

惯性权值对粒子群算法收敛性的影响及改进

黄翀鹏, 熊伟丽, 徐保国

(江南大学通信与控制工程学院, 无锡 214122)

摘要: 研究惯性权值对粒子群算法(PSO)收敛性的影响, 在分析线性权值递减策略基础上, 提出一种基于各粒子适应值的递减策略——FDIW。标准测试函数对比实验表明, 该策略可以使粒子在搜索初期获得更好的多样性, 从而使粒子具有更强的摆脱局部极值的能力, 在搜索末期可以加快粒子收敛速度以提高 PSO 算法的性能。

关键词: 粒子群优化算法; 惯性权值; 递减策略; 适应值

Influence of Inertia Weight on Astringency of Particle Swarm Algorithm and Its Improvement

HUANG Chong-peng, XIONG Wei-li, XU Bao-guo

(School of Communication and Control Engineering, Jiangnan University, Wuxi 214122)

【Abstract】 Inspired by studying the effect of inertia weight in convergence of Particle Swarm Optimization(PSO) and analyzing advantages and disadvantages of the existing Linear Decreasing Inertia Weight(LDIW), a new modified strategy for inertia weight is proposed based on fitness——FDIW. Comparative experiments of benchmark functions indicate that for most continuous optimization problems, FDIW has more varieties of the swarm at the early stages so that it can escape from local minimum more easily. FDIW can speed up the convergence of particles at the later stages to improve the performance of PSO.

【Key words】 Particle Swarm Optimization(PSO) algorithm; inertia weight; decreasing strategy; fitness

文献[1]提出粒子群算法(Particle Swarm Optimization, PSO), 具有计算简洁、易于实现、需调整参数少等特点。惯性权值 w 的选取影响该算法性能和效率, 先后出现了线性递减 w (LDIW)^[2]、模糊 w (FIW)^[3]和随机 w (RIW)^[4]等改进策略。其中, FIW需要专家知识建立模糊规则, 实现难度较大; RIW用于求解动态系统; 而LDIW相对简单且收敛快, 采用最为广泛。本文在LDIW基础上, 从理论和实验两方面分析了 w 对PSO收敛性的影响, 提出基于各粒子适应值的 w 改进策略。该策略通过区别对待同次迭代中适应值好和适应值差的粒子, 以充分发挥其自身优势。实验证明: 该策略在加大粒子收敛性的同时, 有效地减少早熟现象的发生, 并取得了满意的结果。

1 标准粒子群算法

标准 PSO 算法描述为

$$v_{id}^{(t+1)} = wv_{id}^{(t)} + c_1rand()(pbest_{id}^{(t)} - x_{id}^{(t)}) + c_2Rand()(gbest_{id}^{(t)} - x_{id}^{(t)}) \quad (1)$$

$$x_{id}^{(t+1)} = x_{id}^{(t)} + v_{id}^{(t+1)} \quad (2)$$

其中, $v_{id}^{(t)}$ 和 $x_{id}^{(t)}$ 表示 t 时刻微粒 i 速度向量和位置向量的第 d 维分量; $pbest_{id}^{(t)}$ 和 $gbest_{id}^{(t)}$ 表示微粒 i 和种群历史最优位置向量的第 d 维分量; $c_{1,2}$ 是统计加速度权重, 通常取 2; $rand()$ 和 $Rand()$ 是介于(0,1)间的随机数。

w 惯性权值决定着 $v_{id}^{(t)}$ 对 $v_{id}^{(t+1)}$ 的影响程度。经验认为, w 较大有利于粒子群的全局搜索; w 较小有利于粒子群的局部搜索。本文将从理论角度出发, 辅以微观层面对单个粒子在搜索空间运动过程进行分析, 来对该经验加以验证, 并从中得到启发。

2 惯性权值对粒子群算法性能的影响分析

2.1 理论分析

为便于分析, 笔者将PSO简化为一维、单个粒子的运动。同时可以令 $\varphi_1 = c_1rand()$, $\varphi_2 = c_2Rand()$, $\varphi = \varphi_1 + \varphi_2$, $\bar{\varphi} = 1 - \varphi$, $\varphi_p = \varphi_1pbest + \varphi_2gbest$ 。为加强针对性, 令 $\varphi, \bar{\varphi}, \varphi_p$ 为固定常数, 使得 w 作为式(1)、式(2)中的唯一变量, 利用状态空间将其描述为

$$Y(t+1) = AY(t) + B \quad (3)$$

$$\text{其中, } Y(t) = \begin{bmatrix} x(t) \\ v(t) \end{bmatrix}; A = \begin{bmatrix} \bar{\varphi} & w \\ -\varphi & w \end{bmatrix}; B = \begin{bmatrix} \varphi_p \\ \varphi_p \end{bmatrix}。$$

矩阵 A 的特征方程为

$$\lambda^2 - (w + \bar{\varphi})\lambda + w = 0 \quad (4)$$

矩阵 A 的两个特征值为

$$e_{1,2} = \frac{w + \bar{\varphi} \pm \sqrt{\Delta}}{2}$$

其中,

$$\Delta = (w + \bar{\varphi})^2 - 4w \quad (5)$$

式(3)的显式表示为

基金项目: 国家“863”计划基金资助项目“基于智能无线传感器网络的温室环境监控系统平台及其关键技术研究”(10Z22006AA10Z335); 国家“863”计划基金资助重大项目“车载农田土壤信息快速采集关键技术与产品研发”(10A32006AA10A31); 江苏省自然科学基金资助项目“多肽分离纯化粒子群算法优化的研究”(BK2005012)

作者简介: 黄翀鹏(1982-), 男, 博士研究生, 主研方向: 智能控制与优化; 熊伟丽, 博士; 徐保国, 教授、博士生导师

收稿日期: 2008-01-02 **E-mail:** chongpeng@126.com

$$\begin{cases} x(t) = \frac{\varphi_p - [k_1 e_1^t (e_1 - w) + k_2 e_2^t (e_2 - w)]}{\varphi} \\ v(t) = k_1 e_1^t + k_2 e_2^t \end{cases} \quad (6)$$

由式(6)知, $x(t)$ 和 $v(t)$ 的收敛速率完全取决于 $\|e_{1,2}\|$, 值越小, 收敛越快。显然 $\|e_{1,2}\|$ 的大小与 w 的取值成正比, w 对收敛性的影响与经验结论相一致。

2.2 实验分析

虽然上述理论分析与经验相一致, 但 PSO 中 $pbest_{id}^{(t)}$ 和 $gbest_{id}^{(t)}$ 是不断寻优的变量。同时, 正因为 $\varphi_{1,2}$ 所具有的不确定性才给整个算法带来了多样性和创新。因此, 将它们视为常量的理论具有一定局限性。一旦加入如此多的不确定因素, 基于理论上的分析必将变得格外复杂, 很难用常规手段处理。为能直观了解动态系统中 w 对粒子运动轨迹的影响, 现对 Rose 函数最小寻优。Rose 函数为

$$f(x) = 100(x^2 - x)^2 + (1 - x)^2 \quad (7)$$

图 1、图 2 分别给出了 $w = 0.4$ 和 $w = 0.9$ 时, 粒子在搜索空间的运动过程。可见, 当 $w = 0.4$, 所有粒子收敛, 且收敛速率较快, 但粒子振荡幅值小, 一旦搜索维数较大, 或遇到具有大量局部最优的多峰函数时, 往往易陷入局部最优。当 $w = 0.9$, 少部分粒子收敛, 且收敛速率较慢; 大量粒子始终处于运动状态, 这是由于 $\varphi_{1,2}$ 随机变化的缘故; 同时粒子振荡幅值大, 几乎遍及整个搜索空间。因此, 以上微观层面的分析, 亦与前文的经验结论相一致。

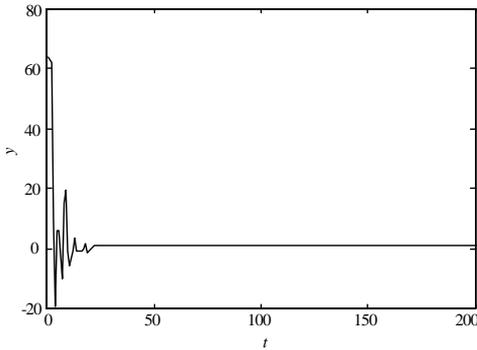


图 1 $w = 0.4$ 时粒子的运动过程

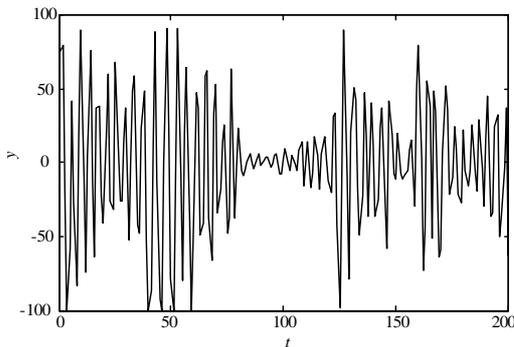


图 2 $w = 0.9$ 时粒子的运动过程

3 惯性权值取值的改进策略

3.1 线性递减策略

目前, 对 w 的改进策略使用最为广泛的是: 文献[2]提出 LDIW 策略, 表示为

$$w = w_{end} + (w_{start} - w_{end})(1 - t/t_{max}) \quad (8)$$

其中, t_{max} 为最大允许迭代次数; t 为当前的迭代次数; w_{start} 和

w_{end} 是初始惯性权值和进化到最大允许迭代次数的取值。

从 $\|e_{1,2}\|$ 角度看, 该策略在迭代初期, 由于 w 较大, $\|e_{1,2}\|$ 也较大, 有利于全局寻优; 在迭代末期, 由于 w 较小, $\|e_{1,2}\|$ 也较小, 有利于局部寻优。从对 Rose 函数寻优看, 见图 3。

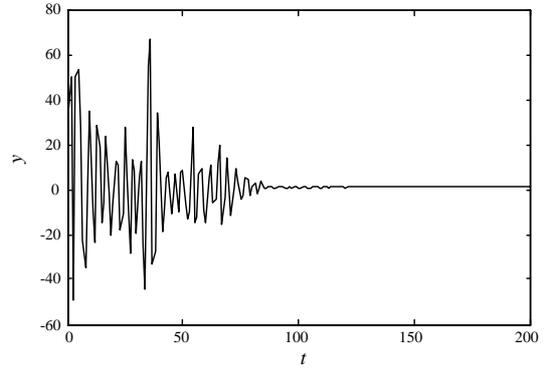


图 3 LDIW 策略

算法存在两个缺陷:

(1) 迭代初期局部搜索能力较弱, 即使初始粒子已接近全局最优, 也往往会被错过, 而在迭代后期, 则因全局搜索能力变弱, 而易陷入局部最优。

(2) 式(8)的最大迭代次数 t_{max} 较难预测, 从而影响算法的性能。

3.2 基于各粒子适应值权值改进策略

综上所述, 为进一步提高 PSO 算法性能, 本文提出一种新的 w 的改进策略——基于各粒子适应值的权值改进策略 (Fitness Decreasing Inertia Weight, FDIW)。

由文献[1]可知, PSO 算法源自鸟群群体运动行为模拟。正如鸟的觅食, 每只鸟都作为独立的个体而存在, 它们有的距食物较近, 放慢速度以确定食物的具体方位; 有的距食物较远, 因此, 提高速度以尽快找出食物的大致方向。同样在 PSO 算法的迭代过程中, 各粒子间亦存在着这种差异。综合对鸟群的分析, 以及“比较优势”可知: 当粒子位置较优, w 应取值较小, 使该粒子以局部搜索为主; 当粒子位置较差, w 应取值较大, 使该粒子以全局搜索为主。

FDIW 策略描述为

$$w = w_{end} + (w_{start} - w_{end})(1 - q) \quad (9)$$

$$q = t/t_{max} \text{ if } (f_g / f_i) < (t/t_{max}) \&\& \text{randosome}_1 = 1 \parallel \text{randosome}_2 = 0 \quad (10)$$

$$q = f_g / f_i \text{ if } (f_g / f_i) > (t/t_{max}) \&\& \text{randosome}_2 = 1 \parallel \text{randosome}_1 = 0 \quad (11)$$

其中, t_{max} , t , w_{start} 和 w_{end} 的定义如前所述; $\text{randosome}_{1,2}$ 为随机判断值, 取 0 或 1; w 为当前第 i 个粒子的惯性权值; q 为第 i 个粒子递减斜率; f_g , f_i 分别对应于 $gbest_{id}^{(t)}$ 和 $pbest_{id}^{(t)}$ 。式(9)反映了迭代过程中种群 w 变化的大势, 与 LDIW 策略类同; 式(10)、式(11)反映了种群中各粒子 w 的变化, 将其嵌入到式(9)中, 是群体学中个体服从群体的诠释。

在 FDIW 策略中 w 的大小由粒子适应值和迭代次数共同决定, 是对缺陷(2)的克服。在迭代初期, 如果某粒子适应值较好, 则 w 由适应值决定, 其值较小, 因此, 在大部分粒子进行全局寻优时, 该粒子可以先行进行局部寻优; 同时, 本文在大部分粒子 w 值参照式(10)、式(11), 随机选取少量粒子, 通过令 $\text{randosome}_1 = 0$ 或 $\text{randosome}_2 = 0$, 使得它的 w 值逆向变化, 以加入扰动, 使迭代后期仍有不少粒子具有一定的速度, 在相对范围较大的空间进行搜索, 是对缺陷(1)进行的策略。

该策略基于多种群、多迭代的 PSO 算法。因此，仅对单一维度的 Rose 函数寻优，很难反映出其优越性。本文将在下面仿真实验中，通过标准函数的来测试这一策略。

4 仿真实验

本文通过两个经典的函数优化问题来测试 FDIW 改进策略对 PSO 性能上的改善。测试函数的具体数学描述如下：

Sphere 函数为

$$f(x) = \sum_{i=1}^n x_i^2 \quad (12)$$

Rosenbrock 函数为

$$f(x) = \sum_{i=1}^n [100 \times (x_{i+1} - x_i^2)^2 + (x_i - 1)^2] \quad (13)$$

其中，算法的其他参数设置如表 1 所示。第 1 代种群采用随机初始化的方法，各函数直接将函数表达式作为适应度函数，结束条件为最大迭代次数。

表 1 函数搜索范围

函数	范围	初始边界	V_{max}	达标值
Sphere	[-100,100]	[50,100]	100	1.00E-01
Rosenbrock	[-100,100]	[15,30]	100	1.00E+03

每个基准函数分别设置维度为 10, 20, 30；相应的迭代次数为 1 000, 1 500, 2 000；相应的粒子数为 20, 40, 80。分别运用 $w = 0.7$ 时的 SPSO、 $w_{start} = 0.9, w_{end} = 0.4$ 的 LPSO 和 FPSO 对其进行测试。图 4、图 5 显示了对最后一种情况的对比。

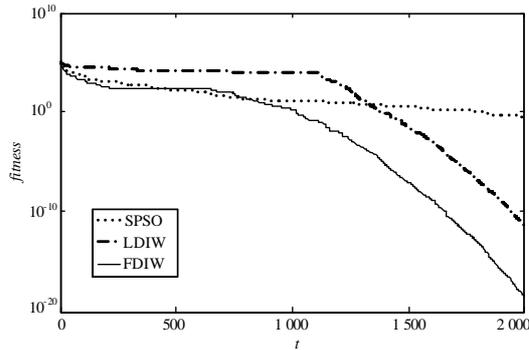


图 4 Sphere 中的对比情况

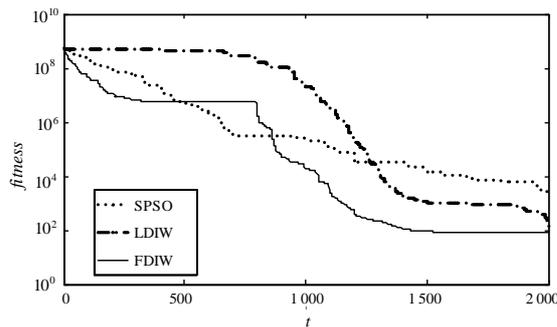


图 5 Rosenbrock 中的对比情况

在运行初始阶段，FPSO 和 SPSO 的收敛速度要优于 LPSO，主要是因为 SPSO 中 w 较小，而在 FPSO 中由适应值决定的较优粒子 w 亦较小，所以这两种方案初始时寻优能力要强于 LPSO；同时，SPSO 到达一定代数后趋于稳定，可能是因为搜索前期 w 相对较小缺乏一定的全局搜索，而后期 w 又相对较大难以收敛的缘故。而 FPSO 由于其迭代初期大部分粒子 w 较大仍以全局搜索，且 w 基本递减；在后期又对部分粒子加入扰动，使得在整体以局部搜索为主时，仍有少部分进行全局搜索，因此，其可以在比 LPSO 迭代次数少的情况下取

得高精度的解。但稍显不足的，在部分函数中 FPSO 和 LPSO 的最终收敛精度趋同，这一方面可能是由于算法本身在多次迭代后具有一定收敛性，另一方面可能是 w_{start} 和 w_{end} 的取值并非 FPSO 的最佳架构，因为作为 FPSO 算法其在迭代过程中均有反向粒子存在，所以在 FPSO 中 $w_{start} \sim w_{end}$ 可取更大范围的值。现取 $w_{start} = 1.0$ 和 $w_{end} = 0.3$ 。表 2、表 3 是 SPSO, LPSO, FPSO 3 种算法对于上述基本函数不同种群、不同维数、不同迭代次数的寻优精度及收敛率。从表 2、表 3 中数据可见，当 FPSO 中 w 具有更大范围时，在迭代次数相同的情况下，往往可以获得更好的取值精度。

表 2 Sphere 中 3 种算法的比较

M	D	G_{max}	SPSO		LPSO		FPSO	
			寻优精度	收敛率/(%)	寻优精度	收敛率/(%)	寻优精度	收敛率/(%)
20	10	1 000	1e-08	84	1e-21	90	1e-24	88
	20	1 500	1e-01	28	1e-10	24	1e-14	30
	30	2 000	—	—	—	—	1e-07	14
40	10	1 000	1e-10	100	1e-25	100	1e-29	100
	20	1 500	1e-02	62	1e-15	52	1e-22	58
	30	2 000	—	—	1e-10	12	1e-16	14
80	10	1 000	1e-12	100	1e-29	100	1e-36	100
	20	1 500	1e-03	90	1e-18	78	1e-27	76
	30	2 000	1e-01	16	1e-13	28	1e-22	28

表 3 Rosenbrock 中 3 种算法的比较

M	D	G_{max}	SPSO		LPSO		FPSO	
			寻优精度	收敛率/(%)	寻优精度	收敛率/(%)	寻优精度	收敛率/(%)
20	10	1 000	9.63	80	9.45	90	7.77	94
	20	1 500	414.77	50	53.68	72	36.56	80
	30	2 000	—	—	257.75	60	137.23	76
40	10	1 000	9.50	86	8.18	94	6.24	98
	20	1 500	217.43	66	42.43	74	25.23	84
	30	2 000	—	—	144.96	60	86.94	78
80	10	1 000	7.28	98	7.77	96	5.29	98
	20	1 500	208.25	72	32.73	84	17.89	90
	30	2 000	—	—	37.80	80	54.43	86

5 结束语

笔者目前在 PSO 算法方面做了较多改进^[5]。本文研究了惯性权值因子大小对粒子群优化算法 (PSO) 收敛性所带来的影响，分析线性权值递减策略 (LDIW)，提出一种基于各粒子适应值的递减策略——FDIW，使得 PSO 算法在寻优能力和精度要求上都有不同程度的提高。但 w_{start} 和 w_{end} 取何值时，PSO 算法的性能会最优，及其在全局和局部间达到最佳平衡，还需要大量实验来验证。同时 PSO 算法本身缺乏理论依据，具有一定的盲动性和随机性，如何解决这一缺陷也是需要着重研究的问题。

参考文献

- [1] Eberhart R C. A New Optimizer Using Particle Swarm Theory[C]// Proceedings of the 6th International Symposium on Micro and Human Science. Nagoya, Japan: [s. n.], 1995: 39-43.
- [2] Eberhart R C. Empirical Study of Particle Swarm Optimization[C]// Proc. of International Conference on Evolutionary Computation. Washington D. C., USA: IEEE Press, 1999: 1945-1950.
- [3] Eberhart R C. Fuzzy Adaptive Particle Swarm Optimization[C]// Proc. of the IEEE Congress on Evolutionary Computation. San Francisco, USA: IEEE Press, 2001: 101-106.
- [4] Shi Y. Tracking and Optimizing Dynamic Systems with Particle Swarms[C]// Proc. of the IEEE Congress on Evolutionary Computation. San Francisco, USA: IEEE Press, 2001: 94-100.
- [5] 熊伟丽, 徐保国, 周其明. 基于改进粒子群算法的 PID 参数优化方法研究[J]. 计算机工程, 2005, 31(24): 41-43.