Eine Monotonieeigenschaft reversibler Markovketten

Seminarausarbeitung zum Seminar „Stochastische Prozesse“ WS 0809

Herr Prof. Hans Daduna

Eingereicht von: Diemo Ruhnow

Matrikelnummer: 5415199

Departement Mathematik der Universität Hamburg

Bundesstraße 55

20146 Hamburg

Inhaltsverzeichnis

1. Einführung

In der Wahrscheinlichkeitstheorie bezeichnet der Hazardrate die Neigung eines Systems oder Bauteils in naher Zukunft auszufallen, wenn es bis zu dieser Zeit überlebt hat.

Der bearbeitete Text „A monotonicity in reversible markov chains" von Robert Lund, Ying Zhao und Peter C. Kiessler identifiziert eine fallende Hazardrate entlang der geraden Indizes (engl. „decreasing hazard rate" (DHR)) in allen reversiblen Markovketten mit abzählbaren Zustandsräumen. Diese Eigenschaft wird dann genutzt, um Konvergenzraten für reversible Markovketten herzuleiten.Diese Eigenschaft der fallenden Hazardrate scheint schwächer als die der klassischen stochastische Monotonie, jedoch gilt diese Eigenschaft für alle Zustände reversibler Markovkette, während nicht alle reversiblen Markovketten auch stochastisch monoton sind.

Die vorliegende Seminarausarbeitung beschäftigt sich mit den Kapiteln 1-3 sowie mit dem Beweis des in Kapitel 3 aufgeführten Theorems, welcher in Kapitel 6 der ursprünglichen Arbeit vorgestellt wird. Es wird also gezeigt, dass die erste Rückkehrzeit eines jeden Zustandes einer reversiblen Markovkette mit abzählbaren Zustandsraum die DHR-Eigenschaft zu geraden Zeitpunkten besitzt. Dies wird zunächst für endliche Zustandsräume bewiesen und dann mittels Stutzung auf abzählbare Zustandsräume erweitert.

Der folgende Teil der Seminarausarbeitung ist wie folgt angeordnet. Zunächst werden grundlegende Notationen und Eigenschaften aus der Erneuerungstheorie sowie der Theorie diskreter Zufallsvariablen in Kapitel 2 dargestellt, die DHR Eigenschaft in Kapitel 3 präsentiert und in Kapitel 4 das Hauptresultat der vorliegenden Arbeit bewiesen.

2. Notationen

Sei $\{\tau_i\}_{i=1}^{\infty}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen (Lebenszeiten) mit Zustandsraum $\{1,2,\ldots\}$. Sei τ eine Zufallsvariable mit identischer Verteilung wie τ_n für $n \geq 1$. Sei $f_k = P[\tau = k]$ für $k \geq 1$ die Zähldichte und sei $\bar{F}_k = P[\tau > k]$ für $k \geq 0$ die komplementäre Verteilungsfunktion. Sei $S_n = \tau_1 + \tau_2 + \tau_3 + \ldots + \tau_n$, $n \geq 1$ die Gesamtlebensdauer der ersten n Zufallsvariablen, mit $S_0 = 0$ und sei die (unverzögerte) Wahrscheinlichkeit eines Austausches zur Zeit n geben durch:

$$u_n = \sum_{k=0}^{\infty} P[S_k = n] \ , n \geq 1$$

Dabei ist $P[S_k = n]$ die Wahrscheinlichkeit, dass genau k Elemente bis genau dem Zeitpunkt n leben (und u_n somit die Wahrscheinlichkeit eines Ausfalles (und damit Austausches) zurzeit n). Wir setzen dabei $u_0 = 1$.

Eine bekannte elementare Rekurrenzgleichung lautet:

$$u_n = \sum_{k=1}^{n} f_k u_{n-k} \ , n \geq 1 \qquad (2.1)$$

Heuristisch ist dabei f_k dabei die Wahrscheinlichkeit eines Ausfalles (und damit Austausches) eines einzelnen Bauteiles, nachdem es k Zeiten gelebt hat und u_{n-k} die Wahrscheinlichkeit eines Ausfalles zur Zeit $n-k$, während die Summe dann alle möglichen dieser Kombination nach einer Zeit n angibt.
Für den Grenzwert gilt: $\lim_{n\to\infty} u_n = \frac{1}{\sum_{k=1}^{\infty} k f_k} =: u_\infty$ (Feller (1968, p.313)) wenn τ aperiodisch (setze $u_\infty = \infty$ wenn $E[\tau] = \infty$).

Gleichung (2.1) kann man in folgender Form darstellen:

$$\bar{F}_n = \sum_{k=1}^{n} (u_{k-1} - u_k)\, \bar{F}_{n-k} \ , n \geq 1 \qquad (2.2)$$

welche später des Öfteren von Nutzen sein wird.

Beweis:

$$u_n = \sum_{k=1}^{n} f_k u_{n-k}$$

$$\underset{u_0=1}{\Longleftrightarrow} u_n = f_n + \sum_{k=1}^{n-1} f_k u_{n-k}$$

$$\underset{+Summe+1}{\Longleftrightarrow} u_n = 1 - \sum_{k=1}^{n-1} f_k + \sum_{k=1}^{n-1} f_k u_{n-k} - 1 + \sum_{k=1}^{n} f_k$$

$$\underset{\bar{F}_n = 1 - F_n}{\Longleftrightarrow} \bar{F}_n + u_n = 1 - \sum_{k=1}^{n-1} f_k + \sum_{k=1}^{n-1} f_k u_{n-k}$$

$$\underset{u_0=1}{\Longleftrightarrow} \quad \bar{F}_n + u_n = u_0 - (\sum_{k=1}^{n-1} f_k - \sum_{k=1}^{n-1} f_k u_{n-k})$$

$$\Leftrightarrow \quad \bar{F}_n + u_n \; = u_0 - (u_0 f_1 + \cdots + u_0 f_{n-1} - u_1 f_{n-1} - \cdots - u_{n-1} f_1)$$

$$\Leftrightarrow \quad \bar{F}_n + u_n = u_0 - ((u_0 f_1 + \cdots + u_0 f_{n-1} - u_1 f_1 - \cdots - u_{n-1} f_{n-1}) + (u_1 f_1 + \cdots + u_1 f_{n-2} - u_2 f_1 - \cdots - u_2 f_{n-2}) + \cdots + (u_{n-2} f_1 - u_{n-1} f_1))$$

$$\Leftrightarrow \quad \bar{F}_n + u_n \; = u_0 - (u_0 - u_1) \sum_{k=1}^{n-1} f_k - (u_1 - u_2) \sum_{k=1}^{n-2} f_k - \cdots - (u_{n-2} - u_{n-1}) \sum_{k=1}^{1} f_k$$

$$\Leftrightarrow \quad \bar{F}_n + u_n \; = (u_0 - u_1)\left(1 - \sum_{k=1}^{n-1} f_k\right) - (u_1 - u_2)\left(1 - \sum_{k=1}^{n-2} f_k\right) - \cdots - (u_{n-2} - u_{n-1})\left(1 - \sum_{k=1}^{1} f_k\right) + u_{n-1}$$

$$\Leftrightarrow \quad \bar{F}_n + u_n \; = u_{n-1} + \sum_{k=1}^{n-1} (u_{k-1} - u_k)\left(1 - \sum_{i=1}^{n-k} f_i\right)$$

$$\underset{-u_n}{\Longleftrightarrow} \quad \bar{F}_n = \; (u_{n-1} - u_n) + \sum_{i=1}^{0} (1 - f_i) + \sum_{k=1}^{n-1} ((u_{k-1} - u_k)\left(1 - \sum_{i=1}^{n-k} f_i\right))$$

$$\Leftrightarrow \quad \bar{F}_n = \; \sum_{k=1}^{n} ((u_{k-1} - u_k)\left(1 - \sum_{i=1}^{n-k} f_i\right))$$

$$\underset{Def.}{\Longleftrightarrow} \quad \bar{F}_n = \; \sum_{k=1}^{n} (u_{k-1} - u_k)\, \bar{F}_{n-k} \quad \blacksquare$$

Definition 2.1 Die Hazardrate h_i von τ zum Index i wird wie folgt definiert:

$h_i = P[\tau = i | \tau \geq i]$ wenn $P[\tau \geq i] > 0$ ist, sonst $h_i = 0, \; i = 1,2, \dots$.

Wir sagen, „τ hat eine fallende Hazardrate (DHR)“, wenn h_i nicht steigend in i ist.
Wir nennen die Folge der Erneuerungszeitpunkte $\{u_n\}_{n=0}^{\infty}$ DHR oder log-konvex, wenn gilt:

$u_{n+1}^2 \leq u_n \, u_{n+2} \, , \quad n \geq 0$.

Korollar 2.2 Es gilt, dass τ DHR ist, genau dann, wenn die folgende Log-Konvex-Beziehung gilt:

$\bar{F}_{n+1}^2 \leq \bar{F}_n \bar{F}_{n+2} \; , n \geq 0 \qquad (2.3)$

Beweis:

Seien $P[\tau \geq n+1] > 0$ und $P[\tau \geq n+2] > 0$.

$$h_{n+1} \geq h_{n+2} \underset{Def.}{\iff} P[\tau = n+1 | \tau \geq n+1] \geq P[\tau = n+2 | \tau \geq n+2]$$

$$\underset{Def.}{\iff} \frac{P[\tau = n+1]}{P[\tau \geq n+1]} \geq \frac{P[\tau = n+2]}{P[\tau \geq n+2]}$$

$$\underset{Aufspalten}{\iff} \frac{P[\tau = n+1]}{P[\tau = n+1] + P[\tau > n+1]} \geq \frac{P[\tau = n+2]}{P[\tau = n+2] + P[\tau > n+2]}$$

$$\Leftrightarrow \; P[\tau = n+1]\,(P[\tau = n+2] + P[\tau > n+2]) \geq P[\tau = n+2]\,(P[\tau = n+1] + P[\tau > n+1])$$

$$\Leftrightarrow \; P[\tau = n+1] P[\tau > n+2] \geq P[\tau = n+2]\, P[\tau > n+1]$$

$$\underset{\tau = ersetzen}{\iff} (P[\tau > n] - P[\tau > n+1]) P[\tau > n+2] \geq (P[\tau > n+1] - P[\tau > n+2])\, P[\tau > n+1]$$

$$\Leftrightarrow \; P[\tau > n]\, P[\tau > n+2] \geq P[\tau > n+1]\, P[\tau > n+1] \underset{Def.}{\iff} \overline{F_n} \bar{F}_{n+2} \geq \bar{F}_{n+1}^2 \quad \blacksquare$$

Definition 2.2.

Sei nun X = $(X_n : (\Omega, \mathrm{A}, \mathrm{P}) \to (\{1,2,\dots\}, 2^{\{1,2,\dots\}}))$ eine ergodische (irreduzibel, aperiodisch, positiv rekurrent) Markovkette mit homogener Übergangsmatrix $P = (p_{ij})_{i,j=0}^{\infty}$ wobei $p_{i,j} = P[X_{n+1} = j | X_n = i]$ für jedes $n \geq 0$ sofern die linke Seite erklärt ist, d.h. $\mathrm{P}[X_n = i] > 0$ gilt.

Ergodische Markovketten haben eine eindeutige stationäre Verteilung bzw. Grenzverteilung, welche mit π bezeichnet wird und unabhängig vom Anfangszustand X_0 ist. Es bezeichnet $\pi_i = \pi(\{i\})$ die stationäre Aufenthaltswahrscheinlichkeit mit der sich die Markovkette im Zustand i aufhält.

$$\pi_j = \lim_{n \to \infty} P[X_n = j | X_0 = i]$$

Erneuerungsfolgen können aus Markovketten konstruiert werden; hierbei können die Zeitpunkte der Besuche der Markovkette in einem festen Zustand k als Erneuerungszeitpunkte angesehen werden.

Für alle Paare von Zuständen j und k betrachten wir mit

$$\tau_{j,k} : (\Omega, \mathcal{A}, \mathcal{P}) \to \mathbb{N} \qquad , \omega \mapsto \inf\{n \geqq 1 : X_n(\omega) = k | X_0 = j\}$$

die Ersteintrittzeit in den Zustand k (bzw. Rückkehrzeit $j = k$).

Im Folgenden werden insbesondere die geraden Rückkehrzeiten von Bedeutung erlangen, die wir folgenermaßen definieren:

$\eta_{j,k}: (\Omega, \mathcal{A}, \mathcal{P}) \to \mathbb{N} \qquad , \omega \ \mapsto \inf\{n \geqq 1: X_{2n}(\omega) = k | X_0 = j\} \qquad (2.4).$

Eine Beziehung, die wiederholt in Beweisen des ursprünglichen 4.Kapitels verwendet wird und hier der Vollständigkeit halber angegeben werden soll, ist folgende:

$$P\left[\tau_{k,k} > n\right] = \prod_{i=1}^{n} (1 - h_i(k)) \qquad (2.5)$$

wobei $h_i(k) = P\left[\tau_{k,k} = i | \tau_{k,k} \geq i\right]$ ist.

(2.5) ergibt sich aus:

$$P\left[\tau_{k,k} > n\right] = \frac{P\left[\tau_{k,k} > 1\right] P\left[\tau_{k,k} > 2\right] \dots P\left[\tau_{k,k} > n\right]}{P\left[\tau_{k,k} > 0\right] P\left[\tau_{k,k} > 1\right] \dots P\left[\tau_{k,k} > n-1\right]}$$

$$= \frac{P\left[\tau_{k,k} > 1\right] P\left[\tau_{k,k} > 2\right] \dots P\left[\tau_{k,k} > n\right]}{P\left[\tau_{k,k} \geq 1\right] P\left[\tau_{k,k} \geq 2\right] \dots P\left[\tau_{k,k} \geq n\right]} = \prod_{i=1}^{n} \left(\frac{P\left[\tau_{k,k} > i\right]}{P\left[\tau_{k,k} \geq i\right]} \right)$$

$$= \prod_{i=1}^{n} \left(\frac{P\left[\tau_{k,k} = i\right] + P\left[\tau_{k,k} > i\right] - P\left[\tau_{k,k} = i\right]}{P\left[\tau_{k,k} \geq i\right]} \right) = \prod_{i=1}^{n} \left(1 - \frac{P\left[\tau_{k,k} = i\right]}{P\left[\tau_{k,k} \geq i\right]} \right) = \prod_{i=1}^{n} (1 - h_i(k)) \qquad \blacksquare$$

Definition 2.3

Eine homogene Markovkette heißt reversibel, wenn für alle Paare von Zuständen j und k die lokale Balancegleichung $\pi_j p_{j,k} = \pi_k p_{k,j}$ gilt, wobei die π_i die Gleichgewichtsverteilung der homogenen Markovkette ist.

Reversible Markovketten sind z.B. GUT-Prozesse, Irrfahrten, Markovketten mit nur zwei Zuständen, sowie viele MCMC-generierte Markovketten.

3. Die DHR Monotonieeigenschaft

Das erste Resultat ist eine DHR Monotonie für reversible Markovketten mit endlichen Zustandsräumen:

Theorem 3.1. Sei $\{X_n\}_{n=0}^{\infty}$ eine reversible ergodische Markovkette mit endlichem Zustandsraum $\{0, \dots, N\}$.

1. Die Zufallsvariable $\eta_{k,k}$ (siehe 2.4) hat eine fallende Hazardrate für jedes k, d.h. $h_n = P\left[\eta_{k,k} = n | \eta_{k,k} \geq n\right]$ ist nicht steigend in n für jedes k.
2. $P[X_{2n} = k | X_0 = k]$ ist nicht steigend und log-konvex in n für jeden Zustand k.

Der Beweis des Theorems 3.1. beruht auf der Spektralzerlegung der Matrix P und wird im Abschnitt 4 präsentiert. Um den sehr technischen Feinheiten der Spektraltheorie für abzählbare Matrizen aus dem Weg zu gehen, nutzen wir die Beweistechnik der Stutzung, um das Theorem 3.1. auf abzählbare Zustandsräume auszudehnen. Dies soll nun im Detail vorgestellt werden.

Sei $\{X_n\}_{n=0}^{\infty}$ eine reversible ergodische Markovkette mit Zustandsraum $\{0,1,\dots\}$ und sei M eine beliebige natürliche Zahl. Wir stutzen $\{X_n\}_{n=0}^{\infty}$ auf den Zustandsraum $\{0,\dots,M\}$ in dem wir Übergänge zu Zuständen, welche größer als M sind, nicht erlauben (z.B. könnte man im jeweiligen Zustand bleiben, bei einem ursprünglichen Übergang außerhalb von $\{0,\dots,M\}$). $\left\{X_n^{(M)}\right\}_{n=0}^{\infty}$ besitzt die Übergangsmatrix P^M mit $p_{i,j}^M = p_{i,j}$ wenn $0 \le i \ne j \le M$ und $p_{i,i}^M = p_{i,i} + \sum_{k=M+1}^{\infty} p_{i,k}$ für $0 \le i \le M$.

Lemma 3.2. Die Markovkette $\left\{X_n^{(M)}\right\}_{n=0}^{\infty}$ ist reversibel und besitzt die Grenzverteilung:

$$\pi_i^{(M)} = \lim_{n\to\infty} P\left[X_n^{(M)} = i\right] = \pi_i / \sum_{i=0}^{M} \pi_i \quad \text{für i mit } 0 \le i \le M.$$

Beweis:
Es ist $\pi_i^{(M)} = \lim_{n\to\infty} P\left[X_n^{(M)} = i\right] = \lim_{n\to\infty} P[X_n = i | X_n \in \{0,\dots,M\}] = \pi_i / \sum_{i=0}^{M} \pi_i$ und es gilt $\pi_i^{(M)} p_{ij}^{(M)} = \pi_i\, p_{ij}^{(M)} \Big/ \sum_{i=0}^{M} \pi_i = \pi_j\, p_{ji}^{(M)} \Big/ \sum_{i=0}^{M} \pi_i = \pi_j^{(M)} p_{ji}^{(M)}$.

Damit ist die Markovkette in lokaler Balance, also reversibel. (heuristisch: Die gestutzte Markovkette hat als stationäre Wahrscheinlichkeit die Wahrscheinlichkeit das der Prozess langfristig in i ist unter der Bedingung, dass der Prozess langfristig in $\{0,\dots,M\}$ ist.

Korollar 3.3. Die Aussagen von Theorem 3.1. gelten auch für reversible ergodische Markovkette mit abzählbar unendlichem Zustandsraum.

Beweis:

Sei $\{X_n\}_{n=0}^{\infty}$ gegeben und $\left\{X_n^{(M)}\right\}_{n=0}^{\infty}$ die oben definierte gestutzte Markovkette. Aus der Reversibilität von $\left\{X_n^{(M)}\right\}_{n=0}^{\infty}$ und Theorem 3.1. folgt, dass

$$(\bar{F}_{n+1}^{(M)})^2 \le \bar{F}_n^{(M)} \bar{F}_{n+2}^{(M)}, \qquad n \ge 0$$

für jedes feste M, wobei $\bar{F}_n^{(M)} = P\left[\eta_{k,k}^{(M)} > n\right]$ und $\eta_{k,k}^{(M)}$ wie in 2.4. definiert ist.

Um Theorem 3.1 für $\{X_n\}_{n=0}^{\infty}$ zu erhalten, ist es ausreichend zu zeigen, dass $\lim_{M\uparrow\infty} \bar{F}_n^{(M)} = \bar{F}_n$ für jedes feste $n \ge 0$ gilt.

Weil X_j eine „echte" Zufallsvariable, d.h. $P\left[X_j \le \infty\right]$ =1 (X_j ist pos. rekurrent) , für jedes feste j mit $1 \le j \le n$ ist, gilt:

$$\lim_{M\uparrow\infty} P\left[X_j \le M, 1 \le j \le n \,|X_0 = k\right] = 1 \qquad (**).$$

Auf der Menge $\left\{X_j \le M, 1 \le j \le n\right\}$ gilt $\eta_{k,k}^{(M)} \le n$ genau dann, wenn auch $\eta_{k,k} \le n$ gilt (*).

Hieraus folgt:

$$\lim_{M\uparrow\infty} P\left[\eta_{k,k}^{(M)} \leq n \,|X_0 = k\right] = \lim_{M\uparrow\infty} P\left[\left\{\eta_{k,k}^{(M)} \leq n\right\} \cap \{X_j \leq M, 1 \leq j \leq n\} |X_0 = k\right]$$

$$=_* \lim_{M\uparrow\infty} P\left[\{\eta_{k,k} \leq n\} \cap \{X_j \leq M, 1 \leq j \leq n\} \,|X_0 = k\right]$$

$$=_{**} P\left[\eta_{k,k} \leq n \,|X_0 = k\right].$$

4. Beweis des Theorems

Um das Theorem zu beweisen, halten wir zunächst eine Reihe von technischen Aussagen fest.

Lemma 4.1.

Sei $\{X_n\}_{n=0}^{\infty}$ eine reversible ergodische Markovkette mit endlichem Zustandsraum $\{0, \dots, N\}$. Dann sind alle Eigenwerte von P^2 reell und nicht negativ.

Beweis:

Sei $\boldsymbol{\pi} = (\pi_i)$ die stationäre Verteilung der Markovkette und Π_D die Diagonalmatrix mit den Einträgen π_i in der Diagonale und $\Pi_\mathrm{D}^{1/2}$ (bzw. $\Pi_\mathrm{D}^{-1/2}$) die Diagonalmatrizen bestehend aus den Einträgen $\pi_i^{1/2}$ (bzw. $\pi_i^{-1/2}$).
Aus $\pi_i p_{ij} = \pi_j p_{ji}$ folgt, dass $\Pi_\mathrm{D}\mathrm{P}$ symmetrisch ist. Vor- und Nachmultiplikation einer symmetrischen Matrix mit einer symmetrischen Matrix erhält die Symmetrie. Daraus folgt, dass auch $\Pi_\mathrm{D}^{-1/2}\Pi_\mathrm{D}\mathrm{P}\Pi_\mathrm{D}^{-1/2} = \Pi_\mathrm{D}^{1/2}\mathrm{P}\Pi_\mathrm{D}^{-1/2}$ symmetrisch ist.
Aus der Linearen Algebra wissen wir, dass die Eigenwerte einer symmetrischen Matrix reell sind und dass Ähnlichkeitstransformation die Eigenwerte erhält. Sei λ ein Eigenwert von P, also $Px = \lambda x$. Die Nichtnegativität folgt aus $P^2x = P(Px) = P(\lambda x) = \lambda\big(P(x)\big) = \lambda^2 x$ ■

Ohne Einschränkung der Allgemeinheit seien $\{\lambda_i\}_{i=0}^{N}$ die Eigenwerte von P, wobei diese entsprechend ihrem Betrage nach folgendermaßen angeordnet seien:

$$1 = \lambda_0 > |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N| \geq 0 \quad (4.1)$$

Nach Kijima (1997, S.63) existiert zu der symmetrischen Matrix $\Pi_\mathrm{D}^{1/2}\mathrm{P}\Pi_\mathrm{D}^{-1/2}$ (siehe Beweis zum Lemma 4.1.) eine Spektralzerlegung der folgenden Art:

$$\Pi_\mathrm{D}^{1/2}\mathrm{P}\Pi_\mathrm{D}^{-1/2} = \textstyle\sum_{j=0}^{N} \lambda_j \boldsymbol{x}_j \boldsymbol{x}_j^t$$

wobei die $\boldsymbol{x}_j$ die jeweiligen Eigenvektoren zu den λ_j sind und $\{\boldsymbol{x}_j\}$ sind orthonormal zu einander sind (d.h. mit $\boldsymbol{x}_i^t\boldsymbol{x}_j = \delta_{ij}$).

Sei A eine symmetrische Matrix, dann hat A dieselben Eigenvektoren wie $\mathrm{A}^\mathrm{k}, \mathrm{k} \geq 1$ und für die Eigenwerte $\lambda_k$von A^k gilt $\lambda_k = \lambda^k$.

Damit ist lässt sich folgendes mit Hilfe der Spektralzerlegung darstellen: $\left(\sum_{j=0}^{N} \lambda_j{}^n \boldsymbol{x}_j \boldsymbol{x}_j^t\right) = \left(\Pi_D^{1/2} P \Pi_D^{-1/2}\right)^n = \Pi_D^{1/2} P^n \Pi_D^{-1/2}$.
Diese formt man um zu: $P^n = \sum_{j=0}^{N} \lambda_j{}^n (\Pi_D^{-1/2} \boldsymbol{x}_j)(\Pi_D^{1/2} \boldsymbol{x}_j^t) = \mathbf{1}\Pi^t + \sum_{j=1}^{N} \lambda_j{}^n (\Pi_D^{-1/2} \boldsymbol{x}_j)(\Pi_D^{1/2} \boldsymbol{x}_j)^t$, wobei $\mathbf{1}$ die Einheitsmatrix bezeichnet.

Betrachtet man nur die i-te Zeile und Spalte hat man damit:

$$p_{ii}^n = \pi_i + \sum_{j=1}^{N} \lambda_j{}^n \boldsymbol{x}_{ji}^2$$

Für P^2 führt dies also zu:

$$P[X_{2n} = k | X_0 = k] = \pi_k + \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 \qquad (4.2)$$

für jeden Zustand k, wobei $x_{j,k}$ die k-te Komponente des Eigenvektors $\boldsymbol{x}_j$ ist.

Nach Kijima (1997, p.63) impliziert (4.2), dass $P[X_{2n} = k | X_0 = k]$ von oben monoton gegen π_k konvergiert (da $\lambda_j \in [0,1)$.

Man kann zudem aber noch folgendes angeben:

Lemma 4.2 Betrachte eine reversible Markovkette $\{X_n\}_{n=0}^{\infty}$ mit Zustandsraum$\{0, \dots, N\}$ sowie einen festen Zustand k. Sei $\Delta_n = u_{2(n-1)} - u_{2n}$, wobei $u_n = P[X_n = k | X_0 = k]$ ist. Dann ist $\{\Delta_n\}_{n=1}^{\infty}$ eine positive, nicht steigende, log-konvexe Folge in n.

Bemerkung zum Lemma 4.2.: Der Originaltext weist an dieser Stelle einen Druckfehler auf ($\Delta_n = u_{2(n-1)} - u_{2(n-1)}$ statt $\Delta_n = u_{2(n-1)} - u_{2n}$).

Beweis:

Einsetzen von (4.2) ergibt:

$$\Delta_n = \left[\pi_k + \sum_{j=0}^{N} \lambda_j^{2(n-1)} x_{j,k}^2\right] - \left[\pi_k + \sum_{j=0}^{N} \lambda_j^{2n} x_{j,k}^2\right] = \sum_{j=0}^{N} \lambda_j^{2(n-1)} (1 - \lambda_j^2) x_{j,k}^2 =$$
$$= \sum_{j=0}^{N} \lambda_j^{2(n-1)} w_j \qquad (4.3)$$

wobei $w_j = (1 - \lambda_j^2) x_{j,k}^2$ nicht negativ ist. Daraus folgt die Positivität. Für die Behauptung des Nicht-Steigens zeigen wir, dass $\Delta_n - \Delta_{n+1} > 0$ gilt:

$$\Delta_n - \Delta_{n+1} = \left[\sum_{j=0}^{N} \lambda_j^{2(n-1)} w_j\right] - \left[\sum_{j=0}^{N} \lambda_j^{2n} w_j\right] = \sum_{j=0}^{N} \lambda_j^{2(n-1)} (1 - \lambda_j^2) w_j \geq 0$$

wobei mit selbiger Argumentation wie im ersten Beweisschritt die Behauptung des Nicht-Steigens bewiesen ist.

Um auch noch die Log-Konvexität zu erhalten, benutzen wir (4.3) und (4.1) und erhalten:

$$\Delta_n^2 = \sum_{j=0}^{N}\sum_{l=0}^{N} \lambda_j^{2(n-1)} w_j \lambda_l^{2(n-1)} w_l =^* \sum_{l=0}^{N} \lambda_l^{2(n-1)} \lambda_l^{2(n-1)} w_l w_l + 2 \sum_{\{l,j:l<j\}} \sum \lambda_j^{2(n-1)} w_j \lambda_l^{2(n-1)} w_l$$

$$\leq \sum_{l=0}^{N} (\lambda_l^2)^{2(n-1)} w_l{}^2 + 2 \sum_{\{l,j:l<j\}} \sum \lambda_j^{2(n-1)} \lambda_l^{2(n-1)} \frac{\lambda_l^2}{\lambda_j^2} w_j w_l \qquad wg. \lambda_l^2 \geq \lambda_j^2$$

$$= \sum_{l=0}^{N} (\lambda_l^2)^{2(n-1)} w_l{}^2 + 2 \sum_{\{l,j:l<j\}} \sum \left(\lambda_j^2\right)^{n-2} (\lambda_l^2)^n w_j w_l$$

$$= \sum_{l=0}^{N} (\lambda_l^2)^{n-2} (\lambda_l^2)^n w_l{}^2 + 2 \sum_{\{l,j:l<j\}} \sum \left(\lambda_j^2\right)^{n-2} (\lambda_l^2)^n w_j w_l$$

$$= \left[\sum_{j=0}^{N} \left(\lambda_j^2\right)^{n-2} w_j\right] \left[\sum_{l=0}^{N} (\lambda_l^2)^n w_l\right]$$

$$= \Delta_{n-1}\Delta_{n+1} \ \blacksquare$$

Lemma 4.3

Sei die Folge $\{C_l\}_{l=1}^{\infty}$ positiv, nicht steigend und log-konvex und die Folge $\{k_n\}_{n=0}^{\infty}$ in der Form (2.2), die rekursiv definiert ist durch

$k_0 = 1$ für n=0

$k_n = \sum_{l=1}^{n} C_l k_{n-l}$ für $n \geq 1$.

Dann folgt, dass ist auch $\{k_n\}_{n=0}^{\infty}$ log-konvex ist.

Beweis:

Mit $k_0 = 1$; $k_1 = C_1$; $k_2 = C_1^2 + C_2$; $k_3 \geq 0$ und aus $C_2 \geq 0$ hat man:

$k_1^2 - k_0 k_2 = C_1^2 - (k_1 C_1 + C_2) = -C_2 \qquad \Rightarrow \quad k_1^2 - k_0 k_2 \leq 0$

also $k_1^2 \leq k_0 k_2$.

Ebenso folgt:

$$k_0 k_3 - k_2 k_1 = (k_2 C_1 + k_1 C_2 + k_0 C_3) - k_2 C_1 = k_1 C_2 + k_0 C_3 \geq 0 \qquad (4.5)$$

Um die Log-Konvexität zu beweisen, nutzen wir folgende Identität:

$$\sum_{l=1}^{n} (k_{n-l} k_{n+2} - k_{n+1} k_{n+1-l})(C_{n+2} C_l - C_{l+1} C_{n+1})$$

$$= \sum_{l=1}^{n} (k_{n-l}k_{n+2}C_{n+2}C_l - k_{n+1}k_{n+1-l}C_{n+2}C_l - k_{n-l}k_{n+2}C_{l+1}C_{n+1} + k_{n+1}k_{n+1-l}C_{l+1}C_{n+1})$$

$$= k_{n+2}C_{n+2}\sum_{l=1}^{n} k_{n-l}C_l - k_{n+1}C_{n+2}\sum_{l=1}^{n} k_{n+1-l}C_l - k_{n+2}C_{n+1}\sum_{l=1}^{n} k_{n-l}C_{l+1}$$
$$+ k_{n+1}C_{n+1}\sum_{l=1}^{n} k_{n+1-l}C_{l+1}$$

$=_{*,**} k_{n+2}C_{n+2}k_n - k_{n+1}C_{n+2}(k_{n+1} + k_0C_{n+1}) - k_{n+2}C_{n+1}(k_{n+1} + k_nC_1) + k_{n+1}C_{n+1}(k_{n+2} - k_{n+1}C_1 - k_0C_{n+2})$

$= (C_{n+2} + C_1C_{n+1})(k_nk_{n+2} - k_{n+1}^2)$ (4.6)

Für $n = 1$ erhält man in (4.6):

$(C_3 + C_1C_2)(k_1k_3 - k_2^2) = \left(C_3C_1 - {C_2}^2\right)(k_0k_3 - k_2k_1) \geq 0.$

Aus (4.5) und der Nichtnegativität und Log-Konvexität von C_l erhält man $k_1k_3 \geq k_2^2$.

Um den Beweis per vollständiger Induktion zu vervollständigen, nehmen wir an, dass:

$k_nk_{n+2} \geq k_{n+1}^2$ gilt für alle $n \leq p-2$ wobei $p > 2$.

Setzt man $n = p - 1$ in (4.6) ein, erhält man:

$\left(C_{p+1} + C_1C_p\right)\left(k_{p-1}k_{p+1} - k_p^2\right) = \sum_{l=1}^{p-1}\left(k_{p-1-l}k_{p+1} - k_pk_{p-l}\right)\left(C_{p+1}C_l - C_{l+1}C_p\right)$ wobei

$\left(C_{p+1}C_l - C_{l+1}C_p\right) \geq 0$ wg. der Log-Konvexität und Positivität v und wg. der Induktionshypothese auch $\left(k_{p-1-l}k_{p+1} - k_pk_{p-l}\right)$ nicht negativ (siehe *** und der Bemerkung 4.4.) ist, folgt aus der linken Seite und der Tatsache, dass $\left(C_{p+1} + C_1C_p\right)$ nicht negativ ist, dass wie benötigt:

$k_{p-1}k_{p+1} \geq k_p^2$ ■

(*) $\sum_{l=1}^{n} k_{n+1-l}C_l = k_nC_1 + k_{n-1}C_2 + \cdots + k_1C_n = k_{n+1} - k_0C_{n+1}$

(**) $\sum_{l=1}^{n} k_{n+1-l}C_{l+1} = k_nC_2 + k_{n-1}C_3 + \cdots + k_1C_{n+1} = k_{n+2} - k_{n+1}C_1 - k_0C_{n+2}$

(***) Nach Induktionsvoraussetzung gilt: $0 < \frac{k_1}{k_0} < \cdots < \frac{k_{p-1}}{k_{p-2}} < \frac{k_p}{k_{p-1}}$. Nicht Negativität des Terms $\left(k_{p-1-l}k_{p+1} - k_pk_{p-l}\right)$ ist gleichbedeutend mit $\frac{k_1}{k_0} < \frac{k_{p+1}}{k_p}, \ldots, \frac{k_{p-1}}{k_{p-2}} < \frac{k_{p+1}}{k_p}$.

Bemerkung 4.4.

Die Aussage, dass die Induktionshypothese auch $\left(k_{p-1-l}k_{p+1} - k_p k_{p-l}\right)$ nicht negativ ist nicht trivial und muss noch gezeigt werden (Dies ist zurzeit noch kein Inhalt dieser Ausarbeitung.)

Beweis von Theorem 3.1:

Sei nun $u_n = P[X_{2n} = k|X_0 = k]$ für ein festes k . Aus (2.2): $\bar{F}_n = \sum_{k=1}^{n}(u_{k-1} - u_k)\, \bar{F}_{n-k}$ folgt, dass:

$$P\left[\eta_{k,k} > n\right] = \sum_{l=1}^{n}(u_{l-1} - u_l)P\left[\eta_{k,k} > n - l\right] \qquad (4.8).$$

Aus Lemma 4.2 folgt, dass $(u_{l-1} - u_l)$ positiv, nicht steigend und log-konvex in l ist.

Anwendung von Lemma 4.3 auf (4.8) ergibt die erste Aussage des Theorems.

Um die zweite Aussage des Nicht-Steigens für $P[X_{2n} = k|X_0 = k]$ folgt aus der Positivität der in Lemma 4.2 definierten Folge Δ_n.

Die Log-Konvexität folgt aus:

$$P[X_{2n} = k|X_0 = k]^2 = \left(\pi_k + \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2\right)^2 = \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \left(\sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2\right)^2$$

$$= \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \sum\nolimits_{j=0}^{N} \sum\nolimits_{l=0}^{N} \lambda_j^{2n} x_{j,k}^2 \lambda_l^{2n} x_{l,k}^2$$

$$= \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \sum\nolimits_{l=0}^{N} \lambda_l^{2n} x_{l,k}^2 \lambda_l^{2n} x_{l,k}^2 + 2 \sum_{\{l,j:l<j\}} \sum \lambda_j^{2n} x_{j,k}^2 \lambda_l^{2n} x_{l,k}^2$$

$$\leq \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \sum\nolimits_{l=0}^{N} \lambda_l^{2n} x_{l,k}^2 \lambda_l^{2n} x_{l,k}^2 + 2 \sum_{\{l,j:l<j\}} \sum \left(\lambda_j^2\right)^{n-1} x_{j,k}^2 (\lambda_l^2)^{n+1} x_{l,k}^2$$

$$= \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \sum\nolimits_{l=0}^{N} (\lambda_l^2)^{n-1} x_{l,k}^2 (\lambda_l^2)^{n+1} x_{l,k}^2 + 2 \sum_{\{l,j:l<j\}} \sum \left(\lambda_j^2\right)^{n-1} x_{j,k}^2 (\lambda_l^2)^{n+1} x_{l,k}^2$$

$$= \pi_k{}^2 + 2\pi_k \sum_{j=1}^{N} \lambda_j^{2n} x_{j,k}^2 + \sum\nolimits_{l=0}^{N} (\lambda_l^2)^{n-1} x_{l,k}^2 (\lambda_l^2)^{n+1} x_{l,k}^2$$

$$\leq^* \pi_k{}^2 + \pi_k \sum_{j=1}^{N} ((\lambda_l^2)^{n-1} + (\lambda_l^2)^{n+1}) x_{j,k}^2 + \sum\nolimits_{l=0}^{N} (\lambda_l^2)^{n-1} x_{l,k}^2 (\lambda_l^2)^{n+1} x_{l,k}^2$$

$$= P\left[X_{2(n+1)} = k|X_0 = k\right] P\left[X_{2(n-1)} = k|X_0 = k\right] \qquad \blacksquare$$

$* \; es \; gilt\!: \; 2x^n \leq x^{n+1} + x^{n-1} \; wg.\, f(x) = x^2 - 2x + 1 \; \geq 0$

5. Literatur

[LZK06] R.Lund, Y.Zhao and P.C. Kiessler. A monoctoniciy in reversible Markov chains. Journal of Apllied Probability. 43: 486-499, 2006.

[MS02] A.Müller and D.Stoyan. Comparison Methods for Stochastic Models and Risks. Wiley, Chichester,2002

[KIJ97] M.Kijima. Markov Processes for Stochastic Modeling. University Press, Cambridge, 1997

[KEL79] F.P.Kelly. Reversibility and Stochastic Networks. Wiley, Chichester, 1979