline{z}\underline{y}} & \text{für } \underline{z}\neq\underline{y} \\ 1-\sum_{\underline{z}\neq\underline{y}} W_{\underline{y}\underline{z}} & \text{für } \underline{z}=\underline{y} \end{cases}$$

Gegebenfalls ($\exists$ nur einen Eigenwert $\lambda = 1$ der Matrix $(Q_{\underline{y}\underline{z}})$, [1]) W.-Verteilung konvergiert gegen stationäre (zeitunabhängige) Verteilung

$$\rho_{ST}(\underline{y}) := \lim_{t\to\infty} P(\underline{y},t)$$

Das ist unabhängig vom gegebenen Startzustand $\underline{y}(0)$, Prozess heisst ergodisch.

$$\boxed{\text{Ziel: Wähle } W_{\underline{y}\underline{z}} \text{ so, dass } \rho_{ST} = \rho}$$

Da $\rho(.)$ zeitunabhängig, (11.3) ⇒

$$0 = \Delta\rho(\underline{y}) = \sum_{\underline{z}} W_{\underline{z}\underline{y}}\rho(\underline{z}) - \sum_{\underline{z}} W_{\underline{y}\underline{z}}\rho(\underline{y}) \quad \forall \underline{y}$$

Z.B. erfüllt, für

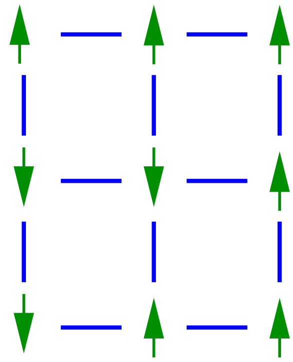
$$W_{\underline{z}\underline{y}}\rho(\underline{z}) - W_{\underline{y}\underline{z}}\rho(\underline{y}) = 0 \quad \forall \underline{y},\underline{z} \tag{11.4}$$

(genannt detaillierte Balance)

Markov Prozess erzeugt im Prinzip Zustände, die gemäß $\rho(.)$ verteilt sind
⇒Mittelwerte nach (11.2)
Achtung: Anfangs hängen die Konfigurationen oft vom Startzustand $\underline{y}(0)$ ab
(mit) Grund: Für viele Algorithmen ($\equiv \underline{\underline{W}}$): $W_{\underline{z},\underline{y}} = 0$ für fast alle $\underline{z},\underline{y}$)
⇒die ersten Zustände $t < t_{equi}$ bei der Mittelwertbildung weglassen ("Equilibrierung")
Außerdem $\underline{y}(t+1)$ zu "ähnlich" zu $\underline{y}(t)$
⇒nur "entferntere" Zustände $\underline{y}(t), \underline{y}(t+\Delta t), \underline{y}(t+2\Delta t), \ldots$ sind unabhängig.
Hinweis: $t_{equi}, \Delta t$ hängen STARK vom Modell, Parametern und Algorithmen (also $\underline{\underline{W}}$) ab.
⇒Test durch Messung von Korrelationsfunktionen

[1]siehe L.E. Reichl, *A Modern Course in Statistical Physics*, (John Wiley & Sons, New York 1998)

11.2 Ising Ferromagnet



Modell:
klassische Ising Spins ($s_i = \pm 1 =$ “hoch”/”runter”) auf regulärem d-dim. Gitter
Konfiguration: $\underline{x} = (s_1, s_2 \ldots, s_N)$

$$H = -J \sum_{\langle i,j \rangle} s_i s_j \quad J > 0. \tag{11.5}$$

$\langle i, j \rangle$: Summe über wechselwirkende Nachbarn (z.B. nächste N.)