

連続世代モデルに基づく微分進化法の分散分析による比較研究

田川 聖治[†]

[†] 近畿大学理工学部

既存の微分進化法 (DE: Differential Evolution) は、すべて離散世代モデルに基づくものであり、現世代の集団は次世代の集団によって一斉に置き換えられる。本稿では、連続世代モデルに基づく DE を示すとともに、分散分析の技法を用いることで、世代モデルの違いが DE の性能に及ぼす効果を解析し、提案した連続世代モデルによる DE の利点を明らかにしている。

A Comparative Study of the Differential Evolution based on Continouous Generation Model By using Analysis of Variance

Kiyoharu TAGAWA[†]

[†] School of Science and Engineering, Kinki University

Conventional Differentiation Evolution (DE) has been based on the discrete generation model in which the current generation's population is replaced by the next generation's population at a time. In this paper, a new DE based on the continuous generation model is presented. Then, by using the techniques of the analysis of variance, the effects of the two generation models on the performance of DE are compared statistically. As results, the advantage of the new DE based on the proposed continuous generation model is clarified.

1 はじめに

進化的計算における集団の更新方法は世代モデルと呼ばれ、離散世代モデルと連続世代モデルに大別される¹⁾。離散世代モデルでは、現世代と次世代の二つの集団が存在し、新たな個体は現世代の集団の個体を用いて作られ、現世代の集団は次世代の集団によって一斉に置き換えられる。一方、連続世代モデルでは、集団は一つであり、その個体は新たな個体と随時比較され入れ換わる。

微分進化法 (DE: Differential Evolution)^{2, 3)} は、決定変数が実数値を取る関数最適化問題を対象とした進化的計算の一種である。また、DE では離散世代モデルを採用している。近年、集団の構成や更新方法を含み、様々な DE の改良や拡張が報告されてるが⁴⁾、著者の知る限り、既存の DE はすべて離散世代モデルに基づいている。

そこで、前回の報告⁵⁾では、連続世代モデルに基づく DE を提案した。また、3 因子実験の分散分析⁶⁾によって、世代モデル、集団サイズ、戦略の違いが、DE の性能に及ぼす効果の評価を試みた。しかし、対象とする問題の特性の違いに関する配慮を欠いたこと、制御パラメータを誤差因子に含めたことなどにより、連続世代モデルによる DE の利点を十分に示すことができなかった。

本稿では、2 因子実験の分散分析により、世代モデル、集団サイズ、両者の交互作用が、DE の性能に及ぼす効果を統計的に評価する。この際、できるだけテスト問題に依存しない普遍的な因果関係を調べるため、テスト問題を易しい問題と難しい問題の二つのグループに分け、各グループで DE の性能を評価する。さらに、分散分析の F 検定と整合性のあるシェフェの多重比較⁶⁾を用いて、実験結果を詳細に解析する。そして、提案した連続世代モデルによれば、易しい問題において、DE の処理時間が短縮されることを明らかにする。

2 微分進化法 (DE)

2.1 個体表現と集団構成

関数最適化問題は、決定変数 x_j ($j = 1, \dots, D$) が実数値を取り、式 (1) のように定式化される。

$$\begin{cases} \text{minimize} & f(\mathbf{x}) = f(x_1, \dots, x_D) \\ \text{sub. to} & \underline{x}_j \leq x_j \leq \bar{x}_j, j = 1, \dots, D \end{cases} \quad (1)$$

DE では、式 (1) の最適化問題の解候補を個体とし、個体の配列を集団とする。すなわち、集団サイズを N とすると、世代 g の集団における i 番目の個体 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) は、式 (2) の浮動小数点表現による D 次元の実数ベクトルとなる。

$$\mathbf{x}_{i,g} = (x_{1,i,g}, \dots, x_{j,i,g}, \dots, x_{D,i,g}) \quad (2)$$

ただし, $x_j \leq x_{j,i,g} \leq \bar{x}_j$ ($j = 1, \dots, D$) とする.

2.2 処理手順

標準的な DE²⁾ の処理手順を以下に示す. 現世代の集団 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) と, 次世代の集団 $\mathbf{x}_{i,g+1}$ ($i = 1, \dots, N$) が存在することから, 明らかに DE は離散世代モデルを採用している.

〔離散世代モデルによる DE〕

手順 1 初期集団の個体 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) をランダムに生成し, 世代を $g = 0$ とする.

手順 2 現世代の集団で最良の個体を $\mathbf{x}_{best}$ とする.

手順 3 所定の終了条件を満たせば, 最良の個体 $\mathbf{x}_{best}$ を出力して終了する.

手順 4 現世代の集団の各個体 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) に関して, 以下の手順 5 から手順 9 を繰返す.

手順 5 個体 $\mathbf{x}_{i,g}$ をターゲットベクトルとする.

手順 6 現世代の集団から一つの個体 $\mathbf{x}_{b,g}$ を複製選択し, ベースベクトルとする.

手順 7 $\mathbf{x}_{b,g}$ に摂動を加え, 変異ベクトル $\mathbf{v}$ とする.

手順 8 $\mathbf{v}$ とターゲットベクトル $\mathbf{x}_{i,g}$ を交叉させ, トライアルベクトル $\mathbf{u}$ を生成する.

手順 9 $f(\mathbf{u}) < f(\mathbf{x}_{i,g})$ ならば $\mathbf{x}_{i,g+1} = \mathbf{u}$ とする. そうでなければ, $\mathbf{x}_{i,g+1} = \mathbf{x}_{i,g}$ とする.

手順 10 $\mathbf{x}_{i,g} = \mathbf{x}_{i,g+1}$ ($i = 1, \dots, N$) とする.

手順 11 $g = g + 1$ として, 手順 2 に戻る.

2.3 微分突然変異

DE の手順 7 において, ベースベクトル $\mathbf{x}_{b,g}$ に摂動を加える操作は, 微分突然変異と呼ばれ, DE の主要な遺伝演算子である. 微分突然変異では, 集団からランダムに K 対の個体 $\mathbf{x}_{r_{k,g}}$ を選び, 式 (3) のように変異ベクトル $\mathbf{v}$ を生成する.

$$\mathbf{v} = \mathbf{x}_{b,g} + \sum_{k=1}^K \Upsilon (\mathbf{x}_{r_{2k-1,g}} - \mathbf{x}_{r_{2k,g}}) \quad (3)$$

ただし, $b \neq r_1 \neq r_2 \neq \dots \neq r_{2K}$ である.

スケール係数 Υ は定数であるが, 集団の収束とともに差ベクトルが小さくなるため, $\mathbf{x}_{b,g}$ に加わる摂動が適切に調節されることが期待できる.

3 連続世代モデルによる DE

連続世代モデルに基づく DE⁵⁾ の処理手順を以下に示す. 集団は一つだけであり, 集団で最良の個体 $\mathbf{x}_{best}$ は, 世代とは独立に随時更新される.

〔連続世代モデルによる DE〕

手順 1 初期集団の個体 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) をランダムに生成し, 世代を $g = 0$ とする.

手順 2 初期集団で最良の個体を $\mathbf{x}_{best}$ とする.

手順 3 所定の終了条件を満たせば, 最良の個体 $\mathbf{x}_{best}$ を出力して終了する.

手順 4 集団の各個体 $\mathbf{x}_{i,g}$ ($i = 1, \dots, N$) に関して, 以下の手順 5 から手順 10 を繰返す.

手順 5 個体 $\mathbf{x}_{i,g}$ をターゲットベクトルとする.

手順 6 集団から一つの個体 $\mathbf{x}_{b,g}$ を複製選択し, ベースベクトルとする.

手順 7 $\mathbf{x}_{b,g}$ に摂動を加え, 変異ベクトル $\mathbf{v}$ とする.

手順 8 $\mathbf{v}$ とターゲットベクトル $\mathbf{x}_{i,g}$ を交叉させ, トライアルベクトル $\mathbf{u}$ を生成する.

手順 9 $f(\mathbf{u}) < f(\mathbf{x}_{i,g})$ ならば $\mathbf{x}_{i,g} = \mathbf{u}$ とする.

手順 10 $f(\mathbf{u}) < f(\mathbf{x}_{best})$ ならば $\mathbf{x}_{best} = \mathbf{u}$ とする.

手順 11 $g = g + 1$ として, 手順 3 に戻る.

4 テスト問題

4.1 易しい問題のグループ

以下の三つのテスト問題⁷⁾ を, 易しい問題のグループとする. 目的関数 $f_1(\mathbf{x}) \sim f_3(\mathbf{x})$ は単峰性である. また, 次数はすべて $D = 10$ とする.

- $f_1(\mathbf{x})$: Sphere 関数
- $f_2(\mathbf{x})$: Discontinuous 関数
- $f_3(\mathbf{x})$: Ridge 関数

4.2 難しい問題のグループ

以下の三つのテスト問題⁷⁾ を, 難しい問題のグループとする. 目的関数 $f_4(\mathbf{x}) \sim f_6(\mathbf{x})$ は多峰性であり, 変数間に依存関係がある. また, 次数はすべて $D = 10$ とする. 通常, Rosenbrock 関数は単峰性に分類されるが⁷⁾, 最近の研究⁸⁾ によれば, $D \geq 4$ で Rosenbrock 関数は多峰性となる.

- $f_4(\mathbf{x})$: Rosenbrock 関数
- $f_5(\mathbf{x})$: Ackley 関数
- $f_6(\mathbf{x})$: Griewank 関数

表 1: 2 因子実験の因子と水準

factor	level-1	level-2	level-3
generation	discrete	continuous	—
model (G)	(G_1)	(G_2)	—
population	12 D	8 D	4 D
size (N)	(N_1)	(N_2)	(N_3)

5 世代モデルの比較

5.1 実験方法

まず、母数因子は、世代モデルと集団サイズとする。表 1 に母数因子と水準を示す。表 1 において、世代モデルは離散世代モデル G_1 と連続世代モデル G_2 の 2 水準である。集団サイズの水準 N_q は、推奨値²⁾の範囲 $N = 5D \sim 10D$ から $N_2 = 8D$ と、その範囲外で大小の値を選び 3 水準とする。

つぎに、初期集団の分布など、DE の確率的な挙動の影響を誤差因子とする。ただし、DE を実行する際の乱数系列は、すべての処理で統一する。

DE の戦略は、「DE/rand/1/bin」の場合と、「DE/best/1/bin」の場合に分けて評価する。また、DE の制御パラメータであるスケール係数 Υ と交叉率 C_R は、 $\Upsilon = 0.9$ と $C_R = 0.9$ に固定する。さらに、DE の終了条件については、最良解 x_{best} の目的関数値 (ERR) が $\varepsilon = 10^{-6}$ 以下になった場合、あるいは、目的関数値の計算回数 (FES) が、その上限の 36×10^4 回に達した場合とする。

DE の性能を評価する指標は、上記の ERR と FES の二つとする。ここで、NFL 定理⁹⁾を引用するまでもなく、DE の性能は対象とする最適化問題に依存する。しかし、世代モデルの違いが DE の性能に及ぼす普遍的な効果を調べるため、前述のテスト問題における ERR と FES のデータを、以下の通り集計し、それらの特性値とする。

まず、前回の報告⁵⁾と同様、各処理 (G_m, N_q) の下で、テスト問題 $f_p(\mathbf{x})$ に DE を 20 回適用し、得られたデータを標準化して $y_{p,m,q,r}$ ($r = 1, \dots, 20$) とする。つぎに、易しい 3 種類のテスト問題において、それらを式 (4) のように平均化し、その処理 (G_m, N_q) の r 番目の特性値 $\hat{y}_{m,q,r}$ とする。

$$\hat{y}_{m,q,r} = \frac{1}{3} \left(\sum_{p=1}^3 y_{p,m,q,r} \right) \quad (4)$$

同様に、難しい問題のグループにおいても、テスト問題ごと 20 回測定したデータを標準化した後、3 種類のテスト問題で平均化して特性値とする。

表 2: FES の分散分析表 (複製選択: rand)

factor	V	F	P	
G	0.130	90.601	3.6E^{-16}	**
N	59.190	41141.44	1.1E^{-163}	**
$G \times N$	0.009	6.596	0.001	**
e	0.001			

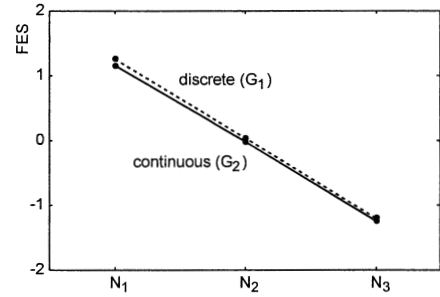


図 1: FES の各水準の母平均 (複製選択: rand)

5.2 易しい問題における実験結果

複製選択が「rand」の場合について、FES に関する分散分析表を表 2 に示す。分散分析表の各列は、要因、不偏分散 V 、分散比 F 、および P 値である。また、要因が危険率 $\alpha = 0.05$ で有意のとき「*」印を、 $\alpha = 0.01$ でも有意なら「**」印を行末に付けている。表 2 によれば、 $G \times N$ を含め、すべての要因が $\alpha = 0.01$ で有意である。

世代モデルと集団サイズの交互作用を詳細に調べるため、 $G \times N$ の各処理の母平均 (特性値の平均) を図 1 に示す。図 1 によれば、集団サイズが小さいほど FES は少なく、集団サイズが等しい場合は、離散世代モデル G_1 と比較し、連続世代モデル G_2 の方が、僅かながら FES は少ない。

図 1 から推察される連続世代モデルの優位性を検証するため、分散分析の F 検定と整合性があるシェフェの多重比較を実施し、その結果を表 3 に示す。表 3 の各列は、比較する二つの処理 (G_m, N_q)、両処理の母平均 $\hat{y}_{m,q}$ 、および、統計量 F である。また、各統計量 F には、両処理の母平均の差が危険率 $\alpha = 0.05$ で有意のとき「*」印を、 $\alpha = 0.01$ でも有意のとき「**」印を付けている。

表 3 の多重比較から、 $N_3 = 4D$ では世代モデルの違いによる FES の差は認められないが、集団サイズが大きな場合、連続世代モデルの方が、明らかに FES を減少させることが確認できる。

表 3: FES の多重比較 (複製選択: rand)

$G_m N_q$	$\hat{y}_{m,q}$		F
$G_1 N_1 \ G_2 N_1$	1.259	1.159	13.890 **
$G_1 N_2 \ G_2 N_2$	0.042	-0.014	4.649 **
$G_1 N_3 \ G_2 N_3$	-1.203	-1.243	2.219

表 4: FES の分散分析表 (複製選択: best)

factor	V	F	P
G	1.575	19.844	1.9E^{-5} **
N	42.409	534.344	1.2E^{-58} **
$G \times N$	0.022	0.287	0.750
e	0.079		

つぎに、複製選択が「best」の場合について、FES に関する分散分析表を表 4 に示す。表 4 によれば、世代モデル、集団サイズともに有意であるが、両者の交互作用はない。そこで、世代モデルごと処理をまとめて比較する。まず、離散世代モデル G_1 の全処理 (G_1, N_q) の母平均は $\hat{y}_{1,q} = 0.114$ 、連続世代モデル G_2 の全処理 (G_2, N_q) の母平均は $\hat{y}_{2,q} = -0.114$ であり、連続世代モデルの方が FES は少ない。さらに、シェフェの多重比較を行ったところ、統計量は $F = 19.844$ となり、両者の FES の差は危険率 $\alpha = 0.01$ で有意であった。

前述の通り、ERR に関しても 2 因子実験を実施した。しかし、何れの複製選択においても、世代モデル、集団サイズ、および、両者の交互作用は、 $\alpha = 0.05$ でも有意とはならず、世代モデルの違いは ERR に影響しないことが確認された。

5.2.1 難しい問題における実験結果

まず、複製選択が「rand」の場合、分散分析の F 検定によれば、FES、ERR とともに世代モデル G 、および、集団サイズとの交互作用 $G \times N$ の効果はなく、集団サイズ N のみが有意であった。

つぎに、複製選択が「best」の場合、FES、ERR とともに世代モデル G の主効果は認められず、集団サイズとの交互作用 $G \times N$ が有意となった。

そこで、シェフェの多重比較により交互作用を解析したところ、 $N_1 = 12D$ と $N_2 = 8D$ では、世代モデルの違いで DE の性能に差はなく、 $N_3 = 4D$ では、離散世代モデルによる DE の方が、FES は少なく、ERR は小さいことが明らかとなった。

6 おわりに

分散分析を用いて得られた統計的な知見によれば、対象とする最適化問題の目的関数が単峰性で、易しい問題に属する場合、適切な集団サイズで連続世代モデルの DE を適用するならば、既存の離散世代モデルの DE によって得られる解の質を損なうことなく、その処理時間を短縮できる。

一方、対象とする最適化問題の目的関数が多峰性で、難しい問題に属する場合、連続世代モデルによる DE の性能は、離散世代モデルの DE と比べて大差はない。ただし、複製選択が「best」で、集団サイズが推奨値²⁾よりも小さな場合、連続世代モデルによる DE の性能は、FES、ERR とともに、離散世代モデルの DE よりも悪くなる。

参考文献

- 1) 北野宏明: 遺伝的アルゴリズム, 産業図書 (1993)
- 2) R. Storn and K. Price: "Differential evolution - a simple and efficient heuristic for global optimization over continuous space," *Journal of Global Optimization*, Vol. 11, No. 4, pp. 341-359 (1997)
- 3) K. V. Price, R. M. Storn and J. A. Lampinen: *Differential Evolution - A Practical Approach to Global Optimization*, Springer (2005)
- 4) N. Noman and H. Iba: "A new generation alternation model for differential evolution," *Proc. of Genetic and Evolutionary Computation Conference*, pp. 1265-1272 (2006)
- 5) 田川聖治: 「微分進化法における連続世代モデルの提案と分散分析による評価」, 情報処理学会研究報告, MPS-69, pp. 45-48 (2008)
- 6) 田中豊: パソコン実験計画法入門, 現代数学社 (1985)
- 7) X. Yao, Y. Liu and G. Lin: "Evolutionary programming made faster," *IEEE Trans. on Evolutionary Computation*, Vol. 3, No. 2, pp. 82-102 (1999)
- 8) "A note of the extended Rosenbrock Function," *Evolutionary Computation*, Vol. 14, No. 1, pp. 119-126 (2006)
- 9) D. H. Wolpert and W. G. Macready: "No free lunch theorems for optimization," *IEEE Trans. on Evolutionary Computation*, Vol. 1, No. 1, pp. 67-82 (1997)