A Convergence Proof and Parameter Analysis of Central Force Optimization Algorithm

[1]Dongsheng Ding, *[1]Xiaoping Luo, [1]Jinfei Chen, [1]Xuejie Wang, [1]Pengying Du, [2]Yunfei Guo
[1]*Key lab of intelligent system, City College, Zhejiang University, Hangzhou 310015, China*
[2] *School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China*
*[1] *E-mail: luoxp@zucc.edu.cn*

Abstract

This paper presents a convergence proof of CFO on the stability theory of discrete-time-linear system. It reveals the necessary convergent conditions of CFO. Stability condition limits the eigenvalues of motion equations inside the unit cycle in complex plane, and a corresponding convergence criterion was deduced related to the probes' parameters. According to above criterion, a qualitative parameter analysis and a general strategy of selecting parameters of CFO is presented. Finally, a simple numerical experiment approved such strategy above completely.

Keywords: *Central force optimization (CFO), Global optimization, Proof of convergence, Parameter analysis, Gravitational force*

1. Introduction

Till now various heuristic optimization algorithms have been proposed and can be divided into biological-inspired ones and physical-inspired ones [1]. Biological-inspired ones are applied in many problems [2]. However, the uncertainty of macro biological theories on a micro individual is evitable, such as randomness [3], which results in heavier cost [4]. Because of the characteristic of general and deterministic, physical-inspired optimization algorithms hit the hot pot recently [1, 4~8].

Among many physical-inspired optimization algorithms, Central Force Optimization (CFO) is a novel optimization algorithm with many merits, for example, fast convergence, deterministic results and convenient to be applied [4]. Therefore CFO attracts researches' attention since it was proposed at 2007 and shows better performances than others [4].

Although CFO has been proved to be effective and convergent empirically, the convergence proof has not been seen yet. Using the stability theory to analyze the equations of CFO's probes, the proof of convergence is presented that reveals the necessary convergent conditions, and a corresponding convergent criterion was deduced related to its parameters in this paper. Accordingly, a qualitative analysis on the trajectory in [9] is given. The direction of analysis and improvement of CFO is provided.

The paper is organized as follows. CFO is described in section 2. The convergence proof is given in section 3. Section 4 carries a parameters' analysis to explain the convergent trajectories and how to select parameters. In section 5, numerical experiments are carried out. Section 6 presents conclusions.

2. CFO Algorithm

Consider an optimization problem:

Max $f\left(X\right),\ \Omega=\left\{X \mid x_d^{\min}<x_d<x_d^{\max}\right\},\ d=1,\cdots,N_d$

CFO is an optimization method by depicting particles' movements and particle's mass is navigated by the metaphorical principle of Newton's law. A general framework is shown in Table 1.

CFO has 3 procedures [4, 9]: (a) *Initialization*. (b) *Acceleration Calculation*. (c) *Motion*. *Initial distribution* is formed by deploying N_p/N_d particles with distribution factor γ [3]. N_p denotes particles number. Initial conditions are 0. In *Acceleration Calculation*, the aggregate acceleration

$$\vec{A}_{j-1}^{p}=G\sum_{\substack{k=1\\k\neq p}}^{N_p}\vec{A}_{j-1}^{k,p}=G\sum_{\substack{k=1\\k\neq p}}^{N_p}\left[U(M_{j-1}^{k}-M_{j-1}^{p})\square(M_{j-1}^{k}-M_{j-1}^{p})^{\alpha}\times\frac{(\vec{R}_{j-1}^{k}-\vec{R}_{j-1}^{p})}{\left\|\vec{R}_{j-1}^{k}-\vec{R}_{j-1}^{p}\right\|^{\beta}}\right] \quad (1)$$

Where $\vec{A}_{j-1}^{k,p}$ is an accelerate vector of particle k toward to particle p. $M_{j-1}^{p}=f(x_1^{p,j-1},x_2^{p,j-1},\cdots,x_{N_d}^{p,j-1})$ is the objective function value of particle p at $j-1$. $j\in\left\{1,\cdots,N_t\right\}$. G, α and β are user-defined.

In *Motion*, particle- p will move from $\vec{R}_{j-1}^{p}$ to $\vec{R}_{j}^{p}$ according to Eq.(2) .

$$\vec{R}_{j}^{p}=\vec{R}_{j-1}^{p}+\frac{1}{2}\vec{A}_{j-1}^{p}\Delta t^2,\ j\geq 1 \quad (2)$$

If some "fly" out, a necessary retrieving mechanism acts by F_{rep}(in Table 1) [11].

More detailed information on CFO can be seen in [4, 9]

Table 1. Description of CFO

Step 1: Initialization A Set N_d, $X_{\min}$, $X_{\max}$, N_p, N_t, G, α, β, γ and ΔF_{rep};
Step 2: Initialization B Compute initial probe distribution X with increment γ, initial fitness matrix M; Assign initial particle acceleration A; Set initial F_{rep}
Step 3: Evaluate earlier termination criterion or less than N_t If reach N_t, return to **Step 2**; If fulfill stop criterion, return ***Step 4***
Step 3.1: Compute probe position vectors
Step 3.2: Retrieve errant particles
Step 3.3: Compute fitness matrix of current particle distribution and finesses
Step 3.4: Increment F_{rep} ***Return Step 3***
Step 4: Stop and putout best solution have reached so far

3. Convergence Analysis

Without loss of generality, consider particle i arbitrarily and one-dimensional, i.e.$N_d = 1$. Assume the feasible region is always positive.

Theorem 1. CFO's motion is an ordinary difference equation in essence.

Proof: According to mass step function [3], define $M_i=\left\{k \mid f\left(X_k\right)>f\left(X_i\right),k=1,\cdots,N_p\right\}$.

Eq. (1) can be written as

$$A_i(j) = G \cdot \sum_{k \in M_i} \frac{\left[f(X_k) - f(X_i)\right]^{\alpha}}{\left\|X_k - X_i\right\|^{\beta}} X_k(j) - G \cdot \sum_{k \in M_i} \frac{\left[f(X_k) - f(X_i)\right]^{\alpha}}{\left\|X_k - X_i\right\|^{\beta}} X_i(j) \quad (3)$$

Define $$\phi_i = G \cdot \sum_{k \in M_i} \frac{\left[f(X_k) - f(X_i)\right]^{\alpha}}{\left\|X_k - X_i\right\|^{\beta}} X_k(j) \quad (4)$$

$$\theta_i = G \cdot \sum_{k \in M_i} \frac{\left[f(X_k) - f(X_i)\right]^{\alpha}}{\left\|X_k - X_i\right\|^{\beta}} \quad (5)$$

Eq. (3) can be updated as $A_i(j) = \phi_i - \theta_i \cdot X_i(j)$ (6)

Combined Eq.(6) with Eq. (2), CFO system equations are obtained as ($\theta_i > 0$ is obvious.)

$$A_i(j) = \phi_i - \theta_i \cdot X_i(j) \quad (7)$$

$$X_i(j+1) = X_i(j) + \frac{1}{2} A_i(j) \cdot \Delta t^2 \quad (8)$$

Substitute Eq. (7) into Eq. (8), the following equation is obtained.

$$X_i(j+1) + \left(\frac{1}{2}\Delta t^2 \cdot \theta_i - 1\right) X_i(j) = \frac{1}{2}\Delta t^2 \cdot \phi_i \quad (9)$$

Because ϕ_i and θ_i are constant at each generation, Eq. (9) is an ordinary difference equation. □

The characteristic equation of Eq. (9) is: $\lambda + \frac{1}{2}\Delta t^2 \cdot \theta_i - 1 = 0$ (10)

Where $\lambda \in R$. The stability condition of discrete-time-linear system is eigenvalues lie inside the unit cycle in the complex plane. The following theorem shows that Eq. (9) is convergent.

Theorem 2. Given $\theta_i > 0$, if and only if $\Delta t \in (-\sqrt{4/\theta_i}, 0) \cup (0, \sqrt{4/\theta_i})$, $\{X(j)\}, j = 0, 1, \cdots$ converges to the limit X_i^{best}.

Proof. According to Eq.(10), $|\lambda| = \left|1 - \frac{1}{2}\Delta t^2 \cdot \theta_i\right| < 1$. (11)

Thus $$-1 < 1 - \frac{1}{2}\Delta t^2 \cdot \theta_i < 1 \Leftrightarrow 0 < \frac{1}{2}\Delta t^2 \cdot \theta_i < 2 \quad (12)$$

Thus the stability condition becomes $\Delta t \in (-\sqrt{4/\theta_i}, 0) \cup (0, \sqrt{4/\theta_i})$ (13)

Substitute Eq. (8) into Eq. (7), we have

$$A_i(j+1) = \left(\frac{1}{2}\theta_i \Delta t^2 - 1\right) A_i(j) \quad (14)$$

Under stability condition Eq. (11), takes limits on both sides of Eq. (14), we have

$$\lim_{j\to\infty} A_i(j+1) = \lim_{j\to\infty}\left(\frac{1}{2}\theta_i \Delta t^2 - 1\right) A_i(j) = 0 \tag{15}$$

To illustrate $\{X(j)\}, j = 0,1,\cdots$ converges to X_{best}, a limit $\lim_{j\to\infty} X_i(j) = X_i^*, i = 1,\cdots,N_p$ is defined based on stability condition and Eq. (9) can expressed accordingly as

$$X_i^*(j+1) + \left(\frac{1}{2}\Delta t^2 \cdot \theta_i - 1\right) X_i^*(j) = \frac{1}{2}\Delta t^2 \cdot \phi_i \tag{16}$$

Simplify Eq. (16), then $\theta_i X_i^* = \phi_i$ (17)

Substitute Eq. (4) and Eq. (5) into Eq. (17), the following equation is obtained.

$$G \cdot \sum_{k\in M_i} \frac{\left[f(X_k^*) - f(X_i^*)\right]^\alpha}{\left\|X_k^* - X_i^*\right\|^\beta} X_i^* = G \cdot \sum_{k\in M_i} \frac{\left[f(X_k^*) - f(X_i^*)\right]^\alpha}{\left\|X_k^* - X_i^*\right\|^\beta} X_k^* \tag{18}$$

Considering the feasible region being positive, we have

$$\sum_{k\in M_i} \left\{\alpha \cdot \ln\left[f(X_k^*) - f(X_i^*)\right] - \beta \ln\left(\left\|X_k^* - X_i^*\right\|\right) + \ln(X_i^*)\right\}$$

$$= \sum_{k\in M_i} \left\{\alpha \cdot \ln\left[f(X_k^*) - f(X_i^*)\right] - \beta \ln\left(\left\|X_k^* - X_i^*\right\|\right) + \ln(X_k^*)\right\} \tag{19}$$

Define the number of elements in set M_i as NM_i. Eq. (19) is simplified as

$$\sum_{k\in M_i} \ln(X_i^*) = \sum_{k\in M_i} \ln(X_k^*) \Rightarrow \ln(X_i^*) = \frac{\sum_{k\in M_i} \ln(X_k^*)}{NM_i} \tag{20}$$

Thus $X_k^*, k \in M_i$ reach the limit X_i^{best} simultaneously, i.e. $X_1^* = X_2^* = \cdots = X_{NM_i}^* = X_i^{best}$. Otherwise, according to the Eq. (8), $\lim_{j\to\infty} A_i(j+1) \neq 0$, which is contradictory to Eq. (15).

Thus Eq. (20) is written as $\ln(X_i^*) = \ln(X_k^*) = \ln(X_i^{best}), k \in M_i$ (21)

Thus $\{X(j)\}, j = 0,1,\cdots$ converges to limit X_i^{best} □

4. Parameters Analysis

A qualitative analysis is carried out based on *Theorem 2* to solve the left question in [9] and point out how to select Δt to enhance the performance. A diagram on θ_i -Δt according to Eq. (13) is drawn in Fig. 1. The region between *stability limit* is stable except $\Delta t = 0$ and the outer is unstable.

Unstable region are divided into

(1) $\lambda > 1$, then $\theta_i < 0$

(2) $\lambda < -1$, then $\Delta t < -\sqrt{4/\theta_i} \vee \Delta t > \sqrt{4/\theta_i}$

(*Unstable region E, F,* emanate amplitude oscillate)

Comparatively, stable region are divided into

(1) $0 < \lambda < 1$, then $-\sqrt{2/\theta_i} < \Delta t < \sqrt{2/\theta_i}$ (*Stable region A, B,* power decay)

(2) $\lambda = 0$, then $\Delta t = \pm\sqrt{2/\theta_i}$ (*Zero line*, persist)

(3) $-1 < \lambda < 0$, then $-\sqrt{4/\theta_i} < \Delta t < -\sqrt{2/\theta_i} \vee \sqrt{2/\theta_i} < \Delta t < \sqrt{4/\theta_i}$

(*Stable region C, D,* attenuate amplitude oscillate)

The remainders are two critical stable limits:

(1) $\lambda = 1$, then $\theta_i = 0$

(2) $\lambda = -1$, then $\Delta t = \pm\sqrt{4/\theta_i}$ (*Stability limit*, equivalent amplitude oscillate)

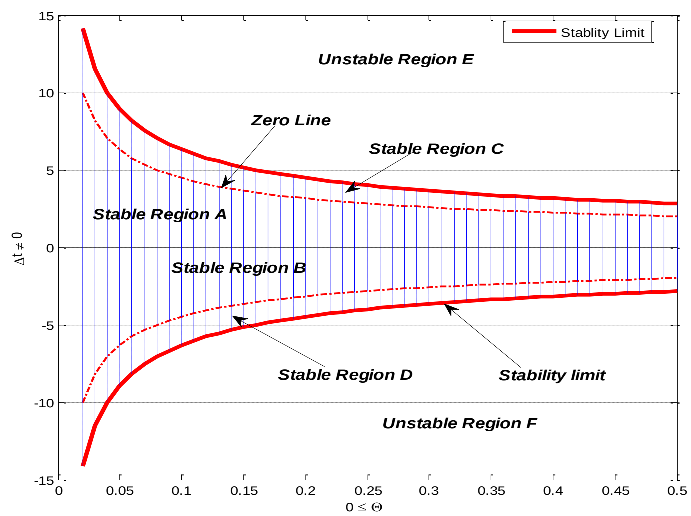

Figure. 1. A diagram on θ_i-Δt

For simplicity, consider a system with two particles. From Fig. 1, It can be seen clearly that when larger Δt is, smaller the stable region becomes. Therefore for a multimodal function optimization, larger Δt is helpful to search an unstable region to enhance the global search ability. While for a unimodal function, smaller Δt is desirable for rapid convergence. Also the necessary condition for CFO entering stable region is θ_i converges to zero, so α need to be set larger than β.

Because θ_i is proportional to *G*, *G* is another key parameter for convergence. A strategy for electing *G* will be deduced based on *Theorem 2* below with supporting experiments in section 5.

Considering definitions $\omega_{\max} = \left(f\left(X_{\max}\right) - f\left(X_{\min}\right)\right)^{\alpha}$, $\chi_{\min} = \min\left\{\left\|X_{\max} - X_{\min}\right\|^{\beta}\right\}$, substitute Eq. (6) into Eq. (13), we have

$$0 < \Delta t^2 \cdot G \cdot \sum_{k \in M_i} \frac{\left[f\left(X_k\right) - f\left(X_i\right) \right]^{\alpha}}{\left\| X_k - X_i \right\|^{\beta}} < \Delta t^2 \cdot G \cdot \sum_{k \in M_i} \frac{w_{\max}}{\chi_{\min}} < \Delta t^2 \cdot G \cdot N_p \frac{w_{\max}}{\chi_{\min}} < 4$$

Given Δt and N_p, so $0 < G < \dfrac{4\chi_{\min}}{\Delta t^2 \cdot N_p \cdot w_{\max}}$ (22)

If G is much larger, θ_i will become larger, particles will make emanating amplitude oscillating for global search, which is necessary at initial. If θ_i is near zero, particles will make power decaying for rapid convergence, and then particles will lose global searching ability later. Generally, we recommend G to be with bell-shaped distribution, not proposed by [10]. Initially, particles exploit entirely the information obtained to converge with a lower G, and then expand for a global landscape by increasing G for better. At last, particles converge to the global optima after completing adaptive global search. The experiments with varying G are shown in section 5.

5. Numerical Experiments

The varying G is investigated by four benchmark functions with 2 variables [11] in Table 2. Differ from scholastic algorithms, CFO gets a deterministic and accurate result under predefined parameters no matter how many times [4, 9]. Total particles are $2 \le N_p \le 15$ with the number of particles being increased 1 each run. γ and F_{rep} are increased by 0.06 with the initial value being 0.06. G is set by the upper bound of Eq. (22).N_t = 1000, α, β =2.

Table 2. Test functions and Performance of CFO

Function	Name	Search range	Maximum (best)	CFO (best)
$F_1(x) = -\sum_{i=1}^{n}\left[x_i^2 - 10\cos\left(2\pi x_i\right) + 10\right]$	*Rastrigin*	$[-5.12, 5.12]^2$	0	**0**
$F_2(x) = -\sum_{i=1}^{n-1}\left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$	*Rosenbrock*	$[-30, 30]^2$	0	**-6.1001e-029**
$F_3(x) = -\max_i \left\{\lvert x_i \rvert, 1 \le i \le n\right\}$	*Schwefel2.21*	$[-100, 100]^2$	0	**-1.2163e-015**
$F_4(x) = -\sum_{i=1}^{n}\left(\lfloor x_i + 0.5 \rfloor\right)^2$	*Step*	$[-100, 100]^2$	0	**-2.4652e-030**

In Fig.2 vertical axis is best function value in right column, left is G. Horizontal axes are the number of iterative times. In Fig. 2, particles exploit the information obtained to converge rapidly with a lower G. If CFO converges close to optima, the population's diversity decreases with a larger G; it will promote CFO's particles to search new space globally. We recommend that G can be adjusted by bell-shaped distribution. However, the elitist preservation is inherent in CFO described in section 2, so particles will converge to the best global optima after completing such adaptive global search. Table 2 presents the results obtained by CFO.

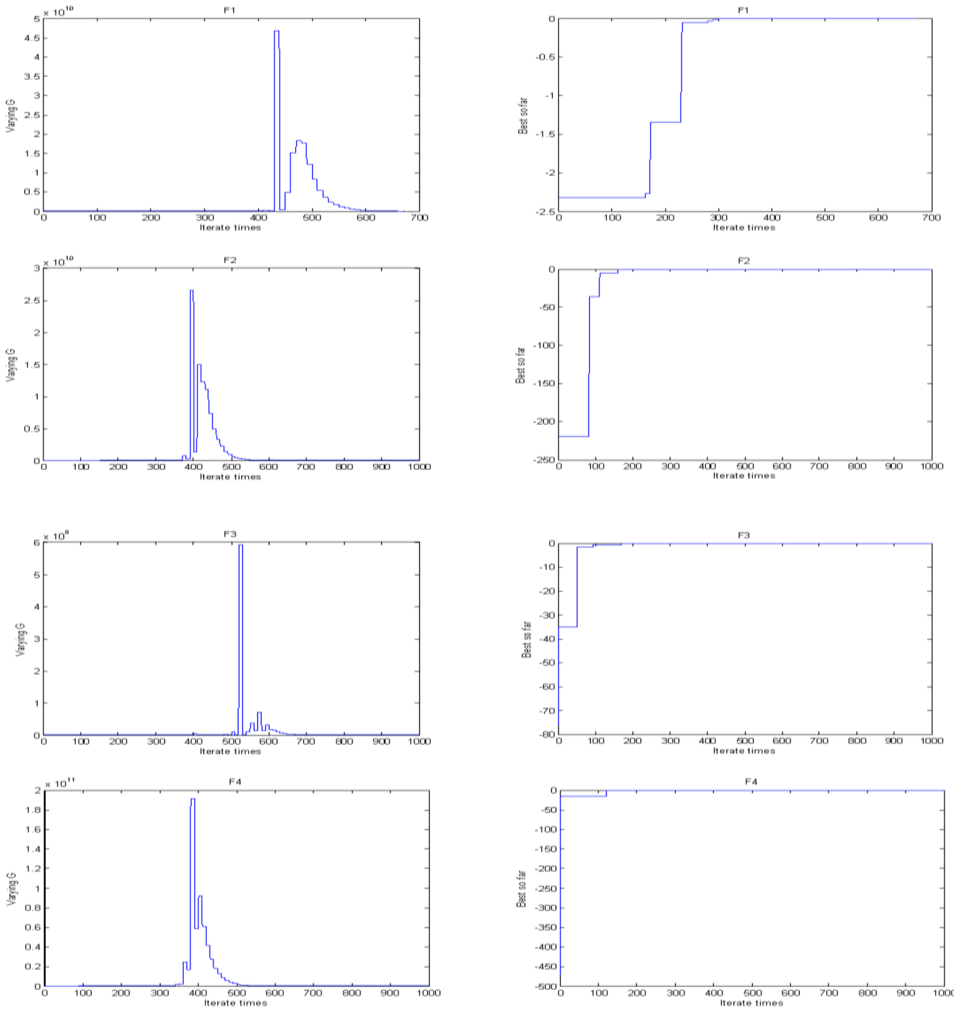

Figure. 2. Results of CFO on F_1-F_4

6. Conclusions

CFO algorithm is a novel deterministic nature heuristic optimization algorithm. A convergence proof of it reveals the necessary convergent conditions under which it is guaranteed to converge to the optima. Thus, a comprehensive parameter analysis with G and Δt is carried out. Based on discrete-time-linear system theory, we give a qualitative analysis of θ_i -Δt and explain the left question in [9]. It directs how to choose proper parameters of CFO. Finally, a simple numerical experiment approves such strategy on G above completely.

Besides, the gravitational metaphor is fascinating. There are many other effects such as dark energy, black hole or supernova, which may implicate global optima. How to simplify principles of them to propose new physical-inspired optimization algorithms is significant.

7. Acknowledgements

The authors give special thanks to Dr. Richard Formato.

The work is supported by Zhejiang Nature Science Foundation under Grant Y1080776 and National Science Foundation of China under Grant 60702023, 60805013.

8. Reference

[1] Ying-Tung Hsiao, Cheng-Long Chuang, Joe-Air Jiang, Cheng-Chih Chien, “A novel optimization algorithm: space gravitational optimization”, IEEE International Conference on Systems, Man and Cybernetics, Vol. 3, pp. 2323-2328, 2005.

[2] Zhiyong Li, Ransikarn Ngambusabongsopa, Esraa Mohammed, Ndatinya Eustache, "A Novel Diversity Guided Particle Swarm Multi-objective Optimization Algorithm", JDCTA: International Journal of Digital Content Technology and its Applications, Vol. 5, No. 1, pp. 269-278, 2011.

[3] JianGang Yang, JinQiu Yang, "Intelligence Optimization Algorithms: A Survey", IJACT : International Journal of Advancements in Computing Technology, Vol. 3, No. 4, pp. 144-152, 2011.

[4] R.A.Formato, “Comparative Results: Group Search Optimizer and Central Force Optimization”, arXiv:1002.2798v2, 2010.

[5] S.ILKER B IRBI L and SHU-CHERNG FANG, “An Electromagnetism-like Mechanism for Global Optimization”, Journal of Global Optimization, Vol. 25, pp.263–282, 2003.

[6] L.Xie, Jianchao Zeng, Richard A. Formato, “Convergence analysis and performance of the extended artificial physics optimization algorithm”(in press), Applied Mathematics and Computation, doi:10.10.1016/j.amc.2011.02.062.

[7] Esmat Rashedia, Hossein Nezamabadi-pour, Saeid Saryazdia, “GSA: A Gravitational Search Algorithm”, Information Sciences, Vol. 179, No. 13, pp.2232-2248, 2009.

[8] Cheng-Long Chuang, Joe-Air Jiang, “Integrated radiation optimization: inspired by the gravitational radiation in the curvature of space-time”, IEEE Congress on Evolutionary Computation (CEC), Singapore, Issue Date: 25-28, pp.3157 – 3164, 2007.

[9] R.A.Formato, “Are Near Earth Objects the Key to Optimization Theory” arXiv:0912.1394, 2009.

[10] Esmat Rashedia, Hossein Nezamabadi-pour, Saeid Saryazdia, “GSA: A Gravitational Search Algorithm”, Information Sciences, Vol. 179, No. 13, pp.2232-2248, 2009.

[11] Xin Yao, Yong Liu, Guangming Lin, “Evolutionary programming made faster”, IEEE Transactions On Evolutionary Computation, Vol. 3, No. 2, pp.82-102, 1999.