A Simple Way for Parameter Selection of Standard Particle Swarm Optimization

Wei Zhang, Ying Jin, Xin Li, and Xin Zhang

College of Chemistry and Chemical Engineering,
Taiyuan University of Technology,
Taiyuan 030024, P.R. China
zhangwei01@tyut.edu.cn

Abstract. A simple way is proposed to estimate the non-negative real parameter tuple {ω, c1, c2} of standard Particle Swarm Optimization algorithm using control theory. The distribution of complex characteristic roots on the convergence region of particles is studied by means of linear discrete-time system analysis method. It is pointed out that the critical factors affecting the modulus value and the phase angle of the complex characteristic roots are the maximum overshoot and angular frequency of damped oscillation. The way shows that the product of the maximum overshoot and the angular frequency of damped oscillation approximately equaling to 1 is the promising guideline for parameter selection in PSO when the angular frequency in the range of (0.65π, 0.35π). Based on this, widely used benchmark problems are employed in series experiments using a stochastic approximation technique, and the results are well back above deduction.

Keywords: particle swarm optimization, statistical experiments, parameter selection.

1 Introduction

Particle swarm optimization (PSO) has been shown to be an efficient, robust and simple optimization algorithm for finding optimal regions of complex search spaces through the interaction of individuals in a population of particles. Accelerating convergence rate and avoiding local minima or prematurely are two main aspects in PSO. Clerc and Kennedy [1] mathematically analyzed the stochastic behavior of the PSO algorithm in stagnation. Trelea [2] analyzed the dynamic behavior and the convergence of the simplified PSO algorithm using standard results from the discrete-time dynamic system theory, and provided a parameter set ($\omega = 0.6, c_1 = c_2 = 1.7$) in the algorithm convergence domain. M. Jiang et al. [3] studied the stochastic convergence property of the standard PSO algorithm, and gave a sufficient condition to ensure the stochastic convergence of the particle swarm system. And then, according to the analysis result, a set of suggested parameters (ω=0.715, $c_1 = c_2$= 1.7) was given in another literature [4]. J. L. Fern´andez Mart´ınez et al. proved the same stability regions under stagnation and with a moving center of attraction. They also

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 436–443, 2011.

pointed out that properties of the second-order moments variance and covariance served to propose some promising parameter sets and proposed a good parameter region of inertia value and acceleration coefficients [5]. Above reports provide insights into how particle swarm system works based on mathematical analyses. Besides, the oscillation properties also have important influence on optimization process, while its analysis in optimization process based on control theory was seldom reported by far.

The rest of the paper is organized as follows. Section 2 surveys the standard PSO in the z-plane according to the control theory. A simple principle to find the best parameter values in particle swarm optimization based is presented in Section 3. Section 4 presents the experimental results using seven benchmark functions. Finally, Section 5 concludes the paper.

2 Analysis of Particle Swarm Optimization

2.1 The Difference Equations of Standard PSO Algorithm

PSO uses a set of particles, representing potential solutions to solve the optimization problem. The particles move around in a multidimensional search space with a position x^t_{id} and a velocity v^t_{id}, where $i=1,2,\ldots,N$ represents the index of the particle, t is the time step, and $d=1,2,\ldots,D$ is the dimensionality of the search space. For each generation, the particle compares its current position with the goal (global best/personal best) position and adjusts its velocity towards the goal with the help of the explicit memory of the best position ever found both globally and individually. Then the updating of velocity and particle position can be obtained by using the two following equations

$$v_{id}^{t+1} = \omega \times v_{id}^{t} + c_1 \times r1_{id}^{t} \times \left(p_{id}^{t} - x_{id}^{t} \right) + c_2 \times r2_{id}^{t} \times \left(p_{gd}^{t} - x_{id}^{t} \right) \tag{1}$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \tag{2}$$

where c_1 and c_2 are positive constants, defined as acceleration coefficients; ω is the inertia weight introduced to accelerate the convergence speed of PSO algorithm; $r1^t_{id}$ and $r2^t_{id}$ are two random functions in the range of [0,1]; p^t_{id} is the best previous position of x^t_{id}; p^t_{gd} is the position of the best particle among the entire population.

During stagnation, each particle behaves independently and each dimension is treated independently too. So updated equations are rewritten as

$$v_{t+1} = \omega v_t + c_1 r_{1,t} \left(p_i - x_t \right) + c_2 r_{2,t} \left(p_g - x_t \right) \tag{3}$$

$$x_{t+1} = x_t + v_{t+1} \tag{4}$$

By substituting Eq. (3) into Eq. (4), the following non-homogeneous recurrence relation is obtained:

$$x_{t+1} + \left(c_1 r_{1,t} + c_2 r_{2,t} - 1 - \omega \right) x_t + \omega x_{t-1} - c_1 r_{1,t} p_i - c_2 r_{2,t} p_g = 0 \tag{5}$$

Applying the expectation operator to both sides of the Eq. (5), obtaining

$$Ex_{t+2}+\left(\frac{c_1+c_2}{2}-1-\omega\right)Ex_{t+1}+\omega Ex_t-\frac{c_1p_i-c_2p_g}{2}=0 \tag{6}$$

According to the z–transform of the second-order difference Eq. (6), the expectation of *x(z)* is

$$Ex(z)=\frac{z^2x_0+zx_1+\left(\frac{c_1+c_2}{2}-\omega-1\right)zx_0+\frac{c_1p_i+c_2p_g}{2}\frac{z}{z-1}}{z^2+\left(\frac{c_1+c_2}{2}-\omega-1\right)z+\omega} \tag{7}$$

and the corresponding characteristic equation is

$$z^2+\left(\frac{c_1+c_2}{2}-\omega-1\right)z+\omega=0 \tag{8}$$

Let $c=c_1=c_2$, the solutions of the corresponding characteristic equation give the eigenvalues

$$z_{1,2}=\frac{1+\omega-c\pm\sqrt{\Delta}}{2} \tag{9}$$

where

$$\Delta=(c-\omega-1)^2-4\omega \tag{10}$$

The positions of eigenvalues in the z-plane affect the dynamic characteristic of Ex_k, and it can be discussed in two cases , both eigenvalues are real ones or complex ones .

2.2 Dynamic Characteristic Analysis of Complex Eigenvalues

According to the time-domain analysis of linear systems, the complex eigenvalues can be expressed as

$$z_{1,2}=e^{(-\sigma\pm j\omega_d)T}=\left|z_{1,2}\right|e^{\pm j\omega_d T} \tag{11}$$

and the model of the complex eigenvalues is

$$\left|z_{1,2}\right|=e^{-\sigma}=e^{-\xi\omega_n} \tag{12}$$

where $\sigma=\xi\omega_n$ is attenuation coefficient, $\omega_d=\omega_n\sqrt{1-\xi^2}$ is angular frequency of damped oscillation, ω_n is natural frequency, $\xi(0<\xi<1)$ is damping ratio in control .

From inverse z-transform, the transient component of the complex eigenvalues can be derived as:

$$Ex(nT) = -\frac{1}{\sqrt{1-\xi^2}} e^{-\xi\omega_n nT} \sin(\omega_d nT + \beta) \quad (13)$$

where $\beta = \arccos\xi$, T is sampling period.

The maximum overshoot

$$M = e^{-\pi\xi / \sqrt{1-\xi^2}} \times 100\% \quad (14)$$

denotes the maximum peak value of the response. It is decided by system's damping degree, and the greater the value of ξ is, the smaller the maximum overshoot will be.

The characteristics of Eqs.(12),(13) and (14) can be summarized as follows:

(1) The complex eigenvalues locate inside the unit circle in the z-plane when $|z_{1,2}| < 1$ and then the dynamic response $Ex(nT)$ is a periodic pulse sequence with damping process. The smaller the value of $|z_{1,2}|$ is, the closer the complex eigenvalues to the origin of the z-plane, and then it will cause the inevitable result of rapid convergence. The convergence rate becomes slower when the value of $|z_{1,2}|$ approaches to 1 and continuous oscillation will occur when $|z_{1,2}| = 1$.

(2) The dynamic characteristic of $Ex(nT)$ decided by the complex eigenvalues is sinusoidal oscillation with the angular frequency ω_d. The fact that the value of ω_d is too small to favor the system overcoming premature convergence and a much greater one causes the system oscillate seriously and even incapable convergence in limited optimization period.

(3) The maximum overshoot is only a function of the damping ratio ξ as shown in Eq.(14).The greater the maximum overshoot is, the bigger the oscillation amplitude will be. The value of M is too great to favor the system fast stabilization, and too small to favor the system optimization.

The relationships between the distribution of complex eigenvalues and their corresponding dynamic responses are shown in Fig.1.

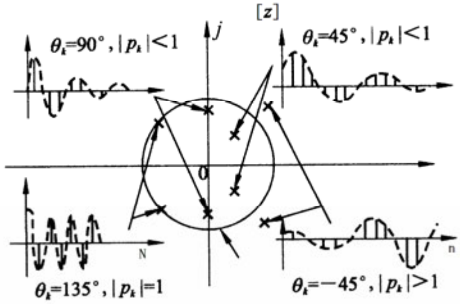

Fig. 1. The distribution of complex characteristic roots and their corresponding dynamic responses

According to the analysis results above, it can be concluded that the maximum overshoot M and the angular frequency of the damped oscillation ω_d affect the optimization behavior mainly. When the value of ω_d is heightened, the maximum overshoot M should be reduced. Conversely when the value of ω_d is decreased, the maximum overshoot M should be increased.

3 A New Simple Parameter Selection Guideline

Based on the theoretical analysis results obtained above and corresponding experimental results, the new guidelines for parameter selection are proposed as following and detailed discussion are made to show the validity of the new guidelines in [6].

$$M \times \omega_d = 1 \qquad \omega_d \in (0.65\pi, 0.35\pi) \tag{15}$$

According to Eq.(15), PSO algorithm will search the solution space thoroughly, and find the optima with higher probability. The way to get the certain values of ω and c from the relationship of M and ω_d in Eq.(15) is shown as following.

As discussed above, the eigenvalues can be easily obtained from Eqs. (9) and (10) when $\Delta < 0$ .The solving formula is

$$z_{1,2} = \frac{1+\omega-c \pm j\sqrt{4\omega-(c-\omega-1)^2}}{2} \tag{16}$$

Let the sampling period $T=1s$, and from Eqs. (16) and (12) the model of eigenvalues can be expressed as

$$\left|z_{1,2}\right| = \left|\frac{1+\omega-c \pm j\sqrt{4\omega-(c-\omega-1)^2}}{2}\right| = \sqrt{\omega} = e^{-\xi\omega_n} \tag{17}$$

The angular frequency of the damped oscillation in the z-plane is

$$\omega_d = \left(180° + \arctan\frac{\sqrt{4\omega-(c-\omega-1)^2}}{1+\omega-c}\right) \times \frac{\pi}{180}(\text{rad}) = \omega_n\sqrt{1-\xi^2} \tag{18}$$

The solving process can be summarized as following.

(1) Choosing the value of ω_d arbitrarily, such as 0.65π, 0.6π, 0.55π, 0.5π, 0.45π, 0.4π, 0.35π listed in table 1.
(2) Getting the corresponding maximum overshoot M according to Eq.(15).
(3) Then the damping ratio ξ can be calculated from Eq.(14).
(4) The corresponding values of ω and c can be calculated respectively according to Eqs. (18) and (17) finally.

The detailed data are listed in table 1 .

Table 1. Data according to the simple parameter selection guidelines

ω_d	ω	c	$M \times \omega_d$
±0.65π	0.398	1.971	1
±0.60π	0.469	1.892	1
±0.55π	0.551	1.783	1
±0.50π	0.640	1.640	1
±0.45π	0.736	1.468	1
±0.40π	0.836	1.271	1
±0.35π	0.821	0.998	1
0.456π [1]	0.729	1.494	1
0.521π [2]	0.600	1.700	1
0.473π [4]	0.715	1.700	1

4 Optimization Strategies and Experiment Results

4.1 Test Conditions

The detailed information of three benchmark functions are summarized in Table 2. A fully connected topology (all particles being neighbors) was used in all cases. For each function, the population sizes were set to 30. Defined the maximum velocity according to equation $v_{max}=x_{max}$. The optimization process will stop when the error goal is reached or the numbers of iterations reach 5000.

Table 2. Typical test functions

Name	Formula	Dim.	Range	Error goal
Rastrigin	$f(x)=\sum_{i=1}^{n}\left(x_i^2-10\cos(2\pi x_i)+10\right)$	30	$[-5.12,5.12]^n$	10^2
Griewank	$f(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	30	$[-600,600]^n$	10^{-1}
Schaffer's f6	$f(x)=0.5+\frac{\left(\sin\sqrt{x_1^2+x_2^2}\right)^2-0.5}{\left(1+0.001\left(x_1^2+x_2^2\right)\right)^2}$	2	$[-100,100]^2$	10^{-5}

4.2 Experimental Results

New simple guidelines for parameter selection in PSO are proposed based on the dynamics characteristic analyses of the eigenvalues in the z-plane according to the control theory. To further explain their validity, experimental comparisons between research results in literatures such as [1, 2, 4] and the new simple guidelines are carried out as following. For each setting, 20 runs are performed. During each run, the operation terminates when the fitness score drops below the cutoff error and it is assumed that the global minimum of the function is reached, henceforth; the score is set to 0. The experimental results are listed as follows:

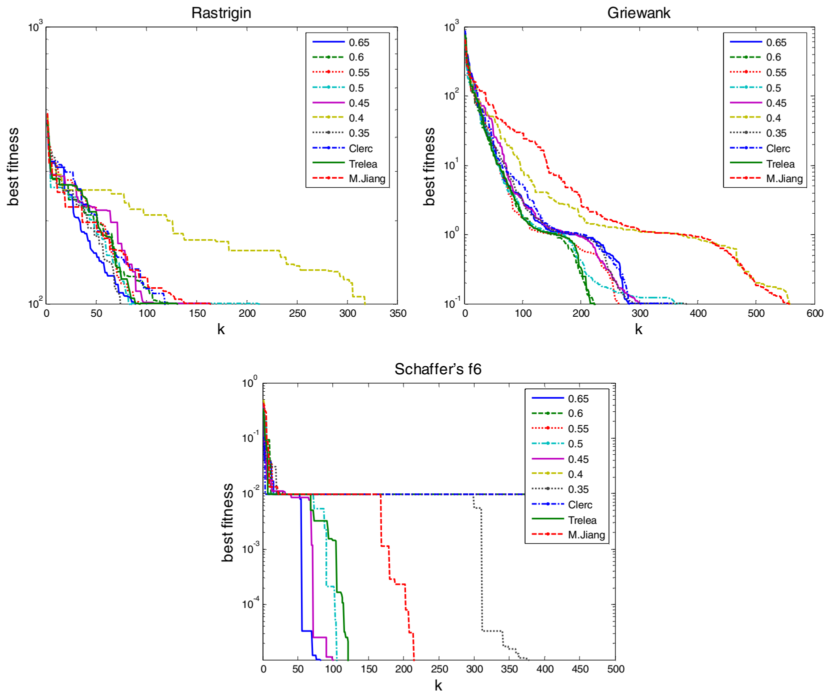

Fig. 2. Iterative comparison with functions of Rastrigin, Griewank and Schaffer's f6

Table 3. Statistical results from 20 runs of the functions

	d	0.65	0.6	0.55	0.5	0.45	0.4	0.35	0.456 [1]	0.521 [2]	0.473 [4]
Rastrigin function	Suc.	1	1	1	1	1	1	1	1	1	1
	Max.	177	141	151	166	215	587	267	343	171	162
	Min.	80	69	64	57	88	159	71	101	67	166
	Mean	115	98	107	102	127	279	122	162	99	178
	St.D	23.5	21.4	21.4	26.4	28.0	103	45.5	62.5	24.8	35.1
Griewank function	Suc.	1	1	0.85	0.95	1	1	0.65	1	0.90	1
	Max.	424	374	339	457	557	724	639	497	337	593
	Min.	243	227	239	225	273	466	216	274	209	423
	Mean	309	289	279	309	328	564	402	356	276	514
	St.D	42.8	39.7	36.5	63.7	64.1	66.8	133	57.4	36.4	50.7
Schaffer's f6 function	Suc.	0.60	0.45	0.60	0.45	0.85	0.75	0.70	0.30	0.55	0.70
	Max.	431	358	440	295	526	554	480	491	460	271
	Min.	4	53	7	9	4	5	5	5	68	5
	Mean	116	184	198	134	145	179	139	200	169	119
	St.D	104	119	133	86	129	182	131	182	140	98

The statistical comparison between the new simple guidelines and research results in literatures are reported in tables 3. Iterative comparisons among new simple way mentioned in this paper and other parameter selection strategies with three functions are showed in Fig.2. The statistics in the tables indicate that the parameters in accordance with the simple way show good performance in speed and reliability. But when ω_d take 0.45π and 0.35π, the values of average and variance are larger in some cases and hence can't avoid getting trapped into local optimum.

5 Conclusion

This paper has presented a simple guideline for parameter selection of standard PSO using control theory. The simple guideline is that the product of the maximum overshoot and the angular frequency of damped oscillation approximately equaling to 1 is the promising guideline for parameter selection in PSO when the angular frequency in the range of (0.65π, 0.35π). The statistical results well back the superiority of the new simple guidelines in terms of time, iterations and convergence.

References

1. Clerc, M., Kennedy, J.: The particle swarm: explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation 6(1), 58–73 (2002)
2. Trelea, I.C.: The particle swarm optimization algorithm: convergence analysis and parameter selection. Information Processing Letters 85, 317–325 (2003)
3. Jiang, M., Luo, Y.P., Yang, S.Y.: Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. Information Processing Letters 102, 8–16 (2007)
4. Jiang, M., Luo, Y.P., Yang, S.Y.: Particle Swarm Optimization-Stochastic Trajectory Analysis and Parameter Selection (2007), Retrieved from `http://www.i-techonline.com`
5. Fernández Martínez, J.L., García Gonzalo, E.: The Generalized PSO: A New Door to PSO Evolution. Journal of Artificial Evolution and Applications, Article ID: 861275, 15 pages (2008), doi:10.1155/2008/861275
6. Zhang, W., Li, H., Zhao, Q., Wang, H.: Guidelines for Parameter Selection in Particle Swarm Optimization According to Control Theory. In: 2009 Fifth International Conference on Natural Computation (ICNC 2009), Tianjin, China, vol. 3, pp. 520–524 (2009)