

Mischungsverhalten und Eigenfunktionen

Alexander Schlüter

18. Januar 2016

1 Einleitung

Wir haben schon die Leitfähigkeit einer Markovkette als eine Möglichkeit zur Abschätzung ihrer Mischzeit kennengelernt. In diesem Vortrag soll die Spektrallücke einer Markovkette anhand der Eigenwerte ihrer Übergangsmatrix definiert werden. Das Hauptresultat ist eine Abschätzung der Spektrallücke mithilfe der Leitfähigkeit.

1.1 Erinnerungen

Im Folgenden sei Ω ein endlicher Zustandsraum, P eine irreduzible Übergangsmatrix auf Ω mit stationärer Verteilung π .

Definition 1.1. Schreibe für $x, y \in \Omega$

$$Q(x, y) := \pi(x)P(x, y).$$

Für $A, B \subset \Omega$ sei das **Randmaß** Q definiert durch

$$Q(A, B) := \sum_{x \in A, y \in B} Q(x, y).$$

Die **Leitfähigkeit** einer Menge $S \subset \Omega$ ist

$$\phi(S) := \frac{Q(S, S^c)}{\pi(S)}$$

und die Leitfähigkeit der ganzen Kette

$$\phi_\star := \min \left\{ \phi(S) \mid S \subset \Omega, 0 < \pi(S) \leq \frac{1}{2} \right\}.$$

Wir haben gesehen, dass die Leitfähigkeit die Mischzeit der Markovkette beeinflusst. Markovketten mit einer geringen Leitfähigkeit mischen langsamer, d.h. sie konvergieren langsamer gegen die stationäre Verteilung. Man kann sich auch vorstellen, dass die Leitfähigkeit die “engste Stelle” der Kette misst, daher der Name “Flaschenhals-Quotient”.

2 Eigenwerte und die Spektrallücke

Die Verteilung einer Markovkette wird im Wesentlichen durch ihre Übergangsmatrix charakterisiert, deshalb ist es sinnvoll, sich diese mithilfe der Linearen Algebra genauer anzuschauen.

Definition 2.1. Seien $f, g \in \mathbb{R}^\Omega$. Definiere

$$\langle f, g \rangle_\pi := \sum_{x \in \Omega} f(x)g(x)\pi(x)$$

$$\|f\|_\pi := \sqrt{\langle f, f \rangle}$$

Für $f, g \in \mathbb{R}^\Omega$ schreibe

$$f \perp_\pi g : \Longleftrightarrow \langle f, g \rangle_\pi = 0.$$

f und g heißen dann π -orthogonal.

Bemerkung 2.2. Da für alle $x \in \Omega$ $\pi(x) > 0$ gilt, wird hierdurch ein Skalarprodukt auf \mathbb{R}^Ω definiert und $\ell^2(\pi) := (\mathbb{R}^\Omega, \langle \cdot, \cdot \rangle_\pi)$ ist ein Hilbertraum.

Lemma 2.3. Sei P eine Übergangsmatrix.

- (i) Ist λ ein Eigenwert von P , dann gilt $|\lambda| \leq 1$.
- (ii) Ist P irreduzibel, dann ist der Eigenraum zum Eigenwert 1 eindimensional und er wird aufgespannt durch $(1, 1, \dots, 1)^T$.
- (iii) Ist P irreduzibel und aperiodisch, so ist -1 kein Eigenwert.

Beweis. (i) Sei f Eigenfunktion zu λ und $x \in \Omega$ mit $|f(x)| = \max_{y \in \Omega} |f(y)|$. Dann gilt

$$|\lambda f(x)| = \left| \sum_{y \in \Omega} P(x, y) f(y) \right| \leq |f(x)| \sum_{y \in \Omega} P(x, y) = |f(x)|$$

und da f nicht konstant 0 ist, folgt $|\lambda| \leq 1$.

- (ii) Da die Zeilensumme 1 ist, ist $(1, 1, \dots, 1)^T$ offensichtlich Eigenfunktion zum Eigenwert 1. Sei f Eigenfunktion zu 1 und $x \in \Omega$ mit $|f(x)| = \max_{y \in \Omega} |f(y)|$. Dann

$$\sum_{y \in \Omega} P(x, y) \frac{f(y)}{f(x)} = 1$$

mit $\frac{f(y)}{f(x)} \leq 1$. Es muss also für alle $y \in \Omega$ mit $P(x, y) > 0$ $\frac{f(y)}{f(x)} = 1$ gelten.

Ist $y \in \Omega$ beliebig, so existiert $t \in \mathbb{N}$ mit $P^t(x, y) > 0$, denn P ist irreduzibel. f ist auch Eigenfunktion von P^t zum Eigenwert 1 und mit dem gleichen Argument wie oben für P^t folgt $f(y) = f(x)$.

- (iii) Zeige die Kontraposition: Sei f Eigenfunktion zum Eigenwert -1 . Sei $x \in \Omega$ mit $|f(x)| = \max_{y \in \Omega} |f(y)|$.

$$|-f(x)| = \left| \sum_{y \in \Omega} P(x, y) f(y) \right| \leq \sum_{y \in \Omega} P(x, y) |f(y)| \leq |f(x)|$$

und da $|f(y)| \leq |f(x)|$ folgt für alle $y \in \Omega$ mit $P(x, y) > 0$, dass $|f(y)| = |f(x)|$.
Ohne Beträge sieht man, dass

$$-f(x) = \sum_{y \in \Omega} P(x, y) f(y)$$

was nur möglich ist, wenn für $y \in \Omega$ mit $P(x, y) > 0$ gilt: $f(y) = -f(x)$.
Sei $t \in \mathcal{T}(x) := \{ s \in \mathbb{N} \mid P^s(x, x) > 0 \}$. Es existiert ein Pfad

$$x = v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_{t-1} \rightarrow v_t = x,$$

mit $v_i \in \Omega$, $P(v_i, v_{i+1}) > 0$ ($0 \leq i \leq t-1$). Nach obiger Betrachtung ist

$$f(x) = -f(v_{t-1}) = \dots = (-1)^t f(x)$$

Also muss t gerade sein und $\text{ggT}(\mathcal{T}(x)) \geq 2$, d.h. P ist nicht aperiodisch. □

Lemma 2.4. Sei P reversibel bezüglich π . Dann hat $\ell^2(\pi)$ eine Orthonormalbasis $f_1, \dots, f_{|\Omega|}$, wobei f_i Eigenfunktion von P zum Eigenwert λ_i ($1 \leq i \leq |\Omega|$) ist. Die Eigenfunktion zum Eigenwert 1 kann als die konstante Einsfunktion $f_1 = \mathbf{1}$ gewählt werden.

Beweis. Aus Reversibilität folgt, dass die Matrix

$$A(x, y) := \pi(x)^{1/2} \pi(y)^{-1/2} P(x, y)$$

symmetrisch ist. Die Behauptung folgt mittels des Spektralsatzes für symmetrische Matrizen. Für den vollen Beweis siehe [LPW08, Lemma 12.2(i)]. □

Bemerkung 2.5. Da A in dem Beweis von Lemma 2.4 symmetrisch ist, sind Eigenfunktionen zu unterschiedlichen Eigenwerten orthogonal. Dies überträgt sich auf die Eigenfunktionen von P : Eigenfunktionen von P zu unterschiedlichen Eigenwerten sind π -orthogonal.

Definition 2.6. Für eine reversible Übergangsmatrix P nummerieren wir ihre Eigenwerte in absteigender Ordnung:

$$1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq -1.$$

Definiere

$$\lambda_\star := \max \{ |\lambda| \mid \lambda \text{ ist Eigenwert von } P, \lambda \neq 1 \}.$$

Die Differenz $\gamma_\star := 1 - \lambda_\star$ heißt **absolute Spektrallücke**.

Die **Spektrallücke** ist definiert durch $\gamma := 1 - \lambda_2$.

Bemerkung 2.7. Nach Lemma 2.3.(iii) ist für eine irreduzible aperiodische Übergangsmatrix $\gamma_\star > 0$.

3 Die Dirichletform

Bis jetzt ist noch nicht klar, was die Spektrallücke mit der Leitfähigkeit zu tun hat. Die Verbindung wird über die Dirichletform hergestellt.

Ab hier sei P eine **reversible** Übergangsmatrix bezüglich der stationären Verteilung π .

Definition 3.1. Sei P eine reversible Übergangsmatrix mit stationärer Verteilung π . Die Abbildung

$$\mathcal{E}: \mathbb{R}^\Omega \times \mathbb{R}^\Omega \rightarrow \mathbb{R}$$

$$(f, g) \longmapsto \langle (I - P)f, g \rangle_\pi$$

heißt **Dirichletform** zu (P, π) .

Lemma 3.2. Definiere für $f \in \mathbb{R}^\Omega$

$$\mathcal{E}(f) := \frac{1}{2} \sum_{x, y \in \Omega} (f(x) - f(y))^2 \pi(x) P(x, y). \quad (1)$$

Dann gilt $\mathcal{E}(f) = \mathcal{E}(f, f)$.

Beweis. Ausmultiplizieren der Klammer in Gleichung (1) gibt

$$\mathcal{E}(f) = \underbrace{\frac{1}{2} \sum_{x, y \in \Omega} f(x)^2 \pi(x) P(x, y)}_{(a)} - \underbrace{\sum_{x, y \in \Omega} f(x) f(y) \pi(x) P(x, y)}_{(b)} + \underbrace{\frac{1}{2} \sum_{x, y \in \Omega} f(y)^2 \pi(x) P(x, y)}_{(c)}$$

Wegen Reversibilität ist $\pi(x)P(x, y) = \pi(y)P(y, x)$ und die Terme (a) und (c) lassen sich zusammenfassen:

$$\begin{aligned} (a) + (c) &= \sum_{x, y \in \Omega} f(x)^2 \pi(x) P(x, y) \\ &= \sum_{x \in \Omega} f(x)^2 \pi(x) \sum_{y \in \Omega} P(x, y) \\ &= \sum_{x \in \Omega} f(x)^2 \pi(x). \end{aligned}$$

Zusammen mit (b) also

$$\begin{aligned} \mathcal{E}(f) &= \sum_{x \in \Omega} f(x)^2 \pi(x) - \sum_{x \in \Omega} \left(\sum_{y \in \Omega} P(x, y) f(y) \right) f(x) \pi(x) \\ &= \langle f, f \rangle_\pi - \langle Pf, f \rangle_\pi \\ &= \langle (I - P)f, f \rangle_\pi \\ &= \mathcal{E}(f, f) \end{aligned}$$

□

Bemerkung 3.3. Es sei $\mathbf{1}$ die konstante Einsfunktion auf Ω . Dann ist

$$E_\pi(g) = \sum_{x \in \Omega} g(x)\pi(x) = \langle g, \mathbf{1} \rangle_\pi,$$

also $E_\pi(g) = 0$ genau dann wenn $g \perp_\pi \mathbf{1}$.

Lemma 3.4. Für die Spektrallücke $\gamma = 1 - \lambda_2$ gilt

$$\gamma = \min \left\{ \mathcal{E}(g) \mid g \in \mathbb{R}^\Omega, g \perp_\pi \mathbf{1}, \|g\|_\pi = 1 \right\} \quad (2)$$

$$= \min \left\{ \frac{\mathcal{E}(g)}{\|g\|_\pi^2} \mid g \in \mathbb{R}^\Omega, g \perp_\pi \mathbf{1}, g \neq 0 \right\} \quad (3)$$

Beweis. Sei $n := |\Omega|$. Laut Lemma 2.4 gibt es in $\ell^2(\pi)$ eine Orthonormalbasis f_1, \dots, f_n von Eigenfunktionen von P , wobei $f_1 = \mathbf{1}$. Ist $g \in \mathbb{R}^\Omega$, so kann g in der ONB entwickelt werden zu

$$g = \sum_{j=1}^n \langle g, f_j \rangle_\pi f_j.$$

Setze $a_j := \langle g, f_j \rangle_\pi$. Ist $g \perp_\pi \mathbf{1}$, d.h. $\langle g, f_1 \rangle_\pi = 0$, so fällt der erste Summand weg. Gilt auch $\|g\|_\pi = 1$, dann ist

$$1 = \|g\|_\pi^2 = \langle g, g \rangle_\pi = \sum_{j=2}^n a_j^2.$$

Insgesamt

$$\begin{aligned} \mathcal{E}(f) &= \langle (I - P)g, g \rangle_\pi = \langle (I - P) \sum_{j=2}^n a_j f_j, g \rangle_\pi \\ &= \langle \sum_{j=2}^n a_j (1 - \lambda_j) f_j, g \rangle_\pi = \sum_{j=2}^n a_j (1 - \lambda_j) \\ &\geq (1 - \lambda_2) \sum_{j=2}^n a_j^2 = (1 - \lambda_2) \end{aligned}$$

mit Gleichheit im vorletzten Schritt für $g = f_2$. Damit ist Gleichung (2) gezeigt.

Die Ungleichung “ \geq ” von (2) zu (3) ist klar. Sei $g \in \mathbb{R}^\Omega$ mit $g \perp_\pi \mathbf{1}$, $\|g\|_\pi \neq 0$. Dann hat $\tilde{g} := g/\|g\|_\pi$ Norm 1 und

$$\mathcal{E}(\tilde{g}) = \langle (I - P) \frac{g}{\|g\|_\pi}, \frac{g}{\|g\|_\pi} \rangle_\pi = \frac{\mathcal{E}(g)}{\|g\|_\pi^2}.$$

Also gilt auch “ \leq ”. □

4 Abschätzung der Spektrallücke über die Leitfähigkeit

Wir formulieren jetzt das Hauptresultat:

Satz 4.1 (Jerrum und Sinclair (1989), Lawler und Sokal (1988))

Sei P eine irreduzible, reversible Übergangsmatrix mit stationärer Verteilung π . Sei λ_2 der zweitgrößte Eigenwert von P und ϕ_\star die Leitfähigkeit. Dann gilt für die Spektrallücke $\gamma = 1 - \lambda_2$:

$$\frac{\phi_\star^2}{2} \leq \gamma \leq 2\phi_\star. \quad (4)$$

Bemerkung 4.1. Die Abschätzung nach oben ist linear in ϕ_\star , die Abschätzung nach unten ist quadratisch. Die folgenden Beispiele zeigen, dass es Fälle gibt in denen die untere Abschätzung scharf ist, und welche, in denen die obere Abschätzung scharf ist.

Beispiel 4.2 (Träge Irrfahrt auf dem n -dimensionalen Hyperwürfel). Sei $\Omega = \{-1, 1\}^n$ der n -dimensionale Hyperwürfel. Betrachte die Menge $S := \{x \in \Omega \mid x_1 = -1\}$ der Ecken, deren erste Koordinaten -1 ist. Die stationäre Verteilung ist die Gleichverteilung $\pi = (2^{-n}, 2^{-n}, \dots, 2^{-n})^T$. Also gilt

$$\pi(S) = 2^{-n}|S| = 2^{-n} \cdot 2^{n-1} = \frac{1}{2}.$$

Es sei E die Menge der Kanten. Es ist $\{x, y\} \in E$ genau dann, wenn sich x und y in genau einer Koordinate unterscheiden. Der Rand von S ist deshalb

$$\begin{aligned} \partial S &= \{ \{x, y\} \in E \mid x \in S \text{ und } y \in S^c \} \\ &= \left\{ \{(-1, v), (1, v)\} \mid v \in \{-1, 1\}^{n-1} \right\}. \end{aligned}$$

Folglich ist $|\partial S| = 2^{n-1}$. Die Leitfähigkeit der Menge S lässt sich nun berechnen:

$$\begin{aligned} \phi(S) &= \frac{Q(S, S^c)}{\pi(S)} = 2 \sum_{x \in S, y \in S^c} \pi(x) P(x, y) \\ &= 2 \cdot |\partial S| \cdot 2^{-n} \cdot \frac{1}{2n} = 2 \cdot 2^{n-1} \cdot 2^{-n} \cdot \frac{1}{2n} = \frac{1}{2n} \end{aligned}$$

Es folgt

$$2\phi_\star \leq 2\phi(S) = 1/n. \quad (5)$$

Für jede Teilmenge $J \subset \{1, \dots, n\}$ definiere $f_J : \Omega \rightarrow \mathbb{R}$

$$f_J(x_1, \dots, x_n) := \prod_{j \in J} x_j$$

Beh: f_J ist Eigenfunktion von P zum Eigenwert $1 - |J|/n$.

Beweis. Sei $x \in \Omega$ beliebig. Partitioniere Ω durch vier Mengen:

$$\begin{aligned} A &:= \{ y \in \Omega \mid y \text{ unterscheidet sich von } x \text{ nur in einer Koordinate aus } J \} \\ B &:= \{ y \in \Omega \mid y \text{ unterscheidet sich von } x \text{ nur in einer Koordinate aus } J^c \} \\ C &:= \{x\} \\ D &:= \{ y \in \Omega \mid y \text{ unterscheidet sich von } x \text{ in mehr als einer Koordinate} \} \end{aligned}$$

Für $y \in D$ ist $P(x, y) = 0$. Außerdem ist $|A| = |J|$ und $|B| = |J^c| = n - |J|$. Ändert man einen Koordinate von x , die in J enthalten ist, so dreht sich das Vorzeichen von f_J um: $f_J(y) = -f_J(x)$. Ändert man aber eine Koordinate aus J^c , so hat dies keinen Einfluss auf f_J , d.h. $f_J(y) = f_J(x)$. Wir können also rechnen

$$\begin{aligned} Pf(x) &= \sum_{y \in \Omega} P(x, y) f(y) \\ &= P(x, x) f(x) + \sum_{y \in A} P(x, y) f(y) + \sum_{y \in B} P(x, y) f(y) \\ &= \frac{1}{2} f(x) + |J| \cdot \frac{1}{2n} \cdot (-f(x)) + (n - |J|) \frac{1}{2n} f(x) \\ &= \left(1 - \frac{|J|}{n}\right) f(x). \end{aligned}$$

□

Beh: Es gibt keine weiteren Eigenwerte.

Beweis. Zeige zunächst, dass für $I, J \subset \Omega$ mit $I \neq J$ gilt: $\langle f_I, f_J \rangle_\pi = 0$. Sei dazu o.B.d.A. $k \in I \setminus J$. Dann haben wir

$$\begin{aligned} \langle f_I, f_J \rangle_\pi &= \sum_{x \in \Omega} f_I(x) f_J(x) \\ &= \sum_{(x_1, \dots, \cancel{x_k}, \dots, x_n)} \sum_{x_k \in \{-1, 1\}} x_k \cdot f_{I \setminus \{k\}}(x) f_J(x) \\ &= 0, \end{aligned}$$

wobei die letzte Gleichheit gilt, da weder $f_{I \setminus \{k\}}$ noch f_J von x_k abhängen. Die Menge

$$Z := \{ f_J \mid J \subset \{1, \dots, n\} \}$$

ist demnach linear unabhängig und $|Z| = 2^n = \dim(\mathbb{R}^\Omega)$, also ist Z eine Basis von \mathbb{R}^Ω .

Wäre nun μ ein anderer Eigenwert von P , so wäre nach Bemerkung 2.5 die zugehörige Eigenfunktion g π -orthogonal auf allen Eigenvektoren zu anderen Eigenwerten, also auch auf allen Elementen aus Z . Dann wäre g linear unabhängig zu Z . Dies ist ein Widerspruch.

□

Demnach ist der zweitgrößte Eigenwert der Übergangsmatrix P nun $\lambda_2 = 1 - 1/n$ und die Spektrallücke $\gamma = 1 - \lambda_2 = 1/n$. Wegen

$$\frac{1}{n} = \gamma \stackrel{4.1}{\leq} 2\phi_\star \stackrel{(5)}{\leq} \frac{1}{n}$$

gilt

$$\gamma = \frac{1}{n} = 2\phi_\star$$

und die obere Abschätzung aus Satz 4.1 ist scharf und von korrekter Ordnung in n .

Literatur

- [LPW08] David A. Levin, Yuval Peres und Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008. URL: <http://pages.uoregon.edu/dlevin/MARKOV/>.