结合了信赖域的粒子群算法及其收敛性证明

刘欧

Abstract

本文给出了一种结合了信赖域算法的粒子群算法 (Tr-PSO). 粒子群算法上定义概率空间, 将粒子在每次迭代中的位置视为一个随机向量, 利用随机过程以及泛函分析理论确定算法的收敛条件, 对于算法的线性递推关系进行了一阶以及二阶稳定性分析. 并且假设在粒子群更新过程中个体最优位置是在运动过程中更新的随机变量 (具有任意均值与方差). 本文确定了算法的收敛条件以及相应的参数选择规则.

Keywords: 粒子群算法; 优化控制; 收敛性分析; 随机过程; 参数选择

1. 介绍

2. Tr-PSO 算法介绍

2.1. PSO 算法介绍

粒子群算法由 Eberhart 与 Kennedy 提出 [**?**], 其核心思想是利用群体中个体对信息的共享使得整个整体在问题求解空间中产生从无序到有序的演化过程, 从而获得问题的可行解.

标准的粒子群算法会维持 M 个粒子的数量, 并且将每个粒子定义为 D-维解空间 $\mathbb{R}^D$ 中的一个解. 粒子群中的所有粒子构成了一个 $M\times D$ 矩阵 X, 其中粒子 i 的位置为矩阵 X 的第 j 个行向量 $X_j=(x_{j1},x_{j2},\cdots,x_{jD})$. 每个粒子都会保持其历史最优位置的记忆, 它们同样构成一个 $M\times D$ 矩阵 P^i, 第 j 个粒子的历史最优位置为矩阵的第 j 个行向量 $(P^i_{j1},P^i_{j2},\cdots,P^i_{jD})$. 将 P^i 的邻域内最优粒子的位置记为 P^g, 这是一个 D 维行向量. 在本文中, 研究星型拓扑结构, 即所有粒子的邻域均为全体粒子.

使用 f 表示需要进行最小化优化的目标函数, 本文研究的是单目标优化, 因此有 $f:\mathbb{R}^n\to\mathbb{R}$. 在不丧失一般性的前提下, 我们把粒子群的时间步长记为其迭代次数. 因此我们可以得到一个与迭代次数 t 相关联的 P^i 更新公式, 考虑第 j 个粒子, 其历史最优位置更新公式如 1所示

$$P^i_j(t+1)=\begin{cases} P^i_j(t), 若 f(X_j(t+1))\geq f(P^i_j(t)) \\ X_j(t+1), 若 f(X_j(t+1))<f(P^i_j(t)) \end{cases} \tag{1}$$

在标准的粒子群算法中, 迭代次数为 t 时, 对于第 j 个粒子的速度 V_j 与位置 X_j 的更新公式分别如公式 (2) 以及 (3) 所示.

$$V_j(t+1)=wV_j(t)+c_1r_{1,t}(P^i_j(t)-X_j(t))+c_2r_{2,t}(P^g(t)-X_j(t)) \tag{2}$$

$$X_j(t+1)=X_j(t)+V_j(t+1) \tag{3}$$

将 (3) 带入 (2) 后得到位置 $X_j(t)$ 的迭代公式

$$X_j(t+1)=(1+w-r_{1,t}c_1-r_{2,t}c_2)X_j(t)-wX_j(t-1)+r_{1,t}c_1P^i_j(t)+r_{2,t}c_2P^g(t) \tag{4}$$

2.2. 信赖域算法

2.2.1. 介绍

信赖域算法是一种求解非线性优化问题的数值方法 [**?**]. 信赖域算法是一种迭代算法, 即从给定的初始解出发, 通过逐步迭代, 不断改进, 直到获得满意的近似最优解为止. 其基本思想是把最优化问题转化为一系列简单的局部寻优问题.

2.2.2. 原理

对于一个优化目标函数 f, 设置搜索初值为 $\boldsymbol{x_0}$, 设置一个信赖域半径 Δ_0, 最高迭代次数为 N, 开始迭代.

对于第 k 次迭代时, 根据此时的信赖域半径 Δ_k 构建信赖域 $\Omega_k = \{x \in R^n | \, \|x - x_k\| \leq \Delta_k\}$.

对于目标函数 f, 其在极值点处近似于二次近似函数 $q^{(k)}(s)$, 因此对于无约束时, 可以利用二次逼近, 构造如下的信赖域子问题：

$$\begin{aligned} \min \quad q^{(k)}(\boldsymbol{s}) &= f(\boldsymbol{x_k}) + grad_k^T \boldsymbol{s} + \frac{1}{2}(\boldsymbol{s}^T H_k \boldsymbol{s}) \\ s.t. \quad \|\boldsymbol{s}\|_2 &\leq \Delta_k \end{aligned}$$

式中,$grad_k$ 是当前目标函数 f 在点 $\boldsymbol{x_k}$ 处的梯度;$\boldsymbol{s} = \boldsymbol{x} - \boldsymbol{x_k}$;$H_k$ 为 f 在点 $\boldsymbol{x_k}$ 处的 Hesse 矩阵.

假设 $\boldsymbol{s_k}$ 为该子问题的解. 目标函数 f 在第 k 步的实际下降量 (观测下降长度) 为

$$Ared_k = f(\boldsymbol{x_k}) - f(\boldsymbol{x_k} + \boldsymbol{s_k})$$

二次近似函数 $q^{(k)}(s)$ 的预测下降量 (预计下降长度) 为

$$Pred_k = q^{(k)}(0) - q^{(k)}(\boldsymbol{s_k})$$

根据观测下降长度与预测下降长度的比值 r_k 得到评价信赖域在下一次迭代时的半径 Δ_{k+1}

若 $r_k \to 1$, 则二次近似函数与目标函数的接近程度好. 这个时候应当适当扩大信赖域范围.

若 $r_k \to 0$, 则二次近似函数与目标函数的接近程度差. 这个时候应当适当缩小信赖域范围.

若 r_k 无显著表现, 则不改变信赖域范围.

一般而言, 当 $r_k \geq 0.75$, 认为 $r_k \to 1$; 当 $r_k \leq 0.25$, 认为 $r_k \to 0$.

2.3. Tr-PSO 算法

Tr-PSO 算法是一种将信赖域算法作为一种针对最优位置进行的优化操作并入 PSO 算法的一种新算法. 其具体过程为在第 t 次迭代中, 当粒子经过更新公式 2和 3得到 $V(t)$ 与 $X(t)$ 后, 计算其最优位置 P^g, 而后对于最优位置使用信赖域算法进行优化.

其特点在于:

- 1 通过影响粒子的全局最优位置而非直接改变粒子位置来影响粒子群的走向.
- 2 全局最优位置在粒子寻找到更优之前相当于是固定的, 利于后文的收敛性分析.
- 3 使用信赖域算法保证算法收敛时粒子必然处于最优解处.

总体过程的伪代码如 Algorithm 1 所示:

Algorithm 1: Tr-PSO

Input: 目标函数 f, 粒子个数 M, 解空间维数 D, 收敛判断阈值 ε
Output: 最优位置 P^g

对于粒子进行初始化;
t = 1;
计算粒子 j 的适应度 $f(X_j)$;
while 满足收敛条件 **do**
  根据 (2) 和 (3) 更新速度与位置计算粒子 j 的适应度 $f(X_j)$;
  根据 (1) 更新 $P_j^i(t)$;
  $P^g(t) =$ 使 $f(P_j^i(t))$ 最小的 $P_j^i(t)$;
  对于 P^g 使用信赖域算法进行优化;
  if $f(P^g) < \varepsilon$ **then**
    满足收敛条件
  end
  t = t+1;
end
return P^g,$f(P^g)$;

3. 收敛性分析

3.1. 收敛性

由于 Tr-PSO 算法本质上仍然为某个线性递归关系, 因此可以使用谱理论 [**?**] 对其收敛性进行分析.

一个线性递推关系被记作 [**?**]

$$x_{t+1} = Mx_t + b \tag{5}$$

其中 x_{t+1}, x_t 与 b 均为一个 $\mathbb{R}^d$ 上的 d 维向量,M 为一个 $d \times d$ 矩阵.

定义 1. *在定义了范数 $\|\bullet\|$ 的向量空间 $\mathbb{R}^d$ 中, 对于序列 $\{x_1, x_2, \cdots\} \subset \mathbb{R}^d$, 若存在一个点 $\hat{x} \in \mathbb{R}^d$ 使得 $\lim_{n\to\infty} \|x_n - \hat{x}\| = 0$ 就称 $\{x_n\}$ 收敛于 $\hat{x}$*

定义 1等价于说明序列 $\{x_n\}$ 是一个收敛列.

引理 1. *一个由 (5) 生成的序列 $\{x_1, x_2, \cdots\}$ 是收敛的, 当且仅当 $\rho(M) = \max(|\gamma|) < 1$ 其中,γ 为 M 的谱内的元素,$\rho(M)$ 为 M 的谱半径.*

引理 1等价于 M 的所有特征值均小于 1.

证明. 见 [**?**] □

因此, 若序列收敛, 则其谱半径 $\rho(M) < 1$, 同时存在一个不动点 $\hat{x}$ 使得 $\lim_{n\to\infty} x_n = \hat{x}$. 那么在收敛时有 $\hat{x} = M\hat{x} + b$ 即 $(I - M)\hat{x} = b$ 因此有 $\hat{x} = (I - M)^{-1}b$.

3.2. 稳定性

3.2.1. 验证 Tr-PSO 为随机过程

将粒子所有可能位置记作一个集合 Ω, 记事件 $A_a := \{粒子会经过点a\}$, 其中 $a \in \Omega$. 不难发现, 粒子的运动轨迹 S 可以表示为若干个 A_a 的并集, 所有可能的 S 的集合记为 $\mathscr{F}$.

命题 2. *$\mathscr{F}$ 是 Ω 的一个 σ 代数.*

证明. 1. $\Omega \in \mathscr{F}$
2. $S \in \mathscr{F} \Leftrightarrow S^c := \Omega \setminus S \in \mathscr{F}$
3. 若轨迹 $S_n \in \mathscr{F}$,$n = 1, 2, \cdots$, 则由于 S_n 可以写为若干个 A_a 的并集, 即 $S^n = \bigcup_{A_a \subset S_n} A_a$, 因此 $\bigcup_{n=1}^{\infty} S_n \in \mathscr{F}$
因此 $\mathscr{F}$ 是 Ω 上的一个 σ 代数. □

从而可以在 $(\Omega, \mathscr{F})$ 上赋予概率 P 使其成为一个概率空间 $(\Omega, \mathscr{F}, P)$.

由于粒子群的更新机制 (2) 以及信赖域算法的存在, 概率 P 的设置与目标函数 f 有关.

因此每个粒子的运动轨迹都可以视为一个随机事件. 而粒子在 t 时刻的位置 $X(t)$ 是一个函数 $F : \Omega \to \mathbb{R}^D$. 因此粒子在 t 时刻的位置可以被视为一个随机向量, 而粒子在维度 d 上的位置 $X^d(t)$ 是一个随机变量.

定义 2. *随机过程是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一族随机变量 $\{X(t) : t \in T\}$, 其中 T 为指标集.*

不难发现, 可以将粒子在 t 时刻在某个维度上的位置 $X^d(t)$ 视为一个随机过程.

*3.2.2. Tr-PSO 收敛性与对应随机过程稳定性 [**?**]*

根据定义 1可以得知, 粒子群收敛于某一个位置 $\hat{X}$ 等价于对于任意的 $\varepsilon > 0$ 存在 T 使得对于任意的 $t > T$ 有 $\|X(t) - \hat{X}\| < \varepsilon$. 因此 Tr-PSO 的收敛过程可以被转变为定义 3.

定义 3. *对于 t 时刻的粒子位置 $X(t)$, 若存在一个 T_0 使得 $t_1, \cdots, t_n \in T + T_0$ 以及任意的 $\varepsilon > 0, h \in T$,$(X(t_1 + h), \cdots, X(t_n + h) \in U((X(t_1), \cdots, X(t_n)), \varepsilon)$, 则认为粒子的位置是收敛的.*

定义 3的意思等价于 $X(t)$ 构成了一个基本列, 由于 Ω 实际上就是 $\mathbb{R}^D$ 因此 $X(t)$ 构成一个收敛列, 即 $X(t)$ 收敛于某点 $\hat{X}$. 不难发现上述条件等价于随机过程 $X^d(t)$ 对于每个 d 都是平稳的. 因此可以使用随机过程中的平稳概念对 Tr-PSO 算法的收敛性进行证明.

定义 4. *[**?**] 设 $\{X(t) : t \in T\}$ 为一个随机过程.*
1. 称 $X(t)$ 的期望 $E[X(t)]$ 为该过程的均值函数.
2. 如果 $\forall t \in T$,$E[X^2(t)]$ 存在, 则称 $\{X(t) : t \in T\}$ 为一个二阶矩过程. 此时, 称函数 $\gamma(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))], t_1, t_2 \in T$ 为该过程的协方差函数; 称 $DX(t) = \gamma(t, t)$ 为方差函数.

由于粒子位置的方差存在, 因此其二阶矩必然存在, 因此根据定义 4可知 Tr-PSO 的收敛过程是一个二阶矩过程.

定义 5. *如果随机过程 $\{X(t):t\in T\}$ 对任意的 $t_1,\cdots,t_n\in T$ 以及任意的 h(使得 $t_i+h\in T$),$(X(t_1+h),\cdots,X(t_n+h)=(X(t_1),\cdots,X(t_n))$ 则称 $\{X(t):t\in T\}$ 是严平稳的.*

结合定义 5与定义 3可以得知, 粒子群在维度 d 上的位置 $X^d(t)$ 收敛的过程在 t 足够大后近似于一个严平稳过程.

定义 6. *如果 $X(t)$ 是一个二阶矩过程, 并且均值函数 $E[X(t)]=\mu$(不依赖于 t), 协方差函数 $\gamma(t,s)$ 只与时间 $t-s$ 有关, 则称 $\{X(t)\}$ 为平稳过程或者二阶平稳过程.*

引理 3. *若 Tr-PSO 是收敛的, 当且仅当对于其每个分量 d 在 $t>T_0$ 时, 有:*
1. 随机过程 $X^d(t)$ 的均值 $EX^d(t)$ 为与 t 无关的常数.
2. 随机过程 $X^d(t)$ 的方差 $DX^d(t)$ 收敛于 0.

证明. (必要性)
根据定义 5以及定义 3可知, 在 t 足够大时, 随机过程 $X^d(t)$ 近似于一个严平稳过程, 而 $X^d(t)$ 同时还是一个二阶矩过程, 且二阶矩有限, 因此得知条件 1 成立, 且此时其方差也为一个常数. 根据定义 3可知, 在粒子位置收敛后, 无论时间如何变化, 其位置一直保持不变, 因此, 此时其方差 $DX^d(t)$ 为 0.
(充分性)
根据条件 1 与条件 2 可知, 若 $X^d(t)$ 的均值为一个常数, 方差为 0 时, 对于 $\forall t>T_0$ 有,$X^d(t)\in U(EX^d(t),\delta)$, 其中 $\delta\to 0$. 因此此时对于任意的 $\varepsilon>0$ 且 $t_1,\cdots,t_n\in T+T_0$,$h\in T$ 都有 $(X^d(t_1+h),\cdots,X^d(t_n+h))\in U((X^d(t_1),\cdots,X^d(t_n)),\varepsilon/\sqrt{n})$, 由于 $\mathbb{R}^n$ 为有限维空间, 因此在其上取 2-范数得到 $(X(t_1+h),\cdots,X(t_n+h))\in U((X(t_1),\cdots,X(t_n)),\varepsilon)$, 根据定义 3知 Tr-PSO 收敛. □

引理 3表明,Tr-PSO 的收敛过程等价于其每个维度 d 上的均值收敛于某个常数, 且方差收敛于 0. 结合引理 1得到

定理 4. *Tr-PSO 收敛的充分必要条件是: 对于每个维度 d,$\{EX^d(t)\}$ 与 $\{DX^d(t)\}$ 所对应线性递推关系的谱半径均小于 1, 且 $\{DX^d(t)\}$ 收敛于 0.*

证明. 由于 $\{X(t)\}$ 满足一个线性递推关系, 因此可以通过矩阵方法将 $\{EX^d(t)\}$ 与 $\{DX^d(t)\}$ 转变为一个线性递推关系, 结合引理 1以及引理 3. 定理立刻得证. □

定理 4表明, 可以将 Tr-PSO 的收敛过程转变为 $\{EX^d(t)\}$ 以及 $\{DX^d(t)\}$ 的收敛过程, 本文将其分别称为一阶稳定与二阶稳定 [**?**].

由于各个维度是相互独立的, 在下文中, 为简便起见, 只考虑一个维度 d 并且将使用 X_t 代替 $X^d(t)$. 因此 $X(t)$ 的均值使用 EX_t 表示, 其方差使用 DX_t 表示.

此外, 由于 Tr-PSO 总是能够将全局最优位置 P^g 保持在局部最优解上, 因此 P^g 在迭代过程中可以被视为是固定的.

3.3. 一阶稳定性分析

由于 Tr-PSO 仅仅对于 P^g 进行信赖域操作, 对于迭代过程并无影响, 并且 $r_{1,t}$ 与 $r_{2,t}$ 为均匀分布, 得到其均值为 $\frac{1}{2}$, 因此可以根据 (4) 得到 $\{EX_t\}$ 的迭代方程

$$EX_{t+1} = \left(1 + w - \frac{c_1 + c_2}{2}\right) EX_t - wEX_{t-1} + \frac{c_1 P^i + c_2 P^g}{2} \tag{6}$$

设 $x_t = [EX_t, EX_{t-1}]^T$. 下面将 (6) 转化为一个线性递推关系, 也就需要找到一个矩阵 M 使得.

$$\begin{bmatrix} EX_{t+1} \\ EX_t \end{bmatrix} = M \begin{bmatrix} EX_t \\ EX_{t-1} \end{bmatrix} + b \tag{7}$$

经计算得到 $M = \begin{bmatrix} 1 + w - \frac{c_1+c_2}{2} & -w \\ 1 & 0 \end{bmatrix}$ 以及 $b = \begin{bmatrix} \frac{c_1P^i+c_2P^g}{2} \\ 0 \end{bmatrix}$

计算出 M 的特征根为

$$\frac{1}{2}\left[(1 + w - \frac{c_1 + c_2}{2} \pm \sqrt{\left(1 + w - \frac{c_1 + c_2}{2}\right)^2 - 4w}\right] \tag{8}$$

结合引理 1当且仅当 (8) 均小于 1 时 Tr-PSO 收敛.

定理 5. *[?] 给定 $w, c_1, c_2 \geq 0$ 当且仅当 $0 \leq w < 1$ 且 $0 < c_1 + c_2 < 4(1 + w)$ 时,(8) 均小于 1, 即 EX_t 收敛.*

因此, 当 $t \to \infty$ 时 $EX_{t-1} = EX_t = EX_{t+1} = EX$,$EX = (1 + w - \frac{c_1P^i+c_2P^g}{2}) - wEX + \frac{c_1P^i+c_2P^g}{2}$ 从而得到 $EX = \frac{c_1P^i+c_2P^g}{c_1+c_2}$.

3.4. 二阶稳定性分析

根据方差的定义可以得到, 随机过程 X_t 的方差表示为

$$DX_t = E(X_t^2) - EX_t^2 \tag{9}$$

下面计算计算 $E(X_t^2)$, 首先需要计算 X_t^2.

为方便计算, 使用一些标记: 令 $C = c_1r_{1,t}+c_2r_{2,t}$,$\kappa := 1+w-C$,$P := c_1r_{1,t}P^i+c_2r_{2,t}P^g$.

根据递推关系 (4), 可以得出

$$X_{t+1}^2 = \kappa^2 X_t^2 - w^2 X_{t-1}^2 - 2\kappa w X_t X_{t-1} + 2\kappa X_t P - 2wX_{t-1}P + P^2 \tag{10}$$

接下来, 对 (10) 求均值得到

$$E(X_{t+1}^2) = E(\kappa^2 X_t^2 - w^2 X_{t-1}^2 - 2\kappa w X_t X_{t-1} + 2\kappa X_t P - 2wX_{t-1}P + P^2)$$

运用均值的线性性质, 将 X_t 有关的项与其它项分离得出

$$E(X_{t+1}^2)=E(\kappa^2)E(X_t^2)-w^2E(X_{t-1}^2)-2wE(\kappa)E(X_tX_{t-1})+2E(\kappa P)E(X_t)+2wE(P)E(X_{t-1})+E(P^2) \tag{11}$$

接下来计算 $E(\kappa^2)$,$E(\kappa)$,$E(\kappa P)$,$E(P)$.

$$\left\{\begin{array}{c} E\left(\kappa^2\right)=1+w^2+E\left(C^2\right)+2w-2E\left(C\right)-2wE\left(C\right) \\ E\left(\kappa\right)=1+w-E\left(C\right) \\ E\left(\kappa P\right)=\left(\frac{c_1}{2}P^i+\frac{c_2}{2}P^g\right)\left(1+w+E\left(C\right)\right) \\ E\left(P\right)=\frac{c_1}{2}P^i+\frac{c_2}{2}P^g \end{array}\right. \tag{12}$$

由于 $C^2=c_1^2r_{1,t}^2+c_2^2r_{2,t}^2+2c_1r_{1,t}c_2r_{2,t}$ 因此得到

$$\left\{\begin{array}{c} E\left(C^2\right)=E(c_1^2)+E(c_2^2)+\frac{c_1c_2}{4} \\ E\left(C\right)=\frac{c_1+c_2}{2} \end{array}\right. \tag{13}$$

其中 $E(c_1^2)=\mu_{r_{1,t}c_1}^2+\sigma_1^2$,$E(c_2^2)=\mu_{r_{2,t}c_2}^2+\sigma_2^2$,$\sigma_1$ 与 σ_2 为方差.

将式 (13) 带回到 (12), 并且计算出 $E(P^2)$ 得到

$$\left\{\begin{array}{c} E\left(\kappa^2\right)=1+w^2+E(c_1^2)+E(c_2^2)+2w-c_1-c_2-w\left(c_1+c_2\right)+\frac{1}{2}c_1c_2 \\ E\left(\kappa\right)=1+w-\frac{c_1+c_2}{2} \\ E\left(\kappa P\right)=\frac{(1+w)(c_1P^i+c_2P^g)}{2}-\frac{c_1c_2(P^i+P^g)}{4}-P^iE(c_1^2)-P^gE(c_2^2) \\ E\left(P\right)=\frac{c_1}{2}P^i+\frac{c_2}{2}P^g \\ E\left(P^2\right)=E(c_1^2)E\left(\left(P^i\right)^2\right)+E(c_2^2)(P^g)^2+\frac{1}{2}c_1c_2P^iP^g \end{array}\right. \tag{14}$$

式 (14) 中,$E((P^i)^2)=(P^i)^2+\sigma_i^2$ 其中 σ_i 为 P^i 的方差.

接下来利用 (4) 计算 $E(X_tX_{t-1})$

$$X_{t+1}X_t=(\kappa X_t-wX_{t-1}+P)X_t=\kappa X_t^2-wX_tX_{t-1}+PX_t \tag{15}$$

将 (15) 转化为均值的形式得到

$$E(X_{t+1}X_t)=E(\kappa)E(X_t^2)-wE(X_tX_{t-1})+E(P)E(X_t) \tag{16}$$

根据 [**?**] 的工作, 可以将前文的过程转变为一个线性递推关系.

命题 6. *(4,11,16) 所给出的关系能够转写为 (5) 所述的线性递推关系形式. 其中,* $x_t=[EX_t,EX_{t-1},E(X_t^2),E(X_tX_{t-1}),E(X_{t-1}^2)]^T$

证明. 注意到

$$x_{t+1}=\begin{bmatrix} EX_{t+1} \\ EX_t \\ E\left({X_{t+1}}^2\right) \\ E\left(X_{t+1}X_t\right) \\ E\left({X_t}^2\right) \end{bmatrix}=\begin{bmatrix} (4) \\ EX_t \\ (11) \\ (16) \\ E\left({X_t}^2\right) \end{bmatrix} \tag{17}$$

因此, 只需要找到 M 和 b 使得.

$$x_{t+1}=Mx_t+b \tag{18}$$

即可. 容易得到

$$M=\begin{bmatrix} E(\kappa) & -w & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 2E(\kappa P) & -2wE(P) & E\left(\kappa^2\right) & -2wE(\kappa) & -w^2 \\ E(P) & 0 & E(\kappa) & -w & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad b=\begin{bmatrix} E(P) \\ 0 \\ E\left(P^2\right) \\ 0 \\ 0 \end{bmatrix}$$

□

若命题 6中的递推关系是收敛的, 自然就有 $E(X_t^2)$ 与 EX_t^2 都是收敛的, 也就说明 DX_t 是收敛的, 也就是说 $\rho(M)<1$ 是 DX_t 收敛的一个必要条件.

但是, 由于矩阵 M 的特征值过于复杂, 难以进行不等式计算, 因此本文使用 [**?**] 的方法进行计算, 令 $c_1=c_2=c$ 后得到.

$$DX=\frac{c(w+1)[8\sigma_i^2+(P^g-P^i)^2]}{4(c(5w-7)+12(1-w^2))} \tag{19}$$

因此, 只需要证明 $P^i(t)$ 能够以 1 的概率收敛于 P^g, 就可以证明该序列是二阶稳定的. 接下来都假设 $c_1=c_2=c$

定理 7. *若迭代过程 DX 保证收敛, 且有 $w,c\geq 0$ 且 $c<(3w^2-2)/(w-2)$,$0<w<1$ 则迭代过程 $\{P^i(t)\}$ 收敛到 P^g 的概率为 1.*

证明. 若迭代过程 $\{DX_t\}$ 收敛, 则迭代过程 $\{EX_t\}$ 也收敛, 且收敛于 $\frac{P^i+P^g}{2}$.$c<(3w^2-2)/(w-2)$ 时

$$(P^g-EX)^2=\frac{(P^g-P^i)^2}{4}<\frac{c(w+1)[8\sigma_i^2+(P^g-P^i)^2]}{4(c(5w-7)+12(1-w^2))}=DX$$

因此, 有 $P(X_t=P^g)>0$ 根据 (1) 可以得到 $P(\lim_{t\to\infty}P^i(t)=P^g)=1$ 并且若 $P^i(t)=P^g$, 它将会保持稳定. 因此显然有迭代过程 $\{P^i(t)\}$ 以 1 的概率收敛于 P^g. □

4. 参考文献

References

[1] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of ICNN'95-international conference on neural networks, Vol. 4, IEEE, 1995, pp. 1942–1948.

[2] Y.-x. Yuan, Recent advances in trust region algorithms, Mathematical Programming 151 (2015) 249–281.

[3] W. Rudin, Functional analysis 2nd ed, 1991.

[4] M. R. Bonyadi, Z. Michalewicz, Stability analysis of the particle swarm optimization without stagnation assumption, IEEE Transactions on Evolutionary Computation 20 (5) (2015) 814–819.

[5] L. J. Allen, An introduction to stochastic processes with applications to biology, CRC press, 2010.

[6] M. Jiang, Y. P. Luo, S. Y. Yang, Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm, Information processing letters 102 (1) (2007) 8–16.

[7] R. Poli, Mean and variance of the sampling distribution of particle swarm optimizers during stagnation, IEEE Transactions on Evolutionary Computation 13 (4) (2009) 712–721.