

実数値型遺伝的アルゴリズムにおける三次統計量の理論解析

A theoretical analysis of the third statistics in real-coded genetic algorithms

染谷 博司*1

Hiroshi Someya

*1統計数理研究所

The Institute of Statistical Mathematics

This paper presents a theoretical analysis of the third statistics in multivariate real-parameter spaces for applying the resultant equations to real-coded genetic algorithms, whereas most related studies have considered only the first and second statistics. Based on the third statistics, asymmetrical properties of two crossover operators are investigated. The theoretically obtained equations have been verified successfully through numerical simulations. Empirical study with benchmark functions has revealed a noticeable relationship between the theory, search behavior, and fitness landscape. The experimental results are discussed from a practical point of view.

1. はじめに

進化計算の挙動解析は構成要素の相互作用の複雑さから難問とされ、実施例の多くは、定性的考察に基づく仮説の提示と実験的解析による妥当性の検証にて構成される。しかし主観性や恣意性の排除には限りがあり、一般性に優れ広範に適用可能な知見を得るには理論解析が不可欠である。特に遺伝子型空間をユークリッド距離にて定義した実数値型進化計算の理論解析は、直感性・客観性の高い定量的知見が期待でき有意義である。

本研究では、実数値型遺伝的アルゴリズム (RCGA) を対象に、多次元遺伝子型空間における個体分布の三次統計量の理論解析を示す。なお既存研究では二次までの統計量を調査対象とすることが一般的である。三次統計量を議論の一部に含む研究報告も極少数あるものの、主観性のある定性的考察や限定的な探索局面での実験的解析例を示すに留まっており、多変量三次統計量の理論解析として一般化された報告は知られていない。

三次統計量は、探索のダイナミクスを理解し最適化効率に優れた RCGA を開発するための重要な指標である。個体集団全体については、ある探索時点までに獲得した有望探索領域の重心を示す一次統計量、その範囲の広がりを表す二次統計量に続き、三次統計量は領域内での有望さの非対称性を表し、目的関数観や探索重点化の方向性を反映する特徴量であると考えられる。交叉設計においては、子個体分布の二次までの統計量が等しい複数の交叉オペレータが提案されており、交叉オペレータ間の性質の差異を定量的に論じ探索特性を明らかにするための有用な統計量である。したがって本研究は、理論的根拠のある RCGA の設計指針の充実に貢献する新たなステップとして意義深く重要な研究課題であると考えられる。

以降、まず、多変量三次統計量の統計学的解析を示し、次に、得られた関係式を RCGA に適用し交叉オペレータの性質を議論する。本稿では“分布”や“統計量”などの表現は遺伝子型空間における幾何学的性質に対し用いる (評価値分布ではない)。

2. 多変量分布の三次統計量の解析

本章の解析は RCGA とは独立に成立する統計学的解析であり、進化戦略や PSO などを含め広範に適用可能である。なお RCGA では、 \mathcal{P} は個体群に、 \mathbf{x} は個体ベクトルに対応する。

連絡先: 染谷 博司 some@ism.ac.jp,

統計数理研究所 〒190-8562 東京都立川市緑町 10-3

2.1 歪度の定量化

多変量の三次統計量を扱った統計学における既存研究の多くは、スカラーにて歪度を定義し、正規性を検定する測度としてのその妥当性を主張するものである [Lütkepohl 91]。一方、本研究では個体群の挙動解析を狙いとしており、例えば親個体群とそこから生成される子個体群といった、分布形状に強い類似性のある分布間の差の定量化に興味がある。そこで、調査対象の分布 \mathcal{P} と \mathcal{P}' の間に次の関係が成立することを仮定する。

$$\mathbb{V}[\mathcal{P}] = \varkappa_2 \mathbb{V}[\mathcal{P}'], \quad \mathbb{T}[\mathcal{P}] = \varkappa_3 \mathbb{T}[\mathcal{P}']. \quad (1)$$

ここで $\mathbb{V}[\star]$ は任意の集合 \star における分布の広がりを表す二次統計量であり、 $\mathbb{T}[\star]$ は非対称性の大きさを表す三次統計量である。 \varkappa_2, \varkappa_3 はスカラーである。本研究では $\mathbb{V}[\mathcal{P}]$ は最も一般的な二次統計量である分散共分散行列とする [喜多 99]。すなわち、 \mathcal{P} の構成要素を \mathbf{x} 、その平均ベクトルを \mathbf{g} とすると

$$\mathbb{V}[\mathcal{P}] = M \left\langle (\mathbf{x} - \mathbf{g})(\mathbf{x} - \mathbf{g})^T \right\rangle \quad (2)$$

である。 $M[\star]$ は平均操作を表し、 T は転置である。

多変量分布の非対称性については、任意の目的に適用可能な広く受け入れられた定量化法は知られていない。本研究では式 (2) に倣い、その自然な拡張として次式にて定義する。

$$\mathbb{T}[\mathcal{P}] = M \left\langle (\mathbf{x} - \mathbf{g})(\mathbf{x} - \mathbf{g})^T (\mathbf{x} - \mathbf{g}) \right\rangle. \quad (3)$$

ただし $\mathbb{T}[\mathcal{P}]$ はスケール変換に依存するため \varkappa_3 を標準化し、 \mathcal{P} と \mathcal{P}' の非対称性の関係を以下にて定義する。

$$\mathbb{S}[\mathcal{P}] = \gamma_3 \mathbb{S}[\mathcal{P}'], \quad \gamma_3 = \frac{\varkappa_3}{\varkappa_2^{\frac{3}{2}}} \quad (4)$$

本研究では便宜的に $\mathbb{S}[\mathcal{P}]$ を \mathcal{P} の“歪度”と呼ぶ。なお、式 (1)~(4) はいずれも座標軸の回転に対して普遍である。

2.2 有限標本分布の三次統計量

標本数 μ (≥ 3) の任意の有限標本分布 \mathcal{P}_μ の三次統計量 $\mathbb{T}[\mathcal{P}_\mu]$ を解析する。構成要素およびその平均ベクトルを $\mathbf{x}_\mu^{(j)}$ および \mathbf{g}_μ と表すと、式 (3) より $\mathbb{T}[\mathcal{P}_\mu]$ は次式にて表される。

$$\begin{aligned} \mathbb{T}[\mathcal{P}_\mu] = & M \left\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)T} \mathbf{x}_\mu^{(j)} \right\rangle + 2\mathbf{g}_\mu \mathbf{g}_\mu^T \mathbf{g}_\mu \\ & - M \left\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)T} \mathbf{g}_\mu + \mathbf{x}_\mu^{(j)} \mathbf{g}_\mu^T \mathbf{x}_\mu^{(j)} + \mathbf{g}_\mu \mathbf{x}_\mu^{(j)T} \mathbf{x}_\mu^{(j)} \right\rangle. \quad (5) \end{aligned}$$

右辺第 3 項内の各項は次式のように表される。

$$\begin{aligned} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{g}_\mu \rangle \\ = \frac{1}{\mu} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle + \frac{\mu-1}{\mu} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} \rangle. \end{aligned} \quad (6)$$

$\mathbf{x}_\mu^{(k)}$ は $\mathbf{x}_\mu^{(j)}$ と異なる別の構成要素である。よって次式を得る。

$$\begin{aligned} \mathbb{T}[\mathcal{P}_\mu] &= \frac{\mu-3}{\mu} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle + 2\mathbf{g}_\mu \mathbf{g}_\mu^\top \mathbf{g}_\mu \\ &\quad - \frac{\mu-1}{\mu} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} + \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(j)} + \mathbf{x}_\mu^{(k)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle. \end{aligned} \quad (7)$$

2.3 母集団分布とその標本分布の歪度の関係

\mathcal{P}_μ を母集団分布とみなし、一様乱数での θ 回 ($3 \leq \theta \leq \mu$) の非復元抽出にて得られる標本分布 \mathcal{P}_θ との歪度の関係性を調べる。

まず、式 (7) を \mathcal{P}_θ に適用し両辺の期待値をとり次式を得る。

$$\begin{aligned} E\langle \mathbb{T}[\mathcal{P}_\theta] \rangle &= \frac{\theta-3}{\theta} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle + 2E\langle \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta \rangle \\ &\quad - \frac{\theta-1}{\theta} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} + \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(j)} + \mathbf{x}_\mu^{(k)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \end{aligned} \quad (8)$$

また、

$$\begin{aligned} \mathbf{g}_\mu \mathbf{g}_\mu^\top \mathbf{g}_\mu &= \frac{1}{\mu^2} \left\{ M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \right. \\ &\quad + (\mu-1) M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} + \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(j)} + \mathbf{x}_\mu^{(k)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \\ &\quad \left. + (\mu-1)(\mu-2) M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(\ell)} \rangle \right\} \end{aligned} \quad (9)$$

であり、 $\mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta$ についても同様に求め期待値をとると

$$\begin{aligned} E\langle \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta \rangle &= \frac{1}{\theta^2} \left\{ M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \right. \\ &\quad + (\theta-1) M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} + \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(j)} + \mathbf{x}_\mu^{(k)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \\ &\quad \left. + (\theta-1)(\theta-2) M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(\ell)} \rangle \right\} \end{aligned} \quad (10)$$

であるので、式 (9)(10) から $M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(\ell)} \rangle$ を消去し

$$\begin{aligned} \frac{\theta^2(\mu-2)}{\theta-1} E\langle \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta \rangle \\ = \frac{(\mu-\theta)(\mu+\theta-3)}{(\mu-1)(\theta-1)} M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \\ + (\mu-\theta) M\langle \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(k)} + \mathbf{x}_\mu^{(j)} \mathbf{x}_\mu^{(k)\top} \mathbf{x}_\mu^{(j)} + \mathbf{x}_\mu^{(k)} \mathbf{x}_\mu^{(j)\top} \mathbf{x}_\mu^{(j)} \rangle \\ + \frac{\mu^2(\theta-2)}{\mu-1} \mathbf{g}_\mu \mathbf{g}_\mu^\top \mathbf{g}_\mu \end{aligned} \quad (11)$$

となる。式 (7)(8) を式 (11) に代入して整理し次式を得る。

$$\mathbb{T}[\mathcal{P}_\mu] = \frac{\theta^2(\mu-1)(\mu-2)}{\mu^2(\theta-1)(\theta-2)} E\langle \mathbb{T}[\mathcal{P}_\theta] \rangle. \quad (12)$$

分散共分散行列については

$$\mathbb{V}[\mathcal{P}_\mu] = \frac{\theta(\mu-1)}{\mu(\theta-1)} E\langle \mathbb{V}[\mathcal{P}_\theta] \rangle \quad (13)$$

であるので ([Someya 08] など)、式 (4) より歪度は

$$\mathbb{S}[\mathcal{P}_\mu] = \frac{\mu-2}{\sqrt{\mu(\mu-1)}} \cdot \frac{\sqrt{\theta(\theta-1)}}{\theta-2} E\langle \mathbb{S}[\mathcal{P}_\theta] \rangle \quad (14)$$

の関係にあることがわかる。

なお、 \mathcal{P}_μ を確率密度関数 $\psi(\mathcal{P}_\mu)$ にしたがる確率分布とみなすと、式 (12) および式 (14) にて $\mu \rightarrow \infty$ とした後に θ を μ に置き換えることで $\psi(\mathcal{P}_\mu)$ の不偏推定量

$$\langle \mathbb{T}[\psi(\mathcal{P}_\mu)] \rangle = \frac{\mu^2}{(\mu-1)(\mu-2)} \mathbb{T}[\mathcal{P}_\mu], \quad (15)$$

$$\langle \mathbb{S}[\psi(\mathcal{P}_\mu)] \rangle = \frac{\sqrt{\mu(\mu-1)}}{\mu-2} \mathbb{S}[\mathcal{P}_\mu] \quad (16)$$

が得られ、スカラー変数の標本群を対象とした正規性検定のための歪度の関係式 (例えば [Waerden 69]) と整合する。

2.4 混合確率密度関数の歪度

重心が \mathbf{g}_θ 、分散共分散行列が $z_2 \mathbb{V}[\mathcal{P}_\theta]$ である任意の確率密度関数を $\psi(\mathcal{P}_\lambda)$ と表す。 z_2 は非負の実数である。また、 \mathcal{P}_μ からの \mathcal{P}_θ の抽出を独立に ω 回行い、それぞれを $\mathcal{P}_\theta^{(j)}$ と表す。各 $\mathcal{P}_\theta^{(j)}$ に対応する $\psi(\mathcal{P}_\lambda)$ を $\psi(\mathcal{P}_\lambda^{(j)})$ と表し、これらの混合確率密度関数 $\frac{1}{\omega} \sum_{j=1}^{\omega} \psi(\mathcal{P}_\lambda^{(j)})$ の $\omega \rightarrow \infty$ における極限を $\psi(\mathcal{P}_\lambda)$ と表す。本節では \mathcal{P}_μ と $\psi(\mathcal{P}_\lambda)$ の歪度の関係性を調査する。

$\psi(\mathcal{P}_\lambda)$ から得られる任意の標本群の構成要素を \mathbf{x}_λ と表し、次式にて $\mathbb{T}'[\psi(\mathcal{P}_\lambda)]$ を定義する。

$$\mathbb{T}'[\psi(\mathcal{P}_\lambda)] = M\langle (\mathbf{x}_\lambda - \mathbf{g}_\mu)(\mathbf{x}_\lambda - \mathbf{g}_\mu)^\top (\mathbf{x}_\lambda - \mathbf{g}_\mu) \rangle. \quad (17)$$

これを用いると $\mathbb{T}[\psi(\mathcal{P}_\lambda)]$ は次のように表される。

$$\mathbb{T}[\psi(\mathcal{P}_\lambda)] = E\langle E\langle \mathbb{T}'[\psi(\mathcal{P}_\lambda)] \rangle \rangle. \quad (18)$$

右辺の期待値演算は、内側は \mathbf{x}_λ のランダム性について、外側は \mathcal{P}_θ のランダム性についてのものである。内側の期待値は、 $\mathbf{x}_\lambda - \mathbf{g}_\mu = (\mathbf{x}_\lambda - \mathbf{g}_\theta) + (\mathbf{g}_\theta - \mathbf{g}_\mu)$ に注意して、

$$\begin{aligned} E\langle \mathbb{T}'[\psi(\mathcal{P}_\lambda)] \rangle \\ = E\langle \mathbb{T}[\psi(\mathcal{P}_\lambda)] \rangle + (\mathbf{g}_\theta - \mathbf{g}_\mu)(\mathbf{g}_\theta - \mathbf{g}_\mu)^\top (\mathbf{g}_\theta - \mathbf{g}_\mu) \\ + E\langle (\mathbf{x}_\lambda - \mathbf{g}_\theta)(\mathbf{x}_\lambda - \mathbf{g}_\theta)^\top (\mathbf{g}_\theta - \mathbf{g}_\mu) \rangle \\ + E\langle (\mathbf{x}_\lambda - \mathbf{g}_\theta)(\mathbf{g}_\theta - \mathbf{g}_\mu)^\top (\mathbf{x}_\lambda - \mathbf{g}_\theta) \rangle \\ + E\langle (\mathbf{g}_\theta - \mathbf{g}_\mu)(\mathbf{x}_\lambda - \mathbf{g}_\theta)^\top (\mathbf{x}_\lambda - \mathbf{g}_\theta) \rangle \end{aligned} \quad (19)$$

であり、 $\mathbb{V}[\psi(\mathcal{P}_\lambda)] = z_2 \mathbb{V}[\mathcal{P}_\theta]$ の仮定より右辺第 3~5 項は

$$\begin{aligned} z_2 \left\{ M\langle \mathbf{x}_\theta^{(j)} \mathbf{x}_\theta^{(j)\top} \mathbf{x}_\theta^{(j)} \rangle - \mathbb{T}[\mathcal{P}_\theta] \right. \\ \left. - M\langle \mathbf{x}_\theta \mathbf{x}_\theta^\top \mathbf{g}_\mu + \mathbf{x}_\theta \mathbf{g}_\mu^\top \mathbf{x}_\theta + \mathbf{g}_\mu \mathbf{x}_\theta^\top \mathbf{x}_\theta \rangle \right. \\ \left. + (\mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\mu + \mathbf{g}_\theta \mathbf{g}_\mu^\top \mathbf{g}_\theta + \mathbf{g}_\mu \mathbf{g}_\theta^\top \mathbf{g}_\theta) - \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta \right\} \end{aligned} \quad (20)$$

であるので、式 (18) は

$$\begin{aligned} \mathbb{T}[\psi(\mathcal{P}_\lambda)] \\ = E\langle \mathbb{T}[\psi(\mathcal{P}_\lambda)] \rangle + 2\mathbf{g}_\mu \mathbf{g}_\mu^\top \mathbf{g}_\mu - z_2 E\langle \mathbb{T}[\mathcal{P}_\theta] \rangle \\ + z_2 M\langle \mathbf{x}_\mu \mathbf{x}_\mu^\top \mathbf{x}_\mu - (\mathbf{x}_\mu \mathbf{x}_\mu^\top \mathbf{g}_\mu + \mathbf{x}_\mu \mathbf{g}_\mu^\top \mathbf{x}_\mu + \mathbf{g}_\mu \mathbf{x}_\mu^\top \mathbf{x}_\mu) \rangle \\ + (1 - z_2) E\langle \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\theta - (\mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\mu + \mathbf{g}_\theta \mathbf{g}_\mu^\top \mathbf{g}_\theta + \mathbf{g}_\mu \mathbf{g}_\theta^\top \mathbf{g}_\theta) \rangle \end{aligned} \quad (21)$$

となる。右辺第 5 項 (最尾項) 内の $\mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\mu$ は $E\langle \mathbf{g}_\theta \mathbf{g}_\theta^\top \mathbf{g}_\mu \rangle = \frac{1}{\theta(\mu-1)} \{ (\mu-\theta) M\langle \mathbf{x}_\mu \mathbf{x}_\mu^\top \mathbf{g}_\mu \rangle + \mu(\theta-1) \mathbf{g}_\mu \mathbf{g}_\mu^\top \mathbf{g}_\mu \}$ であり、

$g_\theta g_\mu^T g_\theta$ と $g_\mu g_\theta^T g_\theta$ も同様に表すことができる [Someya 08]. また $M\langle x_\mu x_\mu^T g_\mu + x_\mu g_\mu^T x_\mu + g_\mu x_\mu^T x_\mu \rangle = M\langle x_\mu x_\mu^T x_\mu \rangle + 2g_\mu g_\mu^T g_\mu - \mathbb{T}[\mathcal{P}_\mu]$ である. これらと式 (7)(11)(12) を式 (21) に代入し次式を得る.

$$\mathbb{T}[\psi(\mathcal{P}_\Lambda)] = E\langle \mathbb{T}[\psi(\mathcal{P}_\Lambda)] \rangle + \frac{(\mu - \theta) \{ \mu - 2\theta + 3\mu(\theta - 1)z_2 \}}{\theta^2(\mu - 1)(\mu - 2)} \mathbb{T}[\mathcal{P}_\mu]. \quad (22)$$

ここで z_3 を実数として $\mathbb{T}[\psi(\mathcal{P}_\Lambda)] = z_3 \mathbb{T}[\mathcal{P}_\theta]$ が成立するとき, 式 (12) を用いて

$$\mathbb{T}[\psi(\mathcal{P}_\Lambda)] = \frac{\mu^2(\theta - 1)(\theta - 2)z_3 + (\mu - \theta) \{ 3\mu(\theta - 1)z_2 + \mu - 2\theta \}}{\theta^2(\mu - 1)(\mu - 2)} \mathbb{T}[\mathcal{P}_\mu] \quad (23)$$

である. また, 式 (15) および

$$\mathbb{V}[\psi(\mathcal{P}_\Lambda)] = \frac{\mu(\theta - 1)z_2 + \mu - \theta}{\mu\theta} \langle \langle \mathbb{V}[\psi(\mathcal{P}_\mu)] \rangle \rangle \quad (24)$$

であることから [Someya 08], 興味対象の関係式は次式となる.

$$\mathbb{S}[\psi(\mathcal{P}_\Lambda)] = \frac{\mu^2(\theta - 1)(\theta - 2)z_3 + (\mu - \theta) \{ 3\mu(\theta - 1)z_2 + \mu - 2\theta \}}{\sqrt{\mu\theta} \{ \mu(\theta - 1)z_2 + \mu - \theta \}^{3/2}} \cdot \langle \langle \mathbb{S}[\psi(\mathcal{P}_\mu)] \rangle \rangle. \quad (25)$$

3. 実数値型遺伝的アルゴリズムへの適用

前章の解析は, 次のように対応付けることで標準的な世代交代モデルにおける RCGA に適用できる.

- $\mathcal{P}_\mu, x_\mu, \mu \dots$ 個体集団, 各個体, 集団サイズ
- $\mathcal{P}_\theta, x_\theta, \theta \dots$ 交叉に参加する親個体群, 各親個体, 親個体数
- $\psi(\mathcal{P}_\Lambda), x_\lambda \dots$ 交叉オペレータの確率密度関数, 各子個体
- $\psi(\mathcal{P}_\Lambda) \dots$ 複数回の交叉により生成される子個体群全体
- $z_2, z_3 \dots$ 交叉の親子間での局所的な分布の拡大縮小率および三次統計量の変化率
- $\varkappa_2, \varkappa_3 \dots$ 複数回の交叉により生成される子個体群全体の拡大縮小率および三次統計量の変化率

本節では, 代表的な交叉オペレータである UNDX- m [喜多00] と SPX [樋口 01] の三次統計量の変化率 z_3 を示し, 交叉オペレータの確率モデルとしての性質を議論する.

3.1 UNDX- m の三次統計量

UNDX- m の確率密度関数は次のように表される.

$$x_\lambda = g_\theta + \sum_{j=1}^{\theta-1} \xi^{(j)} d^{(j)} + \phi, \quad (26)$$

$$d^{(j)} = x_\theta^{(j)} - g_\theta. \quad (27)$$

各 $\xi^{(j)}$ は正規分布 $\mathcal{N}(0, \sigma_\xi^2)$ にしたがう独立な乱数であり, σ_ξ はパラメータである. ϕ は探索次元数に対して親個体数 θ が十分でない時に用いられる補助的な成分であり, 本稿では簡潔にゼロベクトルとして扱う.

重心は g_θ に一致し, 分散共分散行列は $z_2 = (\theta - 1) \sigma_\xi^2$ にて第 2.4 節の $\psi(\mathcal{P}_\Lambda)$ の要件を満たす [Someya 08]. 三次統計量は, σ_ξ の値に関わらず $E\langle \xi^3 \rangle = 0$ であることから $\mathbb{T}[\psi(\mathcal{P}_\Lambda)] = E\langle \xi^3 \rangle \sum_{j=1}^{\theta-1} d^{(j)} d^{(j)T} d^{(j)} = 0$ となり $z_3 = 0$ である.

3.2 SPX の三次統計量

SPX の確率密度関数は次のように表される.

$$x_\lambda = g_\theta + \varepsilon \left(d^{(\theta)} + s^{(\theta)} \right), \quad (28)$$

$$s^{(j)} = \begin{cases} \mathbf{0} & : j = 1 \\ v^{(j-1)} \{ r^{(j)} + s^{(j-1)} \} & : j = 2 \dots \theta, \end{cases} \quad (29)$$

$$r^{(j)} = x_\theta^{(j-1)} - x_\theta^{(j)}, \quad v^{(k)} = \left\{ v^{(k)} \right\}^{\frac{1}{k}}. \quad (30)$$

ε はパラメータであり, 各 $v^{(k)}$ は一様分布 $\mathcal{U}[0, 1]$ にしたがう独立な乱数である.

重心は g_θ に一致し, 分散共分散行列は $z_2 = \frac{\varepsilon^2}{\theta+1}$ にて第 2.4 節の $\psi(\mathcal{P}_\Lambda)$ の要件を満たす [Someya 08]. 三次統計量については, 論理性を損なわずに導出の式展開を書き下すと長大な紙面を要することから, 紙面の制約上, 結果のみを式 (31) に示す. なお, 導出には仮定や近似といった恣意性の余地はなく, 式 (3) の定義に従った率直な手続きのみで可能である.

$$z_3 = \frac{2}{(\theta+1)(\theta+2)} \varepsilon^3. \quad (31)$$

3.3 確率モデルとしての交叉の性質

交叉操作の役割は, 個体集団の確率モデルの構築とリサンピングによる分布の変換であり [喜多 99], UNDX- m や SPX による交叉操作は一次および二次の統計量を保存できる [Someya 08]. しかしこのとき三次統計量に関しては式 (4)(25) にて $\gamma_3 < 1$ となり保存されず, 非対称性が損なわれることを示す.

式 (1)(24) より $z_2 = \frac{\varkappa_2 \theta - 1}{\theta - 1} + \frac{\theta}{\mu(\theta - 1)}$ であることに注意すると, $\gamma_3 \geq 1$ であるための拘束条件は

$$z_3 \geq \frac{(\mu\theta)^2 (\varkappa_2)^{3/2} - (\mu - \theta) (3\mu\varkappa_2\theta - 2\mu + \theta)}{\mu^2(\theta - 1)(\theta - 2)} \quad (32)$$

である. 二次統計量の保存条件 $\varkappa_2 = 1$ を代入し整理すると,

$$z_3 \geq \frac{1}{\mu(1 - \frac{2}{\theta})} \left\{ 3 + \frac{1}{\mu(1 - \frac{1}{\theta})} \right\} + 1 \quad (33)$$

となり, μ や θ の値に関わらず $z_3 > 1$ が必要条件となる.

UNDX- m は $z_3 = 0$ であるためこれを満足しない. SPX では式 (31) の ε を z_2 を介して \varkappa_2 で置き換えると

$$z_3 = \frac{2}{(\theta - 1)(\theta + 2)} \sqrt{\frac{\theta + 1}{\theta - 1}} \left(\varkappa_2 \theta + \frac{\theta}{\mu} - 1 \right)^{3/2} \quad (34)$$

であり, $\varkappa_2 = 1$ では

$$z_3 = G_1(\mu, \theta)^{3/2} \cdot G_2(\theta), \quad G_1(\mu, \theta) = 1 + \frac{1}{\mu} - \frac{1}{\theta}, \quad G_2(\theta) = \frac{2\theta\sqrt{\theta(\theta+1)}}{(\theta+2)(\theta-1)^{3/2}} \quad (35)$$

である. $G_1(\mu, \theta) \leq 1$ であるので, $z_3 > 1$ を満足するには $G_2(\theta) > 1$ が必要条件となる. しかし $G_2(\theta)$ は単調減少関数であり $G_2(5) \lesssim 0.98$ であるので $\theta \geq 5$ では μ の大きさに関わらず $z_3 < 1$ である. $\theta = 3, 4$ については式 (35) は $z_3^{(\theta=3)} \lesssim 1.47 \left(\frac{1}{\mu} + \frac{2}{3} \right)^{3/2}$ および $z_3^{(\theta=4)} \lesssim 1.15 \left(\frac{1}{\mu} + \frac{3}{4} \right)^{3/2}$

となり, 他方, 拘束条件は式 (33) に $\theta = 3, 4$ を代入し変形した

$$z_3^{(\theta=3)} \geq 1.5 \left(\frac{1}{\mu} + \frac{2}{3} \right) + \frac{3}{2\mu} \left(\frac{3}{\mu} + 5 \right), \quad (36)$$

$$z_3^{(\theta=4)} \geq 1.2 \left(\frac{1}{\mu} + \frac{3}{4} \right) + \frac{8}{\mu} \left(\frac{1}{3\mu} + \frac{3}{5} \right) + \frac{1}{10} \quad (37)$$

であるので, いずれも拘束条件を満たさず $\gamma_3 < 1$ となる.

なお, 式 (25) における γ_3 は z_3 に関して単調増加する. よって z_2 が等しい ($=\varkappa_2$ が等しい) ときには, $z_3 > 0$ である SPX に比べ, UNDX- m はより強く歪度を減小させることがわかる.

3.4 実験的確認

数値実験にて理論解析の妥当性を確認する. 適応度選択の影響を排除する狙いから, 次の世代交代モデルを用いた.

Step 1: 集団サイズ μ の初期集団 $\mathcal{P}_\mu^{(0)}$ を準備し, 世代数カウンタを $t = 0$ とする.

Step 2: 子個体群集合 $\mathcal{P}_\lambda^{(t)}$ として空集合を用意する.

Step 3: $\mathcal{P}_\mu^{(t)}$ から一様乱数に基づく非復元抽出にて θ 個体を選び $\mathcal{P}_\theta^{(t)}$ とする.

Step 4: $\mathcal{P}_\theta^{(t)}$ を親個体群とする交叉操作にて λ 個体を生成し子個体群 $\mathcal{P}_\lambda^{(t)}$ を構成する.

Step 5: $\mathcal{P}_\lambda^{(t)}$ を $\mathcal{P}_\lambda^{(t)}$ に加え, $\mathcal{P}_\theta^{(t)}$ を $\mathcal{P}_\mu^{(t)}$ に戻す.

Step 6: Steps 3-5 を ω 回繰り返す. ただし $\lambda\omega \geq \mu$ である.

Step 7: $\mathcal{P}_\lambda^{(t)}$ から一様乱数による非復元抽出にて μ 個体を選び次世代個体集団 $\mathcal{P}_\mu^{(t+1)}$ を構成する.

Step 8: t に 1 を加え step 2 に戻る.

初期集団は, n 次元空間中の領域 $[0, 1]^n$ にて一様ランダムに生成した後に各 i 次元成分を $1 + i/n$ 乗し, 非対称性をもたせた.

歪度の変化率の推移を図 1 に示す. 理論値との高い整合性が確認できる. なお, 対称な初期集団や異なる \varkappa_2, μ, θ を用いた数十種類の実験を行い, いずれも良好な整合性を確認している.

4. 最適化過程における歪度の時系列変化

探索の進行に伴う個体集団の歪度の時系列変化を観察し, 理論解析にて得た知見との関連性の有無を調査する. 目的関数とした \mathcal{F}_{sphere} および \mathcal{F}_{ridge} は最適解を原点に有する単峰性関数である. 多親型 MGG モデル [樋口 01] にて, 初期集団を領域 $[-5.12 - 1.5\delta, 5.12 - 1.5\delta]^n$ に $\delta = 0$ または $\delta = 3$ として一様ランダムに生成する. 前者では初期集団の重心に対して目的関数景観は対称である. 一方, 後者では初期化領域の境界付近に最適解が位置し景観は非対称である. 実験結果を図 2 に示す. 探索領域の景観, 個体集団の歪度, 交叉オペレータの歪度保存特性, の間には一定の関連性があることが確認できる. 個体集団分布の対称性は $\delta = 0$ では保たれているのに対し, $\delta = 3$ では大幅な変動がある. また, UNDX- m に比べ SPX の変動幅は大きく, 前章にて示した歪度の保存特性と整合している.

Steady-state 型の世代交代モデルは部分的な選択淘汰の蓄積により個体集団分布を徐々に変化させる. これを考慮すると, $\delta = 3$ での大幅な変動は, 探索序盤から中盤にて個体集団内における最適解方向の個体密度が高まり非対称性が増大し, その後, 個体集団の重心が最適解付近に概ね到達するに従い対称性を回復するとの挙動が反映されたものと思われる. \mathcal{F}_{sphere} に対する \mathcal{F}_{ridge} での非対称性増大の遅れは, $2 \sim n$ 次元成分の最適化が優先されたことに起因すると考えられる. このように, 今後, 多峰性景観での各探索局面と歪度変化の関係性などの調査にて, より包括的かつ定量的知見を得ることで, 歪度の計測に基づく景観分析や挙動解析の発展が期待できる. また, 能動的な歪度制御に基づく適応的探索法の構築も有望である.

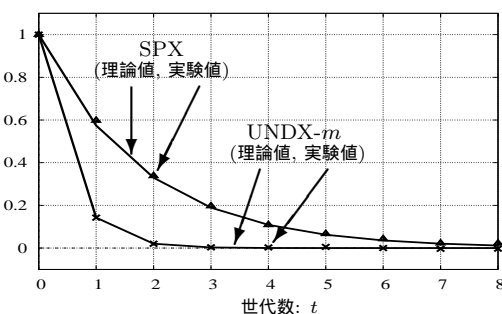
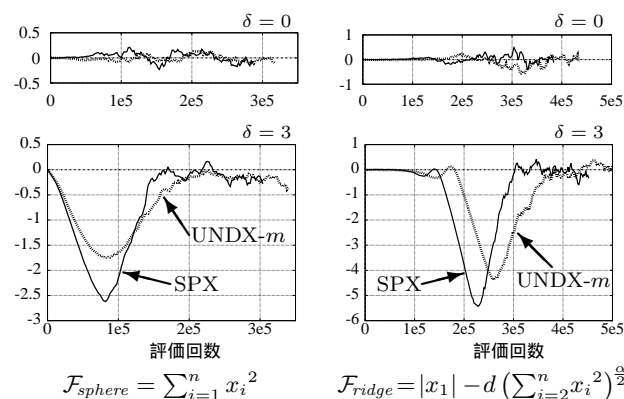


図 1: 初期集団に対する歪度の変化率. 定量的比較と可視化のため $\mathbb{T}[\mathcal{P}_\mu^{(t)}]$ の各要素毎の変化率を平均化している. $n = 12$, $\mu = 20, \theta = \lambda = 10, \varkappa = 1.0, \omega = 30\mu/\theta$ での $3e6$ 試行の平均である. 折れ線グラフは理論値を, 記号グラフは実験値を表す.



$$\mathcal{F}_{sphere} = \sum_{i=1}^n x_i^2 \quad \mathcal{F}_{ridge} = |x_1| - d \left(\sum_{i=2}^n x_i^2 \right)^{\frac{\alpha}{2}}$$

図 2: 最適化過程における第 1 次元成分の歪度の時系列変化. 各曲線は $n = 10, \mu = 500, \theta = n + 1, \lambda = 50, \varkappa_2 = 1.0$ での 1000 試行の平均である. \mathcal{F}_{ridge} は $d = -100, \alpha = 2.5$ である.

5. まとめ

本研究では, 多変量三次統計量の理論解析を行い RCGA に適用した. これにより得られた理論的根拠のある知見として, UNDX- m および SPX による分布の変換操作は歪度を保存せず, 特に前者はその傾向が強いことを示した. また, ベンチマーク関数での最適化実験では, 目的関数景観と歪度変化との関連性が観察され, 景観分析や挙動解析の指標としての有用性が示唆された. 前章にて述べた将来展望に加え, 進化戦略や PSO などの他の実数値型進化計算への理論解析の適用や, 三次統計量の定義自体の性質の調査が今後の課題である.

参考文献

- [Lütkepohl 91] Lütkepohl, H. and Theilen, B.: Measures of multivariate skewness and kurtosis for tests of nonnormality, *Statistical Papers*, Vol. 32, No. 1, pp. 179–193 (1991)
- [Someya 08] Someya, H.: Parameter Tuning of Real-valued Crossover Operators for Statistics Preservation, in *LNCS 5361 (Proceedings of the SEAL-2008)*, pp. 269–278 (2008)
- [Waerden 69] Waerden, van der B.: *Mathematical Statistics, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen Band 156*, Springer (1969)
- [喜多 99] 喜多 一, 山村 雅幸: 機能分担仮説に基づく GA の設計指針, 計測と制御, Vol. 38, No. 10, pp. 612–617 (1999)
- [喜多 00] 喜多 一, 小野 功, 小林 重信: 実数値 GA のための正規分布交叉の多数の親を用いた拡張法の提案, 計測自動制御学会論文集, Vol. 36, No. 10, pp. 875–883 (2000)
- [樋口 01] 樋口 隆英, 筒井 茂義, 山村 雅幸: 実数値 GA におけるシンプлекс交叉の提案, 人工知能学会論文誌, Vol. 16, No. 1, pp. 147–155 (2001)