

Bifurcation and Convergence of Particle Swarm Optimization Dynamics Embedded into Upper and Lower Bound

Hideki Murata Non-member (Keio University)

Eitaro Aiyoshi Member (Keio University)

Keywords: particle swarm optimizaiton, chaotic dynamical systems, global optimization, constraint conditions

1. Introduction

In this paper, two types of dynamics of Particle Swarm Optimization (PSO) embedded into upper and lower bounds are presented. These dynamics have more powerful abilities of searching the global optimal solution in a multi-peaked function with more variables. The presented new type of dynamics called “nonlinear operator model” and “nonlinear variable transformation model” are introduced by discretizing inertial models on continuous time. After certifying stability and instability by the bifurcation diagram, the influences of model parameters settings on the global search are investigated to achieve better performance than the orthodox PSO dynamics in computer simulations.

2. Dynamic Models on Continuous Time

For typical optimization problems with upper and lower bounds formulated as

$$\min_{\mathbf{x}} E(\mathbf{x}) \dots\dots\dots (1a)$$

$$\text{subj.to } p_i \leq x_i \leq q_i, \quad i = 1, \dots, n, \dots\dots\dots (1b)$$

the following new type of dynamic models of PSO embedded into upper and lower bounds.

Nonlinear operator model:

$$\frac{d\mathbf{u}^p(t)}{dt} = -a\mathbf{u}^p(t) + \mathbf{v}^p(t) \dots\dots\dots (2a)$$

$$\frac{d\mathbf{v}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{x}^p(t), t) + \mathbf{C}(\mathbf{x}^p(t), t)] \dots\dots\dots (2b)$$

$$\mathbf{F}^p(\mathbf{x}^p, t) = c_1\{\mathbf{x}^p(T^p(t)) - \mathbf{x}^p\} \dots\dots\dots (2c)$$

$$\mathbf{C}(\mathbf{x}^p, t) = c_2\{\mathbf{x}^{Q(t)}(T^o(t)) - \mathbf{x}^p\} \dots\dots\dots (2d)$$

Nonlinear variable transformation model:

$$\frac{d\mathbf{u}^p(t)}{dt} = -a\mathbf{u}^p(t) + \mathbf{v}^p(t) \dots\dots\dots (3a)$$

$$\frac{d\mathbf{v}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{u}^p(t), t) + \mathbf{C}(\mathbf{u}^p(t), t)] \dots\dots\dots (3b)$$

$$\mathbf{F}^p(\mathbf{u}^p, t) = c_1\{\mathbf{u}^p(T^p(t)) - \mathbf{u}^p\} \dots\dots\dots (3c)$$

$$\mathbf{C}(\mathbf{u}^p, t) = c_2\{\mathbf{u}^{Q(t)}(T^o(t)) - \mathbf{u}^p\} \dots\dots\dots (3d)$$

In both models commonly, the p-best, the g-best and the output function are

$$T^p(t) = \underset{\tau}{\operatorname{argmin}}\{E(\mathbf{x}^p(\tau)) \mid 0 \leq \tau \leq t\} \dots\dots\dots (4a)$$

$$(Q(t), T^o(t)) = \underset{(q, \tau)}{\operatorname{argmin}}\{E(\mathbf{x}^q(\tau))\}$$

$$q = 1, \dots, P, \quad 0 \leq \tau \leq t \dots\dots\dots (4b)$$

$$x_i^p(t) = f_i(u_i^p(t)) = \frac{q_i + p_i \exp(-u_i^p(t))}{1 + \exp(-u_i^p(t))}$$

$$i = 1, \dots, n, \dots\dots\dots (4c)$$

respectively. On the other hand, it is noticed that the driving forces of the particle swarm depend on output variables in (2b) and on state variables in (3b).

3. Discretized Models

By discretizing the above two types model, the discrete model are obtained. The discretised version of Eq.(2a) or (3a)is

$$\mathbf{u}^p(k+1) = (1 - a\Delta T)\mathbf{u}^p(k) + \Delta T\mathbf{v}^p(k), \dots\dots\dots (5)$$

and Eq.(2b)and(3b)are followings repectively.

Nonlinear operator model:

$$\mathbf{v}^p(k+1) = \mathbf{v}^p(k) + c\Delta T[\mathbf{F}^p(\mathbf{x}^p(k), k) + \mathbf{C}(\mathbf{x}^p(k), k)] \dots\dots\dots (6)$$

Nonlinear variable transformation model:

$$\mathbf{v}^p(k+1) = \mathbf{v}^p(k) + c\Delta T[\mathbf{F}^p(\mathbf{u}^p(k), k) + \mathbf{C}(\mathbf{u}^p(k), k)] \dots\dots\dots (7)$$

4. Computational Results

The computational result for Rastrigin function, which is one of bench marks, is shown in Table 1 when the number of variables is increased. The convergence rate of gbest to global optimum through 1000 trials by each model with 20 particles is shown in Table 1. Each computational simulation is ended in 1000 iterations. In case(1), c_1 and c_2 are random and ΔT is relatively big. In case(2), c_1 and c_2 are random and ΔT is relatively small. In case(3), c_1 and c_2 are constant and ΔT is relatively big. Incase(2), c_1 and c_2 are constant and ΔT is relatively small.

Table 1. Convergence rate to the global optimum for increase of the number of variables

model	case	n=3	n=5	n=10
NO	(1)	97.5%	99.9%	99.7%
	(2)	5.4%	0.4%	0.0%
	(3)	54.6%	15.7%	3.0%
	(4)	65.7%	39.2%	15.7%
NT	1	81.3%	51.3%	11.4%
	(2)	21.6%	0.4%	0.0%
	(3)	51.4%	21.5%	9.9%
	(4)	75.2%	50.2%	28.2%
PSO		30.4%	0.0%	0.0%

上下限領域に閉じ込めた Particle Swarm Optimization の力学系の分岐特性と収束特性

非会員 村田 秀樹* 正員 相吉 英太郎*

Bifurcation and Convergence of Particle Swarm Optimization Dynamics Embedded into Upper and Lower Bound

Hideki Murata*, Non-member, Eitaro Aiyoshi*, Member

In this paper, we propose two types of dynamics of Particle Swarm Optimization (PSO) embedded into upper and lower bounds, which have more powerful abilities of searching the global optimal solution on their bounded constraints. The new type of dynamics called "nonlinear operator model" and "nonlinear variable transformation model" are inertial models, which are introduced by discretizing the convolution integral form on continuous time. After certifying stability and/or instability including chaos of dynamics by the bifurcation diagrams, the influences of model parameters settings on the global search are investigated to achieve better performance than the orthodox PSO dynamics in computational simulations for a few benchmark objective functions.

キーワード: particle swarm optimization, カオス力学系, 大域的最適化, 制約条件

Keywords: particle swarm optimization, chaotic dynamical systems, global optimization, constraint conditions

1. はじめに

近年, 生体生命現象や物理現象を模倣し, それらを単純化したモデルに基づいて, 大域的探索をおこなって凹凸状の目的関数の大域的最適解を比較的効率よく探索する最適化手法が提案されている。その大域的探索能力の高さが注目を集めている手法の一つとして, 生物の群の挙動からヒントを得た Particle Swarm Optimization⁽²⁾(以降 PSO と略記する) とよばれる手法がある。この手法は, 目的関数の勾配情報を用いずに複数の探索点を更新し, かつそのパラメータを確率的に揺らす多点型確率的試行探索法であるが, 力学系のモデルとして特徴づけると, 「慣性型移流 (対流) 結合系モデル」と称することができる。

ところで, PSO に関する研究は, この手法自体の特性に関するものから応用研究にいたるまで, すでに数多くなされているが, 一般に, 変数の次元の小さい無制約問題に対しては, 更新式のパラメータを適切に調節すれば, 比較的効率よく大域的最適解が探索可能であるのに対して, 変数次元の大きな問題に対しては, 必ずしも大域的最適化が効率よく求められないことが指摘されている⁽⁶⁾。そこで本論

文では, 変数次元が比較的大きな問題であっても, ある程度狭い有界な領域を制約条件とする問題であるならば, その領域上の大域的最適解を比較的効率よく探索できる制約条件付最適化問題を直接解くことが可能な PSO を提案する。とくに本論文では, まず, この PSO がもつパラメータは確定的な値をとるものとした下で, PSO を上記のような力学モデルとして解釈し, まず上下限制約領域内に束縛した PSO モデルを導出する。このために, 畳み込み積分を用いて慣性型連続時間系 PSO モデルを表し, それに非線形作用素を施すことによって上下限制約領域に束縛したモデルと, 上下限制約を全空間に対応づける変数変換を施したモデルの 2 通りのモデル, およびそれらに対応した内部状態表現モデルを導出する。いずれのモデルにおいても内部状態変数が慣性型連続時間系の PSO モデルで表され, その状態量を出力関数に通すことによって, 上下限制約に束縛された挙動を観測することができる。この出力関数はいわゆるポップフィールド型ニューラルネットワークやカオニューラルネットワークのモデルに用いられるシグモイド関数と同じである。

つぎに, これらの連続系モデルに対してオイラー法で離散化すると, 内部状態に関しては通常の PSO モデルと同様の慣性型離散時間系のモデルと, それに連立する形でシグモイド型出力関数が得られる。このようにいったん上下限領域への束縛操作を連続時間系モデルに取り込み, それ

* 慶應義塾大学

〒 223-8522 横浜市港北区日吉 3-14-1

Keio University

3-14-1, Hiyoshi, Kohoku-ku, Yokohama, 223-8522

を内部状態表現してから離散化を行うことによって、上下制限約条件をつねに満足させる離散時間系を得ることができる。

さらに、このような有界な制約条件の領域内で閉じ込めた離散時間系において、その離散化のサンプリングパラメータを大きくして不安定化させると、カオス現象が現れる可能性があることが知られている⁽⁷⁾⁽⁸⁾。そこで本論文では、その離散化のサンプリングパラメータを大きくすることによって、PSO の力学系を不安定化させ、そのモデルからカオス現象が現れることを確認し、このカオスモデルに対する徐冷操作によって、大域的最適点を得る確定的な PSO モデルを提案する。

ところで、従前の PSO の力学系での一部のパラメータの確率的揺動は、いわば局所的な揺らぎを付与していることであり、変数次元が大きい無制約問題の大域的探索には効果が薄れるものと思われる。一方、本論文で提案する上下制限約に埋め込んだ PSO の力学系モデルは、その探索範囲が有界な制約領域に限定されるため、むしろ従前のパラメータの確率的揺動も効果が期待される。また、上下制限約という有界領域に力学系が閉じ込められるため、個体ごとの収束を必ずしも強制せずに不安定状態のまま放置していても、過去の探索で最良の状態であるいわゆる g-best に注目しているかぎり、大域的最適解の発見割合は比較的高いものと期待される。とくに、カオス特性をもたせて、この有界領域を確定的に走査できれば、その場合の大域的最適解の発見割合も低くはないものと期待される。そこで、本論文では、提案する上下制限約に埋め込んだ PSO 力学系モデルの収束性能について、多峰性関数のベンチマークにより検討した結果を示す。

2. 離散時間慣性型移流 (対流) 結合系 PSO モデル

本章では無制約最適化問題

$$\min_{\mathbf{x}} E(\mathbf{x}) \dots\dots\dots (1)$$

を解くための PSO モデルを扱う。

〈2・1〉 慣性系としての PSO モデル まず、PSO の力学系としての特徴を考えるために、探索点が単一の場合を考え、その更新式は

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \Delta\mathbf{x}(k+1) \dots\dots\dots (2a)$$

$$\Delta\mathbf{x}(k+1) = \lambda\Delta\mathbf{x}(k) + c_1\{\mathbf{x}(l(k)) - \mathbf{x}(k)\} \dots\dots\dots (2b)$$

$$l(k) = \underset{i}{\operatorname{argmin}}\{E(\mathbf{x}(i)) \mid i = 0, 1, \dots, k\} \dots\dots (2c)$$

$$\mathbf{x}(0), \Delta\mathbf{x}(0); \text{ given, } k = 0, 1, 2, \dots$$

となる。ただし、 $l(k)$ は時刻 k の時点までに関数 E を最も小さくした過去の時刻である。ここで (2b) 式の右辺第 2 項を

$$\mathbf{F}(\mathbf{x}, k) = c_1\{\mathbf{x}(l(k)) - \mathbf{x}\} \dots\dots\dots (3)$$

とおくと、(2b) 式は

$$\Delta\mathbf{x}(k+1) = \lambda\Delta\mathbf{x}(k) + \mathbf{F}(\mathbf{x}(k), k) \dots\dots\dots (4)$$

と一般化される。ここで \mathbf{F} の 2 番目の引数 k は $l(k)$ の k であり、関数 \mathbf{F} が時刻 k の変動に直接影響を受けているを表している。(2a) 式と (4) 式から $\Delta\mathbf{x}(k)$ を消去すると、

$$\mathbf{x}(k+1) - (1+\lambda)\mathbf{x}(k) + \lambda\mathbf{x}(k-1) = \mathbf{F}(\mathbf{x}(k), k)$$

$$k = 1, 2, \dots \dots\dots (5)$$

が得られて、特性根として 1 と λ を有する 2 階差分方程式であることがわかり、 $\lambda \leq 0$ のとき振動することがわかる。

さらに、(5) 式を解いて、級数 (和分) 形式で書き直すと、

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \sum_{i=0}^k \lambda^{k-i} \mathbf{F}(\mathbf{x}(i), i) + \lambda^{k+1} \Delta\mathbf{x}(0)$$

$$\dots\dots\dots (6a)$$

$$\mathbf{F}(\mathbf{x}(0), 0) = c_1\{\mathbf{x}(0) - \mathbf{x}(0)\} = \mathbf{0} \dots\dots\dots (6b)$$

となり、明らかに過去の $\mathbf{F}(\mathbf{x}(i), i)$, $i = 0, 1, \dots, k-1$ を重み λ^{k-i} で畳み込んで変位する慣性系であることがわかる。ここで仮に k とは無関係に

$$\mathbf{F}(\mathbf{x}, k) = -c_1 \nabla E(\mathbf{x}) \dots\dots\dots (7)$$

とおくと、勾配法型慣性系である。勾配法における $-\nabla E(\mathbf{x})$ は、 \mathbf{x} において関数 E を局所的に最も減少させる方向であるのに対して、PSO では勾配を用いる代わりに (3) 式を用いていることになるが、 $\mathbf{x}(l(k))$ が過去の時系列でもっとも E を小さくする状態であることから、(3) 式の \mathbf{F} も等号を含めた広義の関数 E の減少方向を与えるものであることがわかる。

さらに、高階の差分方程式は内部状態表現を用いるのが慣例であるから、2 種類の内部状態量 \mathbf{u} , \mathbf{v} を導入し、

$$\mathbf{u}(k) = \mathbf{x}(k) \dots\dots\dots (8a)$$

$$\mathbf{v}(k) = \mathbf{x}(k) - \mathbf{x}(k-1) = \Delta\mathbf{x}(k) \dots\dots\dots (8b)$$

において、(5) 式を内部状態表現式に変換したモデル

$$\mathbf{u}(k+1) = \mathbf{u}(k) + \mathbf{v}(k+1) \dots\dots\dots (9a)$$

$$\mathbf{v}(k+1) = \lambda\mathbf{v}(k) + \mathbf{F}(\mathbf{u}(k), k) \dots\dots\dots (9b)$$

$$\mathbf{u}(0), \mathbf{v}(0); \text{ given, } k = 0, 1, \dots$$

を導入しておく。

〈2・2〉 移流 (対流) 結合系としての PSO モデル 複数の探索点を更新する PSO は、多体結合系のモデルである。ここで、個体数を P ($p = 1, \dots, P$) とし、 $\mathbf{x}^p(k)$ を個体 p の時刻 k における状態とすると、前節の慣性系としての個体 \mathbf{x}^p が、関数 $\mathbf{C}(\mathbf{x}^p, k)$ によって、他の個体からの影響を受けるモデルとして

$$\mathbf{x}^p(k+1) = \mathbf{x}^p(k) + \Delta\mathbf{x}^p(k+1) \dots\dots\dots (10a)$$

$$\Delta\mathbf{x}^p(k+1) = \lambda\Delta\mathbf{x}^p(k) + \mathbf{F}^p(\mathbf{x}^p(k), k)$$

$$+ \mathbf{C}(\mathbf{x}^p(k), k) \dots\dots\dots (10b)$$

と考えることができる。PSO の場合、関数 F^p や結合項を表す関数 C は

$$F^p(\mathbf{x}^p, k) = c_1\{\mathbf{x}^p(l^p(k)) - \mathbf{x}^p\} \dots \dots \dots (11a)$$

$$C(\mathbf{x}^p, k) = c_2\{\mathbf{x}^{Q(k)}(l^o(k)) - \mathbf{x}^p\} \dots \dots \dots (11b)$$

である。ここで $l^p(k)$ や $(Q(k), l^o(k))$ は

$$l^p(k) = \underset{l}{\operatorname{argmin}}\{E(\mathbf{x}^p(l)) \mid l = 0, 1, \dots, k\} \dots \dots \dots (12a)$$

$$(Q(k), l^o(k)) = \underset{(q,l)}{\operatorname{argmin}}\{E(\mathbf{x}^q(l)) \mid q = 1, 2, \dots, P, l = 0, 1, \dots, k\} \dots \dots \dots (12b)$$

である。また $l^p(k)$ は、時刻 k までにおいて、個体 p が関数 E の値をもっとも小さくした時刻を表していて、その状態 $\mathbf{x}^p(l^p(k))$ は PSO で p-best とよばれている。また、 $(Q(k), l^o(k))$ は、時刻 k までにおいて、 P 個の個体の中で関数 E の値をもっとも小さくした個体番号とその時刻を表し、その状態 $\mathbf{x}^{Q(k)}(l^o(k))$ は PSO で g-best とよばれている。ここで、 $l^o(k)$ は個体 p に対応して決まる離散時刻であるから、 $Q(k)$ や $l^o(k)$ との間には

$$Q(k) = \underset{q}{\operatorname{argmin}}\{E(\mathbf{x}^q(l^q(k))) \mid q = 1, 2, \dots, P\} \dots \dots \dots (13a)$$

$$l^o(k) = l^{Q(k)}(k) \dots \dots \dots (13b)$$

という関係がなりたつ。ところで、g-best については、時間を過去に遡らずに、(12) 式において $l = k$ のみとし、

$$Q(k) = \underset{q}{\operatorname{argmin}}\{E(\mathbf{x}^q(k)) \mid q = 1, 2, \dots, P\} \dots \dots \dots (14)$$

とおき、その個体の状態 $\mathbf{x}^{Q(k)}$ を g-best の定義とする場合もある。このとき (11b) 式の関数 C は単に

$$C(\mathbf{x}^p, k) = c_2\{\mathbf{x}^{Q(k)}(k) - \mathbf{x}^p(k)\} \dots \dots \dots (15)$$

となる。 $Q(k) \neq p$ ならば、その時点の時刻 k でもっとも良い状態の個体 $Q(k)$ の状態 $\mathbf{x}^{Q(k)}(k)$ に、個体 p の状態 \mathbf{x}^p が近づくように移流するいわゆる対流結合であり、 $l^o(k) < k$ ならば、(11b) 式は過去の時刻 $l^o(k)$ の個体 $Q(k)$ の状態 $\mathbf{x}^{Q(k)}(l^o(k))$ も記憶していて、その状態へ個体 p の状態 \mathbf{x}^p が近づこうとする、結合項に時間遅れのある特殊な結合構造といえる。(10) 式から $\Delta \mathbf{x}^p(k)$ を消去すると、(5) と同様に

$$\begin{aligned} & \mathbf{x}^p(k+1) - (1+\lambda)\mathbf{x}^p(k) + \lambda\mathbf{x}^p(k-1) \\ &= F^p(\mathbf{x}^p(k), k) + C(\mathbf{x}^p(k), k) \dots \dots \dots (16) \\ & k = 1, 2, \dots, p = 1, 2, \dots, P \end{aligned}$$

が得られ、2 階差分方程式が結合する結合系であることがわかる。また、(17) 式を解いて級数（和分）形式で書き表すと、

$$\begin{aligned} \mathbf{x}^p(k+1) &= \mathbf{x}^p(k) + \sum_{i=0}^k \lambda^{k-i} [F^p(\mathbf{x}^p(i), i) \\ &+ C(\mathbf{x}^p(i), i)] + \lambda^{k+1} \Delta \mathbf{x}^p(0) \dots (17) \end{aligned}$$

となる。ただし

$$F^p(\mathbf{x}^p(0), 0) = c_1\{\mathbf{x}^p(0) - \mathbf{x}^p(0)\} = \mathbf{0} \dots \dots (18)$$

とする。以上より、過去の結合 $C(\mathbf{x}^p(i), i)$, $i = 0, 1, \dots, k$ までもが重み λ^{k-i} で畳み込まれているきわめて複雑な慣性系であることが理解される。

前節と同様に、各個体ごとに 2 種類の内部状態 \mathbf{u}^p , \mathbf{v}^p を用いし、

$$\mathbf{u}^p(k) = \mathbf{x}^p(k) \dots \dots \dots (19a)$$

$$\mathbf{v}^p(k) = \mathbf{x}^p(k) - \mathbf{x}^p(k-1) \dots \dots \dots (19b)$$

とおいて、(17) 式を内部状態表現モデルに変換すると、

$$\mathbf{u}^p(k+1) = \mathbf{u}^p(k) + \mathbf{v}^p(k+1) \dots \dots \dots (20a)$$

$$\begin{aligned} \mathbf{v}^p(k+1) &= \lambda \mathbf{v}^p(k) + F^p(\mathbf{u}^p(k), k) + C(\mathbf{u}^p(k), k) \\ &\dots \dots \dots (20b) \end{aligned}$$

$$\mathbf{u}^p(0), \mathbf{v}^p(0); \text{ given } k = 0, 1, 2, \dots, p = 1, 2, \dots, P$$

となる。

3. 連続時間 PSO モデルとその内部状態表現モデル

〈3・1〉 無制約連続時間 PSO モデル PSO の更新式を力学系モデルとみなし、その連続化を試みると、(17) 式より

$$\frac{d\mathbf{x}^p(t)}{dt} = c \int_0^t e^{-a(t-\tau)} [F^p(\mathbf{x}^p(\tau), \tau) + C(\mathbf{x}^p(\tau), \tau)] d\tau \dots \dots \dots (21)$$

なるモデルを導入することができる。両辺微分することにより、明らかに

$$\frac{d^2\mathbf{x}^p(t)}{dt^2} + a \frac{d\mathbf{x}^p(t)}{dt} = c[F^p(\mathbf{x}^p(t), t) + C(\mathbf{x}^p(t), t)] \dots \dots \dots (22)$$

と等価である。ここで F^p や C はそれぞれ (11) 式に対応して、

$$F^p(\mathbf{x}^p, t) = c_1\{\mathbf{x}^p(T^p(t)) - \mathbf{x}^p\} \dots \dots \dots (23a)$$

$$C(\mathbf{x}^p, t) = c_2\{\mathbf{x}^{Q(t)}(T^o(t)) - \mathbf{x}^p\} \dots \dots \dots (23b)$$

であり、さらに $T^p(t)$ や $(Q(t), T^o(t))$ は (12) 式に対応して

$$T^p(t) = \underset{\tau}{\operatorname{argmin}}\{E(\mathbf{x}^p(\tau)) \mid 0 \leq \tau \leq t\} \dots \dots (24a)$$

$$\begin{aligned} (Q(t), T^o(t)) &= \underset{(q,\tau)}{\operatorname{argmin}}\{E(\mathbf{x}^q(\tau)) \mid q = 1, \dots, P, \\ &0 \leq \tau \leq t\} \dots \dots \dots (24b) \end{aligned}$$

である。 $T^p(t)$ は時間が連続化したことにより、初期時刻 $\tau = 0$ から現時刻 $\tau = t$ の連続時間区間 $[0, t]$ 上の個体 p

の軌道 $\mathbf{x}^p(\bullet)$ に対して関数 E を評価し、現時刻 t までの間に、その最小値を与える時刻 $T^p(t)$ を意味している。現時刻 $\tau = t$ はこれ自体時間変数として時々刻々変化するので、 E の最小値を与える時刻 T^p もこの t の関数として $T^p(t)$ と記してある。

なお、2 階微分方程式で表される連続時間系モデル (22) 式の状態変数表現を

$$\mathbf{u}^p(t) = \mathbf{x}^p(t) \dots\dots\dots (25a)$$

$$\mathbf{v}^p(t) = \frac{d\mathbf{u}^p(t)}{dt} + a\mathbf{u}^p(t) \dots\dots\dots (25b)$$

と導入すると、離散時間系 (20) 式に対応した連続系の内部状態表現モデル

$$\frac{d\mathbf{u}^p(t)}{dt} = -a\mathbf{u}^p(t) + \mathbf{v}^p(t) \dots\dots\dots (26a)$$

$$\frac{d\mathbf{v}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{u}^p(t), t) + \mathbf{C}(\mathbf{u}^p(t), t)] \dots\dots\dots (26b)$$

が得られる。なお (25b) 式の状態変数 \mathbf{v}^p を

$$\mathbf{v}^p(t) = \frac{d\mathbf{u}^p(t)}{dt} \dots\dots\dots (27)$$

と一般的においた内部状態変数モデルも考えられるが、(25b) 式のようにおいた場合、多峰性関数の最適化のベンチマークにおいて大域的最適解の発見率が高かったため、(25b) 式を採用した。

〈3・2〉 上下制限約付連続時間 PSO モデル 上下制限約付最適化問題

$$\min_{\mathbf{x}} E(\mathbf{x}) \dots\dots\dots (28a)$$

$$\text{subj.to } p_i \leq x_i \leq q_i, i = 1, \dots, n \dots\dots\dots (28b)$$

を直接解くために、この上下制限約領域内に系の軌道を閉じ込めた PSO モデルを 2 種類導入する。一つは、制約領域の境界に近づくほどその遷移にブレーキを働かせた「非線形作用素モデル」と、問題 (28) の変数を変換して無制約化した新たな変数空間に無制約 PSO モデルを適用した「変数変換モデル」である。以下では問題 (28) の大域的最適解は上下制限約 (28b) の内部にあるものとする。なお、最適解がこの上下制限約の境界上にある場合を考慮するには、

$$W(\mathbf{x}) = \frac{1}{r} \sum_{i=1}^n \left(\frac{1}{x_i - p} + \frac{1}{q_i - x_i} \right), r \geq 0 \dots\dots\dots (29)$$

のような障壁関数を目的関数 E に付加し、最適解を制約領域内部に摂動させた拡大目的関数に同様のモデルを適用すればよい。

(1) 非線形作用素モデルとその内部状態表現モデル (28b) 式の制約条件に対してブレーキ作用を (21) 式のモデルに施したモデルとして、

$$\begin{aligned} \frac{dx_i^p(t)}{dt} &= c \frac{(x_i^p(t) - p_i)(q_i - x_i^p(t))}{q_i - p_i} \\ &\times \int_0^t e^{-a(t-\tau)} [F_i^p(\mathbf{x}^p(\tau), \tau) + C_i(\mathbf{x}^p(\tau), \tau)] d\tau \\ i &= 1, \dots, n \dots\dots\dots (30) \end{aligned}$$

が得られる。さらに文献 (7) と同様に

$$\frac{dx_i^p(t)}{dt} = \frac{dx_i^p(t)}{du_i^p} \frac{du_i^p(t)}{dt} \dots\dots\dots (31a)$$

$$\frac{dx_i^p}{du_i^p} = \frac{(x_i^p(t) - p_i)(q_i - x_i^p(t))}{q_i - p_i} \dots\dots\dots (31b)$$

と内部状態 \mathbf{u} を導入すると、内部状態モデルとして

$$\begin{aligned} \frac{d^2\mathbf{u}^p(t)}{dt^2} + a \frac{d\mathbf{u}^p(t)}{dt} &= c[\mathbf{F}^p(\mathbf{x}^p(t), t) + \mathbf{C}(\mathbf{x}^p(t), t)] \\ &\dots\dots\dots (32) \end{aligned}$$

が得られ、出力関数として

$$x_i^p(t) = f_i(u_i^p(t)) = \frac{q_i + p_i \exp(-u_i^p(t))}{1 + \exp(-u_i^p(t))} \dots\dots (33)$$

が得られる。ただし、 $\mathbf{F}^p(\mathbf{x}^p, t)$ や $\mathbf{C}(\mathbf{x}^p, t)$ は (23a) 式や (23b) 式と同じである。また、 $p_i = 0, q_i = 1$ の場合、(33) 式は、ニューラルネットワークの出力関数であるいわゆるシグモイド関数であり、 $f_i: [-\infty, \infty] \rightarrow (0, 1)$ である。このモデルの導出過程は、文献 (7) と同様である。

内部状態表現モデル (32) 式をさらに 1 階微分方程式の状態変数表現モデルに換するために、(25b) 式と同じ変数 \mathbf{v}^p を導入すると、(26) 式と類似のモデルが得られる。これに (23), (24) 式も付記し、(33) 式の出力関数 f_i を連立させると、

$$\frac{d\mathbf{u}^p(t)}{dt} = -a\mathbf{u}^p(t) + \mathbf{v}^p(t) \dots\dots\dots (34a)$$

$$\frac{d\mathbf{v}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{x}^p(t), t) + \mathbf{C}(\mathbf{x}^p(t), t)] \dots\dots\dots (34b)$$

$$\mathbf{F}^p(\mathbf{x}^p, t) = c_1 \{ \mathbf{x}^p(T^p(t)) - \mathbf{x}^p \} \dots\dots\dots (34c)$$

$$\mathbf{C}(\mathbf{x}^p, t) = c_2 \{ \mathbf{x}^{Q(t)}(T^o(t)) - \mathbf{x}^p \} \dots\dots\dots (34d)$$

$$T^p(t) = \underset{\tau}{\operatorname{argmin}} \{ E(\mathbf{x}^p(\tau)) \mid 0 \leq \tau \leq t \} \dots\dots (34e)$$

$$\begin{aligned} (Q(t), T^o(t)) &= \underset{(q, \tau)}{\operatorname{argmin}} \{ E(\mathbf{x}^q(\tau)) \} \\ q &= 1, \dots, P, 0 \leq \tau \leq t \} \dots\dots\dots (34f) \end{aligned}$$

$$x_i^p(t) = f_i(u_i^p(t)) = \frac{q_i + p_i \exp(-u_i^p(t))}{1 + \exp(-u_i^p(t))} \dots\dots (34g)$$

$$i = 1, \dots, n$$

となる。無制約の場合には (26b) 式の右辺の \mathbf{F}^p や \mathbf{C} の引数が \mathbf{u}^p であるのに対して、この非線形作用素モデルの状態変数表現モデルでは、出力関数 (34h) 式が連立するため、(34b) 式の右辺の \mathbf{F}^p や \mathbf{C} の引数が \mathbf{x}^p のままであることに注意してほしい。

(2) 非線形変数変換モデルとその内部状態表現モデル (1) とは異なる内部状態表現モデルを導出するために、まず (34h) 式と同じ形の式を変数変換に用いるため、

$$x_i = f_i(y_i) = \frac{q_i + p_i \exp(-y_i)}{1 + \exp(-y_i)} \dots\dots\dots (35)$$

とおく。この変換式を制約条件付問題 (28) に代入して変数 \mathbf{x} を消去すると、変数 \mathbf{y} に対する無制約問題

$$\min_{\mathbf{y}} E(\mathbf{f}(\mathbf{y})) \dots\dots\dots (36)$$

が得られる。ただし、

$$\mathbf{f}(\mathbf{y}) = (f_1(y_1), \dots, f_n(y_n))^T \dots\dots\dots (37)$$

とおく。無制約問題 (36) の変数 \mathbf{y} に関して、連続時間系 PSO モデル (22) をそのまま用いると、

$$\frac{d^2 \mathbf{y}^p(t)}{dt^2} + a \frac{d\mathbf{y}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{y}^p(t), t) + \mathbf{C}(\mathbf{y}^p(t), t)] \dots\dots\dots (38)$$

を得る。ただし \mathbf{F}^p や \mathbf{C} は (23) 式に対応して

$$\mathbf{F}^p(\mathbf{y}^p, t) = c_1 \{\mathbf{y}^p(T^p(t)) - \mathbf{y}^p\} \dots\dots\dots (39a)$$

$$\mathbf{C}(\mathbf{y}^p, t) = c_2 \{\mathbf{y}^{Q(t)}(T^o(t)) - \mathbf{y}^p\} \dots\dots\dots (39b)$$

であり、さらに $T^p(t)$ や $(Q(t), T^o(t))$ は (24) 式と同じである。ただし、その場合は (24) 式において、

$$\mathbf{x}^p(\tau) = \mathbf{f}(\mathbf{y}^p(\tau)) \dots\dots\dots (40)$$

であって、(38) 式は \mathbf{y}^p の力学系であるが、右辺の \mathbf{F}^p や \mathbf{C} の値は、変換式 \mathbf{f} を通した \mathbf{x}^p によって (24) 式で決まることに注意してほしい。

前節と同様に状態変数を

$$\mathbf{u}^p(t) = \mathbf{y}^p(t) \dots\dots\dots (41a)$$

$$\mathbf{v}^p(t) = \frac{d\mathbf{u}^p(t)}{dt} + a\mathbf{u}^p(t) \dots\dots\dots (41b)$$

とおくと、(26) 式と同じ状態変数表現が得られる。この式の中の \mathbf{F}^p や \mathbf{C} の定義式も含ませ、さらに (35) 式の出力関数 f_i を連立させて具体的に記述すると、

$$\frac{d\mathbf{u}^p(t)}{dt} = -a\mathbf{u}^p(t) + \mathbf{v}^p(t) \dots\dots\dots (42a)$$

$$\frac{d\mathbf{v}^p(t)}{dt} = c[\mathbf{F}^p(\mathbf{u}^p(t), t) + \mathbf{C}(\mathbf{u}^p(t), t)] \dots\dots\dots (42b)$$

$$\mathbf{F}^p(\mathbf{u}^p, t) = c_1 \{\mathbf{u}^p(T^p(t)) - \mathbf{u}^p\} \dots\dots\dots (42c)$$

$$\mathbf{C}(\mathbf{u}^p, t) = c_2 \{\mathbf{u}^{Q(t)}(T^o(t)) - \mathbf{u}^p\} \dots\dots\dots (42d)$$

$$T^p(t) = \underset{\tau}{\operatorname{argmin}} \{E(\mathbf{x}^p(\tau)) \mid 0 \leq \tau \leq t\} \dots\dots\dots (42e)$$

$$(Q(t), T^o(t)) = \underset{(q, \tau)}{\operatorname{argmin}} \{E(\mathbf{x}^q(\tau)) \mid q = 1, \dots, P, 0 \leq \tau \leq t\} \dots\dots\dots (42f)$$

$$x_i^p(t) = f_i(u_i^p(t)) = \frac{q_i + p_i \exp(-u_i^p(t))}{1 + \exp(-u_i^p(t))} \dots\dots\dots (42g)$$

$$i = 1, \dots, n$$

となる。ここで注意することは (34) 式の非線形作用素モデルの場合は、 \mathbf{F}^p や \mathbf{C} の引数が \mathbf{x}^p であるのに対して、この非線形変数変換モデルの場合のこれらの引数は、 \mathbf{u}^p であることであり、さらに後者のモデルでは、p-best の時刻 $T^p(t)$ や g-best の個体番号 $Q(t)$ と時刻 $T^o(t)$ の算出にのみ、出力関数 f_i が用いられることである。

4. 上下限制約離散化 PSO モデル

有界な領域内に閉じ込めた離散力学系を不安定化させると、カオス現象が現れる可能性があることが知られている。そこで本章では、前章で導入した「非線形作用素モデル」と「非線形変数変換モデル」の二つのモデルを、まずサンプリングパラメータ ΔT で離散化し、この離散力学系を不安定化してカオス現象を創発させることを考える。

(1) 非線形作用素モデルの離散化 PSO モデル (34a), (34b) 式をサンプリングパラメータ ΔT のオイラー法で離散化すると、

$$\mathbf{u}^p(k+1) = \mathbf{u}^p(k) + \Delta T \{\mathbf{v}^p(k) - a\mathbf{u}^p(k)\} \dots\dots\dots (43a)$$

$$\mathbf{v}^p(k+1) = \mathbf{v}^p(k) + c\Delta T [\mathbf{F}^p(\mathbf{x}^p(k), k) + \mathbf{C}(\mathbf{x}^p(k), k)] \dots\dots\dots (43b)$$

が得られる。これを書き換えて、(34c)~(34h) 式に対応する式も加えると、

$$\mathbf{u}^p(k+1) = (1 - a\Delta T)\mathbf{u}^p(k) + \Delta T \mathbf{v}^p(k) \dots\dots\dots (44a)$$

$$\mathbf{v}^p(k+1) = \mathbf{v}^p(k) + c\Delta T [\mathbf{F}^p(\mathbf{x}^p(k), k) + \mathbf{C}(\mathbf{x}^p(k), k)] \dots\dots\dots (44b)$$

$$\mathbf{F}^p(\mathbf{x}^p(k), k) = c_1 \{\mathbf{x}^p(l^p(k)) - \mathbf{x}^p(k)\} \dots\dots\dots (44c)$$

$$\mathbf{C}(\mathbf{x}^p(k), k) = c_2 \{\mathbf{x}^{Q(k)}(l^o(k)) - \mathbf{x}^p(k)\} \dots\dots\dots (44d)$$

$$l^p(k) = \underset{l}{\operatorname{argmin}} \{E(\mathbf{x}^p(l)) \mid l = 0, \dots, k\} \dots\dots\dots (44e)$$

$$(Q(k), l^o(k)) = \underset{(q, l)}{\operatorname{argmin}} \{E(\mathbf{x}^q(l)) \mid q = 1, 2, \dots, P, l = 0, 1, \dots, k\} \dots\dots\dots (44f)$$

$$x_i^p(k) = f_i(u_i^p(k)) = \frac{q_i + p_i \exp(-u_i^p(k))}{1 + \exp(-u_i^p(k))} \dots\dots\dots (44g)$$

$$i = 1, \dots, n$$

となり、(44a), (44b) 式は、(20) 式と類似のモデルとなる。

PSO では、関数 \mathbf{F}^p や \mathbf{C} の中に含まれる係数 c_1, c_2 に対して一定区間内の一様乱数により揺らぎを与えて局所的に不安定化させているが、(44) 式のモデルにおいて ΔT による分岐特性により不安定化や安定化を調節する新しいモデルを構築することができる。すなわち、従来の PSO では g-best 情報を探索点群全体で共有し、さらに p-best, g-best への移流の重みに一様乱数を用いた揺らぎを与えることにより局所的最適解からの離脱を期待するのに対して、本提案手法では、p-best, g-best へ移流する構造を残した上で、さらに比較的大きな ΔT に対する有界領域内全体に渡るカオス的な走査によって、より大域的な探索を期待するのが特徴である。

(2) 非線形変数変換モデルの離散化 PSO モデル 非線形作用素モデルと同様に、(42a), (42b) 式をそのままオイ

ラー法で離散化して (42c)~(42h) 式に対応する式も加えると、

$$\mathbf{u}^p(k+1) = (1 - a\Delta T)\mathbf{u}^p(k) + \Delta T\mathbf{v}^p(k) \quad \dots\dots\dots (45a)$$

$$\mathbf{v}^p(k+1) = \mathbf{v}^p(k) + c\Delta T[\mathbf{F}^p(\mathbf{u}^p(k), k) + \mathbf{C}(\mathbf{u}^p(k), k)] \quad \dots\dots\dots (45b)$$

$$\mathbf{F}^p(\mathbf{u}^p(k), k) = c_1\{\mathbf{u}^p(l^p(k)) - \mathbf{u}^p(k)\} \quad \dots\dots (45c)$$

$$\mathbf{C}(\mathbf{u}^p(k), k) = c_2\{\mathbf{u}^Q(k)(l^Q(k)) - \mathbf{u}^p(k)\} \quad \dots\dots (45d)$$

$$l^p(k) = \underset{l}{\operatorname{argmin}}\{E(\mathbf{x}^p(l)) | l = 0, \dots, k\} \quad \dots\dots (45e)$$

$$(Q(k), l^Q(k)) = \underset{(q,l)}{\operatorname{argmin}}\{E(\mathbf{x}^q(l)) |$$

$$q = 1, 2, \dots, P, l = 0, 1, \dots, k\} \quad \dots\dots\dots (45f)$$

$$x_i^p(k) = f_i(u_i^p(k)) = \frac{q_i + p_i \exp(-u_i^p(k))}{1 + \exp(-u_i^p(k))}$$

$$i = 1, \dots, n \quad \dots\dots\dots (45g)$$

となり、形の上では制約を考慮しない PSO モデル (20) 式と同様のモデルが得られる。なお、(42) 式に対して言及したように \mathbf{F}^p や \mathbf{C} の $l^p(k)$, $(Q(k), l^Q(k))$ は、変換式 \mathbf{f} を通した (45g) 式によって決まることが、制約を考慮しない PSO モデルと大きく異なっている。このモデルにおいても ΔT の調節で有界領域内に閉じ込められた形で不安定化させることにより、その領域内を走査するカオス的な現象を生み出すことができると期待される。

5. シミュレーション結果

この章では、4 章で提案した離散化 PSO モデルの挙動を確認する。2 種類の離散化 PSO モデルのサンプリングパラメータ ΔT に対する、探索軌道の分岐特性を確認したのち、以下の場合を想定したパラメータ設定のもとで、g-best 個体の大域的最適解への収束率について、シミュレーションにより比較検討を行った。これらのシミュレーションは、上下制限約を満たすようにランダムに初期状態を 1000 通り選んで試行を繰り返したときの割合または平均である。

case1: 一定のパラメータ a , c , ΔT の値、および 0 以上 c_1 , c_2 以下の一様乱数を、全個体が必ずしも収束しないように設定した場合

case2: 一定のパラメータ a , c , ΔT の値、および 0 以上 c_1 , c_2 以下の一様乱数を、全個体がそれぞれの定常点に収束するように設定した場合。

case3: 一定のパラメータ a , c , c_1 , c_2 の下で、すべての個体でカオス軌道が生じるよう、 ΔT を一定の値に設定した場合

case4: ある程度大きな値から ΔT を減少させるカオス徐冷により全個体がそれぞれの定常点へ収束するよう、パラメータ a , c , c_1 , c_2 の値を一定値に設定した場合

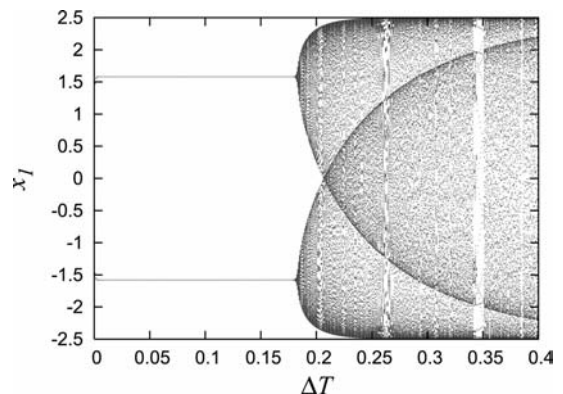
〈5・1〉 Wood-Colville 問題に対する結果 多峰性関数の大域的最適化のベンチマークである Wood-Colville 問題

$$\begin{aligned} \min_{\mathbf{x}} E(\mathbf{x}) &= 100.0(2.5 - x_1^2)^2 + (1.0 - x_1)^2 \\ &\quad + 90.0(2.5 - x_2^2)^2 + (1.0 - x_2)^2 \\ &\quad \dots\dots\dots (46a) \end{aligned}$$

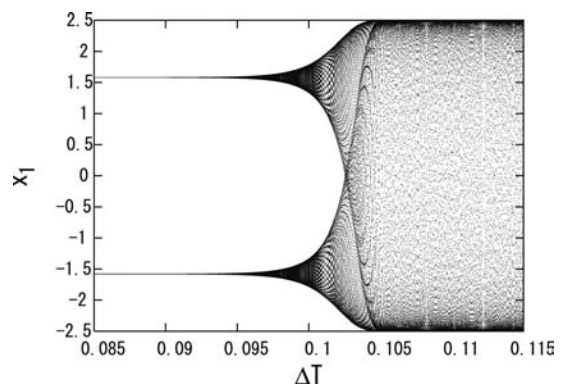
$$\text{subj.to} \quad -2.5 \leq x_i \leq 2.5 \quad i = 1, 2, \dots\dots\dots (46b)$$

に対するシミュレーション結果を示す。4 個の局所解, $A : (-1.58, -1.58)$, $B : (-1.58, 1.58)$, $C : (1.58, -1.58)$, $D : (1.58, 1.58)$ をもち、 D が大域的最適解である。探索個体数は $P = 5$ とする。

まず、非線形作用素モデルおよび非線形変数変換モデルの分岐特性を変数 x_1 に関して示したのが図 1 である。各モデルにおいて設定したパラメータは、図のサブタイトルに示してある。これらは、探索点の初期位置を局所的最適解の近傍にとり、十分時間経過したのちの点列を ΔT ごとにプロットしたものである。非線形作用素モデルでは、 $\Delta T \approx 0.17$ から不安定化し、 ΔT の増加に伴い探索領域を広げ、ときおり狭い範囲で倍周期軌道を取りながら、制約領域内全域を探索する軌道に発展している。また、変数変換モデルでは $\Delta T \approx 0.095$ から、 ΔT の増加に伴い探索領域を広げ、 $\Delta T \approx 0.102 \sim 0.104$ にかけての倍周期軌道を経て、制約領域内全域を探索する軌道に発展している。なお、このモ



(a) Nolinear operator model
($c = 1.0, a = 0.55, c_1 = 3.99, c_2 = 0.01$)



(b) Nonlinear variable transformation model
($c = 0.05, a = 0.4, c_1 = 3.99, c_2 = 0.01$)

図 1 問題 (46) に対する大域的分岐図

Fig.1. Global bifurcation diagram for problem (46).

デルでは変数 v^p の発散による計算機のオーバーフローと、無駄な探索をなくすために、

$$v_i^p = \text{sgn}(v_i^p) \min\{|v_i^p|, V_{\max}\}, i = 1, \dots, n \quad (47)$$

とした⁽⁴⁾⁽⁶⁾。 V_{\max} はあらかじめ与えられた定数とする。本問題に対して $V_{\max} = 300$ とした。

つぎに、前述した case1~4 の 4 通りのパラメータ設定の下での g-best, つまり $x^{Q(k)}(l^o(k))$ が与える目的関数値 $E(x^{Q(k)}(l^o(k)))$ の変化の過程を調べた結果を図 2 に示す。なお、各場合において $c_1 = 4.0 - c_2$ とする。case4 のカオス徐冷の場合における ΔT の更新式は

$$\Delta T(k+1) = \Delta T(k) - d \dots \dots \dots (48)$$

とし、非線形作用素モデルでは $\Delta T(0) = 0.4$, $d = 3.0 \times 10^{-4}$ とし、変数変換モデルでは $\Delta T(0) = 0.2$, $d = 2.0 \times 10^{-4}$ とした。1000 イテレーションで PSO の更新を終了した。非線形変数変換モデルの case3, 4 では最初の 200 ステップまで g-best があまり更新されないが、200 ~ 400 ステップにかけて急速に改善され、探索の最終段階でも比較的効率よく g-best を更新している。これは、探索の初期段階では制約境界付近を中心に探索を行っているためではないかと考えられる。その他の場合においては、最初の 100 ステップで急速に g-best を更新し、その後、冗長な改善を

続けている。また、表 1 は g-best がどの局所的最適解に収束するかを 1000 回の試行の割合で示している。この際、各最適解と有効数字 2 桁一致すれば収束していると判定した。非線形作用素モデルおよび非線形変数変換モデルをそれぞれ model(a), (b) で表している。最後の行の PSO は、上下限制約を外点ペナルティ法⁽¹⁰⁾で目的関数に組み込んだ拡大目的関数

$$\hat{E}(x) = E(x) + r \sum_{i=1}^n \{\min(0, x_i - p_i)\}^2 + r \sum_{i=1}^n \{\min(0, q_i - x_i)\}^2 \dots (49)$$

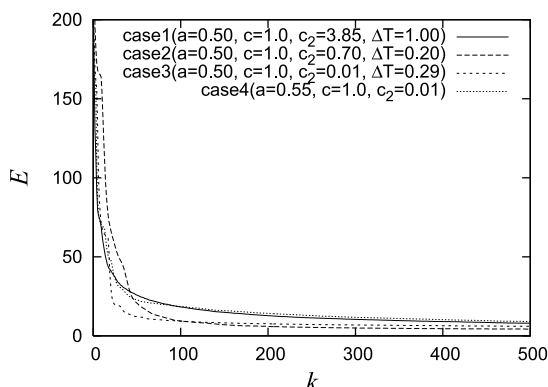
の最小化問題に、IWA-PSO を適用した場合である。ただし、 r はペナルティパラメータであり、 $r = 1.0 \times 10^{10}$ とした。IWA-PSO のパラメータを $c_1 = c_2 = 2.0$ とし、 λ を $0.9 \rightarrow 0.4$ と線形に減少させた。両モデルにおいて、case2 では、大域的最適解への収束率が低い。また、非線形変数変換モデル case1, 3, 4 において 9 割程度、大域的最適解へ収束しており、case2 を除いて、IWA-PSO を上回っていることから、変数変換によって系の軌道を閉じ込めることにより、効率よく探索が行われたと考えられる。

〈5・2〉 Rastrigin 関数に対する結果 多峰性関数の大域的最適化のベンチマークである Rastrigin 関数最小化問題

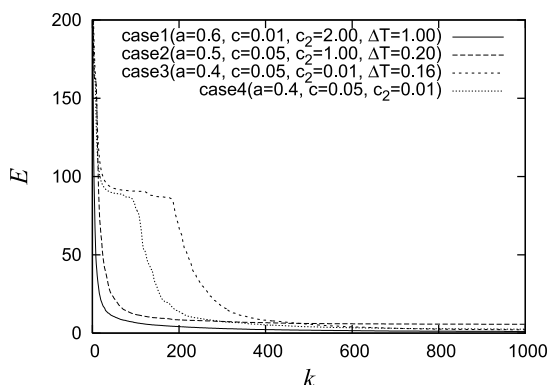
$$\min_x E(x) = 10n + \sum_{i=1}^n x_i^2 - 10.0 \cos(10\pi x_i) \dots \dots \dots (50a)$$

$$\text{subj.to} \quad -1.0 \leq x_i \leq 1.0 \quad i = 1, \dots, n \dots \dots \dots (50b)$$

に対するシミュレーション結果を示す。この問題に対しては、 $n = 2$ の場合に対して〈5・1〉節と同様のシミュレーション結果を示したあと、変数の次元に対する収束性能を確認するために、 $n = 3, 5, 10, 20, 30$ の場合に対する結果を示す。 $n = 2$ の場合、A:(0.0, 0.0), B:(±0.2, 0.0), (0.0, ±0.2), C:(±0.2, ±0.2), (±0.2, ∓0.2) などの局所的最適解が多数あり、A が大域的最適解である。探索個体数は $P = 8$ とする。本問題に対する分岐特性を図 3 に示す。非線形作用素モデルでは $\Delta T \approx 0.33$ から不安定化し、 ΔT の増加に伴い探索領域を広げ、ときおり狭い範囲で倍周期軌道を取りなが



(a) Nonlinear operator model



(b) Nonlinear variable transformation model

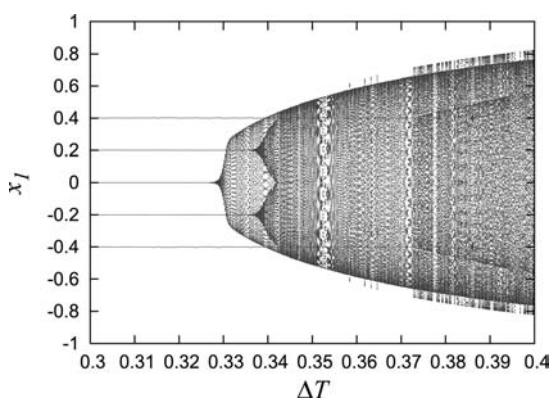
図 2 g-best の履歴

Fig. 2. History plot of g-best.

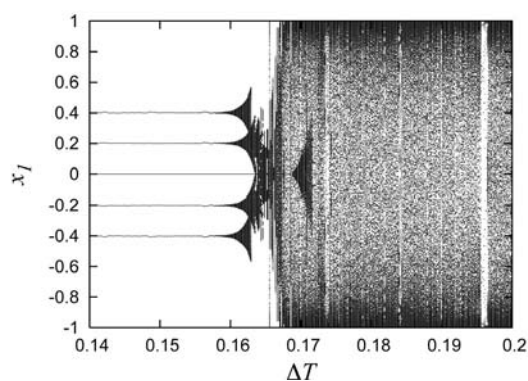
表 1 大域的小および局所的最適解収束率

Table 1. Convergence rate to the global and the local optimum.

model	case	A	B	C	D
(a)	1	0.3%	10.8%	13.1%	70.7%
	2	1.1%	15.4%	17.5%	64.7%
	3	0.5%	11.9%	12.1%	70.6%
	4	0.2%	10.1%	8.8%	80.6%
(b)	1	0.0%	0.3%	0.2%	99.0%
	2	6.2%	17.0%	19.3%	54.6%
	3	0.0%	5.8%	5.4%	88.7%
	4	0.0%	7.3%	7.1%	85.3%
PSO		1.7%	14.7%	12.5%	71.1%



(a) Nonlinear operator model
($c = 1.0, a = 0.65, c_1 = 3.994, c_2 = 0.006$)



(b) Nonlinear variable transformation model
($c = 0.05, a = 0.64, c_1 = 3.99, c_2 = 0.01$)

図3 問題(50)に対する大域的分岐図

Fig. 3. Global bifurcation diagram for problem(50).

ら、制約領域全域を走査する軌道に発展している。非線形変数変換モデルでは、 $\Delta T \approx 0.158$ から不安定化し、制約領域全域にわたる軌道に発展している。その軌道は制約領域内全体をくまなく探索している。そして、制約境界付近で、分岐図が濃くなっていることから、制約境界付近が比較的安定であると考えられる。また、 $\Delta T \approx 0.162 \sim 0.165$ にバウンダリクライシスを生じていることや、 $x_1 = 0$ に対する分岐が、他の局所的最適解と比較して遅いことから、 $x_1 = 0$ に引き込まれやすい大域的構造を有していると考えられる。case4 の徐冷操作に関しては、非線形作用素モデルは $\Delta T(0) = 0.7, d = 7.0 \times 10^{-4}$ とし、非線形変数変換モデルでは $\Delta T(0) = 0.2, d = 2.0 \times 10^{-4}$ とした。1000 イテレーションで更新を終了した。表2は g-best がどの局所的最適解に収束するかを1000回の試行の割合で示している。この際、各最適解と有効数字2桁まで一致すれば収束していると判定した。最後の行の PSO は、上下制限約を外点ペナルティ法で取り込んだ拡大目的関数に対して IWA-PSO を適用した結果である。IWA-PSO のパラメータを $c_1 = c_2 = 2.0$ とし、 λ を $0.9 \rightarrow 0.4$ と線形に減少させた。また、表3は、各モデル、各場合のパラメータを示している。非線形作用素モデルの case1 と、非線形変数変換モデルの case1,3,4 において IWA-PSO を上回る探索性

表2 大域のおよび局所的最適解収束率

Table 2. Convergence rate to the global and the local optimum.

model	case	A	B	C
(a)	1	97.5%	1.0%	0.7%
	2	17.8%	32.1%	12.6%
	3	60.7%	12.4%	5.9%
	4	52.7%	11.5%	10.0%
(b)	1	79.7%	1.3%	0.0%
	2	44.2%	37.6%	8.2%
	3	63.2%	14.1%	6.8%
	4	75.1%	4.1%	5.1%
PSO		61.5%	33.6%	2.8%

表3 パラメータの設定

Table 3. Parameter setting of proposed models.

model	case	a	c	c_2	ΔT
(a)	1	0.8	1.0	0.006	1.0
	2	0.8	1.0	2.0	0.7
	3	0.75	1.0	0.006	0.45
	4	0.65	1.0	0.006	
(b)	1	0.8	0.05	0.01	1.0
	2	0.5	0.05	1.0	0.2
	3	0.4	0.05	0.01	0.2
	4	0.64	0.05	0.01	

表4 変数の増加に対する大域的最適解収束率

Table 4. Convergence rate to the global optimum for increase of the number of variables.

model	case	n=3	n=5	n=10	n=20	n=30
(a)	1	97.5%	99.9%	99.7%	98.0%	93.6%
	2	5.4%	0.4%	0.0%	0.0%	0.0%
	3	54.6%	15.7%	3.0%	0.2%	0.7%
	4	65.7%	39.2%	15.7%	4.3%	1.9%
(b)	1	81.3%	51.3%	11.4%	8.2%	9.0%
	2	21.6%	0.4%	0.0%	0.0%	0.0%
	3	51.4%	21.5%	9.9%	2.7%	2.1%
	4	75.2%	50.2%	28.2%	15.2%	12.9%
PSO		30.4%	0.0%	0.0%	0.0%	0.0%

能を示した。なお、表2の結果より、提案手法の各種モデルや各 case の収束率の合計値は、62.5% から 99.2% までかなりのばらつきがあることがわかる。合計値が低いほど A, B, C 以外の局所的最適解に収束した可能性が高いことを示しているが、とくに非線形作用素モデルで case2 の場合に収束率が低いのは、定常点に収束するパラメータを設定したことで、イテレーションの早い段階で局所的最適解に収束する、いわゆる初期収束現象が生じたものと推定される。

つぎに $n = 3, 5, 10, 20, 30$ とした問題を考える。初期点を制約領域内に乱数で生成し、解を得る作業を1000回繰り返した際の、大域的最適解への収束率を表4に示す。両モデル共に、個体数を20とし、各パラメータと徐冷方法は $n = 2$ の場合と同じである。IWA-PSO では、変数の増加に伴い、探索性能が低下するが、とくに非線形作用素モデルの case1 では、変数の増加によらず、高い探索性能を示

した。

6. おわりに

本論文では、上下限制約条件付き最適化問題に対して直接適用可能な PSO を提案した。とくに、このような PSO の力学系を提案するに当たって、まず連続系の力学系を想定し、それを上下限制約内に閉じ込めた力学系として、「非線形作用素モデル」と「非線形変数変換モデル」の2種類のモデルを提案したあと、それを離散化することによって、内部状態表現モデルが PSO の力学系となる、上下限制約を直接考慮した離散時間系モデルを導出した。さらに、この離散化パラメータを大きくとり、上下限制約に閉じ込めたままそれらの力学系を不安定化させると、カオス特性が創発することも確認した。

いずれの離散化モデルも、IWA-PSO にペナルティ法を適用した場合と比べて、とくに変数次元が大きい場合に、大域的最適解への収束効率が高くなることが、シミュレーションにより確認された。このことから、大域的最適解のおおよその位置がわかっている場合や、ある程度限定した領域内で大域的最適解を探索する場合には、本論文で提案した上下限制約を直接考慮できる PSO が有力であることが確認された。また、このシミュレーションでは、

- (1) Wood-Colville 問題 (<5・1>) には、「非線形変数変換モデル」が、Rastrigin 関数 (<5・2>) には、「非線形作用素モデル」が、とくに case1 の場合に優位な結果が得られたこと
- (2) すべての個体に対してそれぞれ収束させるようなパラメータを必ずしも選ばない方が、g-best の遷移過程に注目している限りでは、その収束特性が良好であること
- (3) カオス特性を持たせることは、上下限制約で探索範囲を限定しているとはいえ、無駄な遷移を許容すると予想され、必ずしも収束性の改善には寄与しないこと

なども確認された。とくに (1) については、軌道の閉じ込め方の違いから、軌道が有界領域の境界付近に偏りやすい非線形変数変換モデルの方が、局所的最適解が領域の四隅に分散している<5・1>節の関数に親和性があり、非線形作用素モデルの場合には、これとはまったく逆のことを示しているものと考えられる。

(平成 17 年 6 月 16 日受付, 平成 18 年 2 月 13 日再受付)

文 献

- (1) J. Kennedy and R.C. Eberhart: "Particle Swarm Optimization", Proc. IEEE Int'l Conf. on Neural Networks, 4, pp.1942-1948, Perth, Australia (1995)
- (2) J. Kennedy and R.C. Eberhart: "Swarm Intelligence", Morgan Kaufmann Publ (2000)
- (3) K.E. Parsopoulos and M.N. Vrahatis: "Recent Approach to Global Optimization Problems through Particle Swarm Optimization", Natural Computing, Vol.1, pp.235-306 (2002)

- (4) J.F. Schutte and A.A. Groenwold: "Optimal Sizing Design of Truss Structures Using the Particle Swarm Optimization Algorithm", Proc. of the 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, pp.2237-2246 (2002)
- (5) Y. Shi: "Particle Swarm Optimization", IEEE Neural Networks Society, pp.8-13 (2004)
- (6) 岩崎信弘・安田恵一郎・井出 東:「Particle Swarm Optimization の高次元問題への適用に関する検討」, 電学研資, 産業計測制御研究会, IIC-04-47~66, pp.87-92 (2004)
- (7) K. Masuda and E. Aiyoshi: "Hybrid Type of Global Optimization Methods with Discretized Chaos Mappings and Increase Accepting Methods", Trans. IEE Japan, Vol.122-C, No.5, pp.892-829, (2002-5) (in Japanese)
増田和明・相吉英太郎:「離散化カオス写像と悪化受理法によるハイブリット型大域的最適化手法」, 電学論 C, **122-C**, 5, pp.892-899 (2002-5)
- (8) K. Masuda and E. Aiyoshi: "Chaotic Dynamics on a Simplex and Global Optimization Methods with Normalized Constraints", Trans. IEE Japan, Vol.123-C, No.6, pp.892-829 (2003-6) (in Japanese)
増田和明・相吉英太郎:「シンプレックス上のカオス力学系と正規化制約付大域的最適化手法」, 電学論 C, **123**, 6, pp.1147-1154 (2003-6)
- (9) R. Horie, E. Aiyoshi, and H. Ishii: "The Neural Networks Realization of a Gradient Method with a Convolution Integral by Using Neurons with "Memory" (Inertial Neurons)", Trans. of the Inst. of Systems, Control and Inf. Eng., Vol.11, No.3, pp.112-119 (1998) (in Japanese)
堀江亮太・相吉英太郎・石井秀教:「"メモリー"付きニューロン (慣性ニューロン) を用いた畳み込み積分型勾配法のニューラルネットワーク実現」, システム制御情報学論, **11**, 3, pp.112-119 (1998-3)
- (10) 相吉英太郎・志水清孝: 数理計画法演習, 朝倉書店 (1985)
- (11) A. Ide and K. Yasuda: "A Basic Study of the Adaptive Particle Swarm Optimization", Trans. IEE Japan, Vol.124-C, No.2, pp.550-557 (2004-2)

村 田 秀 樹 (非会員) 1979 年 8 月 26 日生。2004 年 3 月 慶應義塾大学理工学部物理情報工学科卒業。同年 4 月同大学大学院理工学研究科基礎理工学専攻入学。在学中は、大域的最適化手法、とくに、多点型探索手法の研究を行い、2006 年 3 月修士課程修了。同年 4 月より、横河電機 (株) に勤務。



相 吉 英 太 郎 (正員) 1950 年 8 月 23 日生。1975 年慶應義塾大学工学部計測工学科卒業。1980 年 3 月同大学大学院工学研究科博士課程修了。助手・専任講師、同理工学部助教授を経て、1996 年 4 月より同教授。工学博士。最適化の理論と計算方法に関する基礎研究に従事。

