Proseminar Optimierung, Graphentheorie und Stochastik

Michael Kalab

19. Dezember 2011

Inhaltsverzeichnis

1 Markov-Kette

Definition 1. *Seien* $X_1, X_2, \ldots X_n$ *eine Folge von Zufallsvariablen,* $s_1, \ldots, s_j$ *eine Menge an Zuständen. Wir nennen die Folge* $(X_i)_{i\in\mathbb{N}}$ *eine Markov-Kette, falls*

$$P(X_{n+1} = s_{j_{n+1}} \mid X_n = s_{j_n}) =$$
$$P(X_{n+1} = s_{j_{n+1}} \mid X_n = s_{j_n},\ X_{n-1} = s_{j_{n-1}}, \ldots, X_0 = s_{j_0}).$$

Die Zukunft des Prozesses ist also nur vom aktuellen Zustand abhängig und nicht von der Vergangenheit.

[1]

2 stochastische Matrix

Definition 2. *Wir nennen* $P \in \mathbb{R}^{n\times n}$ *stochastisch, falls* $P_{ij} \geq 0\ \forall i, j$ *und* $\sum_{i=1}^{n} P_{ij} = 1$*, doppelt stochastisch, wenn zusätzlich noch* $\sum_{j=1}^{n} P_{ij} = 1$*. Dabei ist* P_{ij} *die Wahrscheinlichkeit für den Zustand* j*, in Zustand* i *überzugehen.*

[2]

3 Vektoriteration

Theorem 1. *Sei* $A \in \mathbb{C}^{n\times n}$ *diagonalisierbar mit Eigenwerten* $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$*, und zugehörigen Eigenräumen* $V_1, V_2, \ldots, V_n$*. Sei* $\mathbb{C}^n \ni v = \sum_{i=1}^{n} c_i v_i$ *mit* $\|v\| = 1$*,* $v_i \in V_i$ *und* $c_1 \neq 0$*. Sei* $K(v) = \frac{Av}{\|Av\|}$*. Dann konvergiert* $K^n(v)$ *gegen* v_1 *für* $n \to \infty$*. Die Konvergenzgeschwindigkeit hängt dabei nur maßgeblich von* λ_2 *ab.*

Beweis. siehe Tafel. □

[3]

4 Problemstellung

Problem 1. *Gegeben sei ein Pfad* $v_1 - v_2 - \cdots - v_n$*. Die Übergangswahrscheinlichkeit von Knoten* v_i *nach Knoten* v_j *sei* P_{ij}*. Wir betrachten symmetrische*

Übergangswahrscheinlichkeiten, d.h. $P_{ij} = P_{ji}$. Ferner gilt $P_{ij} = 0$ für $|i-j| > 1$. Daher ist

$$P = \begin{pmatrix} P_{11} & P_{12} & 0 & \dots & 0 & 0 \\ P_{21} & P_{22} & \ddots & \dots & 0 & 0 \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & & & \\ & & & & & \\ 0 & 0 & \dots & & P_{n-1,n-1} & P_{n-1,n} \\ 0 & 0 & \dots & & P_{n,n-1} & P_{n,n} \end{pmatrix}$$

eine doppelt stochastische, symmetrische, tridiagonale Matrix. Für welche P konvergiert $P^n x$ am schnellsten gegen einen Vektor $\bar{x}$? Laut Theorem 1 hängt die Konvergenzgeschwindigkeit nur maßgeblich vom betragsmäßig zweitgrößten Eigenwert ab.

[4]

Beispiel 1. *Wir betrachten eine Vielzahl von Individuen eines Mikroorganismusses mit Generationenwechsel. Seien $g_1, \dots, g_n$ die unterschiedlichen Erscheinungsformen. Die Wahrscheinlichkeit des Wechsels von g_i auf g_j sei symmtrisch, d.h. $P(g_i|g_j) = P(g_j|g_i)$. Ferner sei ein Wechsel von g_i auf g_j unmöglich für $|i-j| > 1$.*
Leider ist nur eine Erscheinungsform transportfähig, das Labor braucht zur Durchführung des Experiments aber alle im selben Ausmaß. Welche Bedingungen sollen gewählt werden, um eine möglichst schnelle Gleichverteilung zu gewährleisten?

Lemma (1)**.** Sei P stochastisch und diagonalisierbar mit Eigenwerten $\lambda_1, \dots, \lambda_n$. Dann ist $\lambda_1 = 1$ und $|\lambda_i| \leq 1$ für alle i.

Beweis. Sei $\mathbb{1}$ der Vektor mit allen Einträgen gleich 1. Damit ist

$$P\mathbb{1} = \begin{pmatrix} \sum_{i=1}^n P_{i1} \cdot 1 \\ \vdots \\ \sum_{i=1}^n P_{in} \cdot 1 \end{pmatrix} = \mathbb{1}$$

1 ist ein also ein Eigenwert. Sei nun λ ein Eigenwert und $P(x_1, \dots, x_n)^T = \lambda(x_1, \dots, x_n)^T$. Sei i_0 ein Index mit $|x_{i_0}| \geq |x_i|$ für alle i. Dann gilt

$$|\lambda||x_{i_0}| = |\lambda x_{i_0}| = |\sum_{j=1}^n P_{i_0 j} x_j| \leq \sum_{j=1}^n P_{i_0 j}|x_j| \leq \sum_{j=1}^n P_{i_0 j}|x_{i_0}| = |x_{i_0}|$$

Damit ist $|\lambda| \leq 1$. $\square$

[5]

5 Spektralnorm

Definition 3. *Sei $A \in \mathbb{C}^{n \times n}$. Die Spektralnorm $\|A\|_S$ ist*

$$\|A\|_S = \sqrt{max_j |\lambda_j^2|} \text{ für } \lambda_j^2 \text{ Eigenwert von } A^*A.$$

Satz 1. *Sei $A \in \mathbb{C}^{n \times n}$ hermitesch. Dann entspricht die Spektralnorm dem betragsmäßig größten Eigenwert.*

Beweis. durch Nachrechnen $\square$

Satz 2. *Sei $\|\cdot\|$ die Supremumsnorm, $A \in \mathbb{C}^{n \times n}$. Dann gilt $\|A\| \leq \|A\|_S$.*

Beweis. A^*A ist hermitesch und besitzt daher eine Basis aus orthogonalen Eigenvektoren $u_1,\ u_2, \dots, u_n$. Sei $x \in \mathbb{C}^n$ mit $\|x\| = 1$. Es ist

$$\|Ax\|_2^2 = x^*A^*Ax = (\sum_{k=1}^n c_k u_k)^* \ (\sum_{j=1}^n c_j (A^*A) u_j) =$$

$$= (\sum_{k=1}^n c_k u_k)^* \ (\sum_{j=1}^n \lambda_j^2 c_j u_j) = \sum_{k=1}^n \lambda_k^2 |c_k|^2 \leq$$

$$\leq max_k(|\lambda_k^2|) \sum_{k=1}^n |c_k|^2 = \|A\|_S^2 \|x\|^2$$

$\square$

[6]

Korollar 1. *Seien $P \in \mathbb{R}^{n \times n}$ eine symmetrische, stochastische Matrix, $x \in \mathbb{R}^{n \times n}$ mit $\|x\| = 1$. Dann gilt $<x, Ax> \leq \ \|A\|_S$.*

Beweis. Schreibe anstatt A^*A nur A und ersetze λ_k^2 durch λ_k. $\square$

Lemma **(2).** Sei $P \in \mathbb{R}^{n \times n}$ eine symmetrische, stochastische Matrix mit Eigenwerten $1 > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Dann hat $P - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ die Eigenwerte $0,\ \lambda_2, \dots,\ \lambda_n$.

Beweis. P und $P - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ haben gleiche Eigenräume. $\square$

Lemma (3)***.*** Sei $P \in \mathbb{R}^{n \times n}$ eine symmetrische, stochastische Matrix, $y,\ z \in \mathbb{R}^n$ mit

$$\mathbf{1}^T y = 0,\ \|y\| = 1,$$
$$Pij \neq 0 \Longrightarrow \frac{z_i + z_j}{2} \leq y_i y_j$$

Sei $\mu(P)$ der betragsmäßig zweitgrößte Eigenwert. Dann ist $\mu(P) \geq \mathbf{1}^T z$.

Beweis.

$$\begin{aligned}
&\mu(P) \\
&= \|P - (\frac{1}{n})\mathbf{1}\mathbf{1}^T\|_S \geq \\
&\geq y^T(P - \frac{1}{n}\mathbf{1}\mathbf{1}^T)y = y^T P y = \sum_{i,j} P_{ij} y_i y_j \geq \\
&\geq \sum_{i,j} \frac{1}{2}(z_i + z_j) P_{ij} = \frac{1}{2}(z^T P \mathbf{1} + \mathbf{1}^T P z) = \mathbf{1}^T z
\end{aligned}$$

□

6 Lösung der Problemstellung

Theorem 2. *Die Matrix*

$$\mathcal{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \ddots & & \\ & \ddots & \ddots & & \\ & & & & \\ & & & 0 & \frac{1}{2} \\ & & & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

hat unter allen symmetrischen, stochastische Matrizen den kleinsten betragsmäßig zweitgrößten Eigenwert.

Beweis. Die Eigenwerte und Eigenvektoren von $\mathcal{P}$ sind

$$\lambda_1 = 1,\ v_1 = \frac{1}{\sqrt{n}}\mathbf{1}$$
$$\lambda_j = cos\Big(\frac{(j-1)\pi}{n}\Big)$$
$$v_j(k) = \sqrt{\frac{2}{n}}\ cos\Big(\frac{(2k-1)(j-1)\pi}{2n}\Big)\text{für} j = 2, \ldots,\ n, k = 1, \ldots,\ n$$

Wir wählen $z \in \mathbb{R}^n$ mit

$$z_i = \frac{1}{n}\Big(cos\Big(\frac{\pi}{n}\Big) + cos\Big(\frac{(2i-1)\pi}{n}\Big)/cos\Big(\frac{\pi}{n}\Big)\Big)$$

Damit ist $\mathbf{1}^T z = cos(\frac{\pi}{n})$. Die Vektoren v_2 und z erfüllen die Bedingungen in Lemma 3. Nachdem also $\mu(P) \leq cos(\frac{\pi}{n})$ für alle stochastischen, symmetrischen Tridiagonalmatrizen gilt und $\mu(\mathcal{P}) = cos(\frac{\pi}{n})$, ist $\mathcal{P}$ die gesuchte Matrix. □

[4]

Literatur

[1] http://de.wikipedia.org/wiki/Markow-Kette
11.12.2011

[2] http://de.wikipedia.org/wiki/%C3%9Cbergangswahrscheinlichkeit
11.12.2011

[3] http://www.tm-mathe.de/Themen/html/vektoriteration_nach_von_mises.html
13.12.2011

[4] Boyd, Stephen, Persi Diaconis, Jun Sun und Lin Xiao: Fastest mixing Markov chain on a path. Amer. Math. Monthly, 113(1):70–74, 2006.

[5] http://liinwww.ira.uka.de/courses/vl/rand-alg/skript-07.pdf
11.12.2011

[6] http://de.wikiversity.org/wiki/
Kurs:Numerik_I/2_Normen_und_Fehlerabsch%C3%A4tzungen
#2.2_Spezielle_Matrixnormen
11.12.2011