

分散遺伝的アルゴリズムにおけるパラメータの検討

The examination of parameter settings for Distributed Genetic Algorithms

上浦 二郎

Jiro KAMIURA

Abstract: This paper is about the examination of parameter settings for Distributed Genetic Algorithms (DGA). Each parameter have big effect on GA's performance. The research that searching the good parameter settings is very important. So, I research the relationship between 12 parameters' setting and DGA's performance on 4 objective functions.

1 はじめに

分散遺伝的アルゴリズム (Distributed Genetic Algorithms : DGA) には多数のパラメータが存在し、これらの設定によって DGA の解探索性能は大きく異なる。このため、パラメータチューニングは非常に重要な研究対象となる。しかし、パラメータの組み合わせが膨大になることや、単一母集団の GA に関してはすでに同様の研究が行われているために、これまで DGA の複数のパラメータについて細かく研究がなされることはなかった。

そこで本研究では、複数の対象問題についてパラメータの検討を行った。

2 分散遺伝的アルゴリズムのパラメータ

本研究で対象としたパラメータを以下に示す。

- 個体数：探索点の数。個体数が多いほど広い範囲を探索できるが、収束が遅くなると考えられる。
- 選択手法：次世代を形成する個体を選択する方法。一般に、適合度の高い個体ほど選択されやすくする。
- 交叉手法：個体間の染色体の交換方法。
- 交叉率：個体集団のうち交叉に参加する個体の割合。交叉率が高いほど、個体集団内で染色体情報が交換される頻度が高くなる。
- 突然変異手法：個体の染色体の突然変異方法。
- 突然変異率：個体のもつ染色体のうち突然変異する遺伝子座の割合。突然変異率が高いほど染色体情報が破壊されやすくなる。
- 島数：個体集団の分割数¹。
- 移住トポロジー：個体の移住先の決定方法。

- 移住率：島内の個体のうち移住する個体の割合。移住率が高いほど移住する個体が多くなるため、個体集団内で個体情報が交換される可能性が高くなると考えられる。
- 移住間隔：移住が行われる世代間隔のこと。移住間隔が長いほど、島間での個体情報の交換が行われにくくなるために、各島で異なった探索が進む可能性が高くなると考えられる。
- 移住個体の選択方法：移住個体の選択方法。移住個体の選択方法はさらに、ある島から移住する個体の抽出方法と、他の島から移住してきた個体の挿入方法に分かれる。
- 移住操作の位置：移住操作をどの遺伝的操作の後に行うか。

3 パラメータの分類

DGA のパラメータは多く、すべてのパラメータの組み合わせを調べることは現実的ではない。しかし、各パラメータを独立に変化させる方法ではパラメータ間の依存関係を把握することが困難となる。このため本研究では、分散遺伝的アルゴリズムが選択・交叉・突然変異・移住の 4 つの遺伝的操作からなることに注目してパラメータを以下のように分類し、各分類に関してパラメータの比較を行った。

- 選択に関係するもの：個体数、島数、選択手法。
- 交叉に関係するもの：交叉手法、交叉率。
- 突然変異に関係するもの：突然変異手法、突然変異率。
- 移住頻度に関係するもの：移住率、移住間隔、島数。
- 移住方法に関係するもの：移住トポロジー、移住個体の抽出方法、移住操作の位置。

¹本研究では個体は各島に均等に配分されるため、個体数と島数により各島内の個体数が決定する。

4 対象問題

最適なパラメータ設定は対象問題によって異なると考えられるため、パラメータの検討は複数の対象問題について行う必要がある。本研究では以下の4種類のテスト関数について検討を行った。これらの関数はそれぞれに異なった性質を持っている。このため、これらの関数についてパラメータの検討を行うことで、対象問題の性質とパラメータの傾向との関係を知ることができると考えられる。本研究では、それぞれの関数の設計変数の数を10とし、1設計変数あたり10ビット用いてコーディングを行った。

- **Rastrigin**: 式(1)で表される関数で、設計変数間に依存関係がない。すべての設計変数の値が0の時最小値0をとり、その周辺に格子状に複数の準最適解を持つ。

$$f(x_1, \dots, x_n) = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)) \quad (1)$$

$$-5.12 < x_i \leq 5.12$$

- **Schwefel**: 式(2)で表される関数で、設計変数間に依存関係がない。すべての設計変数の値が412の時最小値(418.98276403 × 設計変数の数)をとる。本研究では、最小値が0となるように調整した関数を Schwefel 関数として用いた。

$$f(x_1, \dots, x_n) = \sum_{i=1}^n -x_i^2 \sin(\sqrt{|x_i|}) \quad (2)$$

$$-512 < x_i \leq 512$$

- **Ridge**: 式(3)で表される関数で、設計変数間に強い依存関係がある。すべての設計変数の値が0の時最小値0をとる。

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2 \quad (3)$$

$$-64 < x_i \leq 64$$

- **Griewank**: 式(4)で表される関数で、設計変数間に依存関係がある。すべての設計変数の値が0の時最小値0をとる。大域的には単峰性で依存関係が弱い、局所的に見ると多峰性で依存関係の強い関数である。

$$f(x_1, \dots, x_n) = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \left(\cos\left(\frac{x_i}{\sqrt{i}}\right) \right) \quad (4)$$

$$-512 < x_i \leq 512$$

各関数の2設計変数の場合の概観を Fig. 1 に示す。

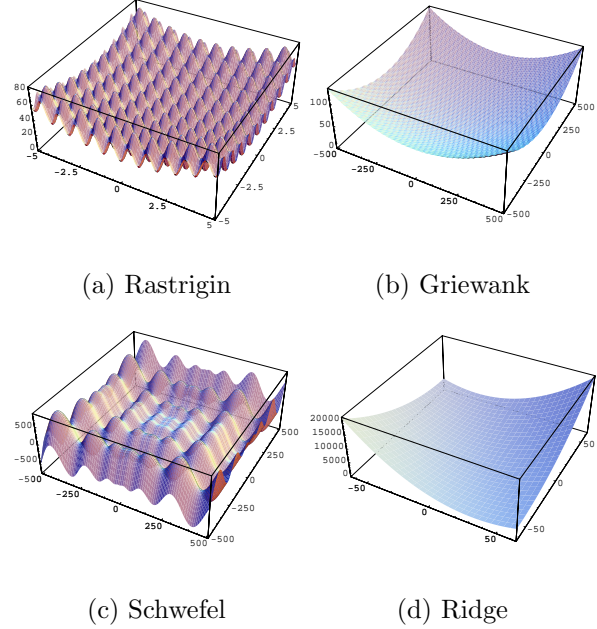


Fig. 1 対象問題の概観(2変数の場合)

5 結果

実験を行った結果、各パラメータが以下のように分類できることが分かった。

- 交叉手法, 交叉率, 突然変異手法, 突然変異率, 移住操作の位置, 鳥数は, 対象問題に依存せず同様の傾向を示す。
- 選択手法, 移住率, 移住間隔, 移住個体の選択方法, 移住トポロジーは Griewank 以外の関数について同様の傾向を示す。
- 最適な個体数は対象問題に依存する。

6 結論

本研究では、DGA についてパラメータの検討を行った。実験の結果、パラメータには対象問題に依存せず同様の傾向を示すものがあることが分かった。このことにより、今後のパラメータチューニングの煩雑性を減少させることができると考えられる。また、その他のパラメータに関しても対象問題の性質に応じた傾向があることが分かった。

7 今後の予定

今回実験を行っていない複数の対象問題についてもパラメータの検討を行い、本研究の結果を検証する必要がある。