実数値型進化計算の選択操作における二次統計量の解析

染谷 博司 モデリング研究系 助教

生物進化の仕組みにヒントを得た確率的最適化手法のひとつである進化計算 (EC) は、性能向上のための工夫においても生物進化とのアナロジーに根拠を求めることが多い。また、メタヒューリスティクスの一手法としても位置づけられるように、設計者が職人技的に工夫を考案し適用することも多く、探索原理の理論的解明は十分ではない。特に実数値型ECでは、一部に、進化戦略におけるステップサイズの自己適応化、確率分布推定アルゴリズムにおけるパラメトリック・モデリング、実数値型遺伝的アルゴリズム (実数値GA) における交叉設計、などがあるものの、単一個体の状態遷移や探索オペレータ個別の理論解析にとどまっており、選択操作を含む挙動全体のダイナミクスの理論解析は極めて稀である。

本研究では、実数値型ECの集団分布のダイナミクスを理論解析し、選択操作の前後において、分布の広がりや多様性の指標のひとつである分散・共分散行列(二次統計量)を定数倍に拡大・縮小するための拘束条件を示した。また、探索中盤以降にてしばしば観察されることが知られているlinear景観における分布の伸縮を制御し初期収束を回避するための生存選択方法の具体例を示した。得られた知見を実数値GAに適用し、子個体生成数などの各種GAパラメータ、生存選択方法、選択圧パラメータ、と伸縮率の関係を示し、望ましい選択操作の設計についての考察を示した。

選択操作は交叉や突然変異と並ぶ実数値型ECの主要な構成要素のひとつであるにも関わらず、現状では、各種の選択操作法をもっとも強く特徴付ける"選択圧の強さ"についてさえも概念的かつ主観的に論じられており、選択圧強度の定量化法、望ましい選択圧強度、選択圧強度と分布変化との対応関係、のいずれも未解明である。また、新たに提案される

選択操作法の有用性の評価は、いくつかの限られた問題設定における計算機実験での最適化性能に論拠が置かれている。このような背景において、本研究は理論的根拠のある設計指針の獲得に向けたひとつのステップとなりうることが期待される重要な研究課題であると考えられる。

解析により得られた二次統計量の理論値と実験値との比較、および、最適化性能との関係性を下図に示す。全体を俯瞰すると一定のパターンが感じられ、示唆された規則性から、定量的な挙動予測への期待感が感じられる。

本研究の成果は文献 [1-3] にて公表している。また、平成23年5月25日(水)に開催された統計数理セミナー「確率的最適化における多変量分布の理論解析」にて紹介した。

公表文献

- [1] Hiroshi Someya. Theoretical analysis of phenotypic diversity in real-valued evolutionary algorithms with more-than-one-element replacement. *IEEE Transactions on Evolutionary Computation*, Vol. 15, No. 2, pp. 248–266, 2011.
- [2] 染谷博司. 実数値型進化計算の選択操作における二次統計量の保存. 計測自動制御学会 システム・情報部門学術講演会 講演論文集, pp. 516-521, 東京工業大学 すずかけ台キャンパス すずかけホール, 11月24日~26日2009. 計測自動制御学会 システム・情報部門 優秀論文賞 受賞.
- [3] 染谷博司. 実数値型遺伝的アルゴリズムのパラメータ設定における可制御性の考察. 第6回 進化計算フロンティア研究会 資料集, pp. 11–18, 名古屋大学, 3月14–15日 (震災の影響により口頭発表はキャンセル) 2011. 人工知能学会.

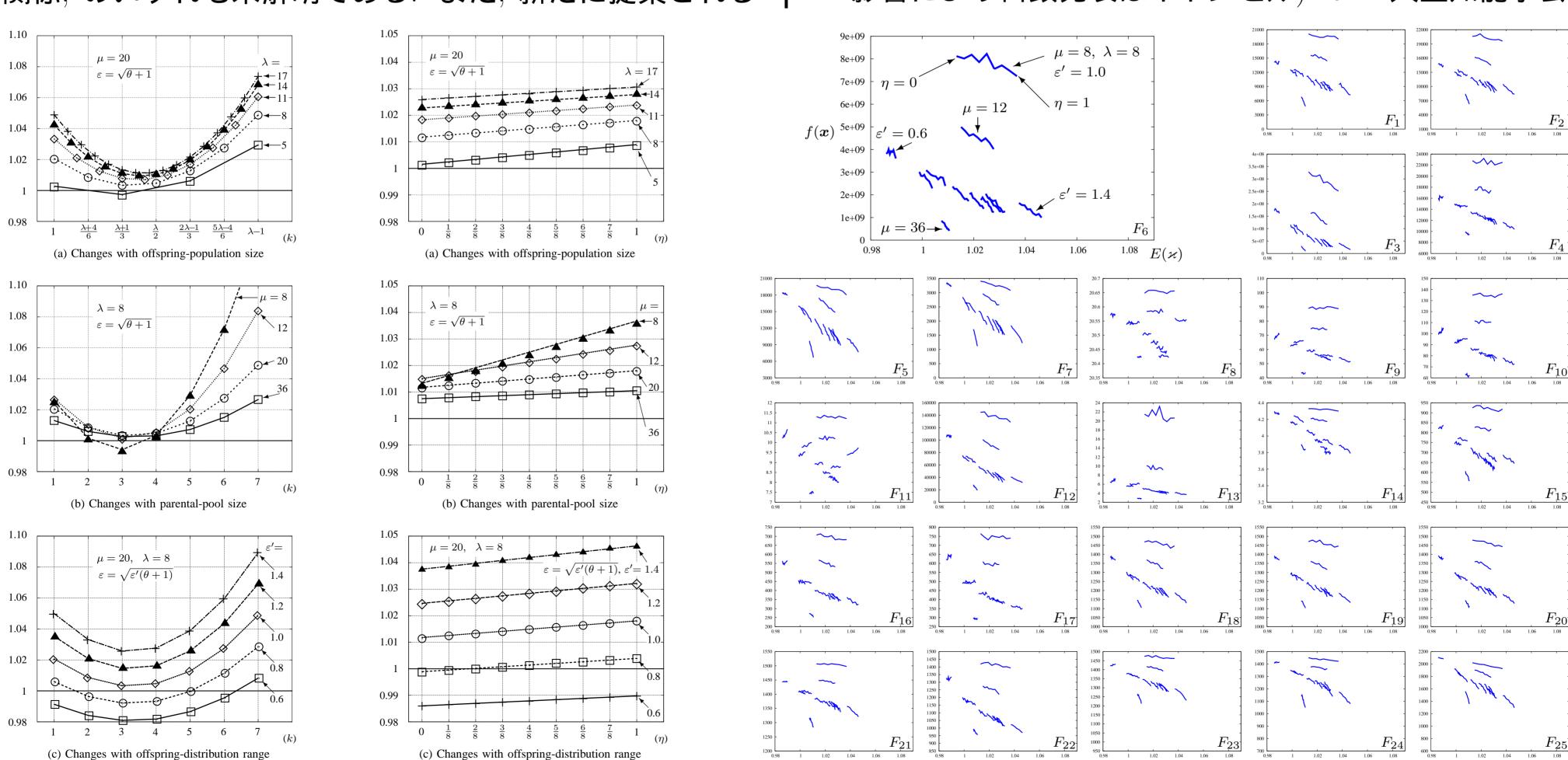


Figure 1: 左図は決定的ランク選択を用いたときの個体集団の二次統計量の変化率である. 折れ線グラフは理論値をあらわし、+などの記号プロットは数値実験の平均値である. 両者の高い整合性が確認できる.中央の図はルーレット選択を用いたときのものである.右図はルーレット選択を採用した実数値GAの二次統計量変化の期待値とCEC 2005 ベンチマークセットでの最適化性能の関係である. 各グラフは、子個体生成数 λ 、集団サイズ μ 、交叉の拡張率パラメータ ε' の組合せに対応する12本の曲線を含んでおり、各曲線は選択圧パラメータ $\eta=0,\frac{1}{8},\frac{2}{8},\ldots,1$ に対応する9点を結んで描かれている.