doi:10.16055/j.issn.1672-058X.2015.0002.010

一种改进的遗传退火算法以及收敛性分析*

高 发 玲
(青岛理工大学 琴岛学院 ,山东 青岛 266100)

摘　要:针对一种约束条件既有 0-1 变量又有整数变量的非线性混合整数规划模型,给出一种改进的遗传退火算法求解,并建立对应的 Markov 链且理论证明其收敛性.

关键词:遗传退火;Markov 链;遍历;收敛性

中图分类号:O221.2　　**文献标识码**:A　　**文章编号**:1672-058X(2015)02-0049-05

0　引　言

针对一种约束条件既有 0-1 变量又有整数变量的非线性混合整数规划模型,数学模型

$$\min f(X,Y)$$
$$\text{s.t}\begin{cases} g_i(X,Y)\geqslant 0, & (i=1,2,\cdots M)\\ h_j(X,Y)=0, & (j=M+1,M+2,\cdots N)\\ X\in \mathbf{Z}^l, Y\in\{0,1\}\end{cases} \tag{1}$$

提出一种改进的遗传退火算法并对其算法对应 Markov 链进行收敛性分析.

1　算法设计

改进的遗传退火算法分为 5 步:编码方案的设计、解除约束与适应度函数的选定、交叉操作设计、变异操作设计以及终止条件的确定.

1.1　编码方案的设计

对于 0-1 决策变量的 Y,采用二进制编码;对于整数变量 X,采用浮点编码.

1.2　解除约束与适应函数的选定

(1) 约束优化问题处理.采用惩罚策略技术将有约束的优化模型转化为无约束优化问题,

$$\text{eval}(X,Y)=f(X,Y)+P(X,Y) \tag{2}$$

其中 $P(X,Y)=\frac{1}{2a}\sum_{i=1}^{N}d_i^2(X,Y)$[1] ,$a$ 是可变惩罚因子,而

$$d_i(X,Y)=\begin{cases}\max\{0,g_i(X,Y)\}, & \text{if}\quad (1\leqslant i\leqslant M)\\ |h_i(X,Y)|, & \text{if}\quad (M+1\leqslant i\leqslant N)\end{cases}$$

此时采用的可变惩罚因子 $a=t_k$,其中 t_k 为第 k 代退火温度.

收稿日期:2014-06-15;**修回日期**:2014-09-10.

作者简介:高发玲(1982-)女,山东青岛人,讲师,硕士,从事网络优化设计研究.

(2) 适应度函数的选定.

$$F_i(t_{k+1}) = \exp\left\{\frac{F_{\min} - F_i(t_k)}{t_k}\right\} \tag{3}$$

1.3 交叉操作设计

由于编码方法采用的是上述混合并行方案,因此在交叉操作中也采用相应的方式,对 0-1 决策变量的 Y 进行单点交叉,对整数变量 X 进行混合交叉.

1.4 变异操作设计

对模型中的两种决策变量采用不同的变异方式,对 0-1 决策变量的 Y 采用基本位变异,对整数变量 X 采用均匀变异.

1.5 终止条件的确定

计算在每一温度下每一代群体染色体的最大适应值和最小适应值,当最大适应值与最小适应值之间的差小于 ε(ε 为参数)时,停止此温度下的种群迭代;设退火温度为 t_k,若满足 $t_k \leqslant \eta$(η 为很小的参数)时(计算迭代到规定的退火温度),则停止计算.

2 算法收敛性分析

2.1 算法对应 Markov 链的建立

离散组合优化问题 $\min f(b_i)$,其中 b_i 为所有变量对应的一个状态,设状态集 $\Omega = \{b_i, i=1,2,\cdots,n\}$,$|\Omega| = n$.

设变量的个数为 k(其中有 l 个 0-1 变量;有 $k-l$ 个整数变量),称 $b_i = a_1^i a_2^i \cdots a_k^i$ 为一个个体或染色体,这里的 $a_j(j=1,2,\cdots,l)$ 为 0 或 1,$a_j(j=l+1,\cdots,k)$ 为整数.

设种群集 Q 中每一种群包含 m 个个体,任一点集 $\boldsymbol{B} = \{b_1, b_2, \cdots b_m\}$ 称为一个种群,则种群总数 $|Q| = C(n,m)$,相当于从 n 个不同个体中可选取 m 个的方案总数,设 $Q = \{\boldsymbol{B}^{(1)}, \boldsymbol{B}^{(2)}, \cdots, \boldsymbol{B}^{|Q|}\}$.

定义 1 矩阵 $\boldsymbol{C} = (c_{B^{(i)}B^{(j)}})$,$\boldsymbol{M} = (m_{B^{(i)}B^{(j)}})$($\boldsymbol{B}^{(i)}, \boldsymbol{B}^{(j)} = 1,2,\cdots|Q|$),若 $c_{B^{(i)}B^{(j)}}$ 是由 $\boldsymbol{B}^{(i)}$ 以交叉概率 $P_c \in [0,1]$ 杂交到 $\boldsymbol{B}^{(j)}$ 对应的一步转移概率,$m_{B^{(i)}B^{(j)}}$ 是由 $\boldsymbol{B}^{(i)}$ 以变异概率 $P_m \in (0,1)$ 变异到 $\boldsymbol{B}^{(j)}$ 对应的一步转移概率,则称 $\boldsymbol{C}$ 和 $\boldsymbol{M}$ 分别为交叉和变异对应的一步转移概率矩阵.

引理 1[2] 若 $\boldsymbol{C}$ 和 $\boldsymbol{M}$ 分别是以交叉概率 $P_c \in [0,1]$ 进行交叉和以变异概率 $P_m \in (0,1)$ 进行变异对应的一步转移概率矩阵,则 $\boldsymbol{C}$ 和 $\boldsymbol{M}$ 均为随机阵.

定义 2 若种群 $\boldsymbol{B}^{(i)}$ 和 $\boldsymbol{B}^{(j)}$ 中仅有某一个个体不同,设 $b_i \in \boldsymbol{B}^{(i)} \neq b_j \in \boldsymbol{B}^{(j)}$ 且满足 $b_j \in \mathbf{N}(b_i)$,则称种群 $B^{(i)}$ 属于种群 $B^{(j)}$ 的邻域 $N(B^{(j)})$.

定义 3 SA 状态产生函数使种群 $\boldsymbol{B}^{(i)}$ 转移到 $\boldsymbol{B}^{(j)}$ 的一步转移概率为

$$g_{B^{(i)}B^{(j)}}(T_h) = \begin{cases} e^{-f(B^{(i)})/T_h} / [\sum\limits_{B^{(j)} \in \mathbf{N}(B^{(i)})} e^{-f(B^{(j)})/T_h}], & \boldsymbol{B}^{(j)} \in \mathbf{N}(\boldsymbol{B}^{(i)}) \\ 0, & \boldsymbol{B}^{(j)} \notin \mathbf{N}(\boldsymbol{B}^{(i)}) \end{cases} \tag{4}$$

其中

$$f(\boldsymbol{B}^{(i)}) = \sum_{b_i \in \boldsymbol{B}^{(i)}} f(b_i) \tag{5}$$

定义 4 对种群中个体作模拟退火操作所引起的种群的一步转移概率:

$$a_{\boldsymbol{B}^{(i)}\boldsymbol{B}^{(j)}}(T_h) = \begin{cases} g_{B^{(i)}B^{(j)}} \min\{1, \exp(-(f(\boldsymbol{B}^{(i)}) - f(\boldsymbol{B}^{(j)}))/T_h\}, & \boldsymbol{B}^{(j)} \in \mathbf{N}(\boldsymbol{B}^{(i)}) \\ 0, & \boldsymbol{B}^{(j)} \notin \mathbf{N}(\boldsymbol{B}^{(i)}) \& \boldsymbol{B}^{(i)} \neq \boldsymbol{B}^{(i)} \\ 1 - \sum\limits_{\boldsymbol{B}^{(j)} \in \mathbf{N}(\boldsymbol{B}^{(i)})} a_{\boldsymbol{B}^{(i)}\boldsymbol{B}^{(j)}}(T_h), & \boldsymbol{B}^{(i)} = \boldsymbol{B}^{(j)} \end{cases} \tag{6}$$

则在温度 T 下对所有个体均作模拟退火操作得到的种群转移概率为

$$A(T) = (a_{ij}(T))i, \qquad j = 1,2,\cdots|Q| \tag{7}$$

此时 $A(T)$ 阵是随机阵.

因此,此种改进的遗传退火算法在每一温度下的状态转移矩阵为

$$P(T)=\mathrm{CMA}(T)(8)$$

从而得到所建立的遗传退火算法对应的 Markov 链可表示如下:

$$V(T_{k+1})=V(T_k)P(T_k)\qquad(k=1,2,\cdots,n)\tag{9}$$

2.2　收敛性分析

定义 5[2]　给定随机阵,其遍历系数定义为

$$\tau(P)=\frac{1}{2}\max_{i,j}\sum_{k=1}^{n}|p_{ik}-p_{jk}|=1-\min_{i,j}\sum_{k=1}^{n}\min(p_{ik},p_{jk})\tag{10}$$

定义 6　给定图 G 的中心 I_c,若 Q 中任意节点到达 I_c 最多经过 $r=\min\limits_{i\in(Q-Q_m)}\{\max\limits_{j\in\mathbf{Q}}d(i,j)\}$ 步,则称 r 为图 G 的半径.其中 $d(i,j)$ 为种群 $B^{(i)}$ 到达 $B^{(j)}$ 的最短路径

$$Q_m=\{B^{(i)}\in\wedge\ |f(B^{(j)})\leqslant f(B^{(i)}),\forall B^{(j)}\in\mathbf{N}(B^{(i)})\}\tag{11}$$

引理 2　若算法对应的有限非齐次 Markov 链在每一温度 T_i 下存在平稳分布,则在每一温度 T_i 下状态转移矩阵 $P(T_i)$ 存在特征值为 1 的左特征向量 $\pi(T_i)$,且 $\|\pi(T_i)\|=1$.其中,$T_0>T_1>\cdots,\lim\limits_{h\to\infty}T_h=0$;$\pi(T_i)=(\pi_1(T_i),\pi_2(T_i),\cdots,\pi_{|Q|}(T_i))$

证明

(1) 若在每一温度 T_i 下 Markov 链存在平稳分布 $\pi(T_i)$,则由平稳分布定义知

$$\pi_k(T_i)=\sum_{j=1}^{|Q|}\pi_j(T_i)P_{jk}(T_i)\,(\pi(T_i)P(T_i)=\pi(T_i))\tag{12}$$

且 $\sum\limits_{k=1}^{|Q|}\pi_k(T_i)=1$ 成立,从而知 $\pi(T_i)$ 为转移矩阵 $P(T_i)$ 的特征值为 1 的左特征向量;

(2) 由引理 1 及式(7)知算法中每一温度 T_i 下的转移矩阵 $P(T_i)$ 均是随机矩阵,必得它的平稳分布 $\pi(T_i)$ 的每一分量均为正数,从而知 $\sum\limits_{k=1}^{|Q|}\pi_k(T_i)=1$ 等价于 $\sum\limits_{k=1}^{|Q|}|\pi_k(T_i)|=1$,再由范数定义[3]知 $\|\pi(T_i)\|=1$.

由(1)、(2)可得证引理 2 成立.

定理 1　当 $T_0>T_1>\cdots,\lim\limits_{h\to\infty}T_h=0$,存在正整数 k_0 满足 $\sum\limits_{k=k_0}^{\infty}\exp(-rD/T_{kr})=\infty$ $(D=\max\limits_{B^{(i)}\in\wedge}\max\limits_{B^{(j)}\in\mathbf{N}(B^{(i)})}|f(B^{(j)})-f(B^{(i)})|)$,且存在 n_1 使得 $n\geqslant n_1$ 时,满足 $\sum\limits_{i=0}^{\infty}(\tau(P(T_i))^n<\infty$,则算法对应的 Markov 链强遍历,即算法具有全局收敛性.

证明　由定理 2 知 Markov 链强遍历的充分必要条件是 Markov 链弱遍历且在每一温度下状态转移矩阵存在特征值为 1 的左特征向量 $\pi(T_i)$,$\|\pi(T_i)\|=1$ 且

$$\sum_{i=0}^{\infty}\|\pi(T_i)-\pi(T_{i+1})\|<\infty\tag{13}$$

所以分为 3 步来证明定理:第一步,证明 Markov 链的弱遍历性;第二步,证明每一温度下状态转移矩阵存在特征值为 1 的左特征向量 $\pi(T_i)$,且 $\|\pi(T_i)\|=1$;第三步,证明 $\sum\limits_{i=0}^{\infty}\|\pi(T_i)-\pi(T_{i+1})\|<\infty$.

第一步:证明 Markov 链的弱遍历性.

令 $\omega=\min\limits_{B^{(i)}\in\wedge}\min\limits_{B^{(j)}\in\mathbf{N}(B^{(i)})}g_{B^{(i)}B^{(j)}}$,由式(6)中 $a_{B^{(i)}B^{(j)}}(T)$ 的表达式知:

$$\forall B^{(i)}\in Q,B^{(j)}\in\mathbf{N}(B^{(i)}),a_{B^{(i)}B^{(j)}}(T_h)\geqslant\omega\exp(-D/T_h)\tag{14}$$

当 $B^{(i)}\in(Q-Q_m)$ 时,若 $B^{(j)}\in\mathbf{N}(B^{(i)})$,则

$$g_{B^{(i)}B^{(j)}}=\frac{e^{-f(B^{(i)})/T_h}}{\sum\limits_{B^{(j)}\in\mathbf{N}(B^{(i)})}e^{-f(B^{(j)})/T_h}}=\frac{1}{\sum\limits_{B^{(j)}\in\mathbf{N}(B^{(i)})}e^{[f(B^{(i)})-f(B^{(j)})]/T_h}}\tag{15}$$

由式(11)知当 $B^{(i)} \in (Q - Q_m)$ 时,$f(B^{(i)}) < f(B^{(j)})$,从而此时的 $g_{B^{(i)}B^{(j)}}$ 是随着 h 的增大而增大,而 $\min\{1,\exp(-(f(B^{(i)})-f(B^{(j)}))/T_h\}$ 是随着 h 的增大而减小,进而得 $a_{B^{(i)}B^{(j)}}(T_h)$ 随 h 的增大而减小;若 $B^{(j)} \notin \mathbf{N}(B^{(i)})$ 且 $B^{(j)} \neq B^{(i)}$,则 $a_{B^{(i)}B^{(j)}}(T_h) = 0$ 从而得当 $B^{(i)} \in (Q - Q_m)$ 时,$a_{B^{(i)}B^{(i)}}(T_h) = 1 - \sum\limits_{B^{(j)} \in \mathbf{N}(B^{(i)})} a_{B^{(i)}B^{(j)}}(T_h)$ 随 h 的增大而增大;即存在整数 $k_0, k_0 < \infty$,对所有的 $B^{(i)} \in (Q - Q_m)$,都有

$$a_{B^{(i)}B^{(j)}}(T_h) \geqslant \omega \exp(-D/T_h), h \geqslant (k_0 - 1)r \tag{16}$$

所以,$\forall B^{(i)} \in Q, h \geqslant k_0 r$

$$a_{B^{(i)}I_c}(T_{h-r}, T_h) \geqslant \prod_{i=h-r}^{h} [\omega \exp(-D/T_i)] \geqslant \omega^r \exp(-rD/T_h) \tag{17}$$

$C, M, A(T)$ 均为随机阵,得每一温度下的状态转移阵 $P(T) = \mathrm{CMA}(T)$ 也是随机阵,且 $\tau(P) \leqslant \tau(A)$,从而

$$\begin{aligned}\tau(P(T_{kr-r}, T_{kr})) &\leqslant \tau(A(T_{kr-r}, T_{kr})) \leqslant 1 - \min_i \min_j (a_{iI_c}, a_{jI_c}) \leqslant \\ &1 - \omega^r \exp(-rD/T_h) \quad (k \geqslant k_0)\end{aligned} \tag{18}$$

即

$$\sum_{k=k_0}^{\infty} (1 - \tau(P(T_{kr-r}, T_{kr}))) \geqslant \sum_{k=k_0}^{\infty} \omega^r \exp(-rD/T_h) = \omega^r \sum_{k=k_0}^{\infty} \exp(-rD/T_h) = \infty \tag{19}$$

因此 $\sum\limits_{k=k_0}^{\infty} (1 - \tau(P(T_{kr-r}, T_{kr}))) = \infty$,即得 Markov 链弱.

第二步:证明每一温度下状态转移矩阵存在特征值为 1 的左特征向量 $\pi(T_i)$,且 $\|\pi(T_i)\| = 1$ 由引理 2 知只要有限非齐次 Markov 链在每一温度 T_i 下存在平稳分布,即平稳分布就是所要找的 $\pi(T_i)$,所以问题就转变为平稳分布的证明,而不可约非周期 Markov 链存在平稳分布[4],下面分为两步来分别证明 Markov 链不可约非周期性.现证明 Markov 链的不可约性及非周期性.

(1) 由设计的算法中的交叉和变异方式知 $\forall B^{(i)}, B^{(j)} \in Q, c_{B^{(i)}B^{(j)}}, m_{B^{(i)}B^{(j)}}$ 均大于 0,由式(4)和(6)知 $\forall B^{(i)}, B^{(j)} \in Q$,存在 $z \geqslant 1$ 使得 $B^{(i)}, B^{(k)}, B^{(k+1)} \cdots B^{(z)}, B^{(j)} \in Q$,满足:

$$a_{B^{(i)}B^{(k)}}(T) > 0, a_{B^{(l)}B^{(l+1)}}(T) > 0, (l = k, k+1, \cdots z), a_{B^{(z)}B^{(j)}}(T) > 0 \tag{20}$$

所以当温度 $T > 0$ 给定时,$\forall B^{(i)}, B^{(j)} \in Q$,算法对应的有限状态的非齐 Markov 链有

$$P_{B^{(i)}B^{(j)}}^{(z)}(T) \geqslant P_{B^{(i)}B^{(k)}}(T) P_{B^{(k)}B^{(k+1)}}(T) \cdots P_{B^{(z)}B^{(j)}}(T) > [\mathrm{CMA}(T)]_{B^{(i)}B^{(k)}} \cdots [\mathrm{CMA}(T)]_{B^{(z)}B^{(j)}} > 0 \tag{21}$$

因此,$B^{(i)} \leftrightarrow B^{(j)}$,进而由状态 $B^{(i)}$ 和 $B^{(j)}$ 的任意性知 $B^{(i)} \leftrightarrow B^{(j)}$.从而,由不可约的充分条件可得到算法对应的 Markov 链是不可约的.

(2) 由式(4)知 $\exists B^{(j)} \neq B^{(i)} \in Q$,且 $B^{(j)} \in \mathbf{N}(B^{(i)})$,使得

$$0 < g_{B^{(i)}B^{(j)}}(T) < 1 \tag{22}$$

而且又满足

$$0 < \min\{1, \exp(-(f(B^{(i)}) - f(B^{(j)}))/T\} < 1 \tag{23}$$

由式(6)得

$$0 < a_{B^{(i)}B^{(j)}}(T) < 1 \tag{24}$$

则状态 $\forall B^{(i)} \in Q$ 到自身的转移概率

$$P_{B^{(i)}B^{(i)}}(T) = \sum_{B^{(j)}} \sum_{B^{(m)} \neq B^{(i)}} c_{B^{(i)}B^{(j)}} m_{B^{(j)}B^{(m)}} (1 - a_{B^{(m)}B^{(i)}}(T)) \geqslant 0 \tag{25}$$

因此,非空集合 $\{n \mid p_{B^{(i)}B^{(i)}}^{n}(T) > 0, n \in \mathbf{Z}^+\}$ 的最大公约数 $d_{B^{(i)}B^{(i)}}(T) = 1$,故状态 $B^{(i)}$ 为非周期的.

由(1)(2)知算法对应的 Markov 链存在平稳分布,进而由引理 2 得出在每一温度下状态转移矩阵存在特征值为 1 的左特征向量 $\pi(T_i)$ 且 $\|\pi(T_i)\| = 1$.

第三步:证明 $\sum\limits_{i=0}^{\infty} \|\pi(T_i) - \pi(T_{i+1})\| < \infty$.由文献定理[5]知存在整数 M,使得 $n \geqslant M$ 时,满足

$$\sum_j |p_{i,j}^{(n)}(T_i) - \pi_j(T_i)| < 2(\tau(P(T_i)))^n \qquad (\forall n \geqslant M) \tag{26}$$

$$\sum_j |p_{i,j}^{(n)}(T_{i+1}) - \pi_j(T_{i+1})| < 2(\tau(P(T_{i+1})))^n \qquad (\forall n \geqslant M) \tag{27}$$

由式(5)及第一步证明过程知当 $j\in Q-Q_m$ 且 $k\in\mathbf{N}(j)$ 时,$a_{j,k}(T_h)$ 随着的增大而减小,所以当 $j\in Q-Q_m$ 且 $k\in\mathbf{N}(j)$ 时,存在 $N\in\mathbf{Z}^+$,使的 $n>N$ 时有

$$p_{j,k}^{(n)}(T_i)\geqslant p_{j,k}^{(n)}(T_i+1)\tag{28}$$

所以,取 $n_0=\max\{M,N\}$,由式(26)、(27)、(28)得,当 $n>n_0$ 时,

$$\begin{aligned}\sum_{k=1}^{|Q|}|\pi_k(T_i)-\pi_k(T_{i+1})|\leqslant&\sum_{k\in\mathbf{N}(j)}|\pi_k(T_{i+1})-p_{j,k}^{(n)}(T_{i+1})+p_{j,k}^{(n)}(T_i)-\pi_k(T_i)|\leqslant\\&\sum_{k\in\mathbf{N}(j)}|p_{j,k}^{(n)}(T_{i+1})-\pi_k(T_{i+1})|+\sum_{k\in\mathbf{N}(j)}|p_{j,k}^{(n)}(T_i)-\pi_k(T_i)|\leqslant\\&2\tau((P(T_{i+1})))^n+2(\tau(P(T_i)))^n\leqslant\\&4(\tau(P(T_i)))^n\end{aligned}\tag{29}$$

成立.又由定理中的假设知存在 n_1,使得 $n\geqslant n_1$ 时,满足

$$\sum_{i=0}^{\infty}(\tau(P(T_i))^n<\infty\tag{30}$$

而由范数定义[3]知

$$\|\pi(T_i)-\pi(T_{i+1})\|=\sum_{k=1}^{|Q|}|\pi_k(T_i)-\pi_k(T_{i+1})|\tag{31}$$

取 $n'=\max\{n_0,n_1\}$,由式(29)、(30)、(31)得,当 $n\geqslant n'$时,

$$\sum_{i=0}^{\infty}\|\pi(T_i)-\pi(T_{i+1})\|=\sum_{i=0}^{\infty}\sum_{k=1}^{|Q|}|\pi_k(T_i)-\pi_k(T_{i+1})|\leqslant4\sum_{i=0}^{\infty}(\tau(P(T_i)))^n<\infty\tag{32}$$

成立.因此,由第一、二和三步的证明以及文献定理[2]可证得此定理成立.

3 结　论

对于有约束非线性混合整数规划方程,提出了一种改进的遗传退火算法求解方法.通过建立算法对应的 Markov 链并对其收敛性给出理论证明.

参考文献:

[1] 陈斌,王晓.基于遗传算法的可再用逆向物流网络规划研究[D].上海海事大学,2006

[2] 王凌,郑大钟.一类 GASA 混合策略及其收敛研究[J].控制与决策,1998,13(6):669-672

[3] 龚光鲁,钱敏平.应用随机过程教程及算法和智能计算中的随机模型[M].北京:清华大学出版社,2004

[4] 吴晓敏.随机变量中马氏链的常返与强常返[J].重庆工商大学学报:自然科学版,2011,28(4)343-346

[5] 龚光鲁,钱敏平.应用随机过程教程及算法和智能计算中的随机模型[M].北京:清华大学出版社,2004

An Improved Genetic Annealing Algorithm and Its Convergence Analysis

GAO Fa-ling

(Qindao College, Qingdao Institute of Technology, Shandong Qingdao 266100, China)

Abstract: Under the constraint condition of nonlinear mixed integer programming model with 0-1 variable and integer variable, the solution to an improved genetic annealing algorithm is given, the corresponding Markov Chain is set up and its convergence is theoretically proved.

Key words: genetic annealing; Markov chain; traversal; convergence

责任编辑:田　静