

以函数 f 收敛到唯一固定点 x_∞ , 且与初始信息无关. □

根据中值定理, 如果函数 $f: V \rightarrow V$ 可微, 那么对于任意 $x, y \in V$, 有公式(9)成立:

$$\|f(y) - f(x)\| \leq \|y - x\| \cdot \sup_{z \in [x, y]} \|f'(z)\| \tag{9}$$

其中, $[x, y] = \{cx + (1-c)y : c \in [0, 1]\}$.

因此, 如果 $\sup_{v \in V} \|f'(v)\| < 1$, 那么 f 是压缩函数, 收敛于固定点 $x_\infty \in V$, 且与初始信息无关.

更进一步地, 如果 $\sup_{v \in V} \|f'(v)\| < 1$, 则 BP 算法收敛, 且与初始消息无关. 因此, 关键问题是构造范数 $\|f'(v)\|$.

赋范空间 $(V, \|\cdot\|)$ 可诱导一个关于映射 $A: V \rightarrow V$ 的矩阵范数: $\|A\| = \sup_{\substack{v \in V \\ \|v\| \leq 1}} \|Av\|$.

$$l_1\text{-范数定义为 } \|A\|_1 = \max_{j \in \{1, 2, \dots, N\}} \sum_{i=1}^N |A_{ij}|, \quad l_\infty\text{-范数定义为 } \|A\|_\infty = \max_{i \in \{1, 2, \dots, N\}} \sum_{j=1}^N |A_{ij}|.$$

对于子向量空间 $V^{a \rightarrow i}$, 设其上的范数表示为 $\|\cdot\|_{a \rightarrow i}$. 向量空间 V 上的范数定义为

$$\|v\| = \sum_{a \rightarrow i} \|v^{a \rightarrow i}\|_{a \rightarrow i} \tag{10}$$

设 $\lambda \in V, f'(\lambda): V \rightarrow V$ 为 $f: V \rightarrow V$ 在 λ 处的导函数.

$$\text{在 } l_1\text{-范数下, 则 } \|f'(\lambda)\|_1 = \max_{b \rightarrow j} \sum_{a \rightarrow i} \|f'(\lambda)_{a \rightarrow i, b \rightarrow j}\|_{a \rightarrow i}^{b \rightarrow j}, \text{ 而 } \|f'(\lambda)_{a \rightarrow i, b \rightarrow j}\|_{a \rightarrow i}^{b \rightarrow j} = \sup_{\substack{v \in V^{b \rightarrow j} \\ \|v\|_{b \rightarrow j} \leq 1}} \|f'(\lambda)_{a \rightarrow i, b \rightarrow j} v\|_{a \rightarrow i}.$$

对于二元 $RB(k, n, \alpha, r_c, p)$ 模型实例的每一个约束 $a \in F$, 假设 $\forall i \in a, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_a} > 0$, 那么方程式(7)的导数为

$$\frac{\partial \lambda_{a \rightarrow i}^{(t+1)}(s_i)}{\partial \lambda_{b \rightarrow i}^{(t)}(s_j)} = 1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) \times \frac{\delta_{(s_i, s_j) \in Q_a} h^{a \setminus i}(s_j)}{\sum_{s'_j \in D} \delta_{(s_i, s'_j) \in Q_a} h^{a \setminus i}(s'_j)} \tag{11}$$

其中, 标识函数 $1_X(x) = \begin{cases} 1, & \text{if } x \in X \\ 0, & \text{o.w.} \end{cases}$. 因此,

$$\|f'(\lambda)\|_1 = \max_{b \rightarrow j} \sum_{a \rightarrow i} 1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) B_{a \rightarrow i, b \rightarrow j}(h^{a \setminus i}(\lambda)) \tag{12}$$

其中, $B_{a \rightarrow i, b \rightarrow j}(h^{a \setminus i}(\lambda)) = \left\| \frac{\delta_{(s_i, s_j) \in Q_a} h^{a \setminus i}(\lambda)}{\sum_{s'_j \in D} \delta_{(s_i, s'_j) \in Q_a} h^{a \setminus i}(\lambda)} \right\|_{a \rightarrow i}^{b \rightarrow j}$. 因此, 只要满足 $\sup_{\lambda \in V} \|f'(\lambda)\|_1 < 1$, BP 算法就收敛.

定义矩阵 A 中的元素为

$$A_{a \rightarrow i, b \rightarrow j} = \sup_{h^{a \setminus i} > 0} B_{a \rightarrow i, b \rightarrow j}(h^{a \setminus i}(\lambda)) \tag{13}$$

假设 $\|\cdot\|_{a \rightarrow i}$ 是 $V^{a \rightarrow i} = \mathbb{R}^{|D|}$ 上的 l_∞ 范数, 由代数知识, $A_{a \rightarrow i, b \rightarrow j}$ 可表示为如下形式^[29]:

$$A_{a \rightarrow i, b \rightarrow j} = N(\delta, i, j) = \sup_{\substack{s_i \neq s'_i \\ s_i, s'_i \in D}} \sup_{\substack{s_j \neq s'_j \\ s_j, s'_j \in D}} \frac{\sqrt{\delta_{(s_i, s_j) \in Q_a} \delta_{(s'_i, s'_j) \in Q_a} - \sqrt{\delta_{(s'_i, s_j) \in Q_a} \delta_{(s_i, s'_j) \in Q_a}}}}{\sqrt{\delta_{(s_i, s_j) \in Q_a} \delta_{(s'_i, s'_j) \in Q_a} + \sqrt{\delta_{(s'_i, s_j) \in Q_a} \delta_{(s_i, s'_j) \in Q_a}}} \tag{14}$$

根据函数的压缩映射原理及公式(9)、公式(12)~公式(14), 有如下推论:

推论 1. 设 $\forall i \in a, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_a} > 0$. 若 $\max_{b \rightarrow j} \sum_{a \rightarrow i} 1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) N(\delta, i, j) < 1$, 则 BP 算法收敛到唯一固定点, 且与初始信息无关.

定义矩阵 M 的元素为 $M_{a \rightarrow i, b \rightarrow j} = 1_{V(j) \setminus a}(b) 1_{V(a) \setminus i}(j) N(\delta, i, j)$, $\lambda_1, \dots, \lambda_n$ 是矩阵 M 的特征值, $\lambda(M) = \{\lambda_1, \dots, \lambda_n\}$ 为矩阵 M 的谱, $\rho(M) = \sup\{|\lambda_i| : \lambda_i \in \lambda(M)\}$ 为矩阵 M 的谱半径.

定理 3. 设函数 $f: V \rightarrow V$ 可微, 进一步假设 $|f'(x)| \leq M$, M 是一个不依赖于 x 的非负矩阵. 如果 $\rho(M) < 1$, 那么对于任意 $x \in V, f$ 收敛到唯一固定点 x_∞ . 也即 $x, f(x), f^2(x), \dots, x_\infty$ (注意, $|f'(x)| \leq M$ 表示由 $f'(x)$ 构成的雅可比矩阵中的每个元素的绝对值不大于 M 中对应元素的值).

证明:对于任意的 $x \in V'$, 假设 $n=1,2,\dots$, 我们有:

$$\|(f^n)'(x)\| = \left\| \prod_{i=1}^n f'(f^{i-1}(x)) \right\| = \left\| \prod_{i=1}^n f'(f^{i-1}(x)) \right\| \leq \left\| \prod_{i=1}^n f'(f^{i-1}(x)) \right\| \leq \left\| \prod_{i=1}^n M \right\| = \|M^n\| \quad (15)$$

由 Gelfand 引理^[35], $\lim_{n \rightarrow \infty} \|M^n\|^{1/n} = \rho(M)$. 我们选取 $\varepsilon > 0$, 使得 $\rho(M) + \varepsilon < 1$, 对于某个 $N \in \mathbb{N}$, $\|M^N\| = (\rho(M) + \varepsilon)^N < 1$. 因此, $\|(f^N)'(x)\| < 1$. 所以, $\|(f^N)(x)\|$ 是压缩函数.

由引理 1 可知, 函数 $f(x)$ 收敛到唯一固定点 x_∞ , 且与初始信息无关. □

因此, 根据定理 3 及公式(11), 有如下结论:

推论 2. 设 $\forall i \in a, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_a} > 0$. 如果矩阵 M 的谱半径 $\rho(M)$ 严格小于 1, 则 BP 算法收敛到唯一固定点, 且与初始信息无关.

对于一个具体的二元 $RB(k, n, \alpha, r_c, p)$ 模型的随机实例 F , 依据 $RB(k, n, \alpha, r_c, p)$ 模型实例的生成方法, 每个约束 $I \in F$ 有 $n^{2\alpha}p$ 个不协调赋值和 $n^{2\alpha}(1-p)$ 个协调赋值. 也即, 有 $n^{2\alpha}p$ 个 $\delta_{(s_i, s_j) \in Q_I} = 0$ 的赋值和 $n^{2\alpha}(1-p)$ 个 $\delta_{(s_i, s_j) \in Q_I} = 1$ 的赋值. 下面给出 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_I} > 0$ 高概率成立的 p 的取值界, 有如下引理:

引理 2. 设 $p = 1 - \frac{c}{n^\alpha}$, c 为常数. 则 $c = \alpha$ 是 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_I} > 0$ 高概率成立的临界点.

证明: 设约束 I 的一个赋值为 $(s_i, s_j) \in D \times D$, 所有的赋值构成赋值矩阵 S , 其元素为 $s_{ij} = (s_i, s_j)$. 位于矩阵 S 的不同行、不同列元素构成一个排列 $s_{1i_1}, s_{2i_2}, \dots, s_{d_n i_{d_n}}$, 这样的排列共有 $d_n!$ 个. 如果在 $n^{2\alpha}(1-p)$ 个协调赋值中至少含有一个排列 $s_{1i_1}, s_{2i_2}, \dots, s_{d_n i_{d_n}}$ 中的元素, 则都有 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_I} > 0$ 成立. 设随机变量 x 表示在 $n^{2\alpha}(1-p)$ 个协调赋值中含有元素 $s_{1i_1}, s_{2i_2}, \dots, s_{d_n i_{d_n}}$ 的事件发生的数目, 则

$$\begin{aligned} \lim_{n \rightarrow \infty} E(x) &= \lim_{n \rightarrow \infty} \frac{n^\alpha! \binom{n^{2\alpha} - n^\alpha}{(1-p)n^{2\alpha} - n^\alpha}}{\binom{n^{2\alpha}}{(1-p)n^{2\alpha}}} \\ &= \lim_{n \rightarrow \infty} \frac{n^\alpha! ((1-p)n^{2\alpha})((1-p)n^{2\alpha} - 1) \dots ((1-p)n^{2\alpha} - n^\alpha + 1)}{n^{2\alpha} (n^{2\alpha} - 1) \dots (n^{2\alpha} - n^\alpha + 1)} \\ &= \lim_{n \rightarrow \infty} \frac{n^\alpha! ((1-p)) \left((1-p) - \frac{1}{n^{2\alpha}} \right) \dots \left((1-p) - \frac{n^\alpha - 1}{n^{2\alpha}} \right)}{\left(1 - \frac{1}{n^{2\alpha}} \right) \left(1 - \frac{2}{n^{2\alpha}} \right) \dots \left(1 - \frac{n^\alpha - 1}{n^{2\alpha}} \right)} \\ &= \lim_{n \rightarrow \infty} n^\alpha! (1-p)^{n^\alpha} \\ &\approx \lim_{n \rightarrow \infty} \frac{n^{\alpha \left(\frac{n^\alpha + 1}{2} \right)}}{e^{n^\alpha}} \cdot (1-p)^{n^\alpha + 1} \\ &\leq \lim_{n \rightarrow \infty} \frac{n^{\alpha(n^\alpha + 1)}}{e^{n^\alpha}} \cdot (1-p)^{n^\alpha + 1} \quad (\text{由于 } \alpha > \frac{1}{2}) \end{aligned} \quad (16)$$

因此, 当 $c > \alpha$ 时, $\lim_{n \rightarrow \infty} E(x) = 0$. 由一阶矩方法可知, $\Pr(x \geq 1) = 0$.

不妨设 $v_1 = (s_i, s_j), v_2 = (s_k, s_l)$ 是约束 I 的两个赋值, $\Pr(\langle v_1, v_2 \rangle)$ 表示赋值 v_1 和 v_2 同时满足约束 I 的概率.

- 如果赋值 v_1 和 v_2 相同, 则 $\Pr(\langle v_1, v_2 \rangle) = \frac{\binom{d_n^k - 1}{|Q_I|}}{\binom{d_n^k}{|Q_I|}} = 1 - p$;

- 如果赋值 v_1 和 v_2 不同, 则 $\Pr(\langle v_1, v_2 \rangle) = \binom{d_n^k - 2}{|Q_i|} / \binom{d_n^k}{|Q_i|} = (1-p)^2 + O\left(\frac{1}{n^{2\alpha}}\right)$.

设 x_i, x_j 是 $n^{2\alpha}(1-p)$ 个协调赋值中分别含有元素 $s_{1i_1}, s_{2i_2}, \dots, s_{d_n i_{d_n}}$ 和 $s_{1j_1}, s_{2j_2}, \dots, s_{d_n j_{d_n}}$ 的两组赋值, 所以,

$$\begin{aligned} E(x^2) &= \sum_{i=j} E(x_i, x_j) + \sum_{i \neq j} E(x_i, x_j) \\ &= E(x) + ((n^\alpha)!^2 - n^\alpha!) \left[(1-p)^2 + O\left(\frac{1}{n^{2\alpha}}\right) \right]^{n^\alpha} \\ &= E(x) + ((n^\alpha)!^2 - n^\alpha!) (1-p)^{2n^\alpha} + O\left(\frac{1}{n^{2\alpha}}\right) \end{aligned} \quad (17)$$

进而有:

$$\lim_{n \rightarrow \infty} \frac{E(x^2)}{E^2(x)} = \lim_{n \rightarrow \infty} \frac{E(x) + ((n^\alpha)!^2 - n^\alpha!) (1-p)^{2n^\alpha}}{E^2(x)} = \lim_{n \rightarrow \infty} \left(\frac{1}{E(x)} + 1 - \frac{1}{n^\alpha!} \right) \quad (18)$$

当 $c < \alpha$ 时, $\lim_{n \rightarrow \infty} E(x) = \infty$. 因此,

$$\lim_{n \rightarrow \infty} \frac{E(x^2)}{E^2(x)} \leq \lim_{n \rightarrow \infty} \left[\frac{1}{E(x)} + 1 - \frac{1}{n^\alpha!} \right] = 1 + o(1) \quad (19)$$

由二阶矩方法, $\Pr(x=0)=0$.

所以, $c=\alpha$ 是 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_i} > 0$ 高概率成立的临界点. \square

引理 2 表明, 当 $p < 1 - en^{-\alpha}$ 时, 则 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_i} > 0$ 高概率成立. 注意到, 在 $RB(k, n, \alpha, r_c, p)$ 模型中的可满足相变点为 $p_c = 1 - e^{-\frac{\alpha}{c}}$, 而 $1 - e^{-\frac{\alpha}{c}} < 1 - en^{-\alpha}$. 根据定理 1 和引理 2, 在 $RB(k, n, \alpha, r_c, p)$ 模型的可满足区域, 总有 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_i} > 0$ 高概率成立. 因此, 推论 2 可用作为判定 BP 算法在 $RB(k, n, \alpha, r_c, p)$ 模型的可满足实例上是否高概率收敛的充分条件. 我们有如下推论:

推论 3. 设 $p_c = 1 - e^{-\frac{\alpha}{c}}$, 其中, $\alpha > \frac{1}{k}, r_c > 0$ 均为常数, 且满足 $ke^{-\frac{\alpha}{c}} \geq 1$. 当 $p < p_c$ 时, 对于二元 $RB(k, n, \alpha, r_c, p)$ 模型实例 F , 如果矩阵 M 的谱半径 $\rho(M)$ 严格小于 1, 则 BP 算法高概率收敛, 且与初始信息无关 (注意, 这里的高概率指引理 2 中的高概率).

注意到, 矩阵 M 中的元素由 $N(\delta, i, j)$ 决定, 为此给出 $N(\delta, i, j)=0$ 高概率成立的条件.

引理 3. 设 $p=n^{-c}, c$ 为常数. 当 $c > 2\alpha$ 时, $N(\delta, i, j)=0$ 高概率成立.

证明: 设 x 表示约束 I 不协调赋值的数目的随机变量. 由马尔可夫不等式得: $\Pr\{x \geq 1\} \leq E[x]$. 而

$$E[x] = n^{2\alpha} p = n^{2\alpha} n^{-c} = n^{2\alpha - c} \quad (20)$$

当 $c > 2\alpha$ 时, $\lim_{n \rightarrow \infty} E(x) = 0$. 所以, $\Pr(x \geq 1) = 0$, 进一步有 $\Pr(x < 1) = 1$. 因此, 根据公式 (14) 有 $\Pr\{N(\delta, i, j)=0\} = 1$. \square

在矩阵分析中, 我们知道: 关于非负方阵 M 的谱半径 $\rho(M)$, 如果存在一个矩阵范数 $\|\cdot\|$ 使得 $\|M\| < 1$, 那么 $\lim_{n \rightarrow \infty} M^n = 0$; 如果 $\lim_{n \rightarrow \infty} M^n = 0$ 当且仅当 $\rho(M) < 1$. 下面给出 BP 算法在二元 $RB(k, n, \alpha, r_c, p)$ 模型上高概率收敛的条件.

定理 4. 设 $p=n^{-c}, c$ 为常数. 当 $c > 2\alpha$ 时, BP 算法在二元 $RB(k, n, \alpha, r_c, p)$ 模型实例集上高概率收敛.

证明: 由引理 2 可知: 当 $p_1 < 1 - en^{-\alpha}$ 时, 高概率地有 $\forall i \in I, \forall s_i \in D, \exists s_j \in D: \delta_{(s_i, s_j) \in Q_i} > 0$ 成立. 而引理 3 中, 当 $p_2 < n^{-2\alpha}$ 时, 高概率地有 $\|M\|_\infty = \max_{i \in \{1, 2, \dots, N\}} \sum_{j=1}^N |M_{ij}| < 1$. 所以, 高概率地有 $\rho(M) < 1$ 成立. 容易看出, $p_2 \leq p_1$. 因此, 当 $p < n^{-2\alpha}$ 时,

BP 算法在二元 $RB(k, n, \alpha, r_c, p)$ 模型实例集上高概率收敛. 至此, 定理 2 得证. \square

定理 2 表明, 如果 $p \in (0, n^{-2\alpha})$, 则基于变量熵的 BP 算法在二元 $RB(k, n, \alpha, r_c, p)$ 模型产生的随机实例集上高概率收敛.

上述有关证明过程及结论容易扩展到多元 RB 模型实例上.

3 数值实验及分析

为了对本文定理 2 中的结论作出实验验证,数值实验中选用 RB 模型实例.在 RB 模型中,我们选择了两组不同的参数,分别为 $\alpha=0.8, r_c=3$ 和 $\alpha=0.6, r_c=1$,用来生成随机 CSP 实例.对于每组参数取 5 个不同规模 n 的 CSP 随机实例,见表 1 和表 2(表 1 中,取 $\alpha=0.8, r_c=3$,对于不同的问题规模 n ,相应的 d_n 和 m 的值;表 2 中,取 $\alpha=0.6, r_c=1$,对于不同的问题规模 n ,相应的 d_n 和 m 的值).在算法中,取 $t_{\max}=10^3, \epsilon=10^{-4}$.实验运行的软、硬件环境为: Intel Core i7-2600 CPU,主频 3.40GHZ,内存 4.00GB,操作系统 WIN7,采用 VC++6.0 集成开发工具,使用 C 和 MATLAB 语言编程.

Table 1 First data of numerical experiment

表 1 数值实验中的第 1 组数据

n	α	d_n	r_c	m
20	0.8	11	3	180
40	0.8	19	3	443
60	0.8	26	3	737
80	0.8	33	3	1 052
100	0.8	40	3	1 382

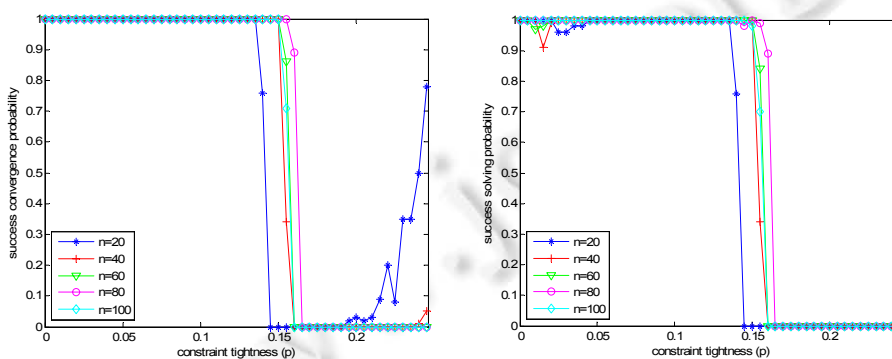
Table 2 Second data of numerical experiment

表 2 数值实验中的第 2 组数据

n	α	d_n	r_c	m
100	0.6	16	1	461
150	0.6	20	1	752
200	0.6	24	1	1 060
250	0.6	27	1	1 380
300	0.6	31	1	1 711

在 $RB(k, n, \alpha, r_c, p)$ 模型中,可满足相变点 $p_c = 1 - e^{-\frac{\alpha}{r_c}}$. 对于第 1 组数据, $p_c=0.234$. 对于第 2 组数据, $p_c=0.451$. 数值实验中,每个数据点由 100 个 $RB(k, n, \alpha, r_c, p)$ 模型的随机实例构成. 第 1 组数据和第 2 组数据的运行结果分别如图 2 和图 3 所示.

图 2(a)是基于变量熵的 BP 算法在第 1 组数据上成功收敛的概率随着参数 p 变化的情形. 当 $p < 0.145$ 时,算法高概率收敛. 当 $p > 0.145$ 时,算法从收敛突变到几乎不收敛. 并且在 $p > 0.195$ 时,算法成功收敛的概率逐渐增加. 我们发现:在 $0.160 < p < 0.195$ 区间内,算法几乎高概率不收敛,即算法失效. 事实上,此区间恰好为 $RB(k, n, \alpha, r_c, p)$ 模型的难解区间,实例的求解难度随着变量规模的增加成指数形式增长.



(a) BP 算法成功收敛的运行结果

(b) BP 算法有效求解的运行结果

Fig.2 Results of the first data

图 2 第 1 组数据的运行结果

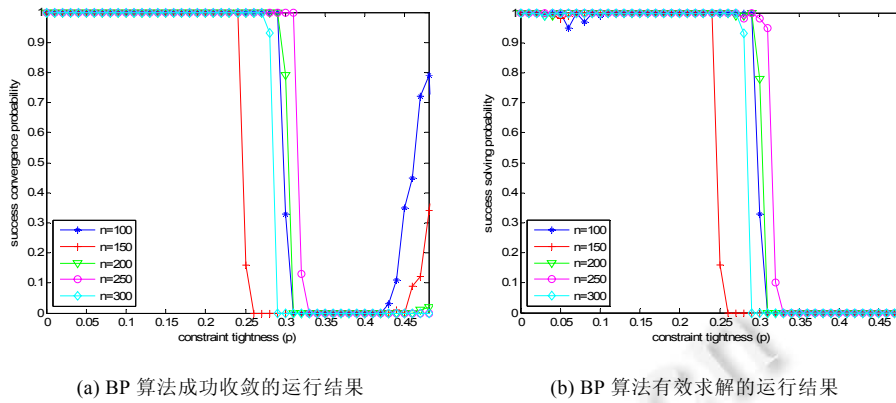


Fig.3 Results of the second data

图3 第2组数据的运行结果

图 2(b)是基于变量熵的 BP 算法在第 1 组数据上能够有效求解实例的概率随着参数 p 变化的情形.当 $p < 0.145$ 时,算法几乎能够高概率地给出实例的解.当 $p > 0.145$ 时,算法不能给出实例的解:一方面,算法本身不收敛;另一方面,算法收敛但不能正确给出实例的解.

对比图 2(a)与图 2(b),我们发现,当 $p > 0.195$ 时,对于部分实例该算法收敛,但未能正确给出实例的解.

对于第 2 组数据,基于变量熵的 BP 算法的运行结果如图 3 所示.当 $p < 0.245$ 时,算法高概率地收敛.当 $p > 0.245$ 时,算法从收敛突变到几乎不收敛.并且在 $p > 0.420$ 时,算法成功收敛的概率逐渐增加.在 $0.325 < p < 0.420$ 区间内,算法几乎高概率不收敛,即算法失效,如图 3(a)所示.

图 3(b)是算法在第 2 组数据上能够解的运行结果.当 $p < 0.245$ 时,算法几乎能够高概率地给出实例的解.当 $p > 0.245$ 时,算法不能给出实例的解.我们发现,当 $p > 0.420$ 时,算法在部分实例收敛,但未能给出实例的解.

为了进一步分析和验证本文结论,并与其他结论进行比较,我们选取 RB 模型中的参数见表 3.对于表 3 中的第 1 组数据 $(n, \alpha, d_n, r_c, m) = (20, 0.4, 3, 4, 240)$,根据定理 2 可计算出算法收敛的条件 p 的取值范围为 $(0, 0.091)$.同理,对于第 2 组数据 $(n, \alpha, d_n, r_c, m) = (40, 0.4, 4, 7, 1033)$ 和第 3 组数据 $(n, \alpha, d_n, r_c, m) = (60, 0.4, 5, 10, 2457)$, p 的取值范围分别为 $(0, 0.052)$ 和 $(0, 0.038)$,这 3 组数据对应实例的相变点 p_c 分别为 0.095, 0.056, 0.039,具体参数值见表 4.不难看出,在定理 2 的收敛性判定条件中, p 的取值上界几乎接近相变点 p_c .进一步说明,对于接近相变点的复杂因子图结构的二元 RB 模型实例,定理 2 的判定条件几乎有效.针对表 3 中的第 1 组数据,选取 p 为 0.078,由 $RB(k, n, \alpha, r_c, p)$ 模型产生 1 000 个二元 CSP 实例进行数值实验,本文能够有效判定的概率为 0.87,文献[29]中能够有效判定的概率为 0.91;同样,对于表 3 中的第 2 组数据和第 3 组数据,本文能够有效判定的概率为 1.00 和 0.93,而文献[29]中分别为 0.84 和 0.79.表 4 中给出了本文结论与文献[29]中结论的对比结果(每组参数生成 1 000 个 RB 模型实例).

Table 3 Values of parameter in RB model

表 3 RB 模型中的相关参数值

n	α	d_n	r_c	m
20	0.4	3	4	240
40	0.4	4	7	1 033
60	0.4	5	10	2 457

Table 4 Comparison of relevant experimental results

表 4 相关实验结果比较

$(0, n^{-2\alpha})$	p_c	p	本文结果	文献[29]结果
$(0, 0.091)$	0.095	0.078	0.87	0.91
$(0, 0.052)$	0.056	0.049	1.00	0.84
$(0, 0.038)$	0.039	0.037	0.93	0.79

当问题规模 n 增大时,理论收敛区间逐渐减小.原因在于, $RB(k,n,\alpha,r_c,p)$ 模型是一个具有增长定义域的随机 CSP 实例产生模型,不协调赋值的数目与参数 p 及问题规模 n 有关.

用难解实例测试算法性能及相关结论是非常必要的,我们使用国际上常用的由 Lecoutre 提供的 benchmark 案例进行测试^[34].数值实验中选取了两组实例集 frb30-15 和 frb40-19 进行测试,结果表明,基于变量熵的置信传播算法在这些实例上基本不收敛.事实上,在这些实例集上,定理 2 中的条件已经不能满足.也即,定理 2 中判定算法收敛的条件失效.当然,不能仅凭某些局部的实验数据就断定某种方法的优劣,每种方法都有其优越性和局限性.

4 结束语

本文分析了基于变量熵的置信传播算法在二元 $RB(k,n,\alpha,r_c,p)$ 模型随机实例集上的收敛性,利用置信传播算法的信息更新函数在实向量空间中的压缩性原理,及更新函数的系数矩阵的谱半径等理论,证明了当 $p \in (0, n^{-2\alpha})$ 时,该算法在二元 $RB(k,n,\alpha,r_c,p)$ 模型产生的随机 CSP 实例集上高概率收敛,给出了置信传播算法能够有效求解二元 $RB(k,n,\alpha,r_c,p)$ 模型实例的理论解释.我们希望这些方法的引入可以提供一种新的思路,用来分析其他信息传播算法的收敛性及性能.进一步工作是:(1) 扩展基于变量熵的 BP 算法在 $RB(k,n,\alpha,r_c,p)$ 模型实例集上的收敛区间;(2) 分析基于变量熵的 BP 算法在 $RB(k,n,\alpha,r_c,p)$ 模型实例集上的有效性,也即,算法收敛后能否给出问题的有效解.

References:

- [1] Martin D, Alan F, Michael M. A probabilistic analysis of randomly generated binary constraint satisfaction. *Theory Computer Science*, 2003,290(3):1815–1828. [doi: 10.1016/S0304-3975(02)00317-1]
- [2] Creignou N, Auade H. The SAT-UNSAT transition for random constraint satisfaction problems. *Discrete Mathematics*, 2009,309(8): 2085–2099. [doi: 10.1016/j.disc.2008.04.025]
- [3] Martin OC, Monasson R, Zecchina R. Statistical mechanics methods and phase transitions in optimization. *Theory Computer Science*, 2001,265(1-2):3–67. [doi: 10.1016/S0304-3975(01)00149-9]
- [4] Tsang E. A glimpse of constraint satisfaction. *Artificial Intelligence Review*, 1999,13(3):215–277. [doi: 10.1023/A:1006558104682]
- [5] Xu K, Li W. The SAT phase transition. *Science in China, Series E*, 1999,42(5):494–501. [doi: 10.1007/BF02917402]
- [6] Fan Y, Shen J. On the phase transitions of random k -constraint satisfaction problems. *Artificial Intelligence*, 2011,175(3-4): 914–927. [doi: 10.1016/j.artint.2010.11.004]
- [7] Achlioptas D, Naro A, Peres Y. Rigorous location of phase transitions in hard optimization problems. *Nature*, 2005,435(7043): 759–764. [doi: 10.1038/nature03602]
- [8] Mertens S, Mézard M, Zecchina R. Threshold values of random k -SAT from the cavity method. *Random Structure Algorithms*, 2006,28(3):340–373. [doi:10.1002/rsa.20090]
- [9] Moskewicz MW, Madigan CF, Zhao Y. Chaff: Engineering an efficient SAT solver. In: *Proc. of the 38th Annual Design Automation Conf.* New York: ACM Press, 2001. 530–535. [doi: 10.1145/378239.379017]
- [10] Aurell E, Gordon U, Kirkpatrick S. Comparing beliefs, surveys, and random walks. *Advances in Neural Information Processing Systems*, 2004,17(1):1–8.
- [11] Mezard M, Parisi G. The cavity method at zero temperature. *Journal of Statistical Physics*, 2003,111(1-2):1–34. [doi: 10.1023/A:102221005097]
- [12] Mezard M, Zecchina R. Random k -satisfiability problem: From an analytic solution to an efficient algorithm. *Physical Review E*, 2002,66(5):No.056126. [doi:10.1103/PhysRevE.66.056126]
- [13] Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations. *Journal of the ACM*, 2007,54(4): 1089–1098. [doi:10.1145/1255443.1255445]
- [14] Braunstein A, Mezard M, Zecchina R. Survey propagation: An algorithm for satisfiability. *Random Structures and Algorithms*, 2005,27(2):201–226. [doi:10.1002/rsa.20057]
- [15] Braunstein A, Zecchina R. Survey and belief propagation on random k -SAT. *Lecture Notes in Computer Science*, 2004,2919(1): 519–528. [doi:10.1007/978-3-540-24605-3_38]

- [16] Yedidia JS, Freeman WT, Weiss Y. Understanding belief propagation and its generalizations. *Artificial Intelligence*, 2003,8(1): 239–269.
- [17] Mezard M, Parisi G, Zecchina R. Analytic and algorithmic solution of random satisfiability problems. *Science*, 2002,297(5582): 812–815. [doi: 10.1126/science.1073287]
- [18] Xu K, Li W. Exact phase transitions in random constraint satisfaction problems. *Journal of Artificial Intelligence Research*, 2000, 12(1):93–103.
- [19] Liu T, Lin XX, Wang CY, Su KL, Xu K. Large hinge width on sparse random hyper graphs. In: *Proc. of the 22nd Int'l Joint Conf. on Artificial Intelligence*. Barcelona: AAAI, 2011. 611–616. [doi: 10.5591/978-1-57735-516-8/IJCAI11-109]
- [20] Xu K, Li W. Many hard examples in exact phase transitions. *Theory Computer Science*, 2006,355(1):291–302. [doi: 10.1016/j.tcs.2006.01.001]
- [21] Xu K, Boussemart F, Hemery F, Lecoutre C. Random constraint satisfaction: Easy generation of hard instances. *Artificial Intelligence*, 2007,171(1):514–534. [doi: 10.1016/j.artint.2007.04.001]
- [22] Zhao CY, Zheng ZM. A belief propagation algorithm based on variable entropy for constraint satisfaction problems. *Science China Information Sciences*, 2012,42(9):1170–1180 (in Chinese).
- [23] Kschischang F, Frey B, Loeliger H. Factor graph and the sum-product algorithm. *Information Theory*, 2001,47(1):498–519. [doi: 10.1109/18.910572]
- [24] Weiss Y. Correctness of local probability propagation in graphical models with loops. *Neural Computation*, 2000,12(1):1–41. [doi: 10.1162/089976600300015880]
- [25] Weiss Y, Freeman WT. Correctness of belief propagation in Gaussian graphical models of arbitrary topology. *Neural Computation*, 2001,13(10):2173–2200. [doi: 10.1162/089976601750541769]
- [26] Tatikonda S, Jordan MI. Loopy belief propagation and Gibbs measure. In: *Proc. of the 18th Annual Conf. on Uncertainty in Artificial Intelligence*. 2002. 493–500.
- [27] Heskes T. On the uniqueness of loopy belief propagation fixed points. *Neural Computation*, 2004,16(11):2379–2413. [doi: 10.1162/0899766041941943]
- [28] Ihler AT, Fisher JW, Willsky AS. Loopy belief propagation: Convergence and effects of message errors. *Machine Learning Research*, 2005,6(1):905–936.
- [29] Mooij JM, Kappen HJ. Sufficient conditions for convergence of the sum-product algorithm. *IEEE Trans. on Information Theory*, 2007,53(12):4422–4437. [doi: 10.1109/TIT.2007.909166]
- [30] Shi XQ, Schonfeld D, Tuninetti D. Message error analysis of loopy belief propagation for the sum-product algorithm. *Computer and Information Science*, 2010,1009(2):1–30. [doi: 10.1109/ICASSP.2010.5495075]
- [31] Gamarnik D, Shah D, Wei Y. Belief propagation for min-cost network flow: Convergence and correctness. *Operations Research*, 2012,60(2):410–428. [doi: 10.1287/opre.1110.1025]
- [32] Brunsch T, Cornelissen K, Manthey B, Roglin H. Smoothed analysis of BP for minimum cost flow and matching. *Journal of Graph Algorithms and Applications*, 2013,17(6):647–670. [doi: 10.7155/jgaa.00310]
- [33] Norshams N, Wainwright MJ. Stochastic belief propagation: A low complexity alternative to the sum product algorithm. *IEEE Trans. on Information Theory*, 2013,59(4):1981–2000. [doi: 10.1109/TIT.2012.2231464]
- [34] Lecoute. <http://www.cril.univ-artois.fr/~lecoute/benchmarks.html>
- [35] Dieudonné J. *Foundations of modern analysis*. Vol.10-I, New York: Academic Press, 1969.

附中文参考文献:

- [22] 赵春艳,郑志明.一种基于变量熵求解约束满足问题的置信传播算法. *中国科学:信息科学*,2012,42(9):1170–1180.



王晓峰(1980—),男,甘肃会宁人,博士,讲师,CCF 专业会员,主要研究领域为人工智能,机器学习,可计算性,计算复杂性.



许道云(1959—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为人工智能,可计算性,计算复杂性.