(iii) Markov-Ketten-Monte Carlo, Metropolis-Algorithmus

Idee: erzeuge Kette $\vec{r}_1, \vec{r}_2, \ldots \vec{r}_n, \ldots$ von Zuständen, wobei die Wahrscheinlichkeitsverteilung $p(\vec{r}_n)$ asymptotisch gegen eine gewünschte Verteilung $p(\vec{r})$ konvergiert.

Definition: Ein Stochastischer Prozess ist eine Familie von Zufallsvariablen $(X_t)_{t\in I}$, wobei die Indexmenge geordnet ist: $I = \{t_0 < t_1 < \ldots\}$. Für eine abzählbare Menge I (typisch $I = \mathbb{N}$) heisst der Prozess zeitdiskret, sonst zeitstetig (z.B. für $I = \mathbb{R}_0^+$).

Bemerkung: Die Wertemengen der Zufallsvariablen können in beiden Fällen diskret oder stetig sein.

Wichtig: Zur vollständigen Charakterisierung eines Stochastischen Prozesses müssen neben den Wahrscheinlichkeitsverteilungen für festes t (1-Punkt-Verteilungen)

$$p_n(X_{t_n}) \equiv p(x_n, t_n)$$

i.a. auch alle n-Punkt-Verteilungen spezifiziert werden, z.B. $p(x_n, t_n; x_{n-1}, t_{n-1}; \ldots; x_1, t_1)$

Beispiel 3a: Wiederholter Münzwurf; bei „Zahl" steigt der Kontostand des Spielers um 1, bei „Wappen" sinkt er um 1.

(hier $t_n = n$)

$$p_0(x) = \delta_{x,0} \quad ; \quad p_1(x) = \tfrac{1}{2}(\delta_{x,-1} + \delta_{x,1});$$

$$p_2(x) = \frac{1}{4}\left(\delta_{x,-2} + 2\,\delta_{x,0} + \delta_{x,2}\right) \equiv p(x_2, 2) \;\ldots$$

aber: $p(x_n, n \mid x_{n-1}, n-1) = \frac{1}{2}\left(\delta_{x_n, x_{n-1}-1} + \delta_{x_n, x_{n-1}+1}\right)$

↑ bedingte W'

(Allgemein gilt: $\quad p(x_n, n; x_m, m) = p(x_n, n \mid x_m, m)\, p(x_m, m)$)

Beispiel 3b: ein modifiziertes Spiel, in dem nach jedem Wurf das Vorzeichen des Kontostandes invertiert wird, hat zwar die gleichen 1-Punkt-Verteilung wie Bsp. 3a, aber andere n-Punkt-Verteilungen für $n \geq 2$.

Definition: Einen stochastischen Prozess, in dem die bedingte W'verteilung nur vom letzten bekannten Zustand abhängt, nennt man Markov-Prozess:

$$p(x_n, t_n \mid x_{i_1}, t_{i_1}; \ldots x_{i_k}, t_{i_k}) = p(x_n, t_n \mid x_{i_j}, t_{i_j}) \quad \text{für } t_{i_j} = \max_{1 \leq j \leq k}\{t_{i_j}\}$$

Beispiel: Zufallsbewegungen (Random Walks) wie Bsp. 3a sind Markov-Prozesse; auch 3b ist Markov-Prozess.

Ein Markov-Prozess hat also ein minimales „Gedächtnis"; Prozesse ohne Gedächtnis nennt man unkorreliert.

Wichtiger Spezialfall: Markov-Prozesse mit stationären (zeitlich translationsinvarianten) Übergangswahrscheinlichkeiten $\quad p(x_n, t_n \mid x_k, t_k) = p_\tau(x_n \mid x_k); \quad \tau = t_n - t_k$

Jetzt: zeitdiskreter Fall mit $t_n = n$ und stationären Übergangswahrscheinlichkeiten. Betrachte speziell einzelnen Zeitschritt: $p_1(x_{n+1} | x_n) \equiv W(x_{n+1} | x_n)$ „Übergangsrate"

Es gilt die Master-Gleichung

$$\boxed{p_{n+1}(x) = p_n(x) + \sum_{x'} W(x | x') p_n(x') - \sum_{x'} W(x' | x) p_n(x)}$$

Die W'verteilung ist genau dann stationär, falls sich für jedes x die Zu- und Abflüsse die Waage halten. Beispiel: Wasserleitungen + Reservoirs

Hinreichende Bedingung für die Stationarität einer gewünschten Gleichgewichtsverteilung $p_{eq}(x)$ ist

$$\boxed{\text{Detailliertes Gleichgewicht: } \frac{W(x | x')}{W(x' | x)} = \frac{p_{eq}(x)}{p_{eq}(x')} \quad \forall\, x, x'}$$

Markov-Ketten-Monte-Carlo-Algorithmen konstruieren Markov-Ketten, für die (i) eine vorgegebene W'verteilung stationär ist (detailliertes Gleichgewicht wird erfüllt) und die (ii) ergodisch sind (alle Zustände mit $p_{eq}(x) > 0$ sind erreichbar).

Vereinfachung: betrachte in jedem Zeitschritt nur „Umgebung" von gegebenem Zustand x. Teile dazu auf:

$$W(x'|x) = W_{Vorschlag}(x'|x)\, W_{Akzeptanz}(x'|x)$$

und wähle Vorschlagsw' symmetrisch: $W_{Vorschlag}(x'|x) = W_{Vorschlag}(x|x')$

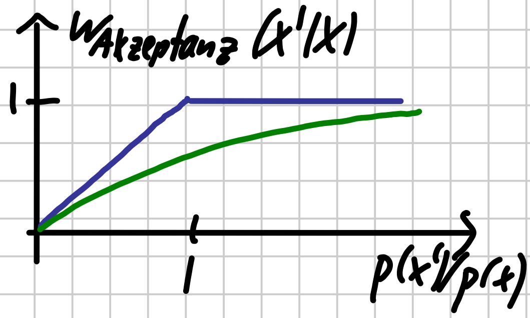
(Beispiele: Intervall/Hyperkubus im kontinuierlichen Fall, z.B. $W_{Vorschlag}(x'|x) = \frac{1}{a}\,\Theta\left(\frac{a}{2} - |x-x'|\right)$ (hier: W'dichte) (oder d-dim Gauß-Vert.) oder Einzel-Spin-Flip im Ising-Modell)

Dann reicht detailliertes Gleichgewicht für die Akzeptanzraten aus. Möglichkeiten:

(i) $W_{Akzeptanz}(x'|x) = \min\left\{1, \frac{p_{eq}(x')}{p_{eq}(x)}\right\}$ Metropolis-Regel

(ii) $W_{Akzeptanz}(x'|x) = \frac{p_{eq}(x')}{p_{eq}(x) + p_{eq}(x')}$ heat bath

Dabei hat der Metropolis-Algorithmus den Vorteil der höchstmöglichen Akzeptanzrate, der heat bath-Algorithmus stellt dagegen instantan das lokale Gleichgewicht her.



Allgemeiner Metropolis-Algorithmus (Vorschlagsw' nicht notwendig symmetrisch):

$$\boxed{W_{Akzeptanz}(x'|x) = \min\{1, r\}, \quad r = \frac{p_{eq}(x')}{p_{eq}(x)}\,\frac{W_{Vorschlag}(x|x')}{W_{Vorschlag}(x'|x)}}$$

Beachte: Im Grenzfall optimaler Vorschlagsw' werden alle Vorschläge akzeptiert.

Typische Realisierung des Metropolis-Algorithmus' in der (klassischen) Statistischen Physik:

Boltzmanngewicht $p_i \propto e^{-\beta E_i}$ für Zustand i mit Energie E_i ($\beta = (k_B T)^{-1}$)

(0) Initialisiere Konfiguration

(i) Wähle Teilchen (in Kontinuum oder auf Gitter) bzw. Spin n aus ($1 \le n \le N$) – entweder zufällig oder bei jedem Durchgang (sweep) alle in fester Reihenfolge.

(ii) Schlage Veränderung von Teilchen n vor: Umklappen oder Verdrehen des Spins, Hüpfen oder Verschieben des Teilchens; berechne Energiedifferenz ΔE

$$r = e^{-\beta \Delta E} \, \frac{W_{Vorschlag}(x_{neu} | x_{alt})}{W_{Vorschlag}(x_{alt} | x_{neu})}$$

(iii) Akzeptiere Update mit W' $\min\{1, r\}$, sonst behalte alte Konfiguration.

(iv) Messe Observablen (z.B. Ausgabe in Datei)

# sweeps erreicht?	
nein	ja

$\rightarrow$ (v) Berechne Mittelwerte mit Standardabweichung, Verteilungen, ...

Wichtig: (i) Die ersten Durchgänge, bei denen die W'-verteilung noch zu weit von der Stationarität entfernt ist, dürfen nicht für Observablenmittelwerte verwendet werden. Typische Wahl: 1% – 10% warm-up sweeps.
(ii) Zustandssumme / Freie Energie nicht messbar!