

進化計算を用いた人間の認知バイアスに適合する乱数の生成手法の提案

大谷 紀子 研究室

1572061 堤 創一朗

1. はじめに

進化計算手法には、多峰的な最適化問題でも大域的な最適解を効率的に得ることができる特徴があり、産業界からも注目を集め、さまざま分野で活用が始まっている。しかし、学术界でアルゴリズムの性能評価に用いられるベンチマーク問題と、産業界が解決を望んでいる問題の間には、少なからずギャップがあることが指摘されている。

以上の状況を踏まえ、進化計算学会実世界ベンチマーク問題分科会は、産業と科学の現場から収集した実問題を定式化して実世界ベンチマーク問題を作成し、毎年 12 月に進化計算コンペティションを開催している。4 回目となる進化計算コンペティション 2020 では、『ゲームを楽しくする乱数の設計問題』と題するベンチマーク問題が出題された[1]。あるゲームの開発において行われた乱数設計の事例に端を発する問題であり、乱数に対する認知バイアスの研究に基づいて定式化されている。

本研究では、人間の認知バイアスに適合する乱数の生成を目的として、実験デザインにより認知バイアスを推定した上で、広域人工蜂コロニーアルゴリズムにより乱数を生成する手法を提案する。進化計算コンペティション 2020 では単目的部門と多目的部門の 2 部門が用意されているが、本研究では単目的部門の問題を対象とする。

2. ゲームを楽しくする乱数の設計問題

野村ら[2]は、擬似乱数についての認知バイアスを表す指標として、全体で出た目の回数の χ^2 値、偶数と奇数が並ぶ部分の数、同じ目が 2 連続する

部分の数などの計 15 個の特徴量を提案している。また、各特徴量の許容範囲を設定し、プレイヤーの不満は許容範囲からの逸脱量に比例するとして、プレイヤーの不満を表すモデルを示している。「ゲームを楽しくする乱数の設計問題」は、野村らの不満モデルを最小化する問題である。同時に、擬似乱数の統計的な偏りを一定の範囲に制限するという制約条件も課されている。

設定された許容範囲はコンペティション中非公開であるため、実験デザインを用いて予測する必要がある。また、コンペティションでは、Small, Medium, Large, Huge という 4 種類の規模の問題がそれぞれ 4 問ずつ、計 16 問が提示された。問題の規模ごとに、生成する乱数列の長さはそれぞれ 60, 840, 2520, 27720, 解の送信回数の上限はそれぞれ 100, 200, 400, 800 と定められている。

3. 提案手法

提案手法における乱数生成手順を以下に記す。

(1)~(4)が実験デザインにより認知バイアスを推定するフェーズ、(5)~(6)が乱数を生成するフェーズである。

(1) 制約条件を満たす乱数列が 100 万個生成されるまで、ランダムに乱数列を生成する処理を繰り返す、実行可能解集合を生成する。

(2) 実行可能解集合から出来るだけ各特徴量の多様性を保ったまま、要素数を 1 万個に減らした部分集合を実験点候補とする。

(3) Maximum Projection 基準[3]に基づき、高速柔軟充填法[4]により実験点候補から実験点を選出する。

表 1: 実験結果

問題	最優良評価値
Small-1	11.963
Small-2	17.243
Small-3	54.373
Small-4	20.680
Medium-4	313.319
Huge-1	5852.115

- (4) 選出された各実験点を不満モデルによって評価し、各特徴量の許容範囲を推定する、具体的には、乱数列の各特徴量を説明変数、不満モデルによる評価値を目的変数として広域人工蜂コロニーアルゴリズムにより、残差二乗和が最小となるような許容範囲の設定を探索する。
- (5) 得られた許容範囲を用いて、不満モデルによる評価値を予測するモデルを作成する。
- (6) 予測モデルによる評価値が、実験点の評価値より小さくなる実行可能解を探索し、最終的な解とする。探索の際はまず、実行可能解集合の中から条件に合う解を探し、十分な解が見つからなかった場合は、貪欲法を用いて条件に合う実行可能解を探す。

4. 評価実験

提示された 16 問のうち、6 問について提案手法により乱数列を生成し、コンペティション用サーバで評価した。得られた最良評価値を表 1 に示す。本結果は、コンペティションにおいて 10 チーム中 3 位となった。

5. 考察

問題 Small-4 の実行可能解集合の一部の要素について、縦軸を不満モデルによる評価値、横軸を認知バイアス推定フェーズで作成された予測モデルに基づいて算出された予測評価値としてプロットすると、図 1 のようになる。ここで、青色の点は実験点、緑色の点は実験点以外の実行可能解を表す。実験点の真の評価値が広く分布しているにもかかわらず、予測評価値の分布が狭いことから、実験点の選出はある程度うまくいっていると考えられる。しかし、実行可能解全体としては真の評

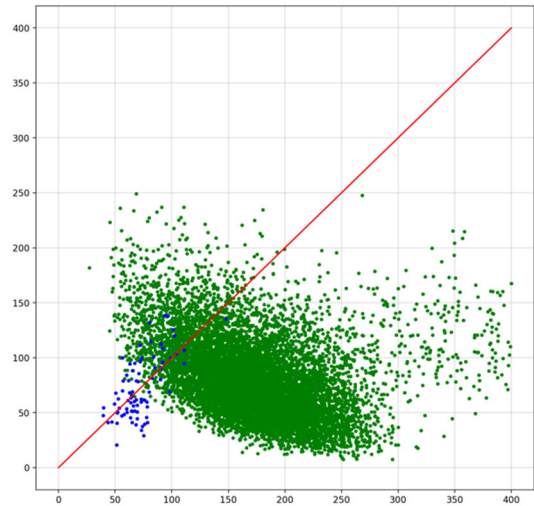


図 1: Small-4 での評価値の分布

価値と予測評価値に負の相関があるように見えるが、実験点にのみ着目すると正の相関があることから、作成した予測モデルは実験点にのみ過剰に適合しているといえる。他の問題についても同様に比較すると、正確な予測モデルが作成されていないことがわかる。以上より、提案手法には予測モデル作成の段階で過剰適合が起りやすいという問題があるといえる。予測モデルの性能を評価したうえで優良解を探索するなどの改良により、精度の向上が見込まれる。

参考文献

- [1] 濱田直希, 於保俊, "進化計算コンペティション 2020 問題解説 ——ゲームを楽しくする乱数の設計問題——," 第 14 回進化計算シンポジウム 2020 講演論文集, pp. 51-58, 2020.
- [2] 野村久光, テンシリリックンシラ, 池田心, "標準的なゲームプレイヤーにとって自然に見える疑似乱数列の生成法," ゲームプログラミングワークショップ 2013 論文集, pp.27-34, 1994.
- [3] Joseph, V. R., Gul, E., Ba, S., "Maximum Projection Designs for Computer Experiments," Biometrika, Vol. Issue 2, pp. 371-380, 2015.
- [4] Lekivetz, R., Jones, B., "Fast Flexible Space-Filling Designs for Nonrectangular Regions," Quality and Reliability Engineering International, Vol. 31, Issue 5, pp. 829-837, 2015.