Vorlesung 11b

Markovketten

Teil 3

0. Wiederholung

Die Wahrscheinlichkeit eines Pfades von n Schritten:

Für eine Markovkette $X = (X_0, X_1, \ldots)$ mit Start in a und Übergangsmatrix P

hat man für jedes $n = 1, 2, \ldots$ die Wahrscheinlichkeiten

$$\mathbf{P}_a(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n)$$
$$= P(a, a_1)P(a_1, a_2) \cdots P(a_{n-1}, a_n).$$

Zerlegung nach dem ersten Schritt.

$$\mathbf{P}_a(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n)$$

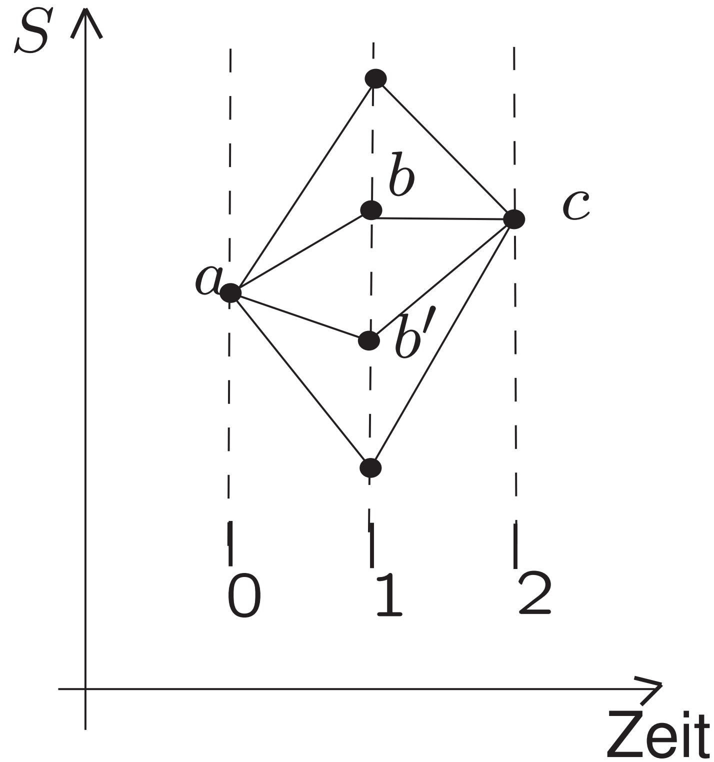
$$= P(a, a_1)P(a_1, a_2) \cdots P(a_{n-1}, a_n)$$

$$= P(a, a_1)\mathbf{P}_{a_1}(X_1 = a_2, \dots, X_{n-1} = a_n)$$

Summation über $a_1, a_2, \dots, a_{n-1}$,

mit b statt a_1 und c statt a_n:

$$\mathbf{P}_a(X_n = c) = \sum_{b \in S} P(a, b)\,\mathbf{P}_b(X_{n-1} = c)\ .$$



hier für $n = 2$

Treffwahrscheinlichkeiten

Die Frage:

P sei eine Übergangsmatrix auf der Menge S

X sei Markovkette mit Übergangsmatrix P.

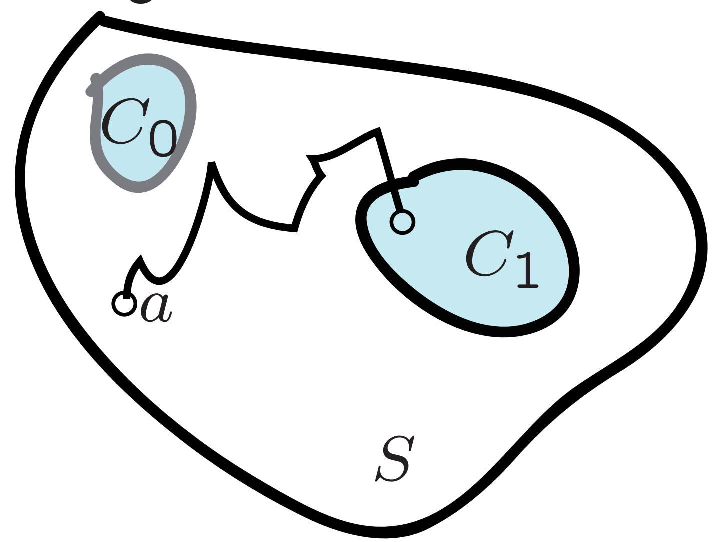
C_0, C_1 seien zwei disjunkte Teilmengen von S

Wie wahrscheinlich ist es,

dass der in $a \in S$ startende Pfad

die Menge C_1

vor der Menge C_0 trifft?



Treffwahrscheinlichkeiten

Die Antwort:

(per Zerlegung nach dem ersten Schritt)

Sei $w(a)$ die gesuchte Wahrscheinlichkeit.

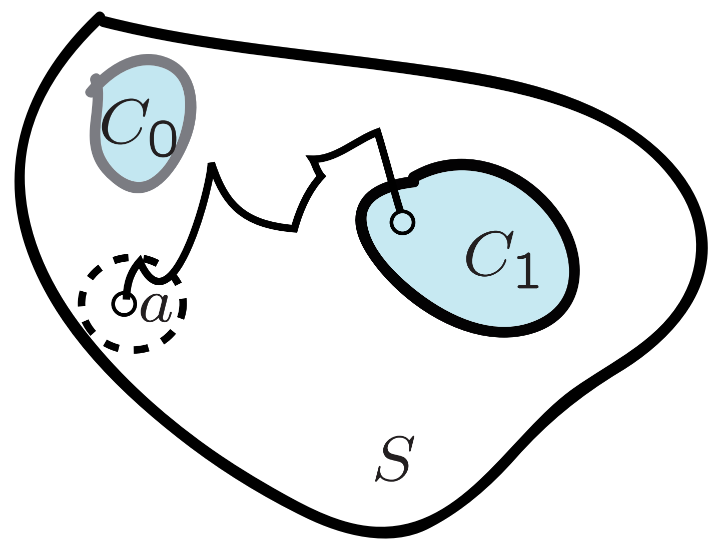
Die Zahlen $w(a)$, $a \in S$, erfüllen das Gleichungssystem

$$\sum_{b \in S} P(a,b)w(b) = w(a)$$

für $a \in S \setminus (C_0 \cup C_1)$,

$w(a) = 1$ für $a \in C_1$,

$w(a) = 0$ für $a \in C_0$.



1. Erwartete Treffzeiten

Sei X eine Markovkette mit Zustandsraum S

und Übergangsmatrix P.

Für eine Teilmenge $C \subset S$ ist

$$T_C := \min\{n : n \geq 0, X_n \in C\}$$

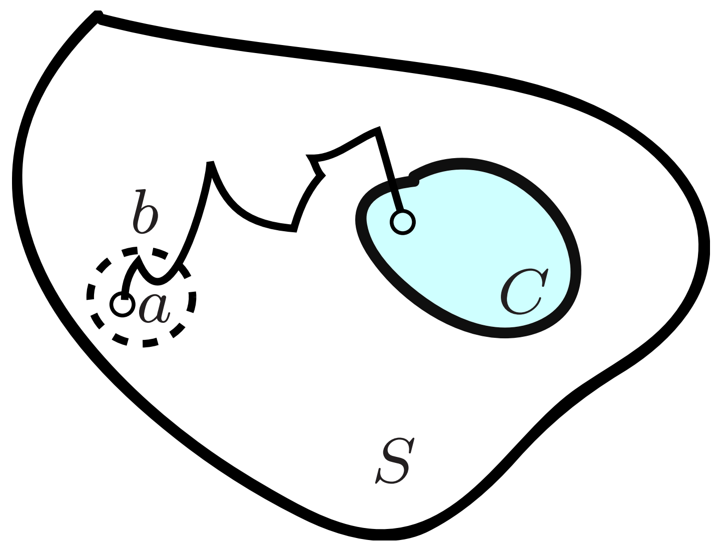
die erste Treffzeit von C.

Es geht um die Berechnung von $\mathbf{E}_a[T_C]$:

Für $a \notin C$

Zerlegung von $\mathbf{E}_a[T_C]$

nach dem ersten Schritt:



Erst ein Schritt

von a nach b gemäß $P(a,b)$,

dann “Neustart” in b:

$$\mathbf{E}_a[T_C] = 1 + \sum_{b \in S} P(a,b)\,\mathbf{E}_b[T_C]$$

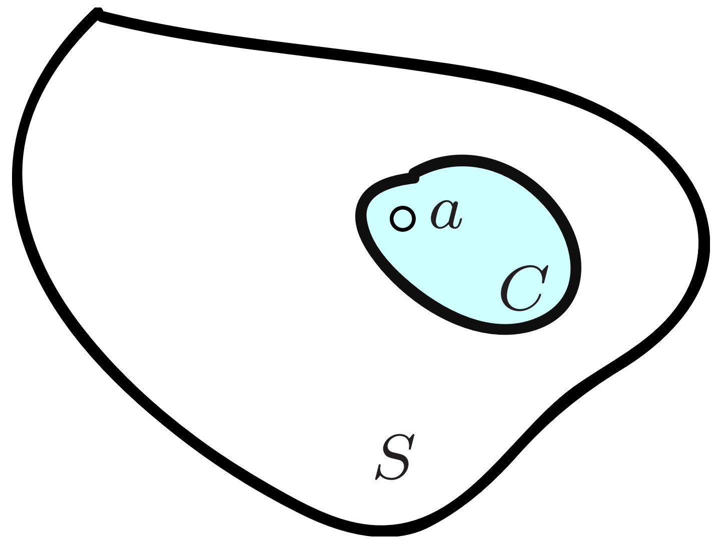
(Formales Argument hierfür: siehe Buch Seite 106.)

Und für $a \in C$

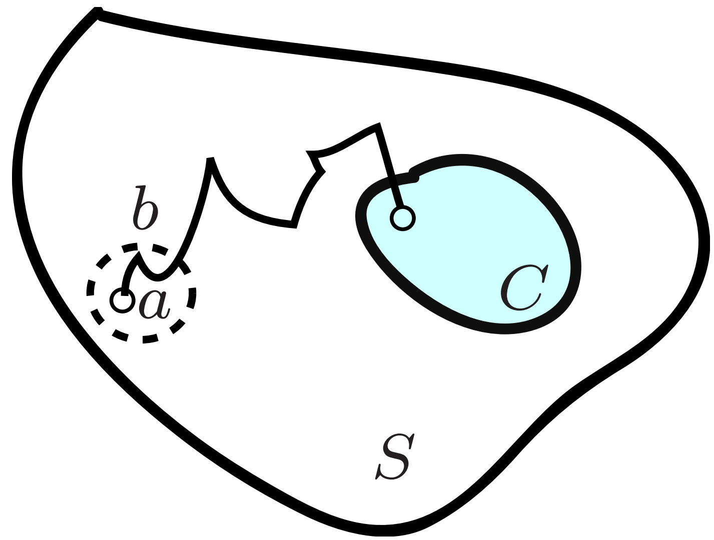
ist $\mathbf{P}_a(T_C = 0) = 1$,

also

$\mathbf{E}_a[T_C] = 0$.



Fazit:



$a \mapsto e(a) := \mathbf{E}_a[T_C]$ erfüllt das Gleichungssystem

$$\mathbf{e}(a) = 1 + \sum_{b \in S} P(a,b)\, e(b) \quad \text{für } a \notin C,$$

$$\mathbf{e}(a) = 0 \quad \text{für } a \in C.$$

Beispiel:

Erwartete Zeit bis zum ersten Erfolg.

$$\begin{tikzcd} f \arrow[loop left, "q"] \arrow[r, "p"] & s \end{tikzcd}$$

(f für *failure*, s für *success*)

$$\mathbf{E}_f[T_s] = 1 + q\mathbf{E}_f[T_s] + p\mathbf{E}_s[T_s]\ .$$

Wegen $\mathbf{E}_s[T_s] = 0$ wird dies gelöst durch

$$\mathbf{E}_f[T_s] = 1/p.$$

2. Transport von Erwartungswerten

Wir betrachten eine Funktion $h : S \to \mathbb{R}_+$

und interessierten uns für

$$u_n(a) := \mathbf{E}_a\big[h(X_n)\big] = \sum_{c \in S} h(c)\mathbf{P}_a(X_n = c)\ .$$

$u_n(a)$ lässt sich als Mittelung über die $u_{n-1}(b)$ ausdrücken,

über eine Zerlegung nach dem ersten Schritt:

$$\sum_{c \in S} h(c)\mathbf{P}_a(X_n = c) = \sum_{c \in S} h(c) \sum_{b \in S} P(a,b)\mathbf{P}_b(X_{n-1} = c)$$

$$\mathbf{E}_a\big[h(X_n)\big] = \sum_{b \in S} P(a,b)\,\mathbf{E}_b\big[h(X_{n-1})\big], \quad a \in S$$

$$\mathbf{E}_a\big[h(X_n)\big] = \sum_{b\in S} P(a,b)\,\mathbf{E}_b\big[h(X_{n-1})\big], \quad a \in S$$

ist gleichbedeutend mit

$$u_n(a) = \sum_{b\in S} P(a,b)\,u_{n-1}(b), \quad a \in S.$$

oder in Vektor-Matrixschreibweise, mit u_n als Spaltenvektor und der Anfangsbedingung $\mathbf{E}_a[h(X_0)] = h(a)$

$$\begin{cases} u_n = P\,u_{n-1}\,, & n \geq 1 \\ u_0 = h\,. \end{cases}$$

3. Transport von Verteilungen

$\mathbf{P}_\rho(X_n = c)$ zerlegt nach X_{n-1} ergibt die Rekursion

$$\mathbf{P}_\rho(X_n = c) = \sum_{b \in S} \mathbf{P}_\rho(X_{n-1} = b)\, P(b,c)$$

Mit

$$\pi_n(\cdot) := \mathbf{P}_\rho(X_n \in \cdot)\,, \quad n = 0, 1, \ldots$$

lautet diese Rekursion

$$\pi_n(c) = \sum_{b \in S} \pi_{n-1}(b) P(b,c)\,, \quad n \geq 1$$

mit der Anfangsbedingung $\pi_0(a) = \rho(a)$

$$\pi_n(c) = \sum_{b \in S} \pi_{n-1}(b) P(b,c) \ , \quad n \geq 1$$

mit der Anfangsbedingung $\pi_0(a) = \rho(a)$

oder in Vektor-Matrix-Schreibweise, mit π_n als Zeilenvektor:

$$\begin{cases} \pi_n = \pi_{n-1} P \ , \quad n \geq 1 \ , \\ \pi_0 = \rho \ . \end{cases}$$

4. Mehrschritt-Übergangswahrscheinlichkeiten

Zweischritt-Übergangswahrscheinlichkeiten

$$P^2(a,c) := \mathbf{P}_a(X_2 = c)\ , \quad n \geq 0\ ,\ a, c \in S$$

Zerlegungen nach dem Zwischenschritt:

$$\mathbf{P}_a(X_2 = c) = \sum_{b \in S} \mathbf{P}_a(X_1 = b)\mathbf{P}_b(X_1 = c)$$

$$P^2(a,c) = \sum_{b \in S} P(a,b)\, P(b,c)$$

$P^2 = (P^2(a,c))_{a,c \in S}$ lässt sich also auffassen als

Produkt der Matrix P mit sich selbst.

Mehrschritt-Übergangswahrscheinlichkeiten

$$P^n(a,c) := \mathbf{P}_a(X_n = c)\ , \quad n \geq 0\ ,\ a, c \in S$$

Zerlegung nach dem ersten Schritt:

$$P^n(a,c) = \sum_{b \in S} P(a,b)\, P^{n-1}(b,c)$$

Zerlegung nach dem letzten Schritt:

$$P^n(a,c) = \sum_{b \in S} P^{n-1}(a,b)\, P(b,c)$$

$P^n = (P^n(a,c))_{a,c \in S}$ lässt sich also auffassen als

n-te Matrixpotenz von P.

5. Gleichgewichtsverteilungen

Eine Verteilung π auf S heißt

Gleichgewichtsverteilung zur Übergangsmatrix P,

wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(G1) $$\sum_{a \in S} \pi(a) P(a,b) = \pi(b)\ , \quad b \in S\ .$$

(G2) $$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b)\ , \quad b \in S$$

(Dann haben auch $X_2, X_3, \ldots$ die Verteilung π.)

Reversible Gleichgewichtsverteilungen

Hinreichend für

(G2)

$$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b)\ , \quad b \in S$$

ist die Bedingung

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a)\ , \quad a, b \in S$$

(Denn dann ist das Paar (X_0, X_1) so verteilt wie (X_1, X_0), also insbesondere X_0 so wie X_1.)

Gleichbedeutend mit

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a)\ , \quad a, b \in S$$

ist

$$\pi(a)P(a,b) = \pi(b)P(b,a)\ , \quad a, b \in S\ .$$

π heißt dann *reversible Gleichgewichtsverteilung* zu P.

6. Beispiele:

Ein Beispiel einer

nicht reversiblen Gleichgewichtsverteilung:

Zyklische Irrfahrt auf $S = \{a, b, c\}$, mit

$$P(a, b) = P(b, c) = P(c, a) := p,$$

$$P(b, a) = P(c, b) = P(a, c) := 1 - p$$

Die uniforme Verteilung auf S

ist eine Gleichgewichtsverteilung zu P.

Nur für $p = 1/2$ ist sie reversibel.

Die Gleichgewichtsverteilung der einfachen Irrfahrt auf dem Würfel $S = \{0, 1\}^3$:

Von jedem $a \in S$ geht man in einem Schritt
zu einem rein zufällig ausgewählten Nachbarn.
(Zwei Elemente von S heißen *benachbart*,
wenn sie sich in genau einer Komponente unterscheiden.)
Für benachbarte Knoten a und b ist hier $P(a, b) = 1/3$.

Die uniforme Verteilung auf S
ist reversible Gleichgewichtsverteilung.

Eine wichtige Beispielklasse:

Die einfache Irrfahrt

auf einem ungerichteten, zusammenhängenden Graphen

mit endlicher Knotenmenge S

Von jedem $a \in S$ geht man in einem Schritt

zu einem rein zufällig ausgewählten Nachbarn:

$$P(a,b) = \frac{1}{g(a)},$$

mit b Nachbar von a, $g(a) := \#$ Nachbarn von a

Ansatz: $\quad \pi(a) := \frac{1}{c} g(a)$

Die Verteilung π erfüllt die die Reversibilitätsbedingung (R),

denn für benachbarte Knoten a, b gilt:

$$\frac{1}{c} g(a) \frac{1}{g(a)} = \frac{1}{c} g(b) \frac{1}{g(b)}$$

Man kann zeigen (hier ohne Beweis):

Es gibt nur **eine** Gleichgewichtsverteilung.

Fazit: Die Gewichte der Knoten

unter der Gleichgewichtsverteilung

sind proportional zur Anzahl der Nachbarn der Knoten.