Handout zu: Top-to-random shuffle und strong stationary times

Manuel Hinz

30.06.2020

Was bisher geschah: Definitionen aus LPW

Definition 1. *A random mapping representation of a transition matrix P on state space $\mathcal{X}$ is*

- *a function $f : \mathcal{X} \times \Lambda \to \mathcal{X}$*
- *along with a Λ-valued random variable Z*
- *satisfying $P\{f(x, Z) = y\} = P(x, y)$*

Definition 2. *The total variation distance between two probability distributions μ and υ on $\mathcal{X}$ is defined by*

$$\|\mu - \upsilon\|_{TV} = \max_{A \subseteq \mathcal{X}} |\mu(A) - \upsilon(A)|. \tag{1}$$

Definition 3.

$$d(t) := \max_{x \in \mathcal{X}} \|P^t(x, \cdot) - \pi\|_{TV} \tag{2}$$

$$\bar{d}(t) := \max_{x,y \in \mathcal{X}} \|P^t(x, \cdot) - P^t(y, \cdot)\|_{TV} \tag{3}$$

Definition 4. *The mixing time is defined by*

$$t_{mix}(\epsilon) := \min\{t : d(t) \leq \epsilon\} \tag{4}$$

and

$$t_{mix} := t_{mix}(\frac{1}{4}) \tag{5}$$

Definition 5. *Suppose a probability distribution π on $\mathcal{X}$ satisfies:*

$$\forall_{x,y \in \mathcal{X}} \pi(x)P(x, y) = \pi(y)P(y, x) \tag{6}$$

The equations are called the detailed balance equations.

Definition 6. *The time reversal of an irreducible Markov chain with transition matrix P and stationary distribution π is the chain with matrix*

$$\hat{P}(x, y) := \frac{\pi(y)P(y, x)}{\pi(x)} \tag{7}$$

The stationary equation $\pi = \pi P$ implies that $\hat{P}$ is a stochastic matrix.

Inhalt

Top-to-random shuffle

- Ein Deck aus n Karten soll gemischt werden.
- Oberste Karte wird an eine zufällige Position im verbleibenden Deck gesteckt.
- Dieser Mischvorgang entspricht einem **random walk** auf der Gruppe S_n.

Proposition 1. *Sei X_t ein random walk auf S_n, welcher dem top-to-random Mischvorgang entspricht. Seien zum Zeitpunkt t k Karten unter der Karte ρ , welche bei $t = 0$ zu unterst liegt. Dann ist jede der $k!$ möglichen Ordnung gleich wahrscheinlich. Sei τ_{top} die Zeit, welche einen Zeitschritt nach dem ρ die oberste Karte ist. Die Verteilung von $X_{\tau_{top}}$ ist dann gleichmäßig über S_n und die Zeit τ_{top} ist unabhängig von $X_{\tau_{top}}$.*

Definition 2. *Für eine Menge Ω ist eine **σ-Algebra** eine Menge $\mathcal{F}$ von Teilmengen mit*

(i) $\Omega \in \mathcal{F}$,

(ii) *wenn* $A_1, A_2, \dots \in \mathcal{F}$, *dann schon* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, *und*

(iii) *wenn* $A \in \mathcal{F}$, *dann* $A^c = \Omega \setminus A \in \mathcal{F}$.

Sei $\mathcal{A}$ eine Menge von Mengen. Wir schreiben $\sigma(\mathcal{A})$ für die kleinste σ-Algebra, welche $\mathcal{A}$ enthält.

Definition 3. *Eine Menge Ω mit einer σ-Algebra $\mathcal{F}$ und einer Funktion $f : \Omega \to \mathbb{R}$ wird **messbar** genannt, wenn $f^{-1}(B) \in \mathcal{F}$ für alle offenen Mengen $B \subset \mathbb{R}$ gilt.*

Definition 4. *Sei $\{\mathcal{F}_t\}$ eine Filtration, d.h. eine Sequenz von σ-Algebraen s.d. $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ für alle t gilt. $\{X_t\}$ heißt adaptiert zu $\{\mathcal{F}_t\}$, wenn X_t für alle t $\mathcal{F}_t$-messbar ist. Wenn $\mathcal{H}_t = \sigma(X_0, X_1, \dots, X_t)$, dann heißt $\{\mathcal{H}_t\}$ die natürliche Filtration. Sei $\{X_t\}$ zu $\{\mathcal{F}_t\}$ adaptiert. Man nennt $\{X_t\}$ eine **Markovkette bezüglich** $\{\mathcal{F}_t\}$, wenn*

$$P_x\{X_{t+1} = y | \mathcal{F}_t\} = P(X_t, y),$$

wobei P die Transitionsmatrix. Eine Markovkette erfüllt die Gleichung, wenn $\{\mathcal{F}_t\}$ die natürliche Filtration ist.

Definition 5. *Eine **Stoppzeit** (stopping time) für eine Filtration $\{\mathcal{F}_\sqcup\}$ ist eine $\{0, 1, \dots\}$-wertige Zufallsgröße τ s.d. $\{\tau = t\} \in \mathcal{F}_t$. Wenn die Filtration einer Stoppzeit nicht angegeben ist, so wird die natürliche Filtration angenommen. Bsp: τ_{top}.*

Sei $A \subset \mathcal{X}$. Dann ist

$$\tau_A = \min\{t \geq 0 : X_t \in A\}$$

eine hitting time un die erste Zeit zu der die Sequenz (X_t) in A ist. τ_A ist für die natural filtration eine stopping time, da $\{X_0 \notin A, X_1 \notin A, \dots X_{t-1} \notin A, X_t \in A\} \subset \sigma(X_0, X_1, \dots, X_t)$.

Definition 6. *Für einen Startwert x nennt man $y \in \mathcal{X}$ einen **halting state** für eine stopping time τ, falls*

$$X_t = y \implies \tau \leq t. \tag{8}$$

(Strong) Stationary Times

Definition 7. *Sei (X_t) eine irreduzible Markovkette mit Gleichgewichtsverteilung π. Sei nun $\{\mathcal{F}_t\}$ eine Filtration und $\{X_t\}$ adaptiert zu $\{\mathcal{F}_t\}$. Eine **stationary time** (Gleichgewichtszeit) τ von (X_t) ist eine $\{\mathcal{F}_t\}$ stopping time, evtl. abhängig von der Startposition x, s.d die Verteilung von X_τ π ist:*

$$\forall y : P_x\{X_\tau = y\} = \pi(y). \tag{9}$$

Definition 8. *Sei (X_t) eine Markovkette bezüglich der Filtration $\{\mathcal{F}_t\}$, mit Gleichgewichtsverteilung π. Eine **strong stationary time** (starke Gleichgewichtszeit) von (X_t) und Anfangsposition x ist eine $\{\mathcal{F}_t\}$-stopping time τ, s.d. für alle t und alle y folgendes gilt:*

$$P_x\{\tau = t, X_\tau = y\} = P_x\{\tau = t\}\pi(y). \tag{10}$$

Mit anderen Worten ist π unabhängig von τ. Beispiel: τ_{top}.

Remark 9. *Sei τ eine strong stationary time für die Startposition x, dann gilt:*

$$\begin{aligned} P_x\{\tau \leq t, X_t = y\} &= \sum_{s \leq t} \sum_z P_x\{\tau = s, X_s = z, X_t = y\} \\ &= \sum_{s \leq t} \sum_z P^{t-s}(z, y) P_x\{\tau = s\}\pi(z). \end{aligned}$$

Es folgt, da π eine Gleichgewichtsverteilung ist, dass $\sum_z \pi(z)P^{t-s}(z,y)P_x\{\tau = s\} = \pi(y)$. Damit gilt für alle $t \geq 0$ und y

$$P_x\{\tau \leq t, X_t = y\} = P_x\{\tau \leq t\}\pi(y). \tag{11}$$

Proposition 10. *Sei τ eine strong stationary time für die Startposition x, dann gilt:*

$$\|P^t(x, \cdot) - \pi\|_{TV} \leq P_x\{\tau > t\}. \tag{12}$$

Definition 11. *Die separation distance (Trennabstand) wird durch*

$$s_x(t) := \max_{y \in \mathcal{X}} \left[1 - \frac{P^t(x,y)}{\pi(y)} \right] \tag{13}$$

definiert. Sei ebenfalls:

$$s(t) := \max_{x \in \mathcal{X}} s_x(t). \tag{14}$$

Lemma 12. *Sei τ eine strong stationary time für die Startposition x, dann gilt:*

$$s_x(t) \leq P_x\{\tau > t\}. \tag{15}$$

Proposition 13. *Wenn ein halting state für den Startwert x existiert, dann ist τ eine optimal strong stationary time für x, d.h.*

$$s_x(t) = P_x\{\tau > t\}$$

außerdem gilt $P_x\{\tau > t\} \leq P_x\{\rho > t\}$ für jede weitere strong stationary time ρ.

Lemma 14. *Für die seperation distance $s_x(t)$ gilt:*

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq s_x(t),$$

und damit auch $d(t) \leq s(t)$.

Proposition 15. *Sei (X_t) die Top-to-random Markovkette mit n Karten. Für jedes $\epsilon > 0$ existiert eine Konstante $\alpha(\epsilon)$ s.d. $\alpha > \alpha(\epsilon)$ impliziert, dass es für alle genügent große n*

$$d_n(n \log n - \alpha n) \geq 1 - \epsilon \tag{16}$$

D.h.

$$t_{mix}(1-\epsilon) \geq n \log n - \alpha n. \tag{17}$$