IAD metody v počítání Markovových řetězců

Ivana Pultarová

Katedra matematiky, Fakulta stavební, ČVUT
Ústav informatiky, AV ČR

SNA 2006

Přehled

Stochastická matice B typu $N \times N$: $\qquad B \geq 0, \quad e^T B = e^T$.

Hledáme vektor stacionárního rozdělení pravděpodobnosti $\hat{x}$ matice B (Perronův - Frobeniův vektor),

$$B\hat{x} = \hat{x}.$$

Perronova - Frobeniova věta. Existence a jednoznačnost $\hat{x}$ pro B ireducibilní.

Např. mocninná metoda pro ireducibilní a necyklickou matici B,

$$\hat{x} = \lim_{k \to \infty} B^k y.$$

Základní označení

n je počet agregačních skupin (G_i),

$G_1, \ldots, G_n$ jsou disjunktní množiny indexů, $\cup_{i=1}^n G_i = \{1, \ldots, N\}$,

I je jednotková matice a e je vektor jedniček,

R je matice redukce typu $n \times N$, $R_{ij} = 1$ pro $j \in G_i$ a 0 jinak,

$S(x)$ je matice prolongace typu $N \times n$, $S_{ji}(x) = x_j / \sum_{k \in G_i} x_k$ pro $j \in G_i$ a 0 jinak,

$P(x)$ je projekce $P(x) = S(x)R$

Obecná iterační agregační - deagregační (IAD) metoda. Opakují se kroky 2 - 5:

1. Zvolí se počáteční aproximace $x^0 > 0$, $||x^0||_1 = 1$, a matice základní iterace T; $k := 0$.
2. Vyřeší se $RBS(x^k)z = z$.
3. Prodloužení z na $y^k = S(x^k)z$.
4. Korekce základní iterací $x^{k+1} = Ty^k$.
5. Test konvergence, $k := k + 1$.

Základní iterace v kroku 4 může odpovídat např. mocninné metodě, blokové Jacobiově nebo Gaussově-Seidelově metodě, atd.

Jako základní iterace používáme násobení maticí $p(B)$, kde p je polynom, $p(1) = 1$. Takovou IAD metodu nazveme Algoritmus 1.

Známé výsledky

Stewart (94) - odhad faktoru konvergence pro téměř úplně rozložitelné matice, kvantitativní odhady při splnění 4 podmínek (matice i řešení) a $\epsilon \longrightarrow 0$.

Mandel, Sekerka (83) - Odhad asymptotického faktoru konvergence pro matici B s pozitivním řádkem.

Courtois, Semal (86) - Konvergence pro cyklickou iterační matici blokové Jacobiovy metody $M^{-1}W$, $I - B = M - W$, je-li prováděna agregace. Otázka četnosti agregačního kroku.

Marek, Mayer (98 a 03) - Odvození tvaru chybové matice. Poprvé uvedena tzv. rychlá konvergence.

Ipsen, Kirkland (04) - Speciální volba agregačních skupin. Studium konvergence na základě jiné formulace algoritmu a použití stochastického doplňku.

Lokální konvergence

Chyba se řídí vztahem

$$x^{k+1} - \hat{x} = J(x^k)(x^k - \hat{x}),$$

kde

$$J(x) = T(I - P(x)Z)^{-1}(I - P(x)),$$

kde Z je dáno spektrálním rozkladem matice B,

$$B = P + Z, \quad P^2 = P, \quad PZ = ZP = 0,$$

P je Perronova-Frobeniova projekce příslušná B.

Studiem $J(\hat{x})$ lze najít některé postačující podmínky pro lokální konvergenci.

Rekurentní formule pro $J(x)$, je-li $T = B$,

$$J(x) = B(I - P(x)) + BP(x)J(x).$$

Věta 1. Je-li $p(t) = \alpha t + (1-\alpha)$ a $\alpha \in (0, 1\rangle$, pak jediným pevným bodem Algoritmu 1 je přesné řešení $\hat{x}$.

Věta 2. Je-li $p(t) = \alpha t + (1-\alpha)$ a $\alpha \in (0, 1)$, pak Algoritmus 1 konverguje lokálně pro každou ireducibilní stochastickou matici B k přesnému řešení.

Věta 3. Je-li $p(t) = t$ a základní iterace je doplněna krokem

$$x^{k+1} := \alpha x^{k+1} + (1-\alpha)x^k,$$

$\alpha \in (0, 1)$, pak Algoritmus 1 konverguje lokálně pro každou ireduciblní stochastickou matici B k přesnému řešení $\hat{x}$.

Věta 4. Je-li $p(t) = t$ a diagonála matice B je kladná, pak Algoritmus 1 konverguje lokálně k přesnému řešení $\hat{x}$.

Věta 5.

Je-li $p(t) = t$ a matice B obsahuje alespoň jeden kladný řádek, pak Algoritmus 1 konverguje lokálně k přesnému řešení $\hat{x}$ a asymptotický faktor konvergence je

$$\sqrt{1-\beta},$$

kde β je minimální prvek tohoto řádku.

Je-li matice B kladná, pak asymptotický faktor konvergence je

$$1-\beta,$$

kde

$$\beta = \max\{||v||_1; B \geq ve^T, v \geq 0\}.$$

Divergence.

Obecně lze tedy ukázat, že

$$\rho(J(p,\hat{x})) < 1$$

pro $p(t) = \alpha t + (1-\alpha)$ a $\alpha \in (0,1)$, ale

$$\rho(J(p,\hat{x})) \le 1$$

pro $\alpha = 0$ nebo 1. Právě pro $\alpha = 1$ lze najít matici takovou, že Algoritmus 1 nekonverguje k řešení ani v lokálním smyslu.

Příklad 1.

$$B = \left(\begin{array}{c|cc} 1/2 & 0 & 1/2 \\ \hline 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{array}\right),$$

agregační skupiny $G_1 = \{1\}$ a $G_2 = \{2,3\}$. Polynom $p(t) = t$. Oscilace. Konvergence pouze pro množinu počátečních přiblížení míry nula.

Zobecnění takové situace.

Příklad 2.

$$B = \left(\begin{array}{ccc|cccc} \times & \dots & \times & \times & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \times & \dots & \times & \times & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & 0 & \dots & \times \\ \times & \dots & \times & \times & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & \times & 0 \end{array}\right).$$

Vysvětlení na základě formulace algoritmu pomocí stochastického doplňku.

Lemma 1. Pro aproximace x^k, $k = 1, 2, \dots$, získané Algoritmem 1 pro $p(t) = t$ platí

$$BP(x^k)x^{k+1} = x^{k+1}.$$

Vyberme některou aggregační skupinu G_i. Potom (po vhodné permutaci) máme B ve tvaru

$$B = \begin{pmatrix} B_{11} & B_{G_i}^C \\ B_{G_i}^R & B_{G_i} \end{pmatrix}.$$

Stejně rozdělíme $P(x^k)$ na blokově diagonální matici

$$P(x^k) = \begin{pmatrix} P(x^k)_1 & 0 \\ 0 & P(x^k)_{G_i} \end{pmatrix}.$$

Potom rovnice $BP(x^k)x^{k+1} = x^{k+1}$ lze vyjádřit ve tvaru $x^{k+1} = \rho_{k+1}\tilde{B}_{G_i}(x^k)e$, kde ρ_{k+1} je faktor zaručující $||x^k||_1 = 1$ a

$$\tilde{B}_{G_i}(x^k) = \begin{pmatrix} 0 & (I - B_{11}P(x^k)_1)^{-1}B_{G_i}^C \\ 0 & S_{G_i}(x^k) \end{pmatrix} \begin{pmatrix} 0 \\ P(x^k)_{G_i} \end{pmatrix},$$

kde S_{G_i} je stochastická matice,

$$S_{G_i}(x^k) = B_{G_i} + B_{G_i}^R P(x^k)_1 (I - B_{11}P(x^k)_1)^{-1} B_{G_i}^C.$$

Věta 6. Je-li matice $S_{G_i}(x)$ cyklická indexu m pro některý vektor x potom posloupnost $\{x^k\}_{k=0}^{\infty}$ získaná Algoritmem 1 konverguje k přesnému řešení $\hat{x}$ pouze v případě, že částečné součty části $x_{G_i}^0$ počátečníh vektoru x^0 odpovídající blokům cykličnosti matice $S_{G_i}(x)$ jsou stejné.

Z toho ihned dostaneme

Věta 7. Je-li matice $S_{G_i}(x)$ cyklická pro některý vektor x, pak možina počátečních přiblížení, pto která Algoritmus 1 konverguje k přesnému řešení, má míru nula v množině všech přípustných počátečních přiblížení.

Vyšší stupeň polynomu p

Příklad 3. Uvažujme primitivní stochastickou matici B_1

$$B_1 = \left(\begin{array}{cc|ccc} 0 & 0 & 0 & 1/2 & 0 \\ 1 & 1/2 & 0 & 1/2 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \end{array}\right).$$

Agregační skupiny jsou $G_1 = \{1, 2\}$ a $G_2 = \{3, 4, 5\}$. Potom máme

$$\hat{x}^1 = (1, 4, 2, 2, 2)^T/11.$$

Nechť $p_1(t) = t$ a $p_2(t) = t^2$. Spektrální poloměry matic $J(., \hat{x}^1)$ jsou

$$\rho(J(p_1, \hat{x}^1)) = 1 \qquad \text{a} \qquad \rho(J(p_2, \hat{x}^1)) = 1.1570 > 1.$$

Příklad 4.

$$B_2 = \left(\begin{array}{cc|ccc} 0 & 0 & 0 & 1/2 & 0 \\ 1 & 1/2 & 1/100 & 1/2 & 1/100 \\ \hline 0 & 0 & 0 & 0 & 99/100 \\ 0 & 0 & 99/100 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \end{array}\right).$$

Agregační skupiny jsou $G_1 = \{1, 2\}$ a $G_2 = \{3, 4, 5\}$. Potom pro stacionální rozdělení pravděpodobnisti $\hat{x}^2$ máme

$$\rho(J(p_1, \hat{x}^2)) = 0.9855 < 1$$

a

$$\rho(J(p_2, \hat{x}^2)) = 1.1271 > 1$$

kde opět $p_1(t) = t$ a $p_2(t) = t^2$.

Všimňeme si, že Algoritmus 1 s $p(t) = t^k$ konverguje loálně pro $k = 1$ ale diverguje (i v lokálním smyslu) pro $k = 2$. Tedy více kroků v základní iteraci poruší konvergenci.

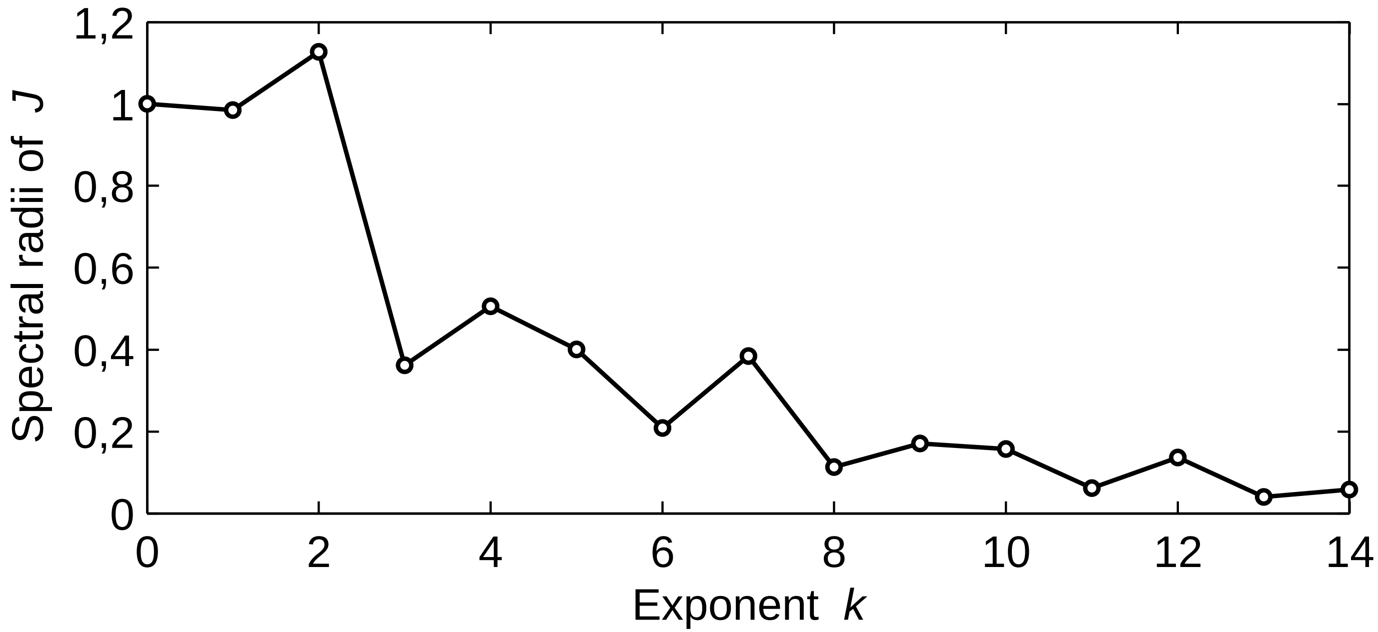

Figure 1: Spektrální poloměr matic $J(p_k, \hat{x})$ pro B_2 a $p_k(t) = t^k$.

Závěr a další otázky

- Lokální konvergence IAD metody není zaručena pro B ireducibilní a necyklickou. Tyto výsledky zatím platí pro $T = B^k$.

- Otázky.

  Jaká mocnina je dostatečná pro lokální konvergenci - na základě struktury nenulových prvků matice B?

  Oscilace v případě divergence?

  Je bloková Jacobiova metoda postačující pro lokální konvergenci?

  Je konvergence v lokálním smyslu vždy postačující pro konvergenci?

Poděkování.

Tato práce byla podpořena projektem Informační společnost č. 1ET400300415.

Literatura

- P. J. Courtois, P. Semal, *Block iterative algorithms for stochastic matrices*, 1986
- I. C. F. Ipsen, S. Kirkland, *Convergence analysis of an improved PageRank algorithm*, 2004
- I. Marek, P. Mayer, *Convergence theory of some classes of iterative aggregation - disaggregation methods for computing stationary probability vectors of stochastic matrices*, 2003
- I. Marek, I. Pultarová, *A note on local and global convergence analysis of iterative aggregation-disaggregation methods*, 2004
- I. Pultarová, *Local convergence analysis of IAD methods with polynomial correction based on sparsity of stochastic matrices*, 2005
- W. J. Stewart, *Introduction to the numerical solutions of Markov chains*, 1994