



Available online at <http://scik.org>

J. Math. Comput. Sci. 11 (2021), No. 4, 4981-4993

<https://doi.org/10.28919/jmcs/5796>

ISSN: 1927-5307

ESTIMATION OF PARAMETERS IN PARTICLE SWARM OPTIMIZATION

BESIANA ÇOBANI^{1,*}, AURORA SIMONI², RIGENA SEMA³, XHENSILDA ALLKA⁴

¹Department of Mathematics, University of Tirana, Albania

²Department of Applied Mathematics, University of Tirana, Albania

³Department of Mathematics, University of Tirana, Albania

⁴Department of Computer Engineering Metropolitan University of Tirana, Albania

Copyright © 2021 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Nowadays the heuristic techniques are becoming one of the most useful tools in solving optimization problems. One of these techniques is Particle Swarm Optimization, PSO algorithm. From the numerical analysis perspective this is a successful method, but many issues are still to be considered regarding the convergence of algorithm. In this paper we deal with the problem of the evaluation of the parameters of the algorithm that assure its convergence. In the previous work we presented some restriction on the parameters of the perturbed dynamical system, that modeled the PSO algorithm. These restrictions are necessary to guarantee the stability of the system. In this paper we present some other restrictions needed to ensure the stability of the system and to advance in the research of the convergence of PSO.

Keywords: PSO algorithm; parameters; convergence; dynamical system.

2010 AMS Subject Classification: 65K10, 37N30.

*Corresponding author

E-mail address: besiana.hamzallari@fshn.edu.al

Received April 1, 2021

1. INTRODUCTION

The particle swarm optimization is an algorithm based on the behavior of a flock of birds, fish, ants, more generally a group of individuals, and considers their behavior to control the area for their necessity, like feeding. There is no surprise that its first origin comes from Social Psychology as a simulation of social and cognitive processes in order to model abstract concepts as intelligence and learning. The model was invented by Kennedy and Eberhart in 1995. The main idea of the algorithm is that the particles play the role of the members of the flock, and the algorithm imitates the behavior of the member (particle) locally and globally which means the flock as a whole. Since the algorithm is based in the behavior of a particle and the interaction of each particle with the ones that are in the neighborhood, then it converges to the global solution. [1-3]. There are many other algorithms which operate in a similar way, known as population-based evolutionary algorithms, however, PSO algorithm is motivated by the simulation of social behavior instead of the best performance made by the strongest individual (natural selection). The challenge in the optimization methods we use for optimization problems is to find the global solution because they mostly converge to the local solution. The heuristic algorithms try to converge to the global solution, which applies even for the method we study, PSO. The PSO algorithm is very efficient numerically but still lacks theoretical foundations. In this work we present a study on the stability of the algorithm. First, we present the basic PSO algorithm and then we consider the dynamical system where this algorithm is modeled. In sections 4 and 5 we introduce the parameters $\alpha, \beta, \gamma, \delta, \eta$ that allow us to operate within intervals where stability holds. Then in section 6 we tested these values on well-known functions and arrived at an optimal solution. Three cases are considered, two of which have the inertia weight vary at each iteration and in one case it is kept constant with value one, the ideal case. Finally, in the conclusion section we give a resume of the evaluation of the parameters theoretically and then the results obtained by applying the restrictions on the parameters in some test functions.

2. PARTICLE SWARM OPTIMIZATION

As it is known the particle swarm optimization (PSO) is an algorithm that finds optimal solution through the interaction of individuals in a population of particles. The algorithm searches a space by adjusting the trajectories of individual vectors, named particles. Each particle is aimed to move toward the positions of their own best position and the best position of their neighbors. More specifically, a particle is identified by its current speed and its position, the most optimal position of each particle and the most optimal position of the neighborhood. We focus our work in the study of the behavior of one particle, of one individual. The speed and the velocity of a particle is given by the formula [3]

$$(1) v_{id}^{k+1} = v_{id}^k + c_1 r_1^k (pbest_{id}^k - x_{id}^k) + c_2 r_2^k (gbest_d^k - x_{id}^k)$$

$$(2) x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}$$

v_{id}^k represents the speed of the particle i in the k time and x_{id}^k represents the d -dimension quantity of its position or more simply the current position of the particle $pbest_{id}^k$ represents the d dimension quantity of the particle i at its most optimal position at its k times and $gbest_d^k$ is the d -dimension quantity of the swarm at its most optimal position. The speed is always between two boundaries, $-v_{dmax}$ and v_{dmax} , in order to avoid the wandering of the point away from the operational searching space. c_1 , c_2 are non-negative constants, called cognitive learning rate or the acceleration coefficients and play an important role in the algorithm process. They represent the particle stochastic acceleration weight toward the personal best (pbest) and the global best (gbest). We should avoid small accelerate constant since the particle can move away from the goal area and large accelerate constant values since the particle can move very quickly to the goal area and leave it totally. How do we determine these acceleration coefficients? In previous work of Kennedy and Eberhart, also Clerc later was stated that high values of the cognitive component c_1 compared to the social component c_2 will bring to extra wandering of the search area (space). If there is a high value of the social component compared to the cognitive one then the particles may hurry prematurely to the local optimum. In many papers the values of the acceleration coefficients are equal to 2.05, other researchers recommend to not take the same value of the

coefficients. All of this discussion is based more heuristically because still there is no mathematical assurance on the convergence of PSO. In this paper we present some values for the acceleration coefficients based on the perturbed dynamical system presented in the work of Clerc and Kennedy.

3. MODELLING THE PROBLEM IN A DYNAMICAL SYSTEM

First, we denote $\varphi_1 = c_1 r_1^k$, $\varphi_2 = c_2 r_2^k$, and $\varphi = \varphi_1 + \varphi_2$. Then we redefine

$$pbest_{id} := \frac{\varphi_1 pbest_{id} + \varphi_2 gbest_d}{\varphi_1 + \varphi_2}$$

The dynamical system obtained by (1) and (2)

$$(3) \quad \begin{cases} v_{t+1} = wv_t + \varphi y_t \\ y_{t+1} = -wv_t + (1 - \varphi)y_t \end{cases}$$

where $y_t = p - x_t$, where p is the best position found so far. From three in an iterating step we have

$$\begin{cases} v_{t+2} = wv_{t+1} + \varphi y_{t+1} \\ y_{t+1} = -wv_t + (1 - \varphi)y_t \end{cases}$$

substituting y_{t+1} at the first equation and adding the first identity of (3) we have

$$v_{t+2} + (\varphi - 1 - w)v_{t+1} + \omega v_t = 0$$

Which is a second order difference equation. Using the Lagrange interpolation, we have a continuous solution in order to study the convergence so the respective differential equation is

$$v_{tt} + \ln(e_1 e_2) v_t + \ln(e_1) \ln(e_2) = 0$$

where e_1, e_2 are the roots of the characteristic equation

$$\lambda^2 + (\varphi - 1 - \omega)\lambda + \omega = 0,$$

respectively

$$(*) \quad e_1 = \frac{\omega+1-\varphi+\sqrt{(\omega+1-\varphi)^2-4\omega}}{2}, \quad e_2 = \frac{\omega+1-\varphi-\sqrt{(\omega+1-\varphi)^2-4\omega}}{2}$$

and the solution of the second order differential equation is $v(t) = l_1 e_1^t + l_2 e_2^t$.

From

$$v_{t+1} = wv_t + \varphi y_t$$

we can derive

$$y(t) = \frac{l_1 e_1^t (e_1 - \omega) + l_2 e_2^t (e_2 - \omega)}{\varphi}.$$

To estimate the parameters l_1, l_2 for $t = 0$

$$\begin{cases} v(0) = l_1 + l_2 \\ y(0) = \frac{l_1(e_1 - \omega) + l_2(e_2 - \omega)}{\varphi} \end{cases}$$

and

$$\begin{cases} v(0) = l_1 + l_2 \\ y(0) = \frac{l_1(e_1 - \omega) + l_2(e_2 - \omega)}{\varphi} \end{cases}$$

We continue our study of convergence taking the inertia weight

$\omega = 1$, and perturbate the dynamical system introducing the parameters $\alpha, \beta, \gamma, \delta, \eta$ as follows

$$(4) \begin{cases} v_{t+1} = \alpha v_t + \beta \varphi y_t \\ y_{t+1} = -\gamma v_t + (\delta - \eta \varphi) y_t \end{cases}$$

where $\varphi \in R^{+*}, \forall t \in N, (y_t, v_t) \in R^2$

In this case we operate with matrix theory as the eigenvalues of the coefficients of a linear dynamical system determine the solution of that system as well.

$$A = \begin{bmatrix} \alpha & \beta \varphi \\ -\gamma & \delta - \eta \varphi \end{bmatrix}$$

If we denote the eigenvalues of the matrix \bar{e}_1, \bar{e}_2 then the solution of (4) is

$$v(t) = l_1 \bar{e}_1^t + l_2 \bar{e}_2^t, \quad y(t) = \frac{l_1 \bar{e}_1^t (\bar{e}_1 - \alpha) + l_2 \bar{e}_2^t (\bar{e}_2 - \alpha)}{\beta \varphi}.$$

As in the previous section we evaluate the parameters l_1, l_2 for the initial time $t = 0$

$$\begin{cases} l_1 = \frac{-\beta \varphi y(0) - (\alpha - \bar{e}_2) v(0)}{\bar{e}_2 - \bar{e}_1} \\ l_2 = \frac{\beta \varphi y(0) + (\alpha - \bar{e}_1) v(0)}{\bar{e}_2 - \bar{e}_1} \end{cases}$$

And also from (*) for $\omega = 1$

$$e_1 = \frac{2 - \varphi + \sqrt{(2 - \varphi)^2 - 4}}{2}, \quad e_2 = \frac{2 - \varphi - \sqrt{(2 - \varphi)^2 - 4}}{2}$$

$$(**) \quad e_1 = 1 - \frac{\varphi}{2} + \frac{\sqrt{\varphi^2 - 4\varphi}}{2}, \quad e_2 = 1 - \frac{\varphi}{2} - \frac{\sqrt{\varphi^2 - 4\varphi}}{2}$$

We have the constriction coefficients k_1, k_2 defined by

$$\begin{cases} \bar{e}_1 = k_1 e_1 \\ \bar{e}_2 = k_2 e_2 \end{cases}$$

From direct computation,

$$\begin{cases} k_1 = \frac{\alpha + \delta - \eta\varphi + \sqrt{(\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2}}{2 - \varphi + \sqrt{\varphi^2 - 4\varphi}} \\ k_2 = \frac{\alpha + \delta - \eta\varphi - \sqrt{(\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2}}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}} \end{cases}$$

4. CHOOSING THE COEFFICIENTS

For k_1 and k_2 to be real for a given value of φ subtracting and adding the two above identities we obtain

$$2(\alpha + \delta - \eta\varphi) = (k_1 + k_2)(2 - \varphi) + (k_1 - k_2)\sqrt{\varphi^2 - 4\varphi}$$

$$2\sqrt{(\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2} = (k_1 + k_2)\sqrt{\varphi^2 - 4\varphi} - (k_1 - k_2)(2 - \varphi)$$

or posing

- $A = \text{sgn}(\varphi^2 - 4\varphi)$
- $B = |\varphi^2 - 4\varphi|$
- $C = (\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2$

we have

$$\begin{aligned} \sqrt{C}(1 - \text{sgn}(C)(2 - \varphi) - (\alpha + \delta - \eta\varphi)\sqrt{B}(1 - A)) &= 0 \\ \sqrt{(|C|)}\sqrt{B}\text{sgn}(C)(1 + A) &= 0 \end{aligned}$$

It is seen the presence of φ so the solution depends on it. As it is obvious there are many ways for the two equations to be zero. A possible choice is

$$\begin{cases} C > 0 \\ A = -1, (\varphi < 4) \\ \alpha + \delta - \eta\varphi = 0 \end{cases}$$

First set of choice $\alpha = \delta$, $\beta\gamma = \eta^2$.

Then

$$\alpha = \frac{1}{4}(2(k_1 + k_2) + (k_1 - k_2)(\sqrt{\varphi^2 - 4\varphi} + \varphi \frac{2-\varphi}{\sqrt{\varphi^2-4\varphi}}))$$

$$\eta = \frac{1}{2}(k_1 + k_2 + \frac{2-\varphi}{\sqrt{\varphi^2-4\varphi}}(k_1 - k_2))$$

Since we want real coefficients and from $A = -1$ then we choose

$$k_1 = k_2 = k \in R$$

And in order for the five coefficients to satisfy the conditions $\alpha = \delta$, $\beta\gamma = \eta^2$, we choose

$$\alpha = \beta = \gamma = \delta = \eta = k$$

In this case the parameters do not depend on φ .

Second set of choice (depend on) $\alpha = \beta$, $\gamma = \delta = \eta$, after the calculations

$$\alpha = \frac{(k_1+k_2)(2-\varphi) + (k_1-k_2)\sqrt{\varphi^2-4\varphi}}{2} + (\varphi-1)$$

$$\gamma = \frac{1}{4(\varphi-1)} \frac{(k_1+k_2)(\varphi-2) - (k_1-k_2)\sqrt{\varphi^2-4\varphi}}{\sqrt{k_1^2(\varphi^2-4\varphi+2-\varphi\sqrt{\varphi^2-4\varphi}) + k_2^2(\varphi^2-4\varphi+2+\varphi\sqrt{\varphi^2-4\varphi}) + 8k_1k_2(2\varphi-1)}}$$

To simplify the result, we take $k_1 = k_2 = k \in R$ and

$$\alpha = (2-\varphi)k + \varphi - 1$$

$$\gamma = k \text{ or } \gamma = \frac{k}{\varphi-1}$$

For the convergence of the first choice of parameters, with the conditions $k_1 = k_2 = k$,

$$(5) \begin{cases} |\bar{e}_1| = k|e_1| \\ |\bar{e}_2| = k|e_2| \end{cases}$$

And for the second choice of coefficients and for the parameter α to be positive, $\varphi \leq 2$

$$(6) \begin{cases} |\bar{e}_1| = \left| k \left(1 - \frac{\varphi}{2}\right) + \frac{\sqrt{k^2(2-\varphi)^2 + 4k(\varphi-2) + 4(\varphi-1)}}{2} \right| \leq k|e_1| = k \\ |\bar{e}_2| = \left| k \left(1 - \frac{\varphi}{2}\right) - \frac{\sqrt{k^2(2-\varphi)^2 + 4k(\varphi-2) + 4(\varphi-1)}}{2} \right| \leq k|e_2| = k \end{cases}$$

(5) and (6) guarantee that the system is stable.

Since the parameters must be real-valued then another restriction is the following

$$(\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2 \geq 0, \forall \varphi \in R^+$$

We can compute the discriminant and obtain

$$(\eta\varphi)^2 - 4\beta\gamma\varphi - 2\varphi\delta\eta + 2\varphi\alpha\eta + (\alpha - \delta)^2 \geq 0$$

Since $(\eta\varphi)^2, (\alpha - \delta)^2$ are nonnegative we have

$$-4\beta\gamma\varphi - 2\varphi\delta\eta + 2\varphi\alpha\eta > 0 \leftrightarrow 2\varphi(-2\beta\gamma - \delta\eta + \alpha\eta) > 0 \leftrightarrow$$

$$\Leftrightarrow (-2\beta\gamma - \delta\eta + \alpha\eta) > 0 \Leftrightarrow (-2\beta\gamma + \eta(-\delta + \alpha)) > 0 \Leftrightarrow \eta(-\delta + \alpha) > 2\beta\gamma > \beta\gamma$$

So, the condition needed is $\eta(\alpha - \delta) > \beta\gamma$. (7)

From the other hand, even if the condition (7) is satisfied one of the values of the eigenvalues \bar{e}_1 , \bar{e}_2 can still have a complex trajectory since $v(t) = l_1\bar{e}_1^t + l_2\bar{e}_2^t$, for $\bar{e}_1 = -1$, $(-1)^t$ where t non integer can be complex. So, we add a stronger condition for the eigenvalues, we require to be positive $\bar{e}_1 > 0$, $\bar{e}_2 > 0$ which is equivalent from (5) that

$$\alpha + \delta - \eta\varphi > 0$$

$$(\alpha + \delta - \eta\varphi)^2 > (\eta\varphi)^2 + 2\varphi(\alpha\eta - \delta\eta - 2\beta\gamma) + (\alpha - \delta)^2$$

From operation on the second inequality we obtain

$$(8) \begin{cases} \alpha + \delta - \eta\varphi > 0 \\ \alpha(\delta - \eta\varphi) + \gamma\beta\varphi > 0 \end{cases}$$

What we can deduce directly from (8) is that both inequalities depend on φ . If the maximum value of φ is known, we denote it by φ_{max} and (8) becomes

$$(9) \begin{cases} \frac{\alpha + \delta}{\eta} > \varphi_{max} \\ \frac{\alpha\delta}{\alpha\eta - \gamma\beta} > \varphi_{max} \end{cases}$$

To conclude this section of restrictions needed for the system to be continuous and real the parameters should satisfy simultaneously (7), (8) and (9).

5. RESULTS OF TESTING VALUES OF PARAMETERS SATISFYING (7), (8), (9) IN PSO ALGORITHM

The theoretical result from the previous sections allow us to propose several variants of PSO. The stability of the perturbed dynamical system depends on the parameters $\alpha, \beta, \gamma, \delta, \eta$ and the continuity of the perturbed system under the conditions (7), (8), (9) of the parameters $\alpha, \beta, \gamma, \delta, \eta$ guarantees that the proposed variants of PSO converge. For different values of the coefficients we obtain different cases of PSO algorithm. We consider some well-known test functions to control the numerical efficiency and stability of the proposed cases. The population taken in consideration for all test functions is 30 particles. To arrive to the optimal value, we ran 10 steps with 1000

iterations each.

We check their numerical results with test functions [5].

The following table presents the test functions and their characteristics related to the variants of PSO.

The function	Dimension	Boundary
$f1(x) = \sum_{i=1}^n x_i^2$	3	± 20
$f2(x) = \sum_{i=1}^n (100 * (x_{i+1} - x_i^2)^2 + (1 - x_i)^2)$	3	± 50
$f3(x) = \sum_{i=1}^n [x_i^2 - 10 * \cos(2\pi x(i)) + 10]$	3	± 10
$f4(x) = \sum_{i=1}^n i * x_i^4$	6	± 20
$f5(x) = \frac{1}{4000} \sum_{i=1}^n (x_i - 100)^2 - \prod_{i=1}^n \cos\left(\frac{x_i - 100}{\sqrt{i}}\right) + 1$	3	± 300
$f6(x) = \sum_{i=1}^n i * x_i^4$	10	± 20

Table 1. Test function used, dimension and boundaries of the respective variables.

We have tested three different cases; in the first and third case we consider the inertia weight variable at each iteration based in the formulas shown in the respective cases. In the second case the inertia weight is kept constant, which is considered an ideal case.

First case considers that the inertia weight changes depending in

$$w_{max} = 0.9; w_{min} = 0.4; max_{iter} = 1000, iteracion$$

while the acceleration coefficients are kept constant.

$$\text{Case 1: } c1 = 2.01, c2 = 2.02 \text{ and } w = w_{max} - \frac{(w_{max} - w_{min}) * iteration}{max_{iter}}$$

In the *second case* we consider the inertia weight equal to one, $w = 1$, and the acceleration coefficients $c1, c2$ equal, more specifically

$$\text{Case 2: } c1 = c2 = 2.05, w = 1$$

In the *third case* we use a different formula to evaluate the inertia weight, and the acceleration coefficients are kept constant.

$$\text{Case 3: } c1 = c2 = 2.01, \omega(iter) = \omega(iter - 1) * 0.99$$

In the following table is shown the optimal value of the objective function and the step in which this value is obtained.

$f_i(x)$	Case.1		Case. 2		Case.3	
	Best-fun	Best-step	Best-fun	Best-step	Best-fun	Best-step
1	1.9518e-67	9	0.0382	3	7.0880e-318	4
2	1.1445e-05	8	2.7578	9	0.0049	9
3	0	1	2.9376	9	0	1
4	7.8302e-79	8	335.5725	2	4.8372e-21	1
5	0	4	0.3315	6	0	1
6	6.7040e-44	1	3.1448e+03	2	2.9317e-68	1

Table 2. Optimal value objective function and best run in *Case 1, Case 2, Case 3*, for each test function taken in consideration.

The convergence of these cases related to the test functions of table 1, is given in the following figures

ESTIMATION OF PARAMETERS IN PARTICLE SWARM OPTIMIZATION

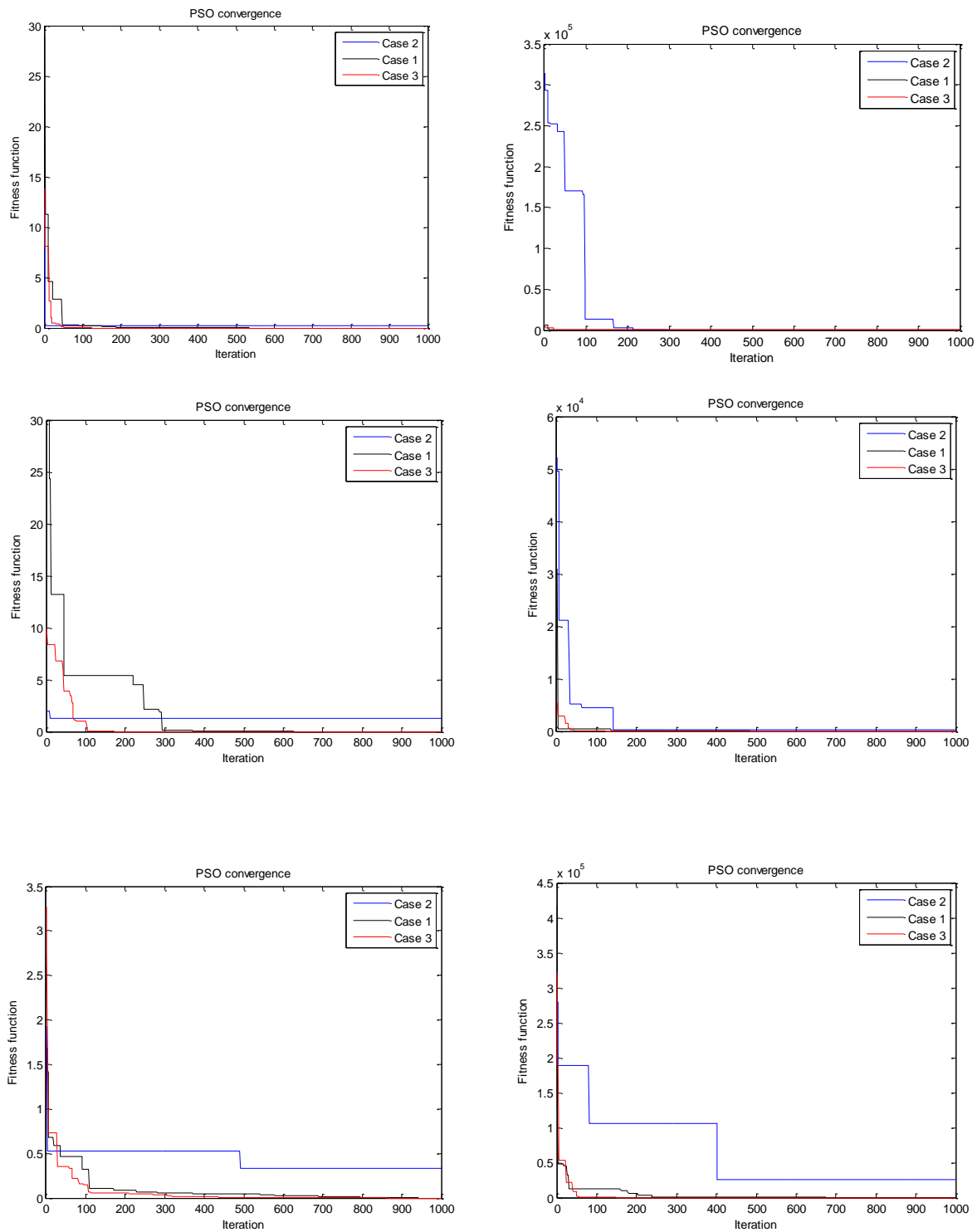


Figure 1. Convergence graph of PSO of three cases, Case 1 Case 2 and Case 3, for each test functions f_1, f_2, \dots, f_6 .

In the figures the black color represents the curve of case 1, blue color the curve of case 2, and the red color the curve of case 3. For each test function we have presented in figure 1, the result of the optimal values which we obtain from the algorithm until we arrive at the best step. From tables 2 we concluded that in case 1 and case 3 we have the best values which is noticed and emphasized also by figure 1. So, we can conclude that the convergence is faster in case 1 and case 3, when the inertia weight changes in each iteration of algorithm.

6. CONCLUSIONS

First, we explained thoroughly how the parameters $\alpha, \beta, \gamma, \delta, \eta$ are chosen to guarantee the convergence of the proposed variants of PSO. This was a very important issue of our work and it was based in the perturbed dynamical systems where we operated with analytic tools. The importance is shown in the request that this system must be stable. Then, for each of the proposed cases we took a group of functions considered as test functions. The results obtained are reasonable and satisfactory, emphasizing that case 1 and case 3 give the best results. This occurs due to the fact that the inertia weight changes in each iteration assuring a faster convergence then the cases when $w = 1$.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

- [1] M. Clerc, J. Kennedy, The particle swarm - explosion, stability, and convergence in a multidimensional complex space, *IEEE Trans. Evol. Computat.* 6 (2002), 58–73.
- [2] T. Ziyu, Z. Dingxue, A modified particle swarm optimization with an adaptive acceleration coefficients, in: 2009 Asia-Pacific Conference on Information Processing, IEEE, Shenzhen, China, 2009: pp. 330–332.
- [3] A. Sahu, S.K. Panigrahi, S. Pattnaik, Fast convergence particle swarm optimization for functions optimization, *Procedia Technol.* 4 (2012), 319–324.

- [4] Q. Bai, Analysis of particle swarm optimization algorithm, *Computer Inform. Sci.* 3 (2010), 180-184.
- [5] D. Tian, A review of convergence analysis of particle swarm optimization, *Int. J. Grid Distr. Comput.* 6 (2013), 117-128.