We are IntechOpen,
the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Particle Swarm Optimization – Stochastic Trajectory Analysis and Parameter Selection

M. Jiang, Y. P. Luo and S. Y. Yang
Tsinghua University
China

1. Introduction

Two important topics in Particle Swarm Optimization (PSO) research filed are trajectory analysis of particles and parameter selection method. Trajectory analysis is important because it can help to determine where the position of each particle is at each evolutionary step, and consequently it can help to clarify the running mechanism of PSO algorithm, so as to explain why and when PSO algorithm can be successful to solve optimization problems. Parameter selection of PSO algorithm is important because the performance of PSO algorithm is sensitive to the chosen parameters.

Till now, some research works have been published in literatures to investigate both of these two topics, but unfortunately, the trajectory analysis is based on simplified deterministic algorithms, regardless of the randomness in real PSO algorithm; and the parameter selection is based on experimental results instead of theoretical results.

This chapter is proposed to investigate both of these two important topics. In this chapter, the trajectory of particle in a general PSO algorithm is theoretically investigated, considering the randomness thoroughly. For arbitrary dimension d of an arbitrary particle i in the general particle swarm system, the update equations investigated in this chapter are given in Eqs. (1) and (2), where t is the evolutionary step, V is the velocity of particle i, X is the position of particle i, P_i is the history best position found by particle i, and P_g is the history best position found by the total swarm.

$$V_i^d(t+1) = \omega V_i^d(t) + c_1 r_{1,i}^d(t)(P_i^d(t) - X_i^d(t)) + c_2 r_{2,i}^d(t)(P_g^d(t) - X_i^d(t)) \tag{1}$$

$$X_i^d(t+1) = a X_i^d(t) + b V_i^d(t+1) \tag{2}$$

By regarding each particle's position on each evolutionary step as a stochastic vector, the general PSO algorithm determined by five-dimensional parameter tuple $\{\omega, c_1, c_2, a, b\}$ is formally analyzed using stochastic process theory. Because the position of particle at each step is stochastic and can not be determined directly, its expected value and variance are investigated instead of the position itself. To make the analysis possible, the particle swarm is supposed to be in stagnation phase.

At the same time, the relationship between convergent speed of particle's trajectory and parameter sets is studied. Those results give some hints on how the chosen parameters can influence the performance of PSO algorithm, and thus parameter selection guideline is

given. According to the analysis result, it is believed that parameters {a, b} would only increase the complexity of PSO algorithm without enhancing performance of PSO algorithm, thus it is suggested to fix both of them to be 1 in PSO algorithm. For more details, see Section 4.

Then, a set of suggested parameter {ω=0.715, c₁=c₂=1.7} is given, which is compared against three sets of parameters which are proposed in literatures.

The structure of this chapter is as follows. Section 2 overviews related works. Stochastic convergence analysis of the general PSO algorithm is investigated in Section 3. Some parameter selection guidelines are given in Section 4; and a set of parameters is suggested. In Section 5, experiments are conducted to compare different parameter sets. Section 6 concludes this chapter.

2. Related Works

Ozcan and Mohan published the first theoretical studies of particle trajectories (Ozcan & Mohan, 1998; Ozcan & Mohan, 1999). In their first study (Ozcan & Mohan, 1998), a simplified PSO system was considered with

- one partcle in the swarm,
- one-dimensional particles,
- $P_i=P_g=P$ keeps constant,
- no inertia weight, i.e. $\omega = 1$,
- no stochastic component, i.e. $\phi_1(t) = \phi_1 = c_1, \phi_2(t) = \phi_2 = c_2$ for all t.

Later, Ozcan and Mohan generalized their findings to a PSO system with multiple, multi-dimensional particles with P_i and P_g not necessarily the same point (Ozcan & Mohan, 1999). For the more general system, Ozcan and Mohan derived the following particle position update equation (Ozcan & Mohan, 1999):

$$X_{ij}(t) - (2 - \phi_1 - \phi_2)X_{ij}(t-1) + X_{ij}(t-2) = \phi_1 P_{ij} + \phi_2 P_{gj} \tag{3}$$

From Eq. (3), when $0 < \phi_1 + \phi_2 < 4$, the following closed form can be easily obtained:

$$X_{ij}(t) = \Lambda_{ij}\sin(\theta_{ij}t) + \Gamma_{ij}\cos(\theta_{ij}t) + \kappa_{ij} \tag{4}$$

where Λ_{ij}, Γ_{ij}, θ_{ij} and κ_{ij} are constants derived from the initial conditions and the value of ϕ_1 and ϕ_2:

$$\delta_{ij} = i\sqrt{4 - (2 - \phi_1 - \phi_2)^2} \tag{5}$$

$$\Lambda_{ij} = \frac{2V_{ij}(0) - (\phi_1 + \phi_2)X_{ij}(0) + \phi_1 P_{ij} + \phi_2 P_{gj}}{\left\|\delta_{ij}\right\|} \tag{6}$$

$$k_{ij} = \frac{\phi_1 P_{ij} + \phi_2 P_{gj}}{\phi_1 + \phi_2} \tag{7}$$

$$\Gamma_{ij} = X_{ij}(0) - k_{ij} \tag{8}$$

$$\theta_{ij} = \operatorname{atan}(\frac{\left\|\delta_{ij}\right\|}{2 - \phi_1 - \phi_2}) \tag{9}$$

The main conclusion from their work is that particle trajectories follow periodic sinusoidal waves. An optimum is searched by randomly 'catching' another wave, and manipulating its frequency and amplitude. In addition to this general finding, Ozcan and Mohan studied the trajectories of particles under a number of special cases.

For the same simple PSO system as given above, Clerc and Kennedy provided a theoretical analysis of particle trajectories to ensure convergence to a stable point (Clerc & Kennedy, 2002),

$$P = \frac{\phi_1 P_i + \phi_2 P_g}{\phi_1 + \phi_2} \tag{10}$$

the main result of this work is the introduction of the constriction coefficient, which is derived to prevent the velocity to grow out of bounds, with the advantage that, theoretically, velocity clamping is no longer required. In addition, Clerc and Kennedy also studied several different classes of constriction models. As a result of this study, the velocity update equation changes to (Clerc & Kennedy, 2002)

$$V_{ij}(t+1) = \chi[V_{ij}(t) + \phi_{1j}(t)(P_{ij}(t) - X_{ij}(t)) + \phi_{2j}(t)(P_{gj}(t) - X_{ij}(t))] \tag{11}$$

where χ is the constriction coefficient calculated as

$$\chi = \frac{2\kappa}{2 - \phi - \sqrt{\phi^2 - 4\phi}} \tag{12}$$

with $\phi = \phi_1 + \phi_2 \geq 4$ and $\kappa \in [0, 1]$. The constant κ controls the speed of convergence. For $\kappa \approx 0$, fast convergence to a stable points is obtained, while a $\kappa \approx 1$ results in slow convergence.

Both above two studies consider a simplified PSO system without an inertia weight, i.e. only the simple situation with $\omega = 1$ is studied. van den Bergh analyzed the convergence of the PSO algorithm with inertia weight (van den Bergh, 2001). Considering the velocity and position of a particle at discrete time steps, the following non-homogeneous recurrence relation is obtained:

$$X_{t+1} = (1 + \omega - \phi_1 - \phi_2)X_t - \omega X_{t-1} + \phi_1 P_i + \phi_2 P_g \tag{13}$$

The characteristic equation corresponding to the recurrence relation is

$$\lambda^2 - (1 + \omega - \phi_1 - \phi_2)\lambda + \omega = 0 \tag{14}$$

The solutions to this equation gives the eigenvalues

$$\lambda_1 = \frac{1 + \omega - \phi_1 - \phi_2 + \gamma}{2} \tag{15}$$

$$\lambda_2 = \frac{1 + \omega - \phi_1 - \phi_2 - \gamma}{2} \tag{16}$$

where

$$\gamma = \sqrt{(1+\omega-\phi_1-\phi_2)^2 - 4\omega} \qquad (17)$$

For initial conditions $X(0)=X_0$ and $X(1)=X_1$, the explicit closed form of the recurrence relation is then given by

$$X_t = k_1 + k_2\lambda_1^t + k_3\lambda_2^t \qquad (18)$$

where

$$k_1 = \frac{\phi_1 P_i + \phi_2 P_g}{\phi_1 + \phi_2} \qquad (19)$$

$$k_2 = \frac{\lambda_2(X_0 - X_1) - X_1 + X_2}{\gamma(\lambda_1 - 1)} \qquad (20)$$

$$k_3 = \frac{\lambda_1(X_1 - X_0) + X_1 - X_2}{\gamma(\lambda_2 - 1)} \qquad (21)$$

and $X_2 = (1+\omega-\phi_1-\phi_2)X_1 - \omega X_0 + \phi_1 P_i + \phi_2 P_g$.

Note that the above equation assumes that $P_i(t)=P_i$ and $P_g(t)=P_g$ for all t. The closed form representation in Eq. (18) therefore remains valid until a better position X (and thus P_i and P_g) is discovered. When a better position is discovered, the above equations can be used again after recalculating the coefficients k_1, k_2 and k_3.

van den Bergh only discussed the situation with imaginary eigenvalues, i.e. $(1+\omega-\phi_1-\phi_2)^2 < 4\omega$. He obtained the conclusion that when $\max\{\|\lambda_1\|, \|\lambda_2\|\} < 1$, the particle's position sequence will converge and

$$\lim_{t\to+\infty} X_t = k_1 = \frac{\phi_1 P_i + \phi_2 P_g}{\phi_1 + \phi_2} \qquad (22)$$

That means, under the condition that $\max\{\|\lambda_1\|, \|\lambda_2\|\} < 1$, a particle converges to a weighted average of its individual best and global best positions.

In the case that ϕ_1 and ϕ_2 are stochastic, the average behavior of the system can then be observed by considering the expected values of ϕ_1 and ϕ_2 (assuming uniform distribution): $E[\phi_1]=c_1/2$, $E[\phi_2]=c_2/2$. Using the expected values, the limit becomes

$$\lim_{t\to+\infty} X_t = \frac{c_1 P_i + c_2 P_g}{c_1 + c_2} = (1-a)P_i + aP_g \qquad (23)$$

where $a = c_2/(c_1+c_2) \in [0,1]$.

Furthermore, with experimental analysis, van den Bergh proposed parameter range to satisfy the condition of $\max\{\|\lambda_1\|, \|\lambda_2\|\} < 1$, i.e., $\omega > (c_1+c_2)/2 - 1$. van den Bergh also discussed on convergence of particle trajectory with certain parameters, but the

experimental results did not conform to his conjecture. This disagreement was informally explained by stochastic probability.

Some other similar results were proposed by Yasuda (Yasuda et al, 2003) and Trelea (Trelea, 2003). Although those results provide insights into how particle swarm system works, all those analysis discard the randomness in the PSO algorithm, and are all based on a simplified deterministic algorithm. Obviously, those analytical results more or less deviate from the real particle swarm system due to the loss of randomness. Recently, researchers have begun to make progress in the analysis of randomness in PSO algorithm. Clerc (Clerc, 2006) mathematically analyzed the stochastic behavior of the particles when the swarm is in stagnation, but he only discussed properties of stochastic coefficients and did not regard the velocity (and position) as stochastic variable, and thus he seemed unaware of the dependent relationship between velocity and the stochastic coefficients. Later, Jiang (Jiang et al, 2007a) extended Clerc's results by regarding each particle's position on each evolutionary step as a stochastic vector. Then the PSO algorithm was analyzed using stochastic process theory. Some stochastic characteristics (including expected value, variance, and auto-variance) of particle's position are obtained, both in explicit and implicit representations, and corresponding properties are analyzed.

Jiang (Jiang et al, 2007b) present the first formal stochastic convergence analysis of the standard particle swarm optimization (PSO) algorithm, which involves with randomness. By regarding each particle's position on each evolutionary step as a stochastic vector, the standard PSO algorithm determined by non-negative real parameter tuple $\{\omega, c_1, c_2\}$ was analyzed using stochastic process theory. The stochastic convergent condition of the particle swarm system and corresponding parameter selection guidelines were also derived.

3. Stochastic Convergence Analysis of PSO in Stagnation

In this section, stochastic convergence analysis of the particle swarm system is conducted, assuming the particle swarm is in stagnation. The particle swarm system is thought to be in stagnation, if arbitrary particle i's history best position P_i and the total swarm's history best position P_g keep constant over some time steps.

There exist many factors that would influence the convergence property and performance of PSO algorithm, including selection of parameter tuple $\{\omega, c_1, c_2, a, b\}$; velocity clamping; position clamping; topology of neighborhood; etc. This chapter focuses on analyzing how the selection of parameter tuple $\{\omega, c_1, c_2, a, b\}$ would influence the trajectories of particles in the PSO algorithm. Factors such as velocity clamping, position clamping, topology of neighborhood may influence the trajectories of particles, but the discussion of those factors is beyond the scope of this chapter. At the same time, the situation with variable parameter values during evolution is also not discussed here. That means, the PSO algorithm studied here is only determined by fixed real-value parameter set $\{\omega, c_1, c_2, a, b\}$. Velocity and position clamping are not considered, and the neighborhood of any particle is the whole swarm.

When the particle swarm system is in stagnation, arbitrary P_i and P_g would keep constant over some time steps, then it's easy to find out that all particles would evolve independently. Thus, only particle i needs to be studied. For i is chosen arbitrarily, the result can be applied to all other particles. At the same time, it appears from Eqs. (1) and (2) that each dimension is updated independently. Thus, without loss of generality, the algorithm

description can be reduced to the one-dimensional case. By omitting particle and dimension notations, and considering discrete time situation, update equations become:

$$V_{t+1} = \omega V_t + c_1 r_{1,t}(P_i - X_t) + c_2 r_{2,t}(P_g - X_t) \qquad (24)$$

$$X_{t+1} = aX_t + bV_{t+1} \qquad (25)$$

It's obviously that the velocity is only an auxiliary variable, and the position is the real significant variable. By substituting Eq. (24) into Eq. (25), the following non-homogeneous recurrence relation is obtained:

$$X_{t+1} = (a + \omega - c_1 b r_{1,t} - c_2 b r_{2,t})X_t - a\omega X_{t-1} + c_1 b r_{1,t} P_i + c_2 b r_{2,t} P_g \qquad (26)$$

From Eq. (26), it's easy to know that single value of coefficients c_1 and b is not important at all. The important factor is the multiple $c_1 b$. Coefficient c_2 and b share the same relationship. Thus, without loss of generality, we can choose $b \equiv 1$ if $b \neq 0$. As a matter of fact, when $b = 0$, it's easy to get $X_i^d(t+1) = a^{t+1}X_i^d(0)$ from Eq. (2), whose convergence condition is $a \leq 1$. This case is not an interesting one, thus it is supposed that $b \neq 0$ in following analysis. Thus we can suppose $b \equiv 1$, and the iteration equation becomes

$$X_{t+1} = (a + \omega - c_1 r_{1,t} - c_2 r_{2,t})X_t - a\omega X_{t-1} + c_1 r_{1,t} P_i + c_2 r_{2,t} P_g \qquad (27)$$

Notice that there exist random numbers in Eq. (27), and that X_0, X_1 are also random numbers, thus each X_t should be regarded as a random variable, and the iterative process $\{X_t\}$ should be regarded as a stochastic process. The expectation and variance of each random variable X_t can then be calculated, and the convergence property of the iterative process can be analyzed.

3.1 Convergence analysis of the expectation of particle's position

In this subsection, the iteration equation of EX_t is obtained, where EX_t is the expectation of random variable X_t. Based on the iteration equation, the convergent condition of sequence $\{EX_t\}$ is analyzed.

According to Eq. (27), iteration equation of sequence $\{EX_t\}$ can be easily obtained.

$$EX_{t+1} = (a + \omega - \frac{c_1 + c_2}{2})EX_t - a\omega EX_{t-1} + \frac{c_1 P_i + c_2 P_g}{2} \qquad (28)$$

The characteristic equation of the iterative process shown in Eq. (28) is

$$\lambda^2 - (a + \omega - \frac{c_1 + c_2}{2})\lambda + a\omega = 0 \qquad (29)$$

Theorem 1. If and only if conditions $-1 < a\omega < 1$ and $2(1-\omega)(a-1) < c_1 + c_2 < 2(1+\omega)(1+a)$ are satisfied together, iterative process $\{EX_t\}$ is guaranteed to converge to $EX = (c_1 P_i + c_2 P_g)/[2(1-\omega)(1-a) + c_1 + c_2]$.

Proof: Let $\psi = a + \omega - (c_1 + c_2)/2$, the convergent condition of iterative process $\{EX_t\}$ is that the absolute values (or complex modulus) of both eigenvalues λ_{E1}, λ_{E2} are less than 1. That is, $\left\| \psi \pm \sqrt{\psi^2 - 4a\omega} \right\| / 2 < 1$. Consider two cases:

1. $\psi^2 < 4a\omega$

Here, both eigenvalues are complex numbers, $\|\lambda_{E1}\|^2 = \|\lambda_{E2}\|^2 = (\psi^2 + 4a\omega - \psi^2)/4 = a\omega$, so $\max\{\|\lambda_1\|, \|\lambda_2\|\} < 1$ requires only $a\omega < 1$. Condition (1) itself requires $a\omega > 0$ and $2(a + \omega - 2\sqrt{a\omega}) < c_1 + c_2 < 2(a + \omega + 2\sqrt{a\omega})$.

2. $\psi^2 \geq 4a\omega$

Here, both eigenvalues are real numbers. Condition (2) is satisfied if

    $a\omega \leq 0$;

or

    $a\omega > 0$, and $c_1 + c_2 \geq 2(a + \omega + 2\sqrt{a\omega})$ or $c_1 + c_2 \leq 2(a + \omega - 2\sqrt{a\omega})$. Consider two more cases.

      $\psi < 0$

         Here $\max\{\|\lambda_{E1}\|, \|\lambda_{E2}\|\} < 1$ requires $c_1 + c_2 < \min\{2(2 + a + \omega), 2(1 + \omega)(1 + a)\}$.
         Conditions (2) and (2.1) together lead to

           $0 < a\omega < 1$ and $2(a + \omega + 2\sqrt{a\omega}) \leq c_1 + c_2 < 2(1 + \omega)(1 + a)$;

         or

           $-1 < a\omega \leq 0$ and $2(a + \omega) \leq c_1 + c_2 < 2(1 + \omega)(1 + a)$;

      $\psi \geq 0$

         Here $\max\{\|\lambda_{E1}\|, \|\lambda_{E2}\|\} < 1$ requires $c_1 + c_2 > \max\{2(\omega + a - 2), 2(1 - \omega)(a - 1)\}$.
         Conditions (2) and (2.2) together lead to

           $0 < a\omega < 1$ and $2(1 - \omega)(a - 1) \leq c_1 + c_2 < 2(a + \omega - 2\sqrt{a\omega})$;

         or

           $-1 < a\omega \leq 0$ and $2(1 - \omega)(a - 1) \leq c_1 + c_2 < 2(a + \omega)$;

         Synthesize case (1) and case (2), the guaranteed convergent condition of iterative process $\{EX_t\}$ is

           $-1 < a\omega < 1$ and $2(1 - \omega)(a - 1) < c_1 + c_2 < 2(1 + \omega)(1 + a)$.

When iterative process $\{EX_t\}$ is convergent, the convergent value EX can be calculated using $EX = (a + \omega - (c_1 + c_2)/2)EX - a\omega EX + (c_1 P_i + c_2 P_g)/2$. That gets $EX = (c_1 P_i + c_2 P_g)/[2(1 - \omega)(1 - a) + c_1 + c_2]$.

It's easy to know that, even if sequence $\{EX_t\}$ can converge, generally speaking, $EX \neq P_g$ is always true.

3.2 Convergence analysis of the variance of particle's position

To further study the convergence property of particle swarm, the variance sequence should be studied. In this subsection, the iteration equation of DX_t is obtained, where DX_t is the

variance of random variable X_t. Based on the iteration equation, the convergent condition of sequence $\{DX_t\}$ is analyzed.

In order to make the procedure of calculating DX_t clear, some symbols should be introduced firstly. Let $\varpi = a\omega$, $\upsilon = (c_1 + c_2)/2$, $\psi = a + \omega - \upsilon$, $R_t = c_1 r_{1,t} + c_2 r_{2,t} - \upsilon$, $Q_t = c_1 r_{1,t} P_i + c_2 r_{2,t} P_g$, from Eq. (27), it is easy to ontain that

$$X_{t+1} = (\psi - R_t)X_t - \varpi X_{t-1} + Q_t \tag{30}$$

Since $r_{1,t}$, $r_{2,t}$ are two independent uniform random number ranged in $[0,1]$, it's obvious that $ER_t = 0$, $EQ_t = (c_1 P_i + c_2 P_g)/2$, $DR_t = (c_1^2 + c_2^2)/12$, $DQ_t = (c_1^2 P_i^2 + c_2^2 P_g^2)/12$, and $E(R_t Q_t) = (c_1^2 P_i + c_2^2 P_g)/12$. Notice that EQ_t, DR_t, DQ_t, and $E(R_t Q_t)$ are all constants, let $Q_E = EQ_t$, $R = DR_t$, $Q_D = DQ_t$, $T = E(R_t Q_t)$.

If $R = 0$, that means $c_1 = c_2 = 0$, which is not an interesting case. Thus we suppose $R > 0$ in all following discussions .

Notice that X_t and X_{t-1} are both independent on R_t and Q_t, but X_t and X_{t-1} are dependent. Thus the following expectation can be obtained.

$$\begin{aligned} EX_{t+1}^2 = (\psi^2 + R)EX_t^2 + \varpi^2 EX_{t-1}^2 + Q_D + Q_E^2 - 2\varpi\psi E(X_t X_{t-1}) \\ + 2(\psi Q_E - T)EX_t - 2\varpi Q_E \cdot EX_{t-1} \end{aligned} \tag{31}$$

$$\begin{aligned} EX_{t+2}^2 = (\psi^2 + R)EX_{t+1}^2 + \varpi^2 EX_t^2 + Q_D + Q_E^2 - 2\varpi\psi E(X_{t+1} X_t) \\ + 2(\psi Q_E - T)EX_{t+1} - 2\varpi Q_E \cdot EX_t \end{aligned} \tag{32}$$

$$E(X_{t+1} X_t) = \psi EX_t^2 - \varpi E(X_t X_{t-1}) + Q_E \cdot EX_t \tag{33}$$

From Eqs. (31)-(33), eliminating $E(X_{t+1} X_t)$ and $E(X_t X_{t-1})$, get

$$\begin{aligned} EX_{t+2}^2 + \varpi EX_{t+1}^2 = (\psi^2 + R)(EX_{t+1}^2 + \varpi EX_t^2) + \varpi^2 (EX_t^2 + \varpi EX_{t-1}^2) \\ + (Q_D + Q_E^2)(1 + \varpi) - 2\varpi\psi(\psi EX_t^2 + Q_E EX_t) \\ + 2(\psi Q_E - T)(EX_{t+1} + \varpi EX_t) - 2\omega Q_E(EX_t + \varpi EX_{t-1}) \end{aligned} \tag{34}$$

Substitute $EX_{t+2} = \psi EX_{t+1} - \varpi EX_t + Q_E$, $\varpi EX_{t-1} = \psi EX_t - EX_{t+1} + Q_E$ and $DX_t = EX_t^2 - (EX_t)^2$ into Eq. (34), obtain

$$\begin{aligned} DX_{t+2} = (\psi^2 + R - \varpi)DX_{t+1} - \varpi(\psi^2 - R - \varpi)DX_t + \varpi^3 DX_{t-1} \\ + R \cdot [(EX_{t+1})^2 + \varpi(EX_t)^2] - 2T \cdot (EX_{t+1} + \varpi EX_t) + Q_D(1 + \varpi) \end{aligned} \tag{35}$$

The characteristic equation of the iterative process shown in Eq. (35) is

$$\lambda^3 - (\psi^2 + R - \varpi)\lambda^2 + \varpi(\psi^2 - R - \varpi)\lambda - \varpi^3 = 0 \tag{36}$$

The iteration equation and characteristic equation of iterative process $\{DX_t\}$ are both quite complex, and it is hard to analyze these two equations directly. Fortunately, the convergent condition of the iterative process $\{DX_t\}$ defined in Eq. (35) is comparatively simple. Before

discussing convergence property of iterative process {DX$_t$}, we can firstly determine the intervals in which the three eigenvalues of the characteristic equation are located. Let

$$f(\lambda) = \lambda^3 - (\psi^2 + R - \varpi)\lambda^2 + \varpi(\psi^2 - R - \varpi)\lambda - \varpi^3 \qquad (37)$$

First of all, consider two special cases.

If $\varpi = 0$, then two among three eigenvalues are zeros. Without loss of generality, let $\lambda_{D2} = \lambda_{D3} = 0$, then $\lambda_{D1} = \psi^2 + R > 0$.

If $\psi = 0$, then $\lambda_{D3} = -\varpi$ and λ_{D1}, λ_{D2} are roots of equation $\lambda^2 - R\lambda - \varpi^2 = 0$. Since $R > 0$,

get $\max\{\|\lambda_{D1}\|, \|\lambda_{D2}\|\} = (R + \sqrt{R^2 + 4\varpi^2})/2 > |\varpi|$. Then $\min\{\|\lambda_{D1}\|, \|\lambda_{D2}\|\} < |\varpi|$ must be satisfied due to $\lambda_{D1}\lambda_{D2}\lambda_{D3} = \varpi^3$.

Next consider two general cases.

When $R > 0$, $\psi \neq 0$ and $\varpi > 0$, it is easily verified that

$$f(0) = -\varpi^3 < 0 \ ; \ f(\varpi) = -2\varpi^2 R < 0; \ f(-\varpi) = -2\psi^2\varpi^2 < 0.$$

According to conclusions in elementary mathematics, because $f(-\varpi)$, $f(0)$ and $f(\varpi)$ have the same sign, the number of roots in the interval $(-\varpi, 0)$ and $(0, \varpi)$ must both be even. Thus there must be at least one root located in interval (ϖ, ∞) to satisfy $\lambda_{D1}\lambda_{D2}\lambda_{D3} = \varpi^3 > 0$.

When $R > 0$, $\psi \neq 0$ and $\varpi < 0$, it is easily verified that

$$f(0) = -\varpi^3 > 0 \ ; \ f(\varpi) = -2\varpi^2 R < 0; \ f(-\varpi) = -2\psi^2\varpi^2 < 0.$$

Likely, according to conclusions in elementary mathematics, there must be one root located in the interval $(\varpi, 0)$ and one root located in the interval $(0, -\varpi)$. The third root must be located in the interval $(-\varpi, \infty)$ to satisfy $\lambda_{D1}\lambda_{D2}\lambda_{D3} = \varpi^3 < 0$.

Without loss of generality, suppose $\|\lambda_{D1}\| \geq \|\lambda_{D2}\| \geq \|\lambda_{D3}\|$, then it is clear that relationship $\lambda_{\max D} = \lambda_{D1} > |\varpi| \geq \|\lambda_{D2}\| \geq \|\lambda_{D3}\|$ exists, and λ_{D1} must be a positive real number.

Theorem 2. Given $R > 0$, if and only if $-1 < \varpi < 1$ and $f(1) > 0$ are satisfied together, it is guaranteed that $\lambda_{\max D} = \max\{\|\lambda_{D1}\|, \|\lambda_{D2}\|, \|\lambda_{D3}\|\} < 1$.

Proof: Obviously, If $|\varpi| \geq 1$, then $\lambda_{\max D} > |\varpi| \geq 1$, which violate $\max\{\|\lambda_{D1}\|, \|\lambda_{D2}\|, \|\lambda_{D3}\|\} < 1$. Thus only cases with $-1 < \varpi < 1$ needs to be discussed. At this point, if and only if $f(1) > 0$, it is guaranteed that $\lambda_{\max D}$ is located in the interval $(|\varpi|, 1)$. □

As a matter of fact, conditions $-1 < \varpi < 1$ and $f(1) > 0$ satisfied together implies that $2(1-\omega)(a-1) < c_1 + c_2 < 2(1+\omega)(1+a)$. Because $f(1) = (1+\varpi)^2(1-\varpi) - R(1+\varpi) - (1-\varpi)\psi^2$, then $-1 < \varpi < 1$ and $f(1) > 0$ lead to $\psi^2 < (1+\varpi)^2 - (1+\varpi)R/(1-\varpi) < (1+\varpi)^2$, that is $2(1-\omega)(a-1) < c_1 + c_2 < 2(1+\omega)(1+a)$.

Theorem 3. Given $R > 0$, if and only if $-1 < \varpi < 1$ and $f(1) > 0$ are satisfied together, iterative process {DX$_t$} is guaranteed to converge to

$$DX = \frac{\frac{1}{12}(1+\varpi)}{f(1)\cdot[2(1-a)(1-\omega)+c_1+c_2]^2}\{2c_1^2c_2^2(P_g-P_i)^2 + 4(1-a)(1-\omega)c_1c_2(c_2P_g-c_1P_i)(P_g-P_i) + (c_1^2P_i^2+c_2^2P_g^2)[2(1-a)(1-\omega)]^2\}$$

where $f(1)$ is defined in Eq. (37).

Proof : The iteration equation of DX_t, Eq. (35), contains items related to EX_t, thus the condition shown in Theorem 1 should be satisfied firstly to make DX_t convergent. As stated above, $-1<\varpi<1$ and $f(1)>0$ implies that $2(1-\omega)(a-1)<c_1+c_2<2(1+\omega)(1+a)$. Thus conditions $-1<\varpi<1$ and $f(1)>0$ together make sure that the conditions stated in Theorem 1 are satisfied.

After EX_t is convergent, the convergent condition of iterative process $\{DX_t\}$ is that the absolute values(or complex modulus) of the three eigenvalues λ_{D1}, λ_{D2}, λ_{D3} are all less than 1. Theorem 2 proves that, $-1<\varpi<1$ and $f(1)>0$ are the necessary and sufficient condition of $\max\{\|\lambda_{D1}\|,\|\lambda_{D2}\|,\|\lambda_{D3}\|\}<1$.

Thus, $-1<\varpi<1$ and $f(1)>0$ together give the necessary and sufficient conditions to guarantee iterative process $\{DX_t\}$ convergent. If iterative process $\{DX_t\}$ is convergent, the convergent value can be easily calculated to be

$$DX = \frac{\frac{1}{12}(1+\varpi)}{f(1)\cdot[2(1-a)(1-\omega)+c_1+c_2]^2}\{2c_1^2c_2^2(P_g-P_i)^2 + 4(1-a)(1-\omega)c_1c_2(c_2P_g-c_1P_i)(P_g-P_i) + (c_1^2P_i^2+c_2^2P_g^2)[2(1-a)(1-\omega)]^2\}$$

It's easy to know that, even if sequence $\{DX_t\}$ can converge, generally speaking, the convergent value can not reach zero.

4. Parameter Selection Guidelines

Holland discussed the balance between *exploration* and *exploitation* that an algorithm must maintain (Holland, 1975). Exploration ability is related to the algorithm's tendency to explore new regions of the search space, while exploitation is the tendency to search a smaller region more thoroughly.

Researchers in PSO community used to believe that inertia weight ω balances exploration and exploitation in PSO algorithm, but new theoretical results give a different new explanation (Jiang et al, 2007a). It is believed that inertia weight can not balance exploration and exploitation by itself in PSO algorithm. The factor to balance exploration and exploitation should be the value of $\lambda_{\max D}$. The larger $\lambda_{\max D}$ is, the stronger is the exploration ability of the PSO algorithm. An empirical evidence can be found in (Xie et al, 2004), which shows that the relationship between exploration ability of PSO algorithm and inertia weight ω is not monotone. For more explanation about this matter, please refer to (Jiang et al, 2007a).

It is widely accepted that PSO algorithm have a good global search ability, while its fine-tune ability is relatively weak. That is to say, PSO algorithm can easily locate the good area

of the solution space in which good solutions are located, while the procedure to find the good solutions is not equally easy for PSO algorithm. Thus it is quite important for PSO algorithm to enhance local search ability. While in PSO algorithm determined by five-dimensional parameter tuple {ω, c_1, c_2, a, b}, generally speaking, the expectation of each particle's position can never reach P_g, and the variance of particle's position can never drop to zero. This is not a desired property for PSO algorithm, for this will make the fine-tune ability of PSO algorithm even worse.

If $(1-a)(1-\omega)=0$, then when P_i is equal to P_g, the expectation of particle's position will converge to P_g, and the variance of particle's position will converge to zero. This may help PSO algorithm to enhance fine-tune ability. Without loss of generality, suppose a=1. Then the PSO algorithm reduces to a simpler version totally determined by three-dimensional parameter tuple {ω, c_1, c_2}. For this simpler version, the following two corollaries apply.

Corollary 1. For PSO algorithm determined by three-dimensional parameter tuple, if and only if conditions $-1<\omega<1$ and $0<c_1+c_2<4(1+\omega)$ are satisfied together, iterative process {EX_t} is guaranteed to converge to $EX=(c_1P_i+c_2P_g)/(c_1+c_2)$.

Corollary 2. For PSO algorithm determined by three-dimensional parameter tuple, given $R>0$, if and only if $-1<\omega<1$ and $f(1)>0$ are satisfied together, iterative process {DX_t} is guaranteed to converge to $DX=(c_1c_2)^2(P_g-P_i)^2(1+\omega)/(c_1+c_2)^2/f(1)/6$.

It is easily verified that, in the simpler PSO version, when P_i is equal to P_g, the expectation of particle's position will gradually converge to P_g, and the variance will gradually converge to zero. This means that the PSO algorithm can thoroughly search around the best position found so far. If the convergent speed of variance is slow, it is more likely to find good solutions in the good area.

Below is the graphical illustrations of parameter ranges, which are determined by conditions shown in above two corollaries.

The parameter range to guarantee the convergence of iterative process {EX_t} is illustrated in Fig. 1. The cyan(light) area in Fig. 1 corresponds to case (1) discussed in Theorem 1, and the blue(dark) area in Fig. 1 corresponds to case (2) discussed in Theorem 1.
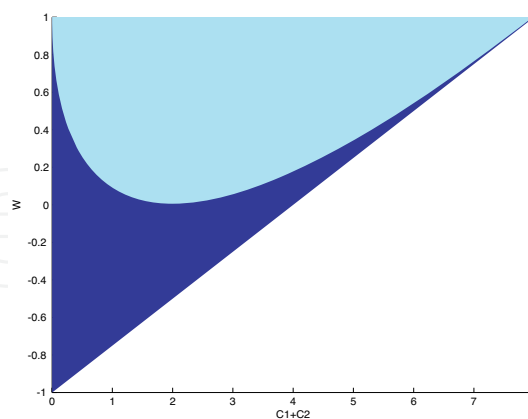

Figure 1 Parameter range to guarantee the convergence of iterative process {EX_t}. The cyan(light) area corresponds to case with complex eigenvalues, the blue(dark) area corresponds to case with real eigenvalues

The parameter ranges to guarantee the convergence of iterative process $\{DX_t\}$ are illustrated in Fig. 2-3. In Fig. 2, the relationship between c_1 and c_2 is illustrated. The relationship between lower and higher range of ω and c_1, c_2 are illustrated in Fig. 3.

The parameter selection of PSO algorithm in literatures favors $c_1=c_2=c$, so more detailed discussion on this condition is given. The relationship between ω and c is illustrated in Fig. 4, in which blue(dark) area is the parameter range to guarantee the convergence of expectation sequence of particle's position, and the cyan(light) area is the parameter range to guarantee the convergence of variance sequence of particle's position.
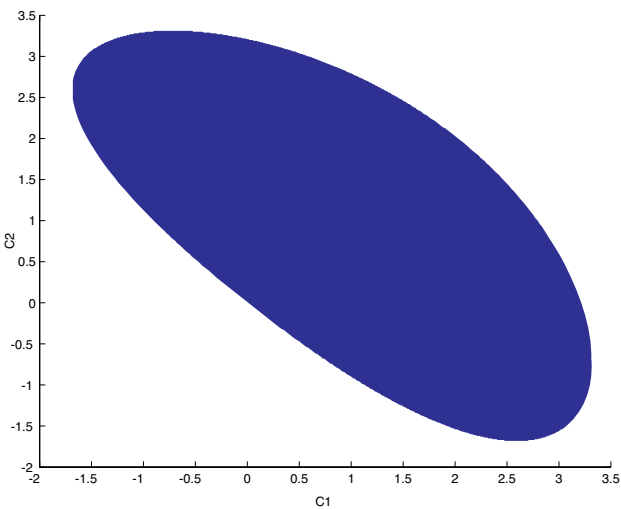

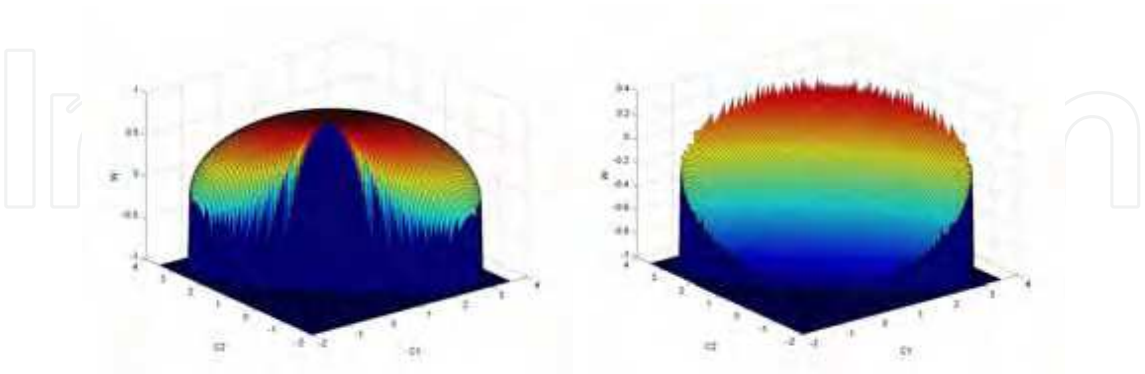Figure 2 Relationship between c_1 and c_2 to guarantee the convergence of iterative process $\{DX_t\}$

Figure 3 Relationship between lower(left) and higher(right) range of ω and c_1, c_2 to guarantee the convergence of iterative process $\{DX_t\}$
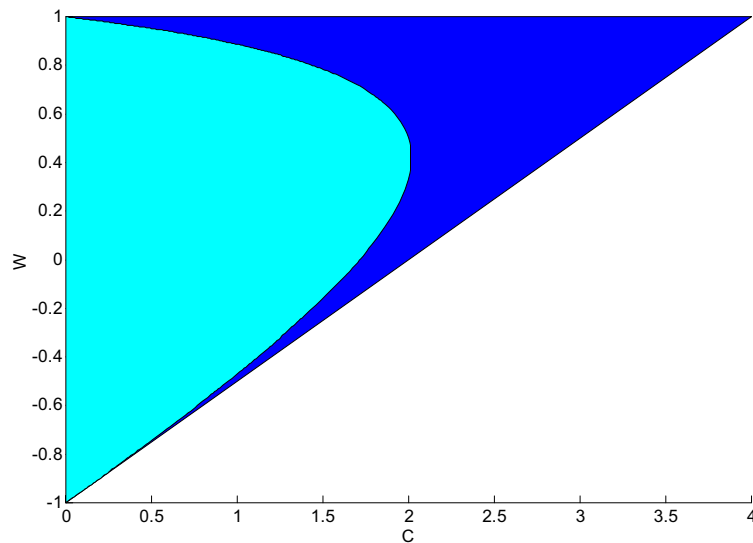
Figure 4 Relationship between ω and c when $c_1=c_2=c$ to simultaneously guarantee the convergence of iterative processes $\{EX_t\}$ and $\{DX_t\}$

Based on the theoretical analysis results obtained above and related experimental results, in order to help PSO algorithm to search the solution space more thoroughly, it is necessary to determine parameters to make $\lambda_{\max D} \approx 1$. Thus we propose a new set of parameters: {ω=0.715, c_1=c_2=1.7}. This set of parameters can help PSO algorithm to search the solution space thoroughly, slowly converge to the best position found so far, so as to find the optima with higher probability. In next section, the PSO algorithm using our suggested parameters will be compared against PSO algorithms using parameters suggested in literatures.

5. Performance Comparison

Parameter selection of PSO algorithm is quite important and have drawn attention from many researchers. Different researchers have proposed many different sets of parameters, such as $\omega = 0.729$, $c_1 = c_2 = 1.494$ (Clerc & Kennedy, 2002); $\omega = 0.6$, $c_1 = c_2 = 1.7$ (Trelea, 2003); $\omega = 0.729$, $c_1 = 2.041$, $c_2 = 0.948$ (Carlisle & Dozier, 2001). For those parameters, corresponding $\lambda_{\max D}$ can be calculated. When $\omega = 0.729$ and $c_1 = c_2 = 1.494$, $\lambda_{\max D} = 0.942$; when $\omega = 0.6$ and $c_1 = c_2 = 1.7$, $\lambda_{\max D} = 0.889$; and when $\omega = 0.729$, $c_1 = 2.041$, $c_2 = 0.948$, $\lambda_{\max D} = 0.975$. Obviously, $\lambda_{\max D}$ corresponding to all those three sets of parameters are all around 1, this may help to explain why the performance of PSO algorithm using those sets of parameters are promising. But it is easy to find out that $\lambda_{\max D}$ corresponding to those parameters is still not large enough to enhance the exploration ability of PSO algorithm, so that the algorithm is often trapped in local optima. For above reasons, we propose a new set of parameters: ω=0.715, c_1=c_2=1.7 (corresponding to $\lambda_{\max D} = 0.995$), and performance comparison is conducted.

No Free Lunch theorem (Wolpert & Macready, 1997) asserts that no algorithm can be better than any other, over all possible functions. Thus it does not seem interesting to demonstrate that the PSO algorithm with suggested parameters is good on some functions and not on others. What we hope for is a problem-solver that can work well with a wide range of problems. Thus, in the current exercises we combine results from a set of test functions, all of which are commonly used in experimentation with optimization algorithms.

Based on three performance measures proposed by Mendes (Mendes et al, 2004), we compare several PSO algorithms with different sets of parameters. Those measures are average standardized optima, average success iterations and average success rate.

1.   Standardized Optima

The first measure is simply the best function result after some arbitrary number of iterations; here we use 2,000. Basically this is a measure of sloppy speed. It does not necessarily indicate whether the algorithm is close to the global optimum; a relatively high score can be obtained on some of these multimodal functions simply by finding the best part of a locally optimal region.

It is not possible to combine raw results from different functions, as they are all scaled differently. For instance, almost any decent algorithm will find a function result less than 0.01 on the sphere function, but a result of 40.0 on Rosenbrock is considered good. In order to combine the function outputs, we standardized the results of each function to a mean of 0.0 and standard deviation of 1.0. All results of all trials for a single function are standardized to the same scale; as all of these problems involve minimization, a lower result is better, and after standardization a negative result means to be better than average. After standardizing each function separately, we can combine them and find the average for each single condition.

2.   Success Iterations

The second measure is the number of iterations required to reach a criterion. This is also a measure of speed, but in this case the criteria are intended to indicate that the searcher has arrived in the region of the global optimum.

There is, however, a problem with this measure, too. That is, some trials might never reach the criteria. Many hours have been lost waiting, trying to give each version a fair chance to find the global optimum, often in vain. Trials where the criteria are not met after a reasonable time — here we use 10,000 iterations — must be coded as infinite, which means among other things that the mean is meaningless.

The proper measure of central tendency for such a data set is the median. If the majority of trials are coded as infinite, then the median is represented as infinity, shown in the results tables with the lemniscus. In order to combine iteration data, mean of the medians is used, with the caveat that if any median were infinite, the mean would be infinite, too.

It is obviously that the first measure is different from the second one. The first measure determines whether the algorithm can get a good solution fast, e.g., after only 2,000 iterations, while the second measure determines how long it takes to find the global optimum if left to run, or whether it can find it at all. Generally speaking, iterations to calculate the second performance should be much larger than iterations to calculate the first measure.

3.   Success Rate

The third measure is perhaps the most important one. This is a simple binary code indicating whether the criteria were met within 10,000 iterations or not. Averaged over all

function trials, this gives the proportion of trials that successfully found the global optimum. There is no trick to this one; the mean of the ones and zeroes, where one indicates success and zero failure, gives the proportion of successes. Iteration used for this measure is the same as that used for the second measure, i.e., 10,000 iterations.

Function	Formula	Optima	Optimal Position
Ackley	$f_1(x) = 20 + e - 20e^{-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}} - e^{\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)}$	0	$(0,\dots,0)$
Griewank	$f_2(\vec{x}) = \frac{1}{4000}\sum_{i=1}^{n}x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1$	0	$(0,\dots,0)$
Rastrigin	$f_3(\vec{x}) = \sum_{i=1}^{n}(x_i^2 - 10\cos(2\pi x_i) + 10)$	0	$(0,\dots,0)$
Rosenbrock	$f_4(\vec{x}) = \sum_{i=1}^{n-1}\left(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right)$	0	$(1,\dots,1)$
Sphere	$f_5(\vec{x}) = \sum_{i=1}^{n}x_i^2$	0	$(0,\dots,0)$

Table 1 Benchmark functions

Function	Dimension n	Value Range [X_{min}, X_{max}]	Criteria
Ackley	10	$[-32,32]^{10}$	1
	30	$[-32,32]^{30}$	2
	50	$[-32,32]^{50}$	2
Griewank	10	$[-600,600]^{10}$	0.1
	30	$[-600,600]^{30}$	0.05
	50	$[-600,600]^{50}$	0.05
Rastrigin	10	$[-5.12,5.12]^{10}$	10
	30	$[-5.12,5.12]^{30}$	60
	50	$[-5.12,5.12]^{50}$	100
Rosenbrock	10	$[-30,30]^{10}$	5
	30	$[-30,30]^{30}$	20
	50	$[-30,30]^{50}$	50
Sphere	10	$[-100,100]^{10}$	1e-5
	30	$[-100,100]^{30}$	1e-5
	50	$[-100,100]^{50}$	1e-5

Table 2 Value range and criteria of benchmark functions

Based on those three measures, we conduct experiments on five benchmark problems, which are commonly used in literature. Those five benchmark functions are Ackley, Griewank, Rastrigin, Rosenbrock and Sphere. Corresponding function formula, optima and optimal positions are shown in Table 1; value range and criteria are listed in Table 2.

Features of those five functions are: Sphere is a simple unimodal function; Rosenbrock is also an unimodal function, but Rosenbrock's variables are strongly dependent and gradient information often misleads algorithms; Ackley, Griewank and Rastrigin are multimodal functions with many local optima. Griewank is strongly multi-modal with significant interaction between its variables, caused by the product term. This function has the interesting property that the number of local minima increases with dimensionality.

The four sets of parameters used to be compared are listed as follows:

Set A (Clerc & Kennedy, 2002): $\omega = 0.729$, $c_1 = c_2 = 1.494$;

Set B (Carlisle & Dozier, 2001): $\omega = 0.729$, $c_1 = 2.041$, $c_2 = 0.948$;

Set C (Trelea, 2003): $\omega = 0.6$, $c_1 = c_2 = 1.7$;

Set D: $\omega = 0.715$, $c_1 = c_2 = 1.7$.

		Set A	Set B	Set C	Set D
Mean (Deviation)	Ackley	0.1139 (0.4041)	0.1155 (0.3483)	0.1765 (0.4460)	0.0511 (0.2553)
	Griewank	0.0944 (0.0583)	0.0606 (0.0357)	0.0876 (0.0457)	0.0803 (0.0411)
	Rastrigin	8.7357 (4.5793)	4.7758 (2.2978)	8.6263 (4.0645)	6.0593 (3.4700)
	Rosenbrock	2.4308 (9.2956)	3.2329 (11.1348)	2.0431 (9.2136)	8.5559 (25.0940)
	Sphere	0 (0)	0 (0)	0 (0)	0 (0)
Success Rate	Ackley	0.92	0.90	0.86	0.96
	Griewank	0.64	0.90	0.67	0.72
	Rastrigin	0.68	1.00	0.69	0.88
	Rosenbrock	1.00	0.99	1.00	1.00
	Sphere	1.00	1.00	1.00	1.00
Success Iteration	Ackley	84	72.5	63	108
	Griewank	265	192	160.5	426
	Rastrigin	179	161	141	201
	Rosenbrock	333.5	319.5	280.5	538.5
	Sphere	186	162	139	259

Table 3. Experimental results on 10-dimensional functions

Experiments are conducted on all five benchmark functions, considered dimensions are 10, 30 and 50, corresponding particle swarm sizes are separately 20, 40 and 100. Each algorithm with each set of parameter is run 100 times, and the final result is the statistical result of all 100 runs.

The intermediate experimental results include mean function value (deviation), success rate and success iteration. Experimental results related to 10-dimensional functions are listed in Table 3. Experimental results related to 30-dimensional functions are listed in Table 4. And experimental results related to 50-dimensional functions are listed in Table 5. Although we don't want to compare the PSO algorithm using different parameters on each single function, it is clearly shown in Table 3-5 that the performance of PSO algorithm using our suggested parameters is promising.

		Set A	Set B	Set C	Set D
Mean (Deviation)	Ackley	1.2639 (0.9417)	1.3589 (0.8030)	1.3556 (0.9833)	0.0250 (0.1761)
	Griewank	0.0155 (0.0196)	0.0208 (0.0242)	0.0160 (0.0199)	0.0113 (0.0144)
	Rastrigin	59.0109 (18.2530)	41.0719 (10.1841)	57.5682 (15.3826)	42.9026 (11.8995)
	Rosenbrock	28.8497 (28.7922)	33.1788 (39.6960)	31.9139 (34.7045)	58.6477 (51.5653)
	Sphere	0 (0)	0 (0)	0 (0)	0 (0)
Success Rate	Ackley	0.82	0.78	0.70	1.00
	Griewank	0.91	0.86	0.90	0.98
	Rastrigin	0.61	0.94	0.59	0.94
	Rosenbrock	0.99	0.90	0.97	0.84
	Sphere	1.00	1.00	1.00	1.00
Success Iteration	Ackley	190	142.5	150.5	280
	Griewank	285	197	237	503.5
	Rastrigin	287.5	182.5	179.5	345.5
	Rosenbrock	1832	2051	1689	5323
	Sphere	435.5	295	363	772

Table 4. Experimental results on 30-dimensional functions

The final synthesized results are listed in Table 6. Experimental data shown in Table 6 clearly indicates that the PSO algorithm using parameter set D outperforms other algorithm in the measures of standardized optima and average success rate, and is outperformed by other algorithms in the measure of success iteration. Both two phenomena can be explained by $\lambda_{\max D}$ corresponding to each set of parameters.

		Set A	Set B	Set C	Set D
Mean (Deviation)	Ackley	1.2454 (0.9683)	1.8311 (0.5866)	1.2278 (0.8901)	0 (0)
	Griewank	0.0130 (0.0229)	0.0108 (0.0200)	0.0140 (0.0219)	0.0079 (0.0118)
	Rastrigin	118.1610 (24.7108)	87.0787 (16.8852)	113.8330 (23.4990)	70.3967 (15.1534)
	Rosenbrock	71.0803 (41.1463)	70.3772 (40.7945)	67.8099 (33.8612)	87.6412 (42.7683)
	Sphere	0 (0)	0 (0)	0 (0)	0 (0)
Success Rate	Ackley	0.77	0.60	0.82	1.00
	Griewank	0.94	0.94	0.93	0.98
	Rastrigin	0.25	0.80	0.27	0.97
	Rosenbrock	0.80	0.60	0.91	0.61
	Sphere	1.00	1.00	1.00	1.00
Success Iteration	Ackley	257	214	214	471
	Griewank	383	214	338	789
	Rastrigin	∞	262	∞	597
	Rosenbrock	2188	3981	2205	5474
	Sphere	603	332	531	1281

Table 5. Experimental results on 50-dimensional functions

	Dimension	Set A	Set B	Set C	Set D
Standardized Optima	10	0.1193	**-0.2070**	0.1137	-0.0260
	30	0.1134	-0.0121	0.1337	**-0.2350**
	50	0.1286	0.0032	0.0882	**-0.2199**
Average Success Rate	10	0.848	**0.958**	0.844	0.912
	30	0.866	0.896	0.832	**0.952**
	50	0.752	0.788	0.786	**0.912**
Average Success Iteration	10	209.5	181.4	**156.8**	306.5
	30	606	573.6	**523.8**	1444.8
	50	∞	**1000.3**	∞	1722.3

Table 6 Synthesized performance comparison results

Parameter set D corresponds to a largest $\lambda_{\max D}$ among all four set of parameters, thus PSO algorithm using parameter set D has the strongest exploration ability, so it is not easy to be trapped into local optima. When the dimension increase, the solution space get more complex, and PSO algorithm gets more likely to be trapped into local optima. At this time, the influence of exploration ability to the performance of algorithm would be more significant. This is why PSO algorithm using parameter set D can outperform other algorithm in the measures of standardized optima and average success rate, and this advantage gets more significant when the dimension increases.

Also due to the strong exploration ability that PSO algorithm using parameter set D has, the algorithm has to waste a lot of time in exploring new search area, so as to influence the speed.

6. Conclusion

The stochastic process theory is applied to analyze the particle swarm optimization algorithm determined by five-dimensional real-value parameter tuple {ω, c_1, c_2, a, b}, considering the randomness thoroughly. Specifically speaking, stochastic convergence analysis is conducted on PSO algorithm when it is in stagnation phase, and the convergent properties of expectation and variance sequence of particle's position are studied. The analysis results determines corresponding parameter ranges, both in formular and graphical form. This result is helpful to understand the mechanism of PSO algorithm and select appropriate parameters to make PSO algorithm more powerful.

After the theoretical stochastic convergence analysis of PSO algorithm in stagnation phase, parameter selection guidelines are discussed, and a set of suggested parameters {ω=0.715, c_1=c_2=1.7} is given, which is compared against other three sets of parameters which are proposed in literatures. Experimental result shows that the PSO algorithm using our suggested parameters can achieve robust performance, but the time expires before reaching optimal area is longer than PSO algorithm using other suggested parameters.

7. References

Carlisle, A. & Dozier, G. (2001). An off-the-shelf pso. *Proceedings of the Workshop on Particle Swarm Optimization*. Indianapolis, USA. 2001.

Clerc, M. & Kennedy, J. (2002). The particle swarm-explosion, stability and convergence in a multidimensional complex space. *IEEE Trans. Evol. Comput.*, Vol. 6, No. 2, April 2002, 58–73.

Clerc M. (2006). *Stagnation analysis in particle swarm optimisation or what happens when nothing happens*. Technical Report CSM-460, Department of Computer Science, University of Essex. August 2006. ISSN: 1744-8050.

Holland, J. (1975). *Adaption in natural and artificial systems*, University of Michigan Press, Ann Arbor, MI, 1975.

Jiang, M.; Luo, Y. P. & Yang, S. Y. (2007a). Stagnation Analysis in Particle Swarm Optimization, *Proceedings of IEEE Swarm Intelligence Symposium 2007*, pp. 14-17, Hawaii, USA. 1-5 April 2007.

Jiang, M.; Luo, Y. P. & Yang, S. Y. (2007b). Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, Vol. 102, No. 1, April 2007, 8-16

Mendes, R.; Kennedy, J. & Neves, J. (2004). The fully informed particle swarm: Simpler, maybe better. *IEEE Trans. Evol. Comput.*, Vol. 8, No. 3, June 2004, 204–210.

Ozcan, E. & Mohan, C. (1998). Analysis of a simple particle swarm optimization system, *Intelligent Engineering Systems Through Artificial Neural Networks*, 1998. 253-258.

Ozcan, E. & Mohan, C. K. (1999). Particle swarm optimization: Surfing the waves. *Proc. IEEE Congr. Evol. Comput.*, Washington, DC. Jul. 1999. 1939–1944.

Trelea, I. C. (2003). The particle swarm optimization algorithm: Convergence analysis and parameter selection. *Inf. Process. Lett.,* Vol. 85, 2003, 317–325.

van den Bergh, F. (2001). *An analysis of particle swarm optimizers*: [Ph.D. dissertation]. University of Pretoria, Pretoria, South Africa, 2001.

Wolpert, D. H. & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.*, Vol. 1, No. 2, April 1997, 67–82.

Xie, X-F. ; Zhang, W-J. & Bi, D-C. (2004). Optimizing semiconductor devices by self-organizing particle swarm. *Congress on Evolutionary Computation (CEC)*, pp. 2017-2022, Oregon, USA, 2004.

Yasuda, K.; Ide, A. & Iwasaki, N. (2003). Adaptive particle swarm optimization. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.* 2003. 1554-1559.
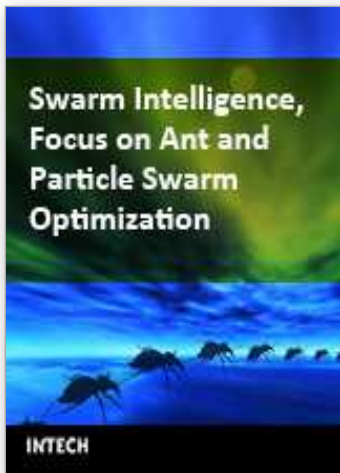
Swarm Intelligence, Focus on Ant and Particle Swarm Optimization

Edited by FelixT.S.Chan and Manoj KumarTiwari

ISBN 978-3-902613-09-7

Hard cover, 532 pages

Publisher I-Tech Education and Publishing

Published online 01, December, 2007

Published in print edition December, 2007

In the era globalisation the emerging technologies are governing engineering industries to a multifaceted state. The escalating complexity has demanded researchers to find the possible ways of easing the solution of the problems. This has motivated the researchers to grasp ideas from the nature and implant it in the engineering sciences. This way of thinking led to emergence of many biologically inspired algorithms that have proven to be efficient in handling the computationally complex problems with competence such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), etc. Motivated by the capability of the biologically inspired algorithms the present book on "Swarm Intelligence: Focus on Ant and Particle Swarm Optimization" aims to present recent developments and applications concerning optimization with swarm intelligence techniques. The papers selected for this book comprise a cross-section of topics that reflect a variety of perspectives and disciplinary backgrounds. In addition to the introduction of new concepts of swarm intelligence, this book also presented some selected representative case studies covering power plant maintenance scheduling; geotechnical engineering; design and machining tolerances; layout problems; manufacturing process plan; job-shop scheduling; structural design; environmental dispatching problems; wireless communication; water distribution systems; multi-plant supply chain; fault diagnosis of airplane engines; and process scheduling. I believe these 27 chapters presented in this book adequately reflect these topics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

M. Jiang, Y. P. Luo and S. Y. Yang (2007). Particle Swarm Optimization - Stochastic Trajectory Analysis and Parameter Selection, Swarm Intelligence, Focus on Ant and Particle Swarm Optimization, FelixT.S.Chan and Manoj KumarTiwari (Ed.), ISBN: 978-3-902613-09-7, InTech, Available from: http://www.intechopen.com/books/swarm_intelligence_focus_on_ant_and_particle_swarm_optimization/particle_swarm_optimization_-_stochastic_trajectory_analysis_and_parameter_selection

INTECH
open science | open minds