Seminar Geometrie, Sommersemester 2010
bei Prof. Dr. C. Bär
Institut für Mathematik der Universität Potsdam

Eigenwerte und Zufallswege I

Clara Wolff, 12.05.2010

Einführung

Definition 1 *Ein Random Walk (RW) auf einem Graphen* $G = (V, E), V = \{v_1, \dots, v_n\}$, *ist eine zufällige Folge von Knoten* $(x_0, x_1, \dots, x_k), x_i \in V$ $\forall i \in \{1 \dots k\}$.

- Aufeinanderfolgende Knoten müssen durch eine Kante des Graphen verbunden sein.
- Der Verlauf des Weges hängt von den Übergangswahrscheinlichkeiten $P(v_j, v_k) = \mathbb{P}(x_{i+1} = v_k \mid x_i = v_j)$ von einem Knoten zum anderen ab.
- Diese Wahrscheinlichkeit hängt nicht vom Zeitpunkt i ab.
- Bei einem gewöhnlichen RW ist $P(v_j, v_k) = \frac{1}{deg\, v_j}$.
- Es ist $\sum_{i=1}^{n} P(v_j, v_i) = 1$ für alle $j \in \{1, \dots, n\}$
- Die Übergangswahrscheinlichkeit kann in einer Übergangsmatrix $P = (p_{i,j})_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$ dargestellt werden, wobei $p_{i,j} = P(v_i, v_j)$.
- Weiterhin definiert man eine Anfangsverteilung $f \in \mathbb{R}^n$, welche die Wahrscheinlichkeit angibt, bei welchem Knoten der Weg beginnt:
  $f = (f_1, \dots, f_n), \quad f_i = \mathbb{P}(x_0 = v_i), \quad \sum_{i=1}^{n} f_i = 1$
- Die Verteilung von x_k, also nach k Schritten, ergibt sich durch k-malige Multiplikation von f mit P:
  $\mathbb{P}(x_k = v_i) = fP^k(i)$

Definition 2 *Ein RW heisst* reversibel, *falls*

$$\forall \{v_{i_0} \dots v_{i_k}\} \subseteq V : \mathbb{P}(x_0 = v_{i_0}, \dots, x_k = v_{i_k}) = \mathbb{P}(x_0 = v_{i_k}, \dots, x_k = v_{i_0}), v_{i_j} \in V$$

also wenn die Wahrscheinlichkeit, den Weg rückwärts zu laufen genauso hoch ist, wie ihn vorwärts zu laufen.

- Dies ist genau dann der Fall, wenn die sogenannte *detailed balance equation* für die Anfangsverteilung f gilt:

$$\forall i, j \in \{1 \dots n\} : f_i P(v_i, v_j) = f_j P(v_j, v_i) \tag{1}$$

Lemma 1 *Jede Verteilung f, welche (1) erfüllt, ist stationär, d.h. $fP = f$.*

- Für RW auf gewichteten Graphen $G = (V, E, \omega), \omega : V \times V \to \mathbb{R}$ setzt man

$$\frac{\omega(v_i, v_j)}{deg v_i} = P(v_i, v_j),$$

wobei hier $deg\, v_i = \sum_{j=1}^{n} \omega(v_i, v_j)$ (gewichteter Grad).

Definition 3 *Man bezeichnet einen RW als* ergodisch*, wenn er für jede beliebige Anfangsverteilung gegen die selbe stationäre Verteilung $\pi = (\pi_1, \ldots, \pi_n) \in \mathbb{R}^n$ konvergiert, d.h. $\lim_{s\to\infty} fP^s(i) = \pi_i$.*

Lemma 2 *Notwendige Bedingungen für Ergodizität sind*

- Irreduzibilität*, d.h. $\forall i, j \leq n, i \neq j : \exists s : P^s(v_i, v_j) > 0$. Dies ist äquivalent dazu, dass G zusammenhängend ist, also $\lambda_1 > 0$.*
- Aperiodizität*, d.h. $\forall i, j \leq n : g.g.T.\{s \mid P^s(v_i, v_j) > 0\} = 1$. Hieraus folgt, dass $\lambda_{n-1} < 2$, denn sonst gäbe es eine bipartite Zusammenhangskomponente und der RW hätte (mindestens) Periode 2.*

Hinreichende Bedingungen für Ergodizität

- Wir wollen nun zeigen, dass Irreduzibilität und Aperiodizität hinreichende Bedingungen für Ergodizität sind.
- Aus den vorhergehenden Vorträgen ist bereits bekannt:

$$P = T^{-1}A = T^{-1/2}(I - \mathcal{L})T^{1/2} \tag{2}$$

Satz 1 *Ein Random Walk mit zugehörigem gewichteten Graphen $G = (V, E, \omega)$ ist genau dann ergodisch, wenn er irreduzibel und aperiodisch ist.*

Beweis

- Notwendigkeit gilt nach Lemma 1.
- Es bleibt zu zeigen: Ist ein RW irreduzibel und aperiodisch, so ist er ergodisch.
- Zuerst wollen wir zeigen, dass $\pi = \frac{\mathbb{1}T}{vol\,G}$ eine stationäre Verteilung für P ist, wobei $vol\,G := \sum_{i=1}^{n} deg\,v_i$ und $\mathbb{1} := (1, \ldots, 1) \in \mathbb{R}^n$:
- Nun wollen wir zeigen, dass der irreduzible, aperiodische RW mit Übergangsmatrix P für beliebige Anfangsverteilungen gegen π in der euklidischen Norm konvergiert.
- Sei $f \in \mathbb{R}^n$ eine beliebige Anfangsverteilung.

- Da die Matrix $\mathcal{L}$ symmetrisch ist, gibt es eine orthonormale Basis aus ihren Eigenvektoren des $\mathbb{R}^n$, diese sei $B = \{\Phi_0, \ldots, \Phi_{n-1}\}$, wobei Φ_i Eigenvektor zum Eigenwert λ_i ist.
- Dann sei $fT^{-1/2} = \sum_{i=0}^{n-1} a_i \Phi_i$ die Basisdarstellung bezgl. B.
- Aus den vorhergehenden Vorträgen ist bekannt, dass $\Phi_0 = \frac{\mathbb{1}T^{1/2}}{\sqrt{Vol\,G}}$ ist(Eigenvektor zu λ_0).
- Dann ist

$$a_0 = \frac{1}{\sqrt{Vol\,G}}$$

- Zeige nun Konvergenz:

$$\begin{aligned}
\|fP^s - \pi\| &= \|fP^s - \frac{\mathbb{1}T}{vol\,G}\| \\
&= \|fP^s - a_0\Phi_0 T^{1/2}\| && (a_0, \Phi_0 \text{ einsetzen}) \\
&\leq (1-\lambda')^s \frac{max_{i=1\ldots n}\sqrt{deg\,v_i}}{min_{i=1\ldots n}\sqrt{deg\,v_i}} \\
&\leq e^{-s\lambda'} \frac{max_{i=1\ldots n}\sqrt{deg\,v_i}}{min_{i=1\ldots n}\sqrt{deg\,v_i}} && (\text{da } (1-\lambda') < 1 \text{ ist } (1-\lambda') \leq e^{-\lambda'})
\end{aligned} \tag{3}$$

mit $\lambda' = \begin{cases} \lambda_1, & \text{wenn } 1-\lambda_1 \geq \lambda_{n-1} - 1 \\ 2 - \lambda_{n-1}, & \text{sonst} \end{cases}$

$\Rightarrow$ für $s \geq \frac{1}{\lambda'} ln \left(\frac{max_{i=1\ldots n}\sqrt{deg\,v_i}}{\epsilon\, min_{i=1\ldots n}\sqrt{deg\,v_i}} \right) =: c_{RW}(\epsilon)$ist $\|fP^s - \pi\| \leq \epsilon$

- Man kann weiterhin zeigen, dass die Konvergenzgeschwindigkeit eigentlich nur von λ_1 abhängt, indem man einen einen modifizierten RW, den sogenannten *lazy walk*(LRW)betrachtet, dieser bleibt mit erhöhter Wahrscheinlichkeit in dem selben Knoten, also

$$\tilde{\omega}(v_i, v_j) = \begin{cases} \omega(v_i, v_i) + c deg\,v_i & \text{falls } i = j \\ \omega(v_i, v_j) & \text{falls } i \neq j \end{cases}$$

$$\Rightarrow deg_{\tilde{G}}\,v_i = (1+c) deg_G\,v_i$$

.

- Dieser LRW muss deshalb langsamer gegen seine stationäre Verteilung konvergieren, also $c_{LRW}(\epsilon) \geq c_{RW}(\epsilon)$
- Für seine Eigenwerte gilt: $\tilde{\lambda}_k = \frac{\lambda_k}{1+c}$
- Wähle nun c so, dass $1 - \tilde{\lambda}_1 = \tilde{\lambda}_{n-1} - 1$, also dass $\tilde{\lambda}' = \tilde{\lambda}_1$.

- Dann folgt für die Konvergenzgrenze des LRW $c_{LRW}(\epsilon)$:

$$c_{RW}(\epsilon) \leq c_{LRW}(\epsilon) \leq \frac{3}{2}\frac{1}{\lambda_1} ln\left(\frac{max_{i=1...n}\sqrt{deg_G\, v_i}}{\epsilon\, min_{i=1...n}\sqrt{deg_G\, v_i}}\right)$$

$\Rightarrow$ Dies ist eine Abschätzung der Konvergenzgeschwindigkeit des RW auf G, die nur von λ_1 abhängt.

$\square$

Relativer punktweiser Abstand

- Neben der Konvergenz in der euklidischen Norm gibt es noch weitere, stärkere Konvergenzbegriffe für RW, zum Beispiel:

Definition 4 *Der* relative punktweise Abstand *(RPD) ist definiert als*

$$\Delta(s) := max_{i,j=1...n}\frac{|P^s(v_i, v_j) - \pi(v_j)|}{\pi(v_j)}$$

Lemma 3 *Der relative punktweise Abstand $\Delta(s)$ ist ein stärkerer Konvergenzbegriff als der Abstand in der euklidischen Norm, also die Konvergenz eines RW gegen eine stationäre Verteilung π in $\Delta(s)$ impliziert die Konvergenz in der euklidischen Norm.*

- Ähnlich wie für die euklidische Norm kann man zeigen, dass für den RPD gilt: $\Delta(s) \leq e^{-s(1-\bar{\lambda})}\frac{vol\, G}{min_{i=1...n} deg\, v_i}$, mit $\bar{\lambda} = max_{i=0...n-1}|1-\lambda_i|$
- Weiterhin kann durch Betrachtung des zugehörigen LRW die Konvergenzgeschwindigkeit im RPD nur in Abhängigkeit λ_1 abgeschätzt werden.
- Es gilt: $\Delta(s) \leq e^{-\frac{2t\lambda_1}{2+\lambda_1}}\frac{vol\, G}{min_{i=1...n} deg\, v_i}$

Literatur

[1] F. Chung: *Spectral Graph Theory*, Amer. Math. Soc. 1997

[2] Davis A. Levin, Yuval Peres, Elizabeth L. Wilmer : *Markov Chains and Mixing Times*