Strong stationary times

Manuel Hinz

30.06.2020

Übersicht

Handout

- Rückblick,
- klassisches Handout.

Intro

- Beispiele: Karten mischen (S_n), layz random walk auf einem Hyperwürfel, ...
- Big Picture: Schranken für τ_{mix}.

Aufteilung

Karten mischen für Menschen mit zu viel Zeit

- 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.

```python
def insert_random(deck:List[int], card:int)->List[int]:
    i=randint(0,len(deck))
    return deck[0:i]+[card]+deck[i:]

def shuffle(deck:List[int], tau:int=0)->Tuple[List[int], int]:
    if deck[0]==len(deck): return insert_random(deck[1:], deck[0]), tau+1
    else: return shuffle(insert_random(deck[1:], deck[0]), tau+1)
```

Abbildung: Algorithmus in python

Top-to-random shuffle

Proposition

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 $X_{\tau_{\text{top}}}$ ist dann gleichmäßig über S_n und die Zeit τ_{top} ist unabhängig von $X_{\tau_{\text{top}}}$.

Beweis der Proposition

Beweis durch Induktion über t

Für $t = 0$ ist die Aussage trivial. Wenn die Aussage zum Zeitpunkt t gilt, so sind zwei Fälle möglich:

- Die zum Zeitpunkt t oberste Karte wird über ρ eingefügt, dann hat sich die Reinfolge der Karten unter ρ nicht geändert.
- Falls die zum Zeitpunkt t oberste Karte unter ρ eingefügt wird, so ist jeder der $(k+1)!$ Möglichkeiten gleichwahrscheinlich.

Aufteilung

σ-Algebra

Definition

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.

Messbar

Definition

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.

Filtration

Definition

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

Stopping time

Definition

Eine Stoppzeit (stopping time) für eine Filtration $\{\mathcal{F}_t\}$ 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}.

Hitting times

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 $\{X_0 \notin A, X_1 \notin A, \dots X_{t-1} \notin A, X_t \in A\} \subset \sigma(X_0, X_1, \dots, X_t)$.

Beispiele für stopping times

- Verkauf ein Future, wenn das underlying für weniger als X gehandelt wird.
- KEIN Beispiel: Verkaufe, wenn der Verkaufspreis am höchsten ist.

Aufteilung

Definition: Stationary Times

Definition

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 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:

$$\forall y : P_x\{X_\tau = y\} = \pi(y). \tag{1}$$

Definition: Strong Stationary Times

Definition

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

$$P_x\{\tau = t, X_\tau = y\} = P_x\{\tau = t\}\pi(y). \tag{2}$$

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

Bemerkung

Bemerkung

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 $\sum_z \pi(z) P^{t-s}(z, y) = \pi(y)$. Damit gilt für alle $t \geq 0$ und y

$$P_x\{\tau \leq t, X_t = y\} = P_x\{\tau \leq t\} \pi(y). \tag{3}$$

Proposition

Proposition

Sei τ eine strong stationary time für die Startposition x, dann gilt:

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq P_x\{\tau > t\}. \tag{4}$$

Definition: Seperation distance

Definition

Die separation distance wird durch

$$s_x(t) := \max_{y \in \mathcal{X}} \left[1 - \frac{P^t(x,y)}{\pi(y)} \right] \tag{5}$$

definiert. Sei ebenfalls:

$$s(t) := \max_{x \in \mathcal{X}} s_x(t). \tag{6}$$

Lemma 1

Lemma

Sei τ eine strong stationary time für die Startposition x, dann gilt:

$$s_x(t) \leq P_x\{\tau > t\}. \tag{7}$$

Beweis: Lemma 1

Beweis.

Nebenrechnung:

$$1 - \frac{P^t(x,y)}{\pi(y)} = 1 - \frac{P_x\{X_t = y\}}{\pi(y)} \leq 1 - \frac{P_x\{X_t = y, \tau \leq t\}}{\pi(y)}$$

Es folgt aus der Bemerkung 3, dass die RHS folgendem gleicht:

$$1 - \frac{\pi(y) P_x\{\tau \leq t\}}{\pi(y)} = P_x\{\tau > t\} \tag{8}$$

□

Halting state

Definition

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

$$X_t = y \implies \tau \leq t. \tag{9}$$

Optimal strong stationary time

Proposition

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.

Beweis und Kommentar

Beweis.

Falls y ein halting state für den Startwert x und die stopping time τ ist, dann wird aus der Ungleichung

$$1 - \frac{P^t(x,y)}{\pi(y)} \leq 1 - \frac{P_x\{X_t = y, \tau \leq t\}}{\pi(y)}$$

eine Gleichung für alle t. Daher impliziert die Existenz eines halting states die geforderte Gleichung. □

Lemma 2

Lemma

Für die seperation distance $s_x(t)$ gilt:

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

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

Beweis: Lemma 2

Beweis.

$$\|P^t(x,\cdot) - \pi\|_{TV} = \sum_{\substack{y \in \mathcal{X} \\ P^t(x,y) < \pi(y)}} [\pi(y) - P^t(x,y)]$$

$$= \sum_{\substack{y \in \mathcal{X} \\ P^t(x,y) < \pi(y)}} \pi(y) \left[1 - \frac{P^t(x,y)}{\pi(y)}\right] \leq \max_y \left[1 - \frac{P^t(x,y)}{\pi(y)}\right] = s_x(t).$$

□

Durch Lemma 1 und Lemma 2 wurde nun die Proposition bewiesen.

Lazy random walk auf einem hypercube

- lazy random walk auf $\{0,1\}^n$.
- τ_{refresh} ist eine strong stationary time.
- Rückblick: coordinate by coordinate coupling und damit auch coupon collector.
- (Proposition 2.4): $P\{\tau > \lceil n\log n + cn \rceil\} \leq e^{-c}$.
- Durch 7 folgt dann $s_x(n\log n + cn) \leq e^{-c}$.
- 10 $:\|P^t(x,\cdot) - \pi\|_{TV} \leq s_x(t)$
- $\|P^{n\log n + cn}(x,\cdot) - \pi\|_{TV} \leq e^{-c}$
- Es folgt $t_{\text{mix}}(\epsilon) \leq n\log n + n\log(\epsilon^{-1})$

Top-to-random shuffle Abschätzung

Anwendungsbeispiel

Wie schon voher bemerkt ist τ_{top} eine strong stationary time. Die Wahrscheinlichkeit, dass eine Karte unter ρ gelegt wird, wenn schon k unter ρ sind, beträgt $(k+1)/n$. Daher hat τ_{top} die gleiche Verteilung wie im Beispiel coupon collector (Seite 63 LPW). Dann folgt aus der Proposition 4 und der Proposition 2.4 (LPW)

$$t_{\text{mix}}(\epsilon) \leq n \log n + n \log(\epsilon^{-1}).$$

Untere Schranke für t_{mix} vom Top-to-random shuffle

Proposition

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

$$d_n(n \log n - \alpha n) \geq 1 - \epsilon \tag{10}$$

D.h.

$$t_{\text{mix}}(1 - \epsilon) \geq n \log n - \alpha n. \tag{11}$$

Untere Schranke für t_{mix} vom Top-to-random shuffle

Beweis 1/4

Seien die A_j die Ereignisse, dass die unteren j Karten unverändert sind. Sei τ_j die Zeit zu welcher die jte Karte von unten die oberste Karte ist.

$$\tau_j = \sum_{i=j}^{n-1} \tau_{j,i}$$

Wobei $\tau_{j,i}$ die Zeiten sind, welche die jte Karte braucht um zur Position i zu kommen. Die $\tau_{j,i}$ sind unabhängig und geometrisch verteilt mit $p = i/n$. Daher $E(\tau_{j,i}) = n/i$ und $Var(\tau_{j,i}) < n^2/i^2$.

$$E(\tau_j) = \sum_{i=j}^{n-1} n/i \geq n \int_j^n \frac{dx}{x} = n(\log n - \log j) \tag{12}$$

Untere Schranke für t_{mix} vom Top-to-random shuffle

Beweis 2/4

$$Var(\tau_j) \leq n^2 \sum_{i=j}^{\infty} \frac{1}{i(i-1)} \leq \frac{n^2}{j-1} \tag{13}$$

12,13 und Chebychev's Ungleichung liefern

$$P\{\tau_j < n \log n - \alpha n\} \leq P\{\tau_j - E(\tau_j) < -n(\alpha - \log j)\} \leq \frac{1}{j-1}$$

Falls $\alpha \geq \log j + 1$. Sei $t_n(\alpha) = n \log n - \alpha n$. Falls $\tau_j \geq t_n(\alpha)$ so bleiben die ursprünglichen j untersten Karten in der ursprünglichen Reinfolge zur Zeit $t_n(\alpha)$:

$$P^{t_n(\alpha)}(A_n, A_j) \geq P\{\tau_j \geq t_n(\alpha)\} \geq 1 - \frac{1}{j-1},$$

für $\alpha \geq \log j + 1$.

Untere Schranke für t_{mix} vom Top-to-random shuffle

Beweis 3/4

Für die gleichverteilte Gleichgewichtsverteilung

$$\pi(A_j) = 1/(j!) \leq (j-1)^{-1}$$

gilt für $\alpha \geq \log j + 1$

$$d_n(t_n(\alpha)) \geq \|P^{t_n(\alpha)}(A_n, A_j) - \pi\|_{TV} \geq P^{t_n(\alpha)}(id, A_j) - \pi(A_j) > 1 - \frac{2}{j-1}. \tag{14}$$

Für $j = \lceil e^{\alpha-1} \rceil$ gilt für $n \geq e^{\alpha-1}$ gilt nun folgendes

$$d_n(t_n(\alpha)) > g(\alpha) := 1 - \frac{2}{\lceil e^{\alpha-1} \rceil - 1}. \tag{15}$$

Untere Schranke für t_{mix} vom Top-to-random shuffle

Beweis 4/4

Es folgt

$$\liminf_{n\to\infty} d_n(t_n(\alpha)) \geq g(\alpha)$$

Und $g(\alpha) \to 1$ wenn $\alpha \to \infty$

Realität

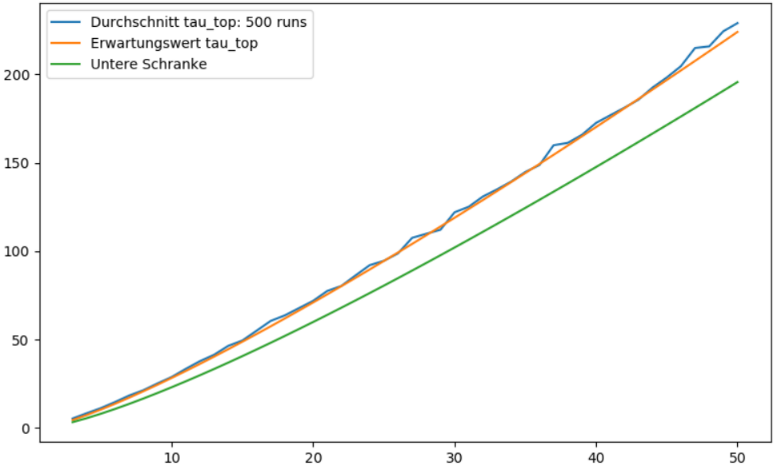

Abbildung: τ_{top}

Des weiteren ...

- $$s(2t) \leq 1 - (1 - \bar{d}(t))^2 \leq 2\bar{d}(t) \leq 4d(t). \qquad \text{(Lemma 6.17)}$$

Des weiteren ...

- $$s(2t) \leq 1 - (1 - \bar{d}(t))^2 \leq 2\bar{d}(t) \leq 4d(t). \qquad \text{(Lemma 6.17)}$$
- Für jede irreduzible und aperiodische Markovkette (X_t). Für jede Startposition x existiert eine optimale strong stationary time. D.h. es existiert eine strong stationary time s.d. für alle $t \geq 0 : s_x(t) = P_x\{\tau > t\}$.

Des weiteren ...

- $$s(2t) \leq 1 - (1 - \bar{d}(t))^2 \leq 2\bar{d}(t) \leq 4d(t). \qquad \text{(Lemma 6.17)}$$
- Für jede irreduzible und aperiodische Markovkette (X_t). Für jede Startposition x existiert eine optimale strong stationary time. D.h. es existiert eine strong stationary time s.d. für alle $t \geq 0 : s_x(t) = P_x\{\tau > t\}$.
- Code: Auf Sciebo.
- Literatur: LPW, H und das Alma 1 Skript von Harbrecht.