Markov-Ketten

Zustandsraum: $\mathcal{X}$

Übergangsmatrix $P = (p_{xy})$:

$$p_{xy} := Prob(X_t = y \mid X_{t-1} = x) \qquad \text{für alle } t \in \mathbb{N} \text{ und } x, y \in \mathcal{X}.$$

Übergangsw'keiten k-ter Stufe:

$$p_{xy}^{(k)} = Prob(X_{t+k} = y \mid X_t = x) \qquad \text{für alle } t \in \mathbb{N}_0.$$

stationäre Verteilung: Zeilenvektor π mit

$$\pi P = \pi.$$

Markov-Ketten (2)

$\mathcal{MC}$ **aperiodisch** $:\Longleftrightarrow$ ggT$\{t \mid p_{xx}^{(t)} > 0\} = 1$ für alle $t \in \mathbb{N}_0$.

$\mathcal{MC}$ **irreduzibel** $:\Longleftrightarrow$ für alle $x, y \in \mathcal{X}$ existiert ein $t \in \mathbb{N}_0$ mit$p_{xy}^{(t)} > 0$.

Satz: Ist $\mathcal{MC}$ eine irreduzible, aperiodische Markov-Kette, so besitzt $\mathcal{MC}$ genau eine stationäre Verteilung.

π **reversibel** $:\Longleftrightarrow$ $p_{xy}\pi_x = p_{yx}\pi_y$ für alle $x, y \in \mathcal{X}$.

Lemma: Ist $\mathcal{MC}$ eine irreduzible, aperiodische Markov-Kette und π eine reversible Verteilung, so ist π eine stationäre Verteilung.

Beispiel 1: Irrfahrt auf dem n-dimensionalen Hyperwürfel

$\mathcal{B}(n) := \{0,1\}^n$

$\mathcal{MC}_1(n)$: Markov-Kette mit Zustandsraum $\mathcal{B}(n)$ und Übergängen wie folgt: in jedem Zustand $x = (x_1, \ldots, x_n)$ wähle man zunächst zufällig und gleichverteilt einen Index $i \in \{1, \ldots, n\}$ und vertausche dann mit Wahrscheinlichkeit $\frac{1}{2}$ den Wert des i-ten Buchstabens.

Dies entspricht genau einer Irrfahrt auf dem n-dimensionalen Würfel mit den Übergangswahrscheinlichkeiten

$$p_{xy} = \begin{cases} \frac{1}{2n} & \text{falls } x \neq y \text{ und } x, y \text{ benachbart sind,} \\ 0 & \text{falls } x \neq y \text{ und } x, y \text{ nicht benachbart sind,} \\ \frac{1}{2} & \text{falls } x = y. \end{cases}$$

$\mathcal{MC}_1(n)$ ist aperiodisch und irreduzibel. $\pi \equiv 1/2^n$ ist stationäre Verteilung.

Beispiel (2): Irrfahrt auf dem $n \times n$ Gitter

$\mathcal{MC}_2(n)$: Markov-Kette, die der Irrfahrt auf dem $n \times n$ Gitter mit den Übergangswahrscheinlichkeiten

$$p_{xy} = \begin{cases} \frac{1}{8} & \text{falls } x \neq y \text{ und } x, y \text{ benachbart sind,} \\ 0 & \text{falls } x \neq y \text{ und } x, y \text{ nicht benachbart sind,} \\ 1 - \sum_{z \neq x} p_{xz} & \text{falls } x = y. \end{cases}$$

entspricht.

$\mathcal{MC}_2(n)$ hat ebenfalls die Gleichverteilung als stationäre Verteilung.

Beispiel (3): Bernoulli-Laplacesches Diffusionsmodell

Um den Austausch zwischen zwei nicht komprimierbaren Flüssigkeiten zu untersuchen, führte Daniel Bernoulli 1769 das folgende Urnenmodell ein. Ausgehend von zwei Urnen, die n rote beziehungsweise n schwarze Kugeln enthalten, werden in jedem Schritt aus jeder Urne eine Kugel ausgewählt und die beiden Kugeln ausgetauscht.

$\mathcal{MC}_3(n,k)$: Markov-Kette $\mathcal{MC}_3(n,k)$, auf den k-elementigen Teilmengen von $\{1,\dots,n\}$. In jedem Schritt wählen wir zufällig ein Element der k-elementigen Teilmenge und eines aus dem Rest und tauschen beide aus:

$$p_{AB} = \begin{cases} \frac{1}{k(n-k)} & \text{falls } |A \triangle B| = 2, \\ 0 & \text{sonst.} \end{cases}$$

$\pi \equiv 1/\binom{n}{k}$ ist stationäre Verteilung – was der Vorstellung vom Mischen zweier Flüssigkeiten entspricht.

Beispiel (4): Matchings in Graphen.

$G = (V, E)$ beliebiger Graph.

$\mathcal{MC}_4(G)$: Markov-Kette mit Zustandsraum $\mathcal{M}(G)$, der Menge aller Matchings von G. In jedem Zustand $M \in \mathcal{M}(G)$ wählen wir eine Kante $e \in E$ zufällig und gleichwahrscheinlich und gehen mit Wahrscheinlichkeit $\frac{1}{2}$ zum Zustand $\tilde{M}$ über, der definiert ist durch

$$\tilde{M} = \begin{cases} M - e & \text{falls } e \in M, \\ M + e & \text{falls } e \notin M \text{ und } M + e \in \mathcal{M}(G), \\ M & \text{sonst.} \end{cases}$$

Auch hier gilt: Gleichverteilung ist die (eindeutig bestimmte) stationäre Verteilung.

Beispiel (5): Spannende Bäume in Graphen

$G = (V, E)$ beliebiger, zusammenhängender Graph

$\mathcal{MC}_5(G)$: Als Zustandsraum wählen wir diesmal die Menge $\mathcal{S}(G)$ aller spannenden Bäume von G. In jedem Zustand $S \in \mathcal{S}(G)$ wählen wir wie oben zwei Kanten e_1 und e_2 zufällig und gleichwahrscheinlich und gehen zum Zustand T über, der definiert ist durch

$$T = \begin{cases} S - e_1 + e_2 & \text{falls } e_1 \in S,\, e_2 \notin S \text{ und } S - e_1 + e_2 \in \mathcal{S}(G), \\ S & \text{sonst.} \end{cases}$$

Auch hier gilt: Gleichverteilung ist die (eindeutig bestimmte) stationäre Verteilung.

Schnell mischende Markov-Ketten

$\mathcal{MC}$ beliebige Markov-Kette:

Für $t \in \mathbb{N}$ und $x \in \mathcal{X}$ bezeichnen wir mit $\mathcal{A}_{MC}(t,x)$, denjenigen (randomisierten) Algorithmus, der ausgehend vom Zustand x genau t Schritte der Markov-Kette simuliert und den dann erreichten Zustand ausgibt.

Als ε-Konvergenzzeit der Markov-Kette bezeichnen wir diejenige Schrittanzahl, ab der sichergestellt ist, daß die Wahrscheinlichkeit mit der ein Zustand y ausgegeben wird um höchstens ε von der Wahrscheinlichkeit π_y von y im stationären Zustand abweicht – und das unabhängig vom Startzustand x. Formal ist die ε-Konvergenzzeit definiert durch

$$\tau(\varepsilon) = \min\{t : (1-\varepsilon)\pi_y \leq Prob(\mathcal{A}_{MC}(t,x) = y) \leq (1+\varepsilon)\pi_y \text{ für alle } x, y \in \mathcal{X}\}.$$

Schnell mischende Markov-Ketten (2)

Ein ε**-uniformer Erzeuger** für eine Menge $\mathcal{S}$ ist ein randomisierter Algorithmus $\mathcal{A}$ mit

$$(1-\varepsilon)\frac{1}{|\mathcal{S}|} \leq Prob(\mathcal{A} = x) \leq (1+\varepsilon)\frac{1}{|\mathcal{S}|} \quad \text{für alle } x \in \mathcal{S}.$$

Satz: Ist $\mathcal{MC}$ eine Markov-Kette mit der Gleichverteilung als (eindeutigem) stationärem Zustand, so gilt: Für alle $x \in \mathcal{X}$ ist $\mathcal{A}_{MC}(\tau(\varepsilon), x)$ ein ε-uniformer Erzeuger.

Ist $\mathcal{MC}(I)$ eine Familie von Markov-Ketten, indiziert durch Probleminstanzen I, so nennen wir $\mathcal{MC}(I)$ **schnell mischend**, falls sich die Konvergenzzeit $\tau(\varepsilon)$ jeder Markov-Kette durch ein Polynom in der Eingabelänge $|I|$ und $\ln \varepsilon^{-1}$ beschränken läßt.

Leitwert einer Markov-Kette

Wir ordnen jeder ergodischen Markov-Kette $\mathcal{MC}$ mit stationärer Verteilung π einen gerichteten Graphen $\mathcal{G}_{MC} = (V, A)$ mit Knotenmenge $V = \mathcal{X}$, Kantenmenge

$$A = \{(x, y) \mid x, y \in \mathcal{X}, x \neq y \text{ und } p_{xy} > 0\}$$

zu und definieren eine Gewichtsfunktion c auf $\mathcal{X} \times \mathcal{X}$ durch

$$c(x, y) = \begin{cases} \pi_x p_{xy} & \text{falls } (x, y) \in A, \\ 0 & \text{sonst.} \end{cases}$$

Für eine Menge $S \subseteq V$ von Zuständen schreiben wir abkürzend $\pi(S)$ für die Summe $\sum_{x \in S} \pi(x)$. Entsprechend bezeichnet $c(S, \bar{S})$ die Summe $\sum_{x \in S} \sum_{y \notin S} c(x, y)$.

Leitwert einer Markov-Kette (2)

Der **Leitwert** der Markov-Kette $\mathcal{MC}$ ist dann definiert durch

$$\Phi_{MC} = \min_{\substack{S \subset \mathcal{X} \\ 0 < \pi(S) \leq \frac{1}{2}}} \frac{c(S, \bar{S})}{\pi(S)}.$$

$\pi_{\min} := \min_{x \in \mathcal{X}} \pi_x$

Satz: Ist $\mathcal{MC}$ eine aperiodische, ergodische und reversible Markov-Kette, so gilt für die ε-Konvergenzzeit

$$\tau(\varepsilon) \leq \frac{2}{\Phi_{MC}} \left(\ln \varepsilon^{-1} + \ln \pi_{min}^{-1} \right).$$

Mehrgüterflüsse in Markov-Ketten

Auf dem gewichteten Graphen $\mathcal{G}_{MC}$ definieren wir folgendes Mehrgüter Flußproblem: Für jedes (geordnete) Paar (x, y) von Zuständen soll von dem Gut g_{xy} die Menge $\pi_x \pi_y$ von x nach y transportiert werden. Mit anderen Worten, wir suchen Flüsse $f_{xy} : A \to \mathbb{R}_{\geq 0}$, so daß

$$\sum_u f_{xy}(u, v) = \sum_u f_{xy}(v, u) \qquad \text{für alle } v \neq x, y$$

und

$$\sum_u f_{xy}(x, u) = \sum_u f_{xy}(u, y) = \pi_x \pi_y .$$

Zu gegebenen Flüssen f_{xy} definieren wir den **Gesamtfluß** f durch

$$f(e) = \sum_{x \neq y} f_{xy}(e)$$

und definieren die relative Kantenauslastung durch

$$\rho(f) = \max_{e \in A} \frac{f(e)}{c(e)} .$$

Mehrgüterflüsse in Markov-Ketten (2)

Lemma: Für jede Markov-Kette $\mathcal{MC}$ mit Fluß f gilt

$$\Phi_{MC} \geq \frac{1}{2\rho(f)}.$$

Satz: Ist $\mathcal{MC}(I)$ eine Familie von aperiodischen, ergodischen und reversiblen Markov-Ketten für die $\ln \pi_{\mathrm{min}}^{-1}$ polynomiell beschränkt ist, dann ist $\mathcal{MC}(I)$ genau dann schnell mischend, wenn es Flüsse f gibt, so daß $\rho(f)$ polynomiell beschränkt ist.

Irrfahrt im Hyperwürfel

$$\pi \equiv 1/2^n \quad \Longrightarrow \quad \ln \pi_{\min}^{-1} = n.$$

Es genügt daher zu zeigen, daß es Flüsse f gibt, so daß $\rho(f)$ polynomiell beschränkt ist.

Für zwei Zustände $x \neq y$ verteilen wir den zu verschickenden Fluß von $\pi_x \pi_y$ gleichmäßig auf alle kürzesten Wege, die x und y verbinden. (Haben x und y den Abstand d, gibt es davon $d!$ viele.)

Wie groß ist nun $\rho(f)$?

Aus Symmetriegründen muß der Fluß auf jeder Kante gleich groß sein. Es genügt daher den "Gesamtfluß" $\sum_{e \in A} f(e)$ zu berechnen und durch die Anzahl Kanten des Graphens zu teilen.

Irrfahrt im Hyperwürfel (2)

Der Gesamtfluß läßt sich einfach berechnen: es gibt $2^n \cdot \binom{n}{d}$ Paare (x, y) mit Abstand d und die für solche Paare in der Definition von f_{xy} verwendeten Wege haben alle die Länge d. Es gilt daher

$$\sum_{e \in A} f_{xy}(e) = d \cdot \pi_x \pi_y.$$

Für den Fluß $f(e)$ gilt daher

$$\begin{aligned} f(e) &= \frac{1}{|A|} \sum_{d=1}^{n} 2^n \binom{n}{d} \cdot d \cdot \pi_x \pi_y \\ &= \frac{1}{n2^n} \sum_{d=1}^{n} 2^n \binom{n-1}{d-1} \cdot \frac{n}{2^{2n}} \\ &= \frac{1}{2^{2n}} \sum_{d=0}^{n-1} \binom{n-1}{d} \\ &= \frac{2^{n-1}}{2^{2n}} \end{aligned}$$

und somit wegen $c(e) \equiv (2n2^n)^{-1}$ also $\rho(f) = n$.