

# 第五章 微粒群算法的收敛性分析

## 5.1 随机算法的收敛准则

为了给出微粒群算法的全局收敛和局部收敛的判据，我们首先给出随机算法的收敛准则。本节主要参照[F.Solis1981]、[F.van den Bergh2002]的内容。为了方便起见，首先介绍一些相关概念。

**定义 5.1**、设有一个以集合为元素的非空类  $\mathfrak{S}$ ，如果它满足下列条件，我们就称  $\mathfrak{S}$  为  $\sigma$ -域。

- 1)  $\Omega \in \mathfrak{S}$ ；
- 2) 如果  $E_j \in \mathfrak{S}$ ， $j=1,2,\dots$ ，则

$$\bigcup_{j=1}^{\infty} E_j \in \mathfrak{S} \quad (5.1)$$

- 3) 如果  $E \in \mathfrak{S}$ ，则  $E^c \in \mathfrak{S}$

其中， $\Omega$  为全集， $E^c$  表示  $E$  的补集。

**定义 5.2**、设  $\mu$  是定义在非空类  $\ell$  上的集函数，并满足条件：

- 1) 对任意的  $E \in \ell$ ，都有

$$0 \leq \mu(E) < +\infty \quad (5.2)$$

- 2) 如果  $\phi \in \ell$ ，则  $\mu(\phi) = 0$

- 3) 若  $E_j \in \ell$ ， $j=1,2,\dots$ ，则  $\bigcup_{j=1}^{\infty} E_j \in \ell$ ，且

$$E_k \cap E_j = \phi, \forall k \neq j \quad (5.3)$$

那么

$$\mu\left(\bigcup_{k=1}^{+\infty} E_k\right) = \sum_{k=1}^{+\infty} \mu(E_k) \quad (5.4)$$

则我们说  $\mu$  是  $\ell$  上的一个测度，若将式子 (5.2) 改为

$$0 \leq \mu(E) \leq 1 \quad (5.5)$$

我们称之为概率测度。

设  $\Omega$  是任一集合， $\mathfrak{S}$  是  $\Omega$  的子集组成的  $\sigma$ -域， $\mu$  是  $\mathfrak{S}$  上的概率测度，称三元组  $(\Omega, \mathfrak{S}, \mu)$  为概率空间。

本节主要考虑下述问题：

$$\min f(x) \quad x \in S \subseteq R^n \quad (5.6)$$

但在该情况下，有可能出现病态形式，比如函数取为

$$f = \begin{cases} x^2, \forall x \neq 1 \\ -10, x = 1 \end{cases} \quad (5.7)$$

这样，函数  $f$  的最优值为  $f(1) = -10$ ，但它是一个间断点，其测度为 0，因而对于进化算法而言，几乎不可能搜索到全局最优解。为此，我们将问题 (5.6) 改为下列形式：

$$\psi = \inf\{x \mid v(z \in S \mid f(z) < x) > 0\} \quad (5.8)$$

在这种情形下，由于 Lebesgue 测度  $v(x \mid f(z) < x) > 0$ ，因而可以避免上述病态情形。

为了具体给出随机算法的收敛准则，下面给出其基本框架：

Step0：随机选择初始点  $z_0 \in S$ ，并置  $k=0$ ；

Step1：在样本空间  $(R^n, B, \mu_k)$  上生成向量  $\xi_k$ ；

Step2：计算  $z_{k+1} = D(z_k, \xi_k)$ ，选择  $\mu_{k+1}$ ，令  $k=k+1$ ，并转 Step1。

其中， $(R^n, B, \mu_k)$  表示算法在第  $k$  代的概率空间， $\mu_k$  为  $B$  上的概率测度 [J.C.Taylor1997]， $B$  为  $R^n$  的某个子集的  $\sigma$ -域。D 是算法迭代方式，用以产生下一代个体。

**定义 5.3.** 设  $M_k$  为  $R^n$  的子集，若它满足下列条件，则称为概率测度  $\mu_k$  的支撑集。

- 1)  $\mu_k(M_k) = 1$ ；
- 2) 任取点列  $\{y_k\}_{k=1}^{+\infty} \subseteq M_k$ ，对于其任意收敛子列  $\{y_j^k\}_{j=1}^{+\infty}$ ，有  $\lim_{j \rightarrow +\infty} y_j^k \in M_k$ ；
- 3) 若  $N \subseteq R^n$ ，且满足 1) 及 2)，则有  $M_k \subseteq N$ 。

为了保证随机算法的有效性，应保证 D 所产生的新个体应优于当前个体，因此，随机算法应满足下列假设：

**假设 5.1.**  $f(D(z, \xi)) \leq f(z)$ ，并且如果  $\xi \in S$ ，则  $f(D(z, \xi)) \leq f(\xi)$ 。

随机算法的全局收敛意味着序列  $\{f(z_k)\}_{k=1}^{\infty}$  应收敛于  $\psi$ 。

**定义 5.4.** 定义算法的  $\varepsilon$ -可接受区域为：

$$R_\varepsilon = \{z \in S \mid f(z) < \psi + \varepsilon\} \quad (5.9)$$

其中， $\varepsilon > 0$ 。若算法发现  $R_\varepsilon$  中的点，则称算法找到了误差为  $\varepsilon$  的可接受点。

**假设 5.2.** 对于  $S$  的任意 Borel 子集  $A$ ，若其测度  $v(A) > 0$ ，则有

$$\prod_{k=0}^{\infty} (1 - \mu_k(A)) = 0 \quad (5.10)$$

其中， $\mu_k(A)$  是由测度  $\mu_k$  所得到的  $A$  的概率。

下面我们将利用假设 5.1、5.2，给出随机算法为全局收敛算法的一个充要条件。

**定理 5.1**、假设目标函数  $f$  为可测函数，区域  $S$  为可测子集，并且假设 5.1、5.2 满足，设  $\{z_k\}_{k=1}^{+\infty}$  为算法所生成的解序列，

$$\lim_{k \rightarrow +\infty} P[z_k \in R_\varepsilon] = 1 \quad (5.11)$$

其中， $P[z_k \in R_\varepsilon]$  是第  $k$  步算法生成的解  $z_k \in R_\varepsilon$  的概率。

证明：由假设 5.1，如果  $z_k \in R_\varepsilon$  或者  $\xi_k \in R_\varepsilon$ ，则对于任意的  $k' > k$ ，有  $z_{k'} \in R_\varepsilon$ ，从而

$$P(z_k \in R_\varepsilon) = 1 - P(z_k \in S \setminus R_\varepsilon) \geq 1 - \prod_{t=0}^{k-1} (1 - \mu_t(R_\varepsilon)) \quad (5.12)$$

取极限，并考虑  $\mu_k$  为概率测度，得到

$$1 \geq \lim_{k \rightarrow +\infty} P(z_k \in R_\varepsilon) \geq 1 - \lim_{k \rightarrow +\infty} \prod_{t=0}^{k-1} (1 - \mu_t(R_\varepsilon)) \quad (5.13)$$

由假设 2，我们有  $\prod_{k=0}^{\infty} (1 - \mu_k(A)) = 0$ ，因此，

$$1 \geq \lim_{k \rightarrow +\infty} P(z_k \in R_\varepsilon) \geq 1 - 0 = 1$$

证毕。

定理 5.1 给出了全局收敛算法的充要条件，下面将考虑局部搜索算法的收敛准则。

**定义 5.5**、局部收敛算法是指对于测度序列  $\{\mu_k\}_{k=1}^{+\infty}$ ，支撑集序列  $\{M_k\}_{k=1}^{+\infty}$ ，除了有限个集合外，都有界且  $M_k \subset S$ 。

按照定义，局部收敛算法的支撑集序列满足下面的关系：

$$v(S \cap M_k) < v(S) \quad (5.14)$$

因此，一个局部收敛算法应满足下列假设：

**假设 5.3**、对于任意  $z_0 \in S$ ，存在  $\gamma > 0, 0 < \eta < 1$ ，使得：

$$\mu_k((\rho(D(z, \xi), R_\varepsilon) < \rho(z, R_\varepsilon) - \gamma), \text{ or }, (D(z, \xi) \in R_\varepsilon)) \geq \eta \quad (5.15)$$

对于所有的  $k$  和集合  $L_0 = \{z \in S \mid f(z) \leq f(z_0)\}$  中的  $z$  均成立。

其中， $\rho(z, A)$  表示点  $z$  与集合  $A$  之间的距离，定义为

$$\rho(z, A) = \inf_{b \in A} \rho(z, b) \quad (5.16)$$

下面我们将利用假设 5.1、5.3，给出随机算法为局部收敛算法的一个充要条件。

**定理 5.2.** 假设目标函数  $f$  为可测函数, 区域  $S$  为可测子集, 并且假设 5.1、5.3 满足, 设  $\{z_k\}_{k=1}^{+\infty}$  为算法所生成的解序列, 则

$$\lim_{k \rightarrow +\infty} P(z_k \in R'_\varepsilon) = 1 \quad (5.17)$$

其中,  $P(z_k \in R'_\varepsilon)$  是第  $k$  步算法生成的解  $z_k \in R'_\varepsilon$  的概率, 区域  $R'_\varepsilon$  表示某一局部极值点的  $\varepsilon$ -可接受区域。

证明: 设  $z_0$  为算法初始点, 由假设 3 的条件及欧几里得定理, 存在整数  $p$ , 使得

$$\eta p > \rho(a, b), \quad \forall a, b \in L_0 \quad (5.18)$$

由假设 3, 有

$$P[z_1 \in R'_\varepsilon] \geq \eta \quad (5.19)$$

和

$$P[z_2 \in R'_\varepsilon] \geq \eta \times P[z_1 \in R'_\varepsilon] \geq \eta^2 \quad (5.20)$$

因此, 继续  $p$  次, 得到

$$P[z_p \in R'_\varepsilon] \geq \eta^p \quad (5.21)$$

继续, 得到

$$P[z_{kp} \in R'_\varepsilon] = 1 - P[z_{kp} \notin R'_\varepsilon] \geq 1 - (1 - \eta^p)^k \quad (5.22)$$

由假设 1,  $z_1, z_2, \dots, z_{p-1} \in L_0$ , 因此得到

$$P(z_{kp+t} \in R'_\varepsilon) \geq 1 - (1 - \eta^p)^k \quad (5.23)$$

$t = 0, 1, \dots, p-1$ 。这表明解序列在  $kp$  与  $(k+1)p$  之间满足假设 3, 由于  $(1 - \eta^p)^k \rightarrow 0$ , 因而定理得证。

## 5.2 基本微粒群算法的收敛性分析

本节将对基本微粒群算法的收敛性问题进行讨论, 主要参考[F.van den Bergh2002]的内容。通过考虑是否满足假设 5.1、5.2 和 5.3, 按照上节的定理 5.1、5.2, 给出其局部收敛性及全局收敛性的分析。

**定理 5.3.** 基本微粒群算法满足假设 5.1。

证明: 定义函数  $D$  为

$$D(p_{g,k}, x_{i,k}) = \begin{cases} p_{g,k}, & \text{if } (f(g(x_{i,k})) \geq f(p_{g,k})) \\ g(x_{i,k}), & \text{otherwise} \end{cases} \quad (5.24)$$

其中, 符号  $g(x_{i,k})$  表示基本微粒群算法的更新方程, 具体为

$$x_{i,k+1} = g(x_{i,k}) = g_1(x_{i,k}) + g_2(x_{i,k}) + g_3(x_{i,k}) \quad (5.25)$$

其中

$$g_1(x_{i,k}) = x_{i,k} + wv_{i,k} \quad (5.26)$$

$$g_2(x_{i,k}) = c_1 r_1(p_{i,k} - x_{i,k}) \quad (5.27)$$

$$g_3(x_{i,k}) = c_2 r_2(p_{g,k} - x_{i,k}) \quad (5.28)$$

$x_{i,k}$  表示第  $k$  代时的微粒  $i$  的位置。则按照这里定义的函数  $D$ ，基本微粒群算法满足假设 1。

**定理 5.4**、任取  $\varepsilon > 0$ ，存在  $N \geq 1$ ，使得对于任意的  $n \geq N$ ，如果选择  $w, \phi_1, \phi_2$ ，使得  $\max(\|\alpha\|, \|\beta\|) < 1$ ，则有  $\|g^n(x_{i,k}) - g^{n+1}(x_{i,k})\| < \varepsilon$ 。

证明：由于

$$\lim_{t \rightarrow +\infty} X(t) = \lim_{t \rightarrow +\infty} (k_1 + k_2 \alpha^t + k_3 \beta^t) = \frac{\phi_1 P + \phi_2 P_g}{\phi_1 + \phi_2} \quad (5.29)$$

且  $X(t+1) - X(t) = V(t+1)$ ，因此，当  $\max(\|\alpha\|, \|\beta\|) < 1$  时，下式成立

$$\begin{aligned} \lim_{t \rightarrow +\infty} V(t+1) &= \lim_{t \rightarrow +\infty} (X(t+1) - X(t)) \\ &= \lim_{t \rightarrow +\infty} k_2 \alpha^t (\alpha - 1) + k_3 \beta^t (\beta - 1) \\ &= 0 \end{aligned}$$

这表明  $\lim_{t \rightarrow +\infty} X(t+1) = \lim_{t \rightarrow +\infty} X(t)$ 。同时，

$$\begin{aligned} X(t+1) &= X(t) + V(t+1) \\ &= X(t) + wV(t) - X(t)(\phi_1 + \phi_2) + P\phi_1 + P_g\phi_2 \end{aligned} \quad (5.30)$$

两边取极限，从而有

$$\lim_{t \rightarrow +\infty} X(t) = P = P_g$$

通过上面的定理 5.4，可以看出当所有的微粒最终收敛于  $\lim_{t \rightarrow +\infty} X(t) = P = P_g$  的位置时，算法将停止运行。因此，如果算法在收敛之前没有搜索到全局（或局部）最优解，将导致过早收敛。从而表明基本微粒群算法不是局部收敛算法。

**定理 5.5**、基本微粒群算法不满足假设 5.2。

证明：如果假设 2 成立，则等同于证明基本微粒群算法满足下式

$$S \subseteq \bigcup_{i=1}^t M_{i,k} \quad (5.31)$$

其中， $M_{i,k}$  表示在算法第  $k$  代时微粒  $i$  的支撑集。由于

$$X(t+1) = X(t) + wV(t) - X(t)(\phi_1 + \phi_2) + P\phi_1 + P_g\phi_2$$

$$\begin{aligned}
&= X(t) + w(X(t) - X(t-1)) - X(t)(\phi_1 + \phi_2) + P\phi_1 + P_g\phi_2 \\
&= (1 + w - \phi_1 - \phi_2)X(t) - wX(t-1) + P\phi_1 + P_g\phi_2 \quad (5.32)
\end{aligned}$$

因此，代入可以得到  $M_{i,k}$  为

$$\begin{aligned}
M_{i,k} &= (1 + w - \phi_1 - \phi_2)x_{i,j,k-1} - wx_{i,j,k-2} + \phi_1 P + \phi_2 P_m \\
&= x_{i,j,k-1} + w(x_{i,j,k-1} - x_{i,j,k-2}) + \phi_1(P - x_{i,j,k-1}) + \phi_2(P_g - x_{i,j,k-1}) \quad (5.33)
\end{aligned}$$

其中， $0 \leq \phi_1 \leq c_1$ ， $0 \leq \phi_2 \leq c_2$ ， $x_{i,j,k}$  表示微粒  $i$  在第  $k$  代时第  $j$  维分量的值。显然， $M_{i,k}$  表示一个由  $\phi_1$ ， $\phi_2$  所确定的超矩形，其中一个端点为  $\phi_1 = \phi_2 = 0$ ，另一个为  $\phi_1 = c_1$ ， $\phi_2 = c_2$ 。

当  $\max(c_1 |P - x_{i,j,k}|, c_2 |P_g - x_{i,j,k-1}|) < 0.5 \times \text{diam}_j(S)$  成立时，显然有  $v(M_{i,k} \cap S) < v(S)$ 。其中， $\text{diam}_j(S)$  表示  $S$  在第  $j$  维分量的长度。由于  $x_i \rightarrow \frac{c_1 P + c_2 P_g}{c_1 + c_2}$ ，因

而  $\lim_{k \rightarrow +\infty} M_{i,k} = 0$ 。从而，随着迭代次数  $k$  的增加， $v(M_{i,k})$  在不断减少，其并  $v(\bigcup_{i=1}^t M_{i,k})$  也在减

少。从而  $v(\bigcup_{i=1}^t M_{i,k} \cap S) < v(S)$ ，这表明存在整数  $k'$ ，使得当  $k > k'$  时，存在集合  $A \subset S$ ，使

得  $\sum_{i=1}^t \mu_{i,k}(A) = 0$ ，这表明基本微粒群算法不满足假设 5.2。

由于基本微粒群算法满足假设 5.1，不满足假设 5.2，因此不是全局收敛算法。

### 5.3 其它改进微粒群算法的收敛性分析

本节考虑两种改进的微粒群算法：一种是保证收敛的微粒群算法（Guaranteed Convergence Particle Swarm Optimizer, 简称 GCPSO）的。另一种是保证全局收敛的随机微粒群算法（Guaranteed Global Convergence Stochastic Particle Swarm Optimizer, 简称 GGCS PSO）。GCPSO 是一种局部收敛算法，而 GGCS PSO 则是我们提出的一种全局收敛算法。本节主要对这两种改进的微粒群算法的收敛性能进行详细的分析。

**定理 5.6、GCPSO 满足假设 5.1。**

证明：定义函数  $D$  为

$$D(p_{g,k}, x_{i,k}) = \begin{cases} p_{g,k}, & \text{if } (f(g(x_{i,k})) \geq f(p_{g,k})) \\ g(x_{i,k}), & \text{otherwise} \end{cases} \quad (5.34)$$

其中，符号  $g(x_{i,k})$  表示基本微粒群算法的更新方程，具体为对于不在当前最优位置的微粒  $i$ ，进

化方程为（ $x_{i,k}$  表示第  $k$  代时的微粒  $i$  的位置）

$$x_{i,k+1} = x_{i,k} + wv_{i,k} + c_1r_1(p_{i,k} - x_{i,k}) + c_2r_2(p_{g,k} - x_{i,k}) \quad (5.35)$$

对于正处于当前最优位置的微粒  $\tau$  , 进化方程为 :

$$x_{\tau,k+1} = p_{g,k} + wp_{\tau,k} + \rho_k(1 - 2r_k) \quad (5.36)$$

则按照这里定义的函数 D , 基本微粒群算法满足假设 5.1。

**定理 5.7**、GCPSO 满足假设 5.3。

证明 : 首先选择初始种群中的最差个体  $x_0$  , 使得它满足

$$x_0 = \arg \max_{x_j} \{f(x_j)\}, j = 1, \dots, t \quad (5.37)$$

据此 , 我们定义  $L_0 = \{x \in S \mid f(x) \leq f(x_0)\}$

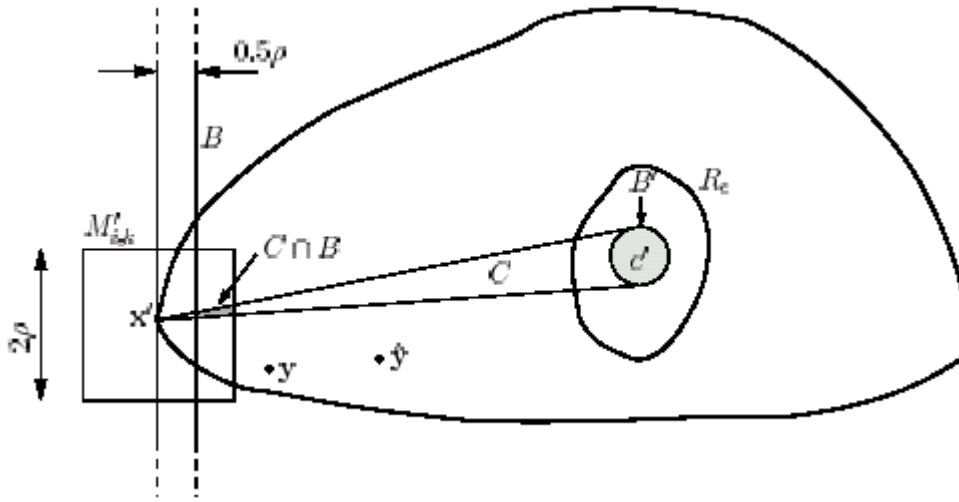


图 5.1、GCPSO 满足假设 5.3 的解释

在图 5.1 中 , 设  $R_\epsilon$  为某一局部极值点的邻域 , 任取  $R_\epsilon$  的内点  $c'$  , 并作以  $c'$  为球心的球  $B'$  , 取点  $x'$  使得其满足

$$x' = \arg \max_x \{dis \tan ce(x, c') \mid x \in L_0\} \quad (5.38)$$

下面令 B 为以  $c'$  为中心 , 长度为  $2[dis \tan ce(c', x') - 0.5\rho]$  的超立方体 , 其中 ,  $dis \tan ce(x, y)$  表示点 x 与 P 的长度。称由点  $x'$  定义的区域 C 为 :  $\forall m \in C, \exists y \in B'$  , 使得 m 位于点  $x'$  与 P 所呈的线段上。显然 ,  $\mu(C \cap B) > 0$  , 且任取点  $x \in L_0$  , 设由点 x 定义的区域为  $C'$  , 则有

$$\mu(C' \cap B) > \mu(C \cap B) > 0 \quad (5.39)$$

因而 , 有

$$\mu_k(dis \tan ce(D(P_g, x_\tau), R_\epsilon)) < dis \tan ce(x, R_\epsilon) - 0.5\rho \geq \eta = \mu = [C \cap B] > 0 \quad (5.40)$$

其中  $\mu_k$  为在在以 x 为中心 , 长度为  $2\rho$  所成的超立方体上的均匀分布。从而满足假设 5.3。

**定理 5.8**、GCPSO 不满足假设 5.2。

证明：由于基本微粒群算法不满足假设 5.2，而 GCP SO 中除了具有全局最优位置的微粒外，其余的进化方程与基本微粒群算法相同。因此，算法 GCP SO 是否满足假设 5.2，将由具有全局最优位置的微粒的进化方程所决定。

设具有全局最优位置的微粒为  $\tau$ ，则它满足

$$x_{\tau,j,k} = P_{g,k} + wv_{\tau,j,k} + \rho_k(1 - 2r_k) \quad (5.41)$$

其中， $0 \leq r_k \leq 1$ ，设其支撑集为超矩形  $M_{\tau,k}$ ，则其边长由  $\rho_k$  决定，因而当  $\rho_k < diam(S)$  时，有  $v(M_{\tau,k} \cap S) < v(S)$ 。按照  $\rho_k$  的进化方程，如果发现不了更好的解，则  $\rho_k$  会持续减少。因此，当 GCP SO 收敛到某个局部最优解时，将会导致

$$\lim_{k \rightarrow +\infty} \rho_k = \rho_{\min} \quad (5.42)$$

$\rho_{\min}$  为一任意小的数，故此，当  $\rho_{\min} \ll diam(S)$ ，有

$$v\left(\bigcap_{j=1}^t M_{j,k} \cap S\right) < v(S) \quad (5.43)$$

从而 GCP SO 不满足假设 5.2。

**定理 5.9**、GGCSP SO 以概率 1 收敛于全局最优解。

证明：在保证全局收敛的随机微粒群算法中，其解序列为  $\{p_{g,t}\}$ ，其中  $t$  为进化代数， $p_{g,t}$  为第  $t$  代时的微粒群最好位置。对 GGCSP SO 算法定义函数  $D$  为

$$D(p_{g,t}, x_i(t)) = \begin{cases} p_{g,t}, f(p_{g,t}) \leq f(x_i(t)) \\ x_i(t), f(p_{g,t}) > f(x_i(t)) \end{cases} \quad (5.44)$$

则很容易证明其满足假设 5.1。

为了满足假设 5.2，规模为  $S$  的微粒群的样本空间的并必须包含  $S$ ，即

$$S \subseteq \bigcup_{i=1}^S M_{i,t} \quad (5.45)$$

其中  $M_{i,t}$  为  $t$  代时微粒  $i$  的样本空间的支撑集。对于满足  $x_i(t) = p_k = p_g$  的微粒  $i$ ， $M_{i,t} = S$ 。而对于其它微粒  $i$ ：

$$M_{i,t} = X_i(t-1) + \varphi_1(P_i - X_i(t-1)) + \varphi_2(P_g - X_i(t-1)) \quad (5.46)$$

由于  $0 \leq \varphi_1 \leq c_1, 0 \leq \varphi_2 \leq c_2$ ，则  $M_{i,t}$  为一具有顶点  $\varphi_1 = \varphi_2 = 0$  和  $\varphi_1 = c_1, \varphi_2 = c_2$  的超矩形体，

且当  $\max(c_1 |P_i - x_i(t-1)|, c_2 |P_g - x_i(t-1)|) < 0.5 \times diam(S)$  时， $v[M_{i,t} \cap S] < v(S)$ ，其中

$diam(S)$  表示  $S$  的长度。由定理 5.1，当  $t \rightarrow \infty$  时， $M_{i,t}$  的长度趋于 0。因此，随着  $t$  的增长，每

一  $M_{i,t}$  的闭包  $v[M_{i,t}]$  在逐渐变小，其并集  $\bigcup_{i \neq j} M_{i,t}$  的闭包  $v[\bigcup_{i \neq j} M_{i,t}]$  也在变小。因而存在  $N$ ，

使得  $t > N$  时， $v[\bigcup_{i \neq j} M_{i,t} \cap S] < v[S]$ 。也就是说仅有进化方程 (2) 的 PSO 算法不满足假设 5.2。



但由于  $M_{i,t} = S$  , 所以  $\bigcup_{i=1}^S M_{i,t} = S$  , 定义  $S$  的 Borel 子集  $A = M_{i,t}$  , 则有  $v[A] > 0$  ,

$\mu_i[A] = \sum_{i=0}^S \mu_{i,t}[A] = 1$  , 从而满足假设 5.2。由定理 5.2 , GGCPSO 算法以概率 1 收敛于全局最

优解。