

目的関数の加法分解性および差分の符号独立性にもとづく 実数値遺伝的アルゴリズムのリンケージ同定

手 塚 大 ††† 棟 朝 雅 晴 ††† 赤 間 清 †††

最適化問題を独立に最適化可能な複数の部分問題に分割可能な場合、各部分問題を独立に最適化することによって効率的に最適化ができる。遺伝的アルゴリズム (GA) による最適化では、この部分問題を構成する遺伝子座の集合をリンケージグループといい、リンケージグループを識別する手法をリンケージ同定という。本論文では、実数値 GA のリンケージとは何かを明確に定義する。この定義に基づいて直接的にリンケージの識別を行う二つのリンケージ同定手法、LINC-R と LIDI-R を提案する。LINC-R は目的関数の加法分解性、LIDI-R は差分の符号独立性にもとづいてリンケージの有無を判定する。これらの手法は直接的にリンケージを識別するため、効率的にリンケージ同定ができる。これらの手法を用いて、リンケージ同定を行い、リンケージグループごとに並列に最適化を行うことで、従来の手法よりも短時間でより良い解を得られる。

Linkage Identification for Real-Coded Genetic Algorithms based on Additive Decomposability and Difference Signature Independency of Objective Function

TEZUKA MASARU^{†,††}, MUNETOMO MASAHARU[†]
and AKAMA KIYOSHI[†]

In the case that a problem is decomposable to a number of sub-problems which can be optimized independently, the problem is solved effectively by optimizing sub-problems separately. In optimization problems by means of genetic algorithms, a set of loci of which each sub-problem consists is called linkage group. Linkage identification is the method which recognizes linkage groups. In this paper, we define the linkage of Real-Coded GAs clearly. Then we propose two new linkage identification methods, LINC-R and LIDI-R, directly based on the definition. LINC-R is based on additive decomposability of a objective function and LIDI-R is based on the independency of the signature of difference. These methods effectively identify linkages. Parallel optimization of the linkage groups has a capability to obtain better solutions in smaller number of function evaluations than conventional GAs.

1. はじめに

最適化問題では、もとの問題を複数の部分問題に分解できる場合、部分問題ごとに独立に最適化を行うことによって効率的な最適化ができる。

遺伝的アルゴリズム (GA) では部分問題の解であるビルディングブロックを交換することによって最適化が行われる。ビルディングブロックを構成する遺伝子のつながりをリンケージという。環状の染色体に特殊

な交叉を適用することで動的にリンケージを生成する LLGA¹⁾、集団内の個体の分布をもとに間接的にリンケージを同定する BOA²⁾、遺伝子座間の非線形性、非単調性、エピスタシス尺度をもとに直接的にリンケージを同定する LINC³⁾、LIMD⁴⁾、LIEM⁵⁾ など多くの研究がされている。

実数値 GA は実数のベクトルを染色体として持つ GA⁶⁾ で、様々な実数最適化問題への適用が研究されている。しかし、実数値 GA のリンケージ同定の研究は Tsutsui らによる PICI (Piecewise Interval Correlation by Iteration)⁷⁾⁸⁾ が報告されているが、まだ数が少ない。実数値を扱う大規模な実用問題を実数値 GA で効率的に解くためには、実数値問題のリンケージ同定が重要となる。

† 北海道大学大学院工学研究科
Graduate School of Engineering, Hokkaido University
†† (株) 日立東日本ソリューションズ
Hitachi East Japan Solutions, Ltd.
††† 北海道大学情報基盤センター
Information Initiative Center, Hokkaido University

2. 新しいリンケージ同定手法の提案

2.1 リンケージの定義

ベクトル $x \in \mathbb{R}^n$ とし、目的関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ の最適化問題を考える。実数値 GA ではこのベクトルが染色体となる。x の i 番目の要素 x_i が遺伝子の値であり、 i が GA の遺伝子座である。 $x_i \in X_i, i = 1, \dots, n$ とし、 $X_i \subseteq \mathbb{R}$ は定義域である。以下では二つの遺伝子座に着目し、これ以外の遺伝子座は省略する。

ある遺伝子座 j の値を x_j とした場合に、 f が最大値をとる遺伝子座 i の値 x_i の集合 $X_i^*(x_j) \subseteq X_i$ は次式のように表される。

$\forall x_i^1, x_i^2 \in X_i^*(x_j), f(x_i^1, x_j) = f(x_i^2, x_j)$ (1)
であり、かつ

$$\begin{aligned} & [X_i \setminus X_i^*(x_j) = \emptyset] \vee \\ & [\forall x_i^* \in X_i^*(x_j), \forall x_i \in X_i \setminus X_i^*(x_j), \\ & \quad f(x_i^*, x_j) > f(x_i, x_j)] \end{aligned} \quad (2)$$

ここで、次式 (3) が成立する場合、遺伝子座 i と遺伝子座 j について別々に最大化し、その値を組み合わせることで f を最大化できる。

$$\begin{aligned} & [\forall x_j^1, x_j^2 \in X_j, X_i^*(x_j^1) = X_i^*(x_j^2)] \\ & \wedge [\forall x_i^1, x_i^2 \in X_i, X_j^*(x_i^1) = X_j^*(x_i^2)] \end{aligned} \quad (3)$$

この式が成立しない場合には、遺伝子座 i と j について同時に f を最大化する必要がある。

このような、値の組合せを同時に最適化しないと解が得られないような遺伝子座の組を、実数値 GA のリンケージと定義する。他の定義と区別する場合には独立最適化可能の意味でのリンケージということにする。

2.2 LINC-R

式 (3) を満たす最も簡単な状況は、目的関数 f が式 (4) のように同じ引数を持たない部分関数に加法的に分解でき、二つの遺伝子座 x_i と x_j が異なる部分関数に属する場合、すなわち加法分解性の意味でリンケージがない場合である。

$$f(x_i, x_j) = g_1(x_i) + g_2(x_j) \quad (4)$$

加法分解性の意味でリンケージが無い場合、ある遺伝子座についての目的関数の差分係数は、もう一方の遺伝子座の値によらない。つまり、 $\forall x_i \in X_i, \forall x_j^1, x_j^2 \in X_j$ について

$$\Delta f(x_i, x_j^1) / \Delta x_i = \Delta f(x_i, x_j^2) / \Delta x_i \quad (5)$$

である。

この関係を用いてリンケージ同定を行う方法として LINC-R を提案する。まず基準となる点 $(x_i^1, x_j^1), x_i^1 \in X_i, x_j^1 \in X_j$ をランダムに選び、変量 Δx_i と Δx_j で

変位させ、 $x_i^2 = x_i^1 + \Delta x_i, x_j^2 = x_j^1 + \Delta x_j$ とする。式 (6) を満たす場合に遺伝子座 i と j にリンケージがあると判断する。 ϵ は許容誤差のパラメータである。

$$\begin{aligned} & \left| \left\{ f(x_i^2, x_j^1) - f(x_i^1, x_j^1) \right\} \right. \\ & \quad \left. - \left\{ f(x_i^2, x_j^2) - f(x_i^1, x_j^2) \right\} \right| > \epsilon \end{aligned} \quad (6)$$

2.3 LIDI-R

遺伝子座の間に加法分解性の意味でリンケージがある場合でも、各遺伝子座を独立に最適化できる場合がある。遺伝子座 i の任意の 2 点間の目的関数の大小関係が、遺伝子座 j の値によらない場合、 x_i は x_j によらず f を最大化できる。すなわち、

$$\begin{aligned} & \forall x_i^1, x_i^2 \in X_i, \forall x_j^1, x_j^2 \in X_j, \\ & \quad \text{sgn}(f(x_i^1, x_j^1) - f(x_i^2, x_j^1)) \\ & \quad = \text{sgn}(f(x_i^1, x_j^2) - f(x_i^2, x_j^2)) \\ & \Rightarrow \forall x_j^1, x_j^2 \in X_j, X_i^*(x_j^1) = X_i^*(x_j^2) \end{aligned} \quad (7)$$

である。ここで $\text{sgn}(x)$ は x が正、0、負の時にそれぞれ 1, 0, -1 となる符号関数である。

式 (7) が二つの遺伝子座の間で相互に成り立つ場合に式 (3) が成立し、各遺伝子座を個別に最適化できる。この関係を用いて定義するリンケージを、目的関数の差分の符号独立性の意味でのリンケージとよぶことにする。

これを同定する手法として LIDI-R を提案する。 $x_i^1, x_i^2 \in X_i, x_j^1, x_j^2 \in X_j$ について式 (8) が成り立たない場合にリンケージがあると判断する。

$$\begin{aligned} & \left[\text{sgn}(f(x_i^1, x_j^1) - f(x_i^2, x_j^1)) \right. \\ & \quad \left. = \text{sgn}(f(x_i^1, x_j^2) - f(x_i^2, x_j^2)) \right] \\ & \wedge \left[\text{sgn}(f(x_i^1, x_j^1) - f(x_i^1, x_j^2)) \right. \\ & \quad \left. = \text{sgn}(f(x_i^2, x_j^1) - f(x_i^2, x_j^2)) \right] \end{aligned} \quad (8)$$

次式の f_{Sp^2} は加法的には分解ができないが、差分の符号独立性がある関数の例である。

$$f_{\text{Sp}^2}(x_1, \dots, x_n) = \left(\sum_{i=1}^n (x_i - 1)^2 \right)^2 \quad (9)$$

差分の符号独立性の意味でのリンケージは、独立最適化可能の意味でのリンケージに含まれる。また、加法独立性の意味でのリンケージは差分の符号独立性の意味でのリンケージに含まれる。

3. 最適化能力の数値実験

LINC-R と LIDI-R によるリンケージ同定を用いて

最適化を行う数値実験を行った。最適化は二つのステージに分かれ、リンケージ同定ステージでは LINC-R または LIDI-R によってリンケージ同定を行い、その後、最適化ステージで並列 GA によりリンケージグループごとに最適化を行う。リンケージ同定ステージは目的関数の評価回数が 100,000 回になるまで実行し、その段階で同定できたリンケージグループを用いた。

目的関数として以下の F_1, F_2, F_3 を用いた。 F_1, F_2 は Tsutsui ら⁷⁾⁸⁾ が用いたものと同じ関数である。

$$F_1(\mathbf{x}) = F_{\text{Rsn}}(x_1, \dots, x_T) + F_{\text{Sp}}(x_{T+1}, \dots, x_{T+L}) \quad (10)$$

$$F_2(\mathbf{x}) = \sum_{i=1}^T F_{\text{Rsn}}(x_{2i-1}, x_{2i}) + F_{\text{Sp}}(x_{2T+1}, \dots, x_{2T+L}) \quad (11)$$

$$F_3(\mathbf{x}) = \sum_{i=1}^T F_{\text{Rsn}}(x_{2i-1}, x_{2i}) + \sum_{i=T+1}^{2T} F_{\text{Sp}^2}(x_{2i-1}, x_{2i}) \quad (12)$$

F_{Rsn} は式 (13) で表される T 次元 Rosenbrock 関数で、第 1 引数と他の引数との間にリンケージがある。また、 F_{Sp} は次の式 (14) で表される L 次元 Sphere 関数で、引数間にリンケージはない。

$$f_{\text{Rsn}}(x_1, \dots, x_n) = \sum_{i=2}^n \left(100(x_1 - x_i^2)^2 + (x_i - 1)^2 \right) \quad (13)$$

$$F_{\text{Sp}}(x_1, \dots, x_n) = \sum_{i=1}^n (x_i - 1)^2 \quad (14)$$

f_{Sp^2} は式 (9) の加法的には分解できないが差分の符号独立性がある関数である。

$L = 20$ に固定し、 $T = 2, \dots, 8$ について最適化を行った。実験結果を表 1, 表 2, 表 3 に示す。それぞれ 100 試行を実行し、最適解に到達した試行の数を #Opt に示した。MNE と STD は最適解に到達した試行での、最適解を得るまでの目的関数の評価回数の平均値および標準偏差である。評価回数にはリンケージ同定に要した評価回数も含んでいる。#LI は、LINC-R では加法分解性の意味で、LIDI-R では差分符号の独立性の意味で正確にリンケージ同定できた試行数である。

表 1, 表 2 にはリンケージ情報を使わない GA と PICI による結果も記載した。これらは Tsutsui⁸⁾ らの報告の値で、全 10 試行のものとなっている。表 3 にはリンケージ情報を使わない通常の GA の結果も記

載した。これは全 100 試行の結果である。

LINC-R を用いた GA では F_1 と F_2 の全てのリンケージを正確に同定し、全試行で最適解を得ている。また、従来手法である PICI よりも短い時間で最適解を得ている。LIDI-R はリンケージが長くなると同定能力が低下しているが、 F_2 のように短いリンケージが多数ある問題の場合には高精度でのリンケージ同定が可能であり、すべての試行で最適解を得ている。

F_3 問題の F_{Sp^2} のように、差分の符号独立性はあるが加法的に分解できないような場合、独立に最適化可能な遺伝子座であっても LINC-R ではリンケージ有りと判断されてしまう。表 3 が示すように、このような問題では LIDI-R を用いてリンケージ同定を行ってから最適化をするほうがよい。

4. おわりに

本論文では、実数値 GA におけるリンケージを、遺伝子座間の独立最適化可能の意味でのリンケージとして明確に定義した。さらに、この定義をもとに直接的にリンケージ同定を行う LINC-R と LIDI-R を提案した。LINC-R は加法分解性をもとに、LIDI-R は目的関数の差分の符号独立性をもとにリンケージ同定を行う。LINC-R のほうが少ない計算時間でリンケージ同定が可能だが、符号の差分独立性の意味でリンケージが無い場合にも、リンケージがあると誤判定してしまう。そのような問題の場合には LIDI-R を用いるほうが良い。

NFL 定理⁹⁾ によって、どのような解探索アルゴリズムでも、全ての問題に対する期待性能は等しいことが示されている。この定理は、対象とする問題群だけに高性能なアルゴリズムを設計することが可能であることも示唆している。実用上は、現実世界に広く存在するような問題群を効率的に解けることが重要である。本論文で報告した LINC-R は目的関数が加法的に分解できるような問題を、LIDI-R は目的関数の差分の符号独立性を持つような問題を効率的に解ける方法である。リンケージグループごとに探索することで、探索空間を大幅に小さくすることが可能である。実問題にはこのように複数の部分問題に分解できる場合が多くあると考える。

参考文献

- 1) Harik, G. R. and Goldberg, D. E.: Learning Linkage, *Foundations of Genetic Algorithms IV* (Belew, R. K. and Vose, M. D.(eds.)), Morgan Kaufmann, San Francisco, CA, pp.247-262

表 1 F_1 関数の最適化結果Table 1 Result of optimization of F_1 function.

T	リンケージ情報無し			PICI(LIMS)		LINC-R				LIDI-R			
	#Opt	MNE		#Opt	MNE	#Opt	MNE	STD	#LI	#Opt	MNE	STD	#LI
2	10/10	200,033		10/10	172,420	100/100	150,255	0.0	100	100/100	150,255	0.0	100
3	3/10	452,392		10/10	204,837	100/100	150,555	1,698	100	100/100	150,555	1,698	100
4	1/10	441,632		10/10	222,771	100/100	156,956	9,645	100	91/100	156,270	7,710	91
5	0/10	-		10/10	246,794	100/100	191,190	13,014	100	73/100	190,960	13,472	73
6	0/10	-		10/10	266,844	100/100	214,471	12,576	100	39/100	215,279	15,081	39
7	0/10	-		10/10	287,307	100/100	256,898	12,621	100	15/100	259,032	14,555	15
8	0/10	-		10/10	326,890	100/100	303,312	13,822	100	4/100	368,384	111,381	4

表 2 F_2 関数の最適化結果Table 2 Result of optimization of F_2 function.

T	リンケージ情報無し			PICI(LIMS)		LINC-R				LIDI-R			
	#Opt	MNE		#Opt	MNE	#Opt	MNE	STD	#LI	#Opt	MNE	STD	#LI
2	0/10	-		10/10	205,808	100/100	150,296	0.1	100	100/100	150,296	0.1	100
3	0/10	-		10/10	252,105	100/100	153,130	11,587	100	100/100	153,130	11,587	100
4	0/10	-		10/10	285,725	100/100	151,752	8,345	100	100/100	152,790	10,465	100
5						100/100	153,979	12,115	100	100/100	153,979	12,115	100
6						100/100	153,700	10,686	100	100/100	153,700	10,686	100
7						100/100	158,919	26,266	100	100/100	163,774	24,011	100
8						100/100	165,250	22,624	100	100/100	165,260	22,624	100

表 3 F_3 関数の最適化結果Table 3 Result of optimization of F_3 function.

T	リンケージ情報無し			LINC-R				LIDI-R			
	#Opt	MNE	STD	#Opt	MNE	STD	#LI	#Opt	MNE	STD	#LI
2	100/100	166,616	12,888	100/100	152,471	13,537	100	100/100	150,137	0.5	100
3	100/100	511,416	26,750	100/100	208,675	24,388	100	100/100	150,384	1,646	100
4	12/100	975,700	21,377	100/100	329,477	68,759	100	100/100	150,567	2,950	100
5	0/100	-	-	95/100	525,286	114,240	100	100/100	151,300	4,285	100
6	0/100	-	-	76/100	741,418	70,657	100	100/100	156,441	22,103	100
7	0/100	-	-	3/100	930,445	25,517	100	100/100	160,450	32,692	100
8	0/100	-	-	0/100	-	-	100	100/100	157,266	15,892	100

- (1996).
- Pelikan, M., Goldberg, D. E. and Cantú-Paz, E.: BOA: The Bayesian Optimization Algorithm, *Proceedings of the 1999 Genetic and Evolutionary Computation Conference*, pp. 525-532 (1999).
 - Munetomo, M. and Goldberg, D.E.: Designing a genetic algorithm using the linkage identification by nonlinearity check, *IlligAL Technical Report* (1998).
 - Munetomo, M. and Goldberg, D. E.: Linkage Identification by Non-monotonicity Detection for Overlapping functions, *Evolutionary Computation*, Vol. 7, No. 4, pp. 377-398 (1999).
 - 棟朝雅晴: エピスタシス尺度に基づくリンケージ同定手法の提案, 情報処理学会論文誌「数理モデル化と応用」, Vol. 43, No. SIG10(TOM7), pp. 6-13 (2002).
 - Fogel, D. B.: Real-valued vectors, *Handbook of Evolutionary Computation* (Bäck, T., Fogel, D. B. and Michalewicz, Z.(eds.)), Institute of Physics Publishing and Oxford University Press, pp. C1.3:1-1 (1997).
 - Tsutsui, S., Goldberg, D. E. and Sastry, K.: Linkage Learning in Real-Coded GAs with Simplex Crossover, *Proceedings of the 5th Intl. Conf. on Artificial Evolution*, pp. 73-84 (2001).
 - Tsutsui, S. and Goldberg, D. E.: Simplex Crossover and Linkage Identification: Single-Stage Evolution VS. Multi-Stage Evolution, *Proceedings of 2002 Congress on Evolutionary Computation*, pp. 974-979 (2002).
 - Wolpert, D. H. and Macready, W. G.: No Free Lunch Theorems for Optimization, *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, pp. 67-82 (1997).