Angewandte Stochastik

Dr. C.J. Luchsinger

6 Langzeitverhalten, stationäre Masse

Multiplikation von Rechts!

Literatur Kapitel 6

* Grimmett & Stirzaker: Kapitel 6.4
* Krengel: § 16.1

6.1 Motivation

Um die nachfolgenden Fragestellungen und Einschränkungen besser verstehen zu können, werden wir eingangs ein paar Fragen anhand von drei Beispielen aufwerfen.

1. In der ersten Situation starten wir in Punkt D. Man überlege sich, mit welchen Wahrscheinlichkeiten man zum Zeitpunkt 0, 1, 2, 3, 4, 10'000, 1'000'000 etc. in Punkten A, B, C oder D sein wird.

Von 3.2.1 her wissen wir, dass man solche Fragen exakt mit Hilfe von Rechtsmultiplikationen berechnet (wie ist hier der Vektor λ; hat er zum Zeitpunkt 10'000, 1'000'000 noch Relevanz?):

Ist D ein "repräsentativer" Startwert? Was wäre repräsentativ?

2. Betrachten Sie folgende Markovkette. Was lässt sich über $p_{11}^{(n)}$ aussagen?

3. Betrachten Sie jetzt diese Markovkette. Wir starten zur Zeit 0 in Punkt A. Wie wird die Verteilung zum Zeitpunkt $10^{100'000!}$ sein? Halten Sie das für eine befriedigende Aussage?

Mit Hilfe des Satzes von Perron-Frobenius haben wir in Kapitel 4 Aussagen über das Langzeitverhalten von Markovketten gemacht. Sie waren ziemlich abschliessend. Welche *Voraussetzungen* haben wir dort sinnvollerweise gemacht (wir werden sie weiterhin oft machen)?

Betrachten Sie obige Beispiele nochmals unter diesen Aspekten.

6.2 Langzeitverhalten

Nach unseren Erfahrungen in 6.1 werden wir bis auf weiteres vorerst einmal nur fordern, dass wir eine einzige, abgeschlossene Klasse haben (die Übergangsmatrix P ist damit irreduzibel). Periodizität ist also noch erlaubt. Wir fragen uns in 6.2 generell, wie oft ein Zustand besucht wird. Unser Untersuchungsobjekt wird also eine *relative Häufigkeit* zum Beispiel der Art

$$H_n(i) := \frac{1}{n}\sum_{k=1}^{n} I[X_k = i]$$

sein (wie oft kommt die Markov-Kette bis zur Zeit n in Zustand i vorbei). Konvergieren diese $H_n(i)$'s mit $n \to \infty$? Gegen was? Gegen den linken Eigenvektor zu Eigenwert 1 von Übergangsmatrix P? Wie steht es hiermit in Beispielen 1., 2. und 3. in 6.1?

Solche (oder einfachere) Ausdrücke haben wir schon früher benutzt (Münzwurf). Wir wollen zum Beispiel herausfinden, wie gross die Wahrscheinlichkeit für Kopf ist (sollte ungefähr 0.5 sein). Intuitiv wird jedeR die Münze möglichst oft werfen und die relative Häufigkeit von Kopf zählen. Bei 100'000 Würfen wird dies vielleicht 49'875 mal Kopf sein. Mathematisch wird dies folgendermassen formalisiert: $X_k = 1$ falls Kopf und $X_k = 0$ falls Zahl im k-ten Wurf. Also haben wir

$$\frac{1}{n}\sum_{k=1}^{n} X_k \tag{6.1}$$

zu untersuchen (bei obigen Zahlen hätten wir also 0.49875 erhalten). Wichtige Nebenbemerkung: Wir fordern *Unabhängigkeit* und *gleiche Verteilung* der Würfe (keine Abnützung an der Münze, Windverhältnisse immer gleich etc.)! Wegen des Satzes von Kolmogoroff (WT-Satz 5.6) folgt jetzt, dass (6.1) fast sicher (und in Wahrscheinlichkeit) gegen $E[X_1]$ konvergiert (also dem Gesetz der grossen Zahlen genügt). Das reicht uns noch nicht ganz: Es gilt:

$$E[X_1] = 1P[X_1 = 1] + 0P[X_1 = 0] = P[X_1 = 1].$$

Dies ist aber genau die Wahrscheinlichkeit für Kopf.

In WT-Satz 5.6 müssen paarweise Unabhängigkeit und gleiche Verteilung gegeben sein. Schauen wir einmal, was alles passieren kann, wenn die Unabhängigkeit aufgegeben wird. **In der Klasse zu lösen**: Suchen Sie ein einfaches Beispiel (nicht rechnen!) von identisch verteilten Zufallsgrössen, welche jedoch nicht unabhängig sind, sodass die Folge dem Gesetz der grossen Zahlen nicht genügt.

Das vom Dozenten vorgestellte Beispiel einer nicht mehr vorhandenen Unabhängigkeit ist zwar sehr extrem. Aber auch bei unseren Markov-Ketten werden wir die Unabhängigkeit nicht mehr fordern können. WT-Satz 5.6 kann man deshalb in 6.2 normalerweise nicht *direkt* anwenden.

Definition 6.1 $[H_n(i), H_n(B)]$ *Wir definieren die relativen Häufigkeiten von Zustand i bzw. Menge B folgendermassen:*

$$H_n(i) := \frac{1}{n} \sum_{k=1}^{n} I[X_k = i]$$

und

$$H_n(B) := \frac{1}{n} \sum_{k=1}^{n} I[X_k \in B].$$

Definition 6.2 $[z_{iB}, z_{ij}]$ *z_{iB} ist die "erwartete Anzahl in B, bevor man wieder in i vorbeikommt"; mathematisch mit $N_{i0} := 0$:*

$$N_{ir} := \inf\{m \geq 1, m > N_{i,r-1} | X_m = i\}$$

ist der Zeitpunkt des r-ten Males, wo man wieder in i vorbeikommt, $r \geq 1$. Dann definieren wir

$$z_{iB} := E_i\Big[\sum_{n \geq 1} I[X_n \in B, n \leq N_{i1}]\Big] \; (\leq \infty).$$

Die z_{ij} sind einfach z_{iB} mit $B = \{j\}$ eine ein-elementige Menge.

Die starke Markov-Eigenschaft (vgl. 3.2.6) wird für die z_{iB} wichtig sein: Wenn wir (wieder einmal) in i vorbeikommen, wird die Markov-Kette immer wieder den gleichen Gesetzen gehorchen.

Der Zusammenhang zwischen den z_{iB} und den $H_n(B), H_n(i)$ kommt in folgendem Satz zum Ausdruck:

Satz 6.3 [Starker Ergodensatz] *Sei der Zustandsraum eine einzige, rekurrente Klasse. Dann gilt für Elemente i und Mengen B:*

$$\frac{H_n(B)}{H_n(i)} \to z_{iB}$$

fast sicher wenn $n \to \infty$.

Skizze:

Beweis von Satz 6.3: Wir definieren (vgl. obige Skizze):

$$W := \sum_{m \geq 1} I[X_m \in B, m \leq N_{i1}],$$

die Anzahl Aufenthalte in B bis zum ersten Wiederbesuch in i.

$$Y_k := \sum_{m \geq 1} I[X_m \in B, N_{ik} < m \leq N_{i,k+1}],$$

die Anzahl Aufenthalte in B nach dem k'ten und bis und mit dem $(k+1)$'ten Aufenthalt in i.

$$W_n' := \sum_{m \geq 1} \sum_{l \geq 0} I[X_m \in B, N_{il} < m \leq n \leq N_{i,l+1}] \, (\leq Y_{nH_n(i)}),$$

die Anzahl Aufenthalte in B bis Zeit n, nach dem letzten Aufenthalt in i, welcher noch vor Zeit n liegt.

$$\frac{H_n(B)}{H_n(i)} = \frac{1}{nH_n(i)} \left[W + Y_1 + \ldots + Y_{nH_n(i)-1} + W_n' \right].$$

Damit gilt:

$$\frac{W}{nH_n(i)} + \frac{Y_1 + \ldots + Y_{nH_n(i)-1}}{nH_n(i)} \leq \frac{H_n(B)}{H_n(i)} \leq \frac{W}{nH_n(i)} + \frac{Y_1 + \ldots + Y_{nH_n(i)}}{nH_n(i)}. \tag{6.2}$$

1. Weil die Klasse rekurrent ist, gilt $nH_n(i) \to \infty$ fast sicher wenn $n \to \infty$ (Satz 4.7 b)). Es gibt also eine Ausnahme-Nullmenge $N^{(1)}$ derart, dass für alle $\omega \in \Omega \backslash N^{(1)}$ gilt: $nH_n(i)(\omega) \to \infty$ als Folge von reellen Zahlen!

2. Des weiteren besteht die Folge der (Y_i) wegen der starken Markov-Eigenschaft aus iid Zufallsgrössen. Falls $E[Y_1] < \infty$, können wir also WT-Satz 5.6 anwenden und folgern:

$$\frac{Y_1 + \ldots + Y_k}{k} \to E[Y_1] \; (= z_{iB}),$$

für $k \to \infty$ fast sicher. Diese Aussage stimmt aber auch (da $Y_1 \geq 0$) für den Fall $E[Y_1] = \infty$ (Aufgabenblatt 6). Damit finden wir eine Ausnahme-Nullmenge $N^{(2)}$ derart, dass für alle $\omega \in \Omega \backslash N^{(2)}$ gilt:

$$\frac{Y_1(\omega) + \ldots + Y_k(\omega)}{k} \to E[Y_1] \; (= z_{iB}),$$

als Folge von reellen Zahlen. Von jetzt an konzentrieren wir uns auf die Menge $A := (\Omega\backslash N^{(1)})\backslash N^{(2)}$. Wegen 1. und 2. erhalten wir:

$$\frac{Y_1 + \ldots + Y_{nH_n(i)}}{nH_n(i)} \to E[Y_1]$$

und

$$\frac{Y_1 + \ldots + Y_{nH_n(i)-1}}{nH_n(i)} = \frac{Y_1 + \ldots + Y_{nH_n(i)-1}}{nH_n(i) - 1} * \frac{nH_n(i) - 1}{nH_n(i)} \to E[Y_1].$$

Der Ausdruck

$$\frac{W}{nH_n(i)}$$

konvergiert in A gegen 0. Damit können wir von (6.2) folgern:

$$0 + z_{iB} \leq \liminf \frac{H_n(B)}{H_n(i)} \leq \limsup \frac{H_n(B)}{H_n(i)} \leq 0 + z_{iB}$$

in A. Damit gilt auch

$$\frac{H_n(B)}{H_n(i)} \to z_{iB}$$

fast sicher.

□

Korollar 6.4 [Quotient-Ergodensatz] *Sei der Zustandsraum eine einzige, rekurrente Klasse. Dann gilt für Elemente i, j:*

$$\frac{H_n(j)}{H_n(i)} \to z_{ij}$$

fast sicher wenn $n \to \infty$. Falls $z_{ij} > 0$ und $z_{ji} > 0$, so haben wir $z_{ij}z_{ji} = 1$.

Beweis von Korollar 6.4: Satz 6.3.

□

Bemerkung: Wir werden später sehen, dass in diesem Setting sowieso $z_{ij} > 0$ und $z_{ji} > 0$ erfüllt sein wird.

Aufmerksame StudentInnen werden festgestellt haben, dass in Satz 6.3 $i \in B$ erlaubt ist (überprüfen Sie den Beweis von Satz 6.3 auf diesen Punkt hin, falls Sie bis jetzt eine falsche Vorstellung hatten). Damit können wir folgendes Korollar formulieren:

Korollar 6.5 [Erwartete Rekurrenzzeit von i**]** *Sei der Zustandsraum eine einzige, rekurrente Klasse. Dann gilt für jedes Element* i *des Zustandsraums:*

$$H_n(i) \to \frac{1}{m_i}$$

fast sicher, wenn $n \to \infty$. $m_i := E_i[N_{i1}]$ *bezeichne dabei die erwartete Rekurrenzzeit von* i.

Beweis von Korollar 6.5: Satz 6.3: wähle $B := \mathbb{Z}$, damit ist $H_n(B) = 1$ für alle n.

□

Warnung: Es ist möglich, dass $m_i = \infty$ (Random Walk auf $\mathbb{Z}$). Dann konvergiert $H_n(i)$ gemäss Korollar 6.5 fast sicher gegen 0.

Wir haben uns in 6.1 anhand von Beispiel 2 gefragt, ob etwa die $p_{11}^{(n)}$ für $n \to \infty$ konvergieren. Dies musste damals schon in dieser naiven Form verworfen werden. Hingegen gilt

Korollar 6.6 [schwacher Ergodensatz] *Sei der Zustandsraum eine einzige, rekurrente Klasse. Dann gilt für* i, k *beliebig:*

$$\frac{1}{n} \sum_{m=1}^{n} p_{ki}^{(m)} \to \frac{1}{m_i}$$

wenn $n \to \infty$. *Man beachte, dass* k *auf der rechten Seite nicht mehr vorkommt!*

Beweis von Korollar 6.6: Wir haben in Definition 6.2 mit N_{i0} den Startpunkt der Markov-Kette jeweils bei i gesetzt und die weiteren N_{is} als *Wieder*besuche nach Zeit 0 definiert (damit definierten wir dann auch z_{iB}). Im Beweis von Satz 6.3 ist jedoch gar nicht benutzt worden, dass die Kette in Punkt i startet. Die Kette kann irgendwo starten;

das "W" wurde ja im Verlauf von $n \to \infty$ sowieso immer unbedeutender. Damit können wir Satz 6.3 bzw. die bisherigen Korollare benutzen, auch wenn der Startpunkt jetzt ein beliebiges k ist.

$$\frac{1}{n}\sum_{m=1}^{n} p_{ki}^{(m)} = \frac{1}{n}\sum_{m=1}^{n} E_k[I[X_m = i]] = E_k[(1/n)\sum_{m=1}^{n} I[X_m = i]] = E_k[H_n(i)].$$

Wegen Korollar 6.5 konvergiert $H_n(i) - 1/m_i$ fast sicher gegen 0 mit $n \to \infty$. Die $H_n(i) - 1/m_i$ sind zudem für jedes n durch 1 gegen oben beschränkt. Damit kann man den Satz von der majorisierten Konvergenz (Lebesgue) anwenden (WT-Satz 4.11) und folgern:

$$E_k[H_n(i) - 1/m_i] \to 0$$

falls $n \to \infty$.

$\square$

Einschub I: Was, wenn wir *nicht* eine einzige, abgeschlossene Klasse haben?

Seien $K_1, K_2, K_3, \dots$ die Menge aller rekurrenten (abgeschlossenen wegen Satz 4.10 a)) Klassen. Es darf unendlich viele davon haben! Mit der Menge V fassen wir alle vergänglichen Klassen zusammen. Diese können auch abgeschlossen sein (Random Walk auf $\mathbb{Z}^3$!). Wir haben also die folgende Situation:

Was lässt sich jetzt über

$$\frac{1}{n}\sum_{m=1}^{n} p_{ki}^{(m)} \tag{6.3}$$

aussagen, $n \to \infty$? Wir unterscheiden 4 Fälle:

* $k \in K_j$ **für ein** j **und** $i \notin K_j$**:** Dann sind alle $p_{ki}^{(m)} = 0$ und (6.3) ist 0 von Anfang an.

* $k \in K_j$ **für ein** j **und** $i \in K_j$**:** Dann sind wir in der Situation von Korollar 6.6 (wir können vergessen, dass es noch weitere Klassen gibt, man kommt sowieso nie dorthin, will auch nicht dorthin ($i \in K_j$!)) und es gilt somit

$$\frac{1}{n}\sum_{m=1}^{n} p_{ki}^{(m)} \to \frac{1}{m_i}$$

für $n \to \infty$.

* $k \in V$ **und** $i \in V$**:** Falls $k \not\rightsquigarrow i$ (möglich, können ja in verschiedenen vergänglichen Klassen sein), haben wir von Anfang an 0 in (6.3). Sonst haben wir vom Beweis von Korollar 6.6 mit $A := \{\exists l < \infty : X_l = i\}$

$$\begin{aligned}
\frac{1}{n}\sum_{m=1}^{n} p_{ki}^{(m)} &= E_k[H_n(i)] \\
&= E_k[H_n(i)|A]P_k[A] + E_k[H_n(i)|A^c]P_k[A^c] \\
&= E_k[H_n(i)|A]P_k[A] \\
&\leq \frac{1}{n} + E_i[H_n(i)] \\
&\leq \frac{1}{n} + \frac{1}{n}\frac{F_{ii}}{1-F_{ii}}
\end{aligned}$$

wegen (4.1). Damit konvergiert dieser Ausdruck aber gegen 0 mit $n \to \infty$.

* $k \in V$ **und** $i \in K_j$ **für ein** j**:** Wir definieren die Menge $C_j := \cup_{n\geq 1} \cap_{m \geq n} \{X_m \in K_j\}$, ob die Markov-Kette in Menge K_j absorbiert wird und $\alpha_{kj} := P_k[C_j]$. Dann gilt:

$$\begin{aligned}
\frac{1}{n}\sum_{m=1}^{n} p_{ki}^{(m)} &= E_k[H_n(i)] \\
&= E_k[H_n(i)|C_j]P_k[C_j] + E_k[H_n(i)|C_j^c]P_k[C_j^c] \\
&= E_k[H_n(i)|C_j]\alpha_{kj}.
\end{aligned}$$

Wenn wir uns in der Menge C_j befinden, wird die Kette in endlicher Zeit in K_j ankommen. Wir können dann mit denselben Argumenten wie in den Beweisen von Satz 6.3

und Korollar 6.6 folgern, dass der Ausdruck $E_k[H_n(i)|C_j]$ gegen $1/m_i$ konvergiert für $n \to \infty$.

Diese 4 Fälle (und streng genommen auch Korollar 6.6) können wir in folgendem Satz zusammenfassen:

Satz 6.7 [schwacher Ergodensatz; Verallgemeinert] *Der Limes von*

$$\frac{1}{n} \sum_{m=1}^{n} p_{ki}^{(m)}$$

für $n \to \infty$ existiert immer; und zwar gilt:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ki}^{(m)} = \begin{cases} 0, & i \text{ vergänglich} \\ \alpha_{kj} \frac{1}{m_i} & i \in K_j \text{ rekurrent.} \end{cases}$$

Dabei ist mit $C_j := \cup_{n\geq 1} \cap_{m\geq n} \{X_m \in K_j\}$ $\alpha_{kj} := P_k[C_j]$. α_{kj} wird dabei mit Hilfe von Satz 5.3 berechnet und ist im Fall von Korollar 6.6 sowieso 1.

Einschub II: Das "(z_{ij}, m_k)-Kalkül"

Von jetzt an haben wir wieder eine einzige Klasse. Dann kann man einige kleine Formeln herleiten. Beginnen wir erstmals mit

Lemma 6.8 [erwartete Rekurrenzzeit als Funktion der z_{ik}] *Sei i ein rekurrenter Zustand. Dann gilt:*

$$m_i = 1 + \sum_{k \neq i} z_{ik} = \sum_k z_{ik},$$

mit $m_i := E_i[N_{i1}] \leq \infty$ und $z_{ik} := E_i[\sum_{n \geq 1} I[X_n = k, n \leq N_{i1}]]$.

Beweis von Lemma 6.8:

$$m_i = E_i[N_{i1}] = E_i[\sum_{n \geq 1} I[n \leq N_{i1}]] = E_i[\sum_{n \geq 1} \sum_k I[X_n = k, n \leq N_{i1}]] = 1 + \sum_{k \neq i} z_{ik}.$$

□

Lemma 6.9 [eine einzige Klasse $\Rightarrow 0 < z_{ij} < \infty$] *Unter der Annahme einer einzigen Klasse (rekurrent oder transient) gilt $0 < z_{ij} < \infty$.*

Beweis von Lemma 6.9: Wir zeigen zuerst, dass $z_{ij} > 0$. Da $i \leftrightsquigarrow j$ finden wir ein $m > 0$, sodass $p_{ij}^{(m)} > 0$. Wir definieren

* $A_1 := \cup_{n \geq 1} \{X_n = j, n \leq N_{i1}\}$
kommt in Zustand j vobeikommt, bevor zum ersten Mal (nach 0) in i

* $A_l := \cup_{n \geq 1} \{X_n = j, n \leq N_{il}\}$
kommt in Zustand j vorbeikommt, bevor zum l-ten Mal (nach 0) in i

* $B_k := \{\omega | \ |\{n : X_n(\omega) = j, n \leq N_{i1}(\omega)\}| = k\}$
kommt genau k mal in Zustand j vobeikommt, bevor zum ersten Mal (nach 0) in i

Es gelten $(A_l)_{l \geq 1} \nearrow$ und $A_1 = \dot{\cup}_{k \geq 1} B_k$. Damit gilt sicher einmal

$$P_i[A_1] \leq 1 P_i[B_1] + 2 P_i[B_2] + 3 P_i[B_3] + \ldots = z_{ij}.$$

Des weiteren haben wir für $l \geq 2$ wegen der FTW:

$$\begin{aligned}
P_i[A_l] &= P_i[A_l|A_1]P_i[A_1] \\
&\quad + P_i[A_l|A_1^c \cap \{N_{i1} < \infty\}]P_i[A_1^c \cap \{N_{i1} < \infty\}] \\
&\quad + P_i[A_l|A_1^c \cap \{N_{i1} = \infty\}]P_i[A_1^c \cap \{N_{i1} = \infty\}] \\
&= P_i[A_1] + P_i[A_l|A_1^c \cap \{N_{i1} < \infty\}]P_i[A_1^c \cap \{N_{i1} < \infty\}] \\
&\leq P_i[A_1] + P_i[A_{l-1}](1 - P_i[A_1]) \leq P_i[A_1] + P_i[A_{l-1}].
\end{aligned}$$

Wir beginnen einen Widerspruchsansatz: Wenn $P_i[A_1] = 0$ dann haben wir $P_i[A_l] \leq P_i[A_{l-1}]$ für alle $l \geq 2$ und damit $P_i[A_l] = 0$ für alle $l \geq 1$. Andererseits haben wir auch

$$P_i[A_m] = P_i[\cup_{n\geq 1}\{X_n = j, n \leq N_{im}\}] \geq P_i[X_m = j] = p_{ij}^{(m)} > 0.$$

Widerspruch! Also muss $0 < P_i[A_1] \leq z_{ij}$ gelten.

Wir müssen noch zeigen, dass diese $z_{ij} < \infty$. Seien dazu zuerst i, j vergänglich. Dann gilt:

$$z_{ij} \leq 1 + E_j[\sum_{n\geq 1} I[X_n = j]] = 1 + \frac{F_{jj}}{1 - F_{jj}} < \infty.$$

Seien jetzt i, j rekurrent: Von Korollar 6.4 her wissen wir, dass dann $z_{ij}z_{ji} = 1$. Da die $z_{ij} > 0$ sind, folgt daraus die Behauptung.

$\square$

Warnung: Man darf in Lemma 6.9 j nicht einfach durch eine mehrelementige Menge B ersetzen! Weshalb nicht?

Wir wollen kompakt nochmals die bisherigen Rechenregeln im Rahmen des "(z_{ij}, m_k)-Kalküls" zusammenfassen und einige weitere angeben. Als einzige Voraussetzung fordern wir immer und zuerst ausschliesslich (vergänglich erlaubt):

eine einzige Klasse:

1. von Lemma 6.9

$$0 < z_{ij} < \infty$$

für alle i, j.

Von jetzt an fordern wir

eine einzige, rekurrente Klasse:

2. Dann erhalten wir zusätzlich zu obigem Resultat

$$z_{ij} z_{ji} = 1.$$

3. Des weiteren haben wir sofort auch (Korollar 6.4)

$$z_{ij} z_{jk} = z_{ik}$$

für alle i, j, k.

4. Fixieren wir einen Zustand a. Es gilt dann $z_{ij} z_{ja} = z_{ia}$ und damit

$$z_{ij} = \frac{z_{ia}}{z_{ja}} = \frac{z_{aj}}{z_{ai}}. \tag{6.4}$$

5. Aus (6.4) durch Summation über alle j, mit Hilfe von Lemma 6.8 ($z_{ii} = 1$):

$$m_i = \frac{m_a}{z_{ai}}$$

und damit auch $z_{ai} = m_a/m_i$. Vektoriell können wir das folgendermassen schreiben:

$$\Big(z_{a1}, z_{a2}, z_{a3}, \ldots\Big) = \left(\frac{1}{m_1}, \frac{1}{m_2}, \frac{1}{m_3}, \ldots\right) m_a.$$

Mit Hilfe von Referenzpunkten (hier a) kann man offenbar kompakte Aussagen machen. Dies wird im kommenden Hauptsatz über stationäre Masse (Satz 6.10) wichtig sein.

6.3 Stationäre Masse

In diesem Teil werden wir die Verbindung zu den Übergangsmatrizen und den linken Eigenvektoren zu Eigenwert 1 schaffen. Es gilt vorerst mal folgender Satz:

Satz 6.10 [Hauptsatz über stationäre Masse] *Sei a ein fester Referenzpunkt im Zustandsraum. Die Markov-Kette bewege sich auf einer einzigen, abgeschlossenen Klasse.*

I. Dann existiert immer eine minimale Lösung der (Un-) Gleichungen

i) $x_a = 1$

ii) $(x^t P)_k = x_k$, *für* $k \neq a$

iii) $x_k \geq 0$, *für alle* k

iv) $(x^t P)_a \leq x_a$.

Die Lösung ist $x_a^* = 1$ *und* $x_i^* := z_{ai}$ *für* $i \neq a$.

II. Des weiteren gilt für diese minimale Lösung x^*

$$(x^{*t} P)_a = F_{aa} \begin{cases} < x_a^* = 1 & \text{wenn vergänglich} \\ = x_a^* = 1 & \text{wenn rekurrent.} \end{cases} \tag{6.5}$$

III. Wenn $F_{aa} = 1$ *(Rekurrenz), ist die minimale Lösung* x^* *die einzige Lösung des Systems* $x_a = 1$, $x_k \geq 0$ *für alle* k *und* $(x^t P)_k \leq x_k$, *für alle* k.

IV. Im Fall von Rekurrenz nennen wir jede Lösung $\mu \geq 0, \mu \neq 0$ *von* $\mu^t P = \mu^t$ *ein stationäres Mass (Gleichgewichtsmass); es gilt dann*

$$z_{ij} = \frac{\mu_j}{\mu_i}$$

und $\mu_i > 0$ *für jedes* i.

Beweis von Satz 6.10: Die 4 Aussagen werden in 4 Punkten bewiesen. Der Beweis ist lehrreich und sollte auf jeden Fall durchgearbeitet werden.

I. Existenz der minimalen Lösung

Mit unserer Wahl $x_a^* = 1$ und $x_i^* := z_{ai}$ für $i \neq a$ sind Bedingungen i) und iii) bereits erfüllt. Wir untersuchen Forderung ii): Es gilt erstmals

$$z_{ai} := E_a\Big[\sum_{n\geq 1} I[X_n = i, n \leq N_{a1}]\Big] = \sum_{n\geq 1} P_a[X_n = i, n \leq N_{a1}].$$

Für ii) erhalten wir damit

$$\begin{aligned}((x^*)^t P)_k &= p_{ak} + \sum_{j\neq a} z_{aj} p_{jk} \\ &= p_{ak} + \sum_{j\neq a}\sum_{n\geq 1} P_a[X_n = j, n \leq N_{a1}] p_{jk} \\ &= p_{ak} + \sum_{n\geq 1} P_a[X_{n+1} = k, n+1 \leq N_{a1}] \\ &= p_{ak} + \sum_{n\geq 2} P_a[X_n = k, n \leq N_{a1}] \\ &= \sum_{n\geq 1} P_a[X_n = k, n \leq N_{a1}] \\ &= z_{ak} \\ &= x_k^*.\end{aligned}$$

Forderung iv):

$$\begin{aligned}((x^*)^t P)_a &= p_{aa} + \sum_{j\neq a} z_{aj} p_{ja} \\ &= p_{aa} + \sum_{j\neq a}\sum_{n\geq 1} P_a[X_n = j, n \leq N_{a1}] p_{ja} \\ &= F_{aa} \\ &\leq 1 \\ &= x_a^*.\end{aligned}$$

Bleibt die Minimalität:

Sei dazu y eine Lösung von $y \geq 0, y^t P \leq y^t, y_a = 1$. Man beachte, dass dieses y weniger Anforderungen erfüllen muss; es gibt also mehr Konkurrenten; trotzdem bleibt unser x^* minimal in jeder Koordinate (wir werden das im 3. Punkt des Beweises benutzen). Wir zeigen also durch Induktion

$$y \geq x^*.$$

Induktionshypothese (Schritt m) sei dabei für $i \neq a$ (bei $i = a$ ist nichts zu beweisen)

$$y_i \geq \sum_{n=1}^{m} P_a[X_n = i, n \leq N_{a1}].$$

Man beachte, dass gilt:

$$x_i^* := z_{ai} := E_a\Big[\sum_{n\geq 1} I[X_n = i, n \leq N_{a1}]\Big] = \sum_{n\geq 1} P_a[X_n = i, n \leq N_{a1}].$$

Induktionsschritt:

$$\begin{aligned}
y_i &\geq \sum_k y_k p_{ki} \\
&\geq p_{ai} + \sum_{k\neq a} \sum_{n=1}^{m} P_a[X_n = k, n \leq N_{a1}] p_{ki} \\
&= p_{ai} + \sum_{n=1}^{m} P_a[X_{n+1} = i, n+1 \leq N_{a1}] \\
&= p_{ai} + \sum_{n\geq 2}^{m+1} P_a[X_n = i, n \leq N_{a1}] \\
&= \sum_{n=1}^{m+1} P_a[X_n = i, n \leq N_{a1}].
\end{aligned}$$

Die Induktionsverankerung ist:

$$y_i \geq P_a[X_1 = i, 1 \leq N_{a1}] = p_{ai}.$$

Dies ist aber sicher wegen

$$y^t \geq y^t P \tag{6.6}$$

erfüllt, denn der Summand p_{ai} kommt als $y_a p_{ai} = 1 p_{ai}$ auf der rechten Seite von (6.6) unter lauter nicht-negativen Summanden vor.

Weil

$$\lim_{m\to\infty} \sum_{n=1}^{m} P_a[X_n = i, n \leq N_{a1}] = z_{ai} = x_i^*$$

sind wir mit der Minimalität fertig.

II. $\underline{({x^*}^t P)_a = F_{aa}}$

Dies wurde bei der Überprüfung von Forderung iv) in Teil 1 schon erledigt.

III. Einzige Lösung von schwächeren Bedingungen bei Rekurrenz

Wenn $F_{aa} = 1$ (Rekurrenz), so erfüllt unsere minimale Lösung x^* exakt $(x^*)^t P = (x^*)^t$ mit Gleichheit in allen (Un-)Gleichungen. Es ist nun überraschend, dass aus $y : y_a = 1, y \geq 0, y^t P \leq y^t$ folgt: $y = x^*$. Dies wird folgendermassen bewiesen: erstmals gilt in allen Koordinaten:

$$y^t \geq y^t P \geq y^t P^2 \geq y^t P^3 \geq y^t P^4 \ldots \geq y^t P^m$$

für alle $m \geq 0$. Dies gilt auch für Koordinate a. Wir haben $x_a^* = 1 = y_a$. Weiter haben wir, da x^* in allen Koordinaten minimal ist,

$$y^t P^m \geq (x^*)^t P^m. \tag{6.7}$$

Dies muss auch für die a-Koordinate gelten. Die a-Koordinate der rechten Seite von (6.7) ist aber 1. Damit ist für alle $m \geq 0$ auch die a-Koordinate der linken Seite von (6.7) 1. Weiter haben wir in allen Koordinaten $y \geq x^*$ wegen der Minimalität. Die Matrix P hat zudem lauter nichtnegative Einträge. Daraus folgt jetzt aber $x_k^* = y_k$ für alle k, sodass $p_{ka}^{(m)} > 0$. Da $k \leftrightsquigarrow a$ für alle k, finde ich immer so ein $m = m(k)$, sodass $x_k^* = y_k$ für alle k gelten muss.

IV. Stationäre Masse

Im Fall von Rekurrenz ist die minimale Lösung ja die einzige Lösung, welche obige Bedingungen inklusive $x_a^* = 1$ erfüllt. Wenn wir eine Lösung μ von $\mu \geq 0$, $\mu \neq 0$ und $\mu^t P = \mu^t$ haben, so muss dieses μ deshalb bis auf eine Konstante (μ_a) gleich x^* sein:

$$\frac{\mu_i}{\mu_a} = x_i^*$$

für alle i. Damit gilt aber

$$\frac{\mu_j}{\mu_i} = \frac{\frac{\mu_j}{\mu_a}}{\frac{\mu_i}{\mu_a}} = \frac{x_j^*}{x_i^*} = \frac{z_{aj}}{z_{ai}} = z_{ij}.$$

$\mu_i > 0$ folgt aus $z_{ij} > 0$ für alle i, j.

$\square$

Satz 6.10 ist selber sehr komplex. Die direkter anwendbaren Korollare werden nachfolgend entwickelt. Wir wollen vorher noch übungshalber den Random Walk auf den natürlichen Zahlen nochmals auf Rekurrenz untersuchen (vgl. Aufgabenblatt), aber jetzt mit Hilfe von Satz 6.10. Sei dazu $p \in (0, 0.5)$, $p + q = 1$ und $p_{i,i-1} = q = 1 - p_{i,i+1}$ wenn $i \geq 1$ und $p_{00} = q = 1 - p_{01}$. Wir haben also folgendes Bild:

Wir wählen jetzt $a = 0$ (wir sind hier frei, weil Rekurrenz eine Klasseneigenschaft ist). Wir wissen von 5.2, dass die allgemeine Lösung dieses Rekurrenzgleichungssystems für Punkt ii) von Satz 6.10 folgende Lösung hat für $k \geq 1$:

$$x_k = A + B(p/q)^k.$$

Von Punkt i) in Satz 6.10 brauchen wir $x_0 = 1$. Daraus folgt $1 = A + B$. Im Fall $p \in (0, 0.5)$ erhalten wir die minimale Lösung aus Satz 5.1 durch die Wahl $B = 1$. Also gilt:

$$x_k = (p/q)^k.$$

Dies ist bis auf eine Konstante *das* stationäre Mass (Gleichgewichtsverteilung) in diesem Beispiel. Da wir eine Drift nach links (zu Null) haben, ist es eine schöne, geometrisch fallende Folge. Wir wollen jetzt noch mit Satz 6.10 die Rekurrenz untersuchen. Wegen (6.5) müssen wir

$$F_{00} = (x^t P)_0 = q + q\frac{p}{q} = 1$$

untersuchen; wir haben also (keine Überraschung) auch so bewiesen Rekurrenz.

In Satz 6.10 ist nicht vorausgesetzt, dass die Klasse rekurrent ist. Wir wollen in folgendem Korollar diesen Fall herausgreifen. Dabei kann der Zustandsraum entweder nur aus einer rekurrenten Klasse bestehen oder eine rekurrente Klasse besitzen (Bsp. 3 in 6.1), neben anderen Klassen.

Korollar 6.11 [Rekurrente Klasse (ausschliesslich oder u.a.)] *Sei C eine rekurrente Klasse (also abgeschlossen wegen Satz 4.10 a)). Sei P^C der $C \times C$-Teil der Matrix P, wobei $X_0 \in C$. Sei weiter $\mu \geq 0$ eine Lösung der Gleichungen $\mu^t P^C = \mu^t \neq 0$. Dann gilt für $k \in C$*

$$m_k = \frac{\sum_{j \in C} \mu_j}{\mu_k} \leq \infty.$$

Beweis von Korollar 6.11: Von Lemma 6.8 haben wir

$$m_k = 1 + \sum_{j \neq k} z_{kj}.$$

Dabei gilt für $k \in C$ und $j \notin C$: $z_{kj} = 0$. Von Satz 6.10 haben wir

$$z_{kj} = \frac{\mu_j}{\mu_k}.$$

Daraus folgt die Behauptung.

□

Bis jetzt haben wir erst Rekurrenz und Transienz unterschieden. Intuitiv gibt es aber einen Unterschied, ob wir einen *zwar rekurrenten*, symmetrischen Random Walk auf $\mathbb{Z}$ haben oder eine rekurrente Kette z.B. auf endlichem Zustandsraum: die erwartete Rückkehrzeit ist im ersten Fall unendlich, im zweiten Fall endlich.

Korollar 6.12 [Null- & Positiv-Rekurrent] *Sei C eine rekurrente Klasse. Dann gilt mit der Notation von Korollar 6.11*

entweder

$$\sum_{j \in C} \mu_j = \infty;$$

daraus folgt dann $m_k = \infty$ für alle $k \in C$, die Klasse heisst dann Null-Rekurrent;

oder

$$\sum_{j \in C} \mu_j < \infty;$$

daraus folgt dann $m_k < \infty$;

$$\frac{1}{m_k} = \frac{\mu_k}{\sum_{j \in C} \mu_j} > 0$$

für alle $k \in C$, die Klasse heisst dann Positiv-Rekurrent. Wir haben dann auch

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} p_{lk}^{(i)} = \frac{1}{m_k} =: \pi_k$$

für $l, k \in C$ beliebig und es gilt:

$$\sum_{k \in C} \pi_k = 1.$$

Beweis von Korollar 6.12: Folgt sofort aus Korollar 6.6, Satz 6.10 und Korollar 6.11.

□

1. Aus Korollar 6.12 folgt, dass positiv- bzw. null-Rekurrenz eine Klasseneigenschaft ist. Warum?

2. Man wird bei positiver Rekurrenz die Normierung bei einem stationären Mass vorteilhafterweise derart wählen, dass der gesamte Vektor auf 1 summiert, mithin eine Wahrscheinlichkeitsverteilung ist (ev. die Limes-Verteilung (vgl. Satz 6.14))

Einer der hilfreichsten (und überraschendsten) Sätze ist

Satz 6.13 [Summierbares Mass] *Sei C eine abgeschlossene Klasse. Sei $\mu \geq 0$, $\mu \neq 0$ eine Lösung der Gleichungen*

$$\mu^t P^C = \mu^t,$$

sodass $\sum_{j \in C} \mu_j < \infty$. Dann ist C positiv rekurrent und somit gilt für alle $k \in C$:

$$\frac{1}{m_k} = \frac{\mu_k}{\sum_{j \in C} \mu_j}.$$

Bemerkung: Im Gegensatz zu Korollar 6.12 fordern wir hier *nicht* von Anfang an Rekurrenz - im Gegenteil: wir können hiermit einfach (sogar positive) Rekurrenz nachweisen.

Beweis von Satz 6.13: 1. Da $\mu \neq 0$ gibt es einen (Referenz-) Punkt a derart, dass $\mu_a > 0$. Wir definieren jetzt

$$y_i := \frac{\mu_i}{\mu_a}.$$

Damit erfüllt y die Bedingungen i) bis iv) von Satz 6.10 und ist damit grösser gleich die *minimale* Lösung $(z_{ai})_{i \in \mathbb{Z}}$. Wir haben insbesondere

$$\frac{\mu_i}{\mu_a} \geq z_{ai}.$$

2.

$$\begin{aligned}
\sum_i z_{ai} &= \sum_i \sum_{n \geq 1} E_a[I[X_n = i, n \leq N_{a1}]] \\
&= \sum_{n \geq 1} E_a[I[n \leq N_{a1}]] \\
&= \sum_{n \geq 1} P_a[n \leq N_{a1}] \\
&= \sum_{n \geq 1} P_a[N_{a1} \geq n] \\
&= E_a[N_{a1}] = m_a.
\end{aligned}$$

3. Damit haben wir

$$m_a = \sum_i z_{ai} \leq \sum_i \frac{\mu_i}{\mu_a} < \infty.$$

Wenn $m_a = E_a[N_{a1}] < \infty$, dann auch $P_a[N_{a1} = \infty] = 0$ und damit $F_{aa} = 1$; also rekurrent und auch positiv rekurrent wegen Korollar 6.11 und 6.12.

$\square$

In Korollar 6.12 haben wir

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} p_{lk}^{(i)} = \frac{1}{m_k} =: \pi_k$$

gehabt. Wegen der nicht ausgeschlossenen Periodizität (Bsp 2 in 6.1) haben wir bis jetzt nie das heiss ersehnte Resultat

$$p_{ij}^{(n)} \to \pi_j$$

erhalten, wo dieses π ein linker Eigenvektor von P zum Eigenwert 1 sein soll. Dies wird zum Schluss von Kapitel 6 jetzt erreicht. Satz 6.14 sagt aus, dass man (unter Bedingungen) im Limes unabhängig vom Startwert in einer (Gleichgewichts-) Verteilung derart ist, dass ein weiterer Schritt

$$\pi^t P = \pi^t P^2 = \pi^t P^n = \pi^t$$

den Vektor stabil lässt, stationär oder eben im Gleichgewicht. In Kapitel 4 haben wir unter Verwendung des Satzes 4.14 von Perron-Frobenius den Fall des *endlichen Zustandsraumes* ziemlich abschliessend behandelt. Nun haben wir den allgemeinen Fall erreicht.

Satz 6.14 [$p_{ij}^{(n)} \to \pi_j$**, von W. Döblin**] *Sei X eine Markov-Kette mit P irreduzibel (eine einzige Klasse) und aperiodisch und positiv rekurrent ($\exists \pi : \pi > 0, \pi^t P = \pi^t, \sum_j \pi_j = 1$). Dann gilt:*

$$p_{ij}^{(n)} \to \pi_j$$

wenn $n \to \infty$ für alle i, j.

Beweis von Satz 6.14: In Anbetracht der besprochenen Beispiele und bisherigen Theorie ist das Resultat einleuchtend. Wir verzichten deshalb auf den Beweis. Er befindet sich z.B. in Grimmett & Stirzaker, "Probability and Random Processes", Kapitel 6.4; Theorem (17).

□

W. Döblin war der Sohn des Schriftstellers Alfred Döblin (1878-1957, "Berlin Alexanderplatz").

Nach dieser Fülle von verschiedenen Sätzen, Lemmatas und Korollaren wollen wir noch die verwendeten Terme und Situationen ordnen.

<u>Ordnung in Kapitel 6</u>

Wir haben in 6.1 mit Beispiel 2 und 3 und mit dem Random Walk auf $\mathbb{Z}$ drei hervorragende Beispiele, anhand derer das gesamte Kapitel 6 entwickelt und danach memoriert werden kann. Zahlreiche Sätze, Fragestellungen und Situationen müssen dabei voneinander unterschieden werden. Die ersten "Checks", welche gemacht werden müssen, sind die folgenden **drei**:

* Habe ich eine einzige Klasse (algebraische Struktur)? (sonst mit Satz 5.3 zuerst Treffwahrscheinlichkeiten und mit Satz 5.4 erwartete Zeit bis zur Absorption berechnen, oder nur Satz 6.7 anwenden)

* Habe ich Periodizität? (die meisten Aussagen stimmen immer noch ausser Satz 6.14)

* Habe ich Transienz oder (Null-/Positive-) Rekurrenz? (entweder Ad hoc bestimmen oder Satz 6.10 oder Satz 6.13)

Wir sind im Verlauf dieses Kapitels u.a. folgenden Ausdrücken begegnet, welche alle "irgendwie zusammenhängen" (siehe jeweilige Sätze, Korollare und Lemmatas weiter oben). Man kann sie in **vier Gruppen** einteilen:

$$H_n(i) \tag{6.8}$$

$$m_i \tag{6.9}$$

$$p_{ki}^{(n)}, \ \frac{1}{n}\sum_{l=1}^{n} p_{ki}^{(l)} \tag{6.10}$$

$$z_{ai}, \ \mu_i, \ x_i^*, \ \pi_i \tag{6.11}$$

1. Zwischen (6.8) und (6.9) ist erstmals deshalb eine Trennlinie zu ziehen, weil wir in (6.8) eine zufällige Grösse haben und danach nur noch deterministische Grössen.

2. (6.9) ist deshalb von allen anderen Termen zu unterscheiden, weil er umgekehrt proportional zu allen anderen Ausdrücken steht: siehe dazu z.B. Korollar 6.5, 6.6, Satz 6.7 und weitere.

3. In (6.10) sind eng miteinander verwandte Ausdrücke aufgeführt: $p_{ki}^{(n)}$ ist unser am einfachsten interpretierbares Untersuchungsobjekt. Es gibt die Wahrscheinlichkeit an, dass wir bei Start in k, in n Schritten in Zustand i sind. Satz 6.14 verbindet diesen Ausdruck mit den linken Eigenvektoren zu Eigenwert 1 unter gewissen Bedingungen. Insbesondere die Periodizität (vgl. Beispiel 2 in Kapitel 6.1) erlaubt uns aber leider eventuell nur die Betrachtung des "Verlegenheitsausdrucks"

$$\frac{1}{n}\sum_{l=1}^{n} p_{ki}^{(l)}.$$

4. Die Ausdrücke in (6.11) unterscheiden sich allenfalls nur noch durch Konstanten, wobei sogar $z_{ai} = x_i^*$ und π auf 1 summiert.