

東京都立大学大学院 理工学研究科 電気電子工学専攻
博士(工学) 学位論文 公聴会

変換不変性とパラメータ調整に
基づくメタヒューリスティクスの
設計に関する研究



TOKYO METROPOLITAN UNIVERSITY
東京都立大学

熊谷 渉

- (1) 電気電子工学専攻 博士後期課程
(2) 横河電機株式会社

2020年8月1日

略歴

■ 熊谷 渉 (29歳)

■ 学歴

- ◆ 2010年3月 北海道北広島高等学校 卒業
- ◆ 2014年3月 茨城大学 工学部 機械工学科 卒業
- ◆ 2016年3月 首都大学東京大学院 理工学研究科 電気電子工学専攻
博士前期課程 修了
- ◆ 2017年10月 首都大学東京大学院 同研究科 博士後期課程 入学
(現 東京都立大学大学院)
- ◆ 2020年9月 東京都立大学大学院 同研究科 修了見込

■ 職歴

- ◆ 2016年4月 横河電機株式会社 入社、現在に至る

■ 業務

- ◆ AIによるモデリングや最適化などの数値計算を中心とする研究開発
- ◆ エネルギー・バイオ分野への応用・事業企画提案

目次

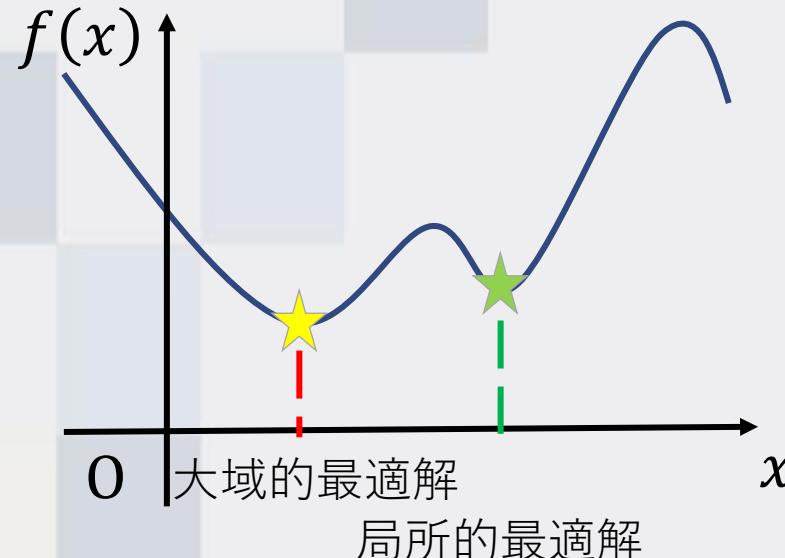
- 1 研究背景と目的** pp.4 – 17
ブラックボックス最適化とメタヒューリスティクス
本研究の位置付け
- 2 研究成果** pp.18 – 58
変換不変性に基づくメタヒューリスティクスのロバスト化(第4章)
パラメータ調整に基づくメタヒューリスティクスの適応化(第5章)
変換不変性を有するメタヒューリスティクスのフレームワーク
と総合比較(第6章)
- 3 本研究のまとめと今後の展望** pp.59 – 61
参考文献 pp.62 – 63

最適化問題と応用例

- 最適化は、ある制約条件の下で、多数の選択肢の中から、目的に対して最適な成果を出す選択肢（解）を選ぶ行為。
- 多くの工学の問題は、最適化問題として数理的に記述でき、最適化技術によって自動的に最適解を得る。

連続・単一目的・無制約最適化問題

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} f(x) \quad \text{式(2.1)}$$

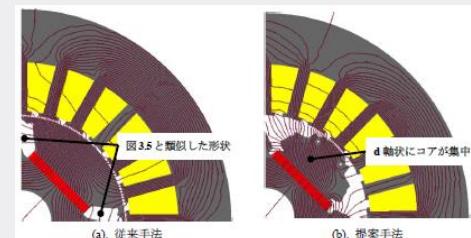


エネルギー設備運用[1]



蒸気量・電力、原料

製品の寸法・形状設計[2]



IPM型モータ

他に、機械学習の学習問題など

最適化技術を取り巻く環境の変化

4つの主要な要素が相互に影響を与えており、最適化技術を取り巻く環境は急速に変化している。

実システムの大規模さ・複雑さ

- ニーズの高度化・多様化に伴い、実システムの大規模化・複雑化が進む



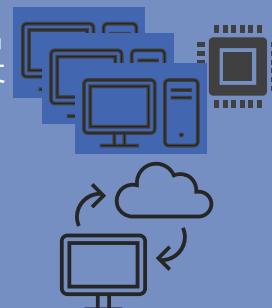
モデリング・シミュレーション技術

- 深層学習が牽引する第3次AIブームの到来（人間を超える威力を達成）
- 模擬による信頼性向上、多様性(特殊環境含む)の対応



コンピュータの計算能力

- 並列計算(HPC)の進展(GPGPU・超スカラ)
- クラウドサービスによる利便性向上



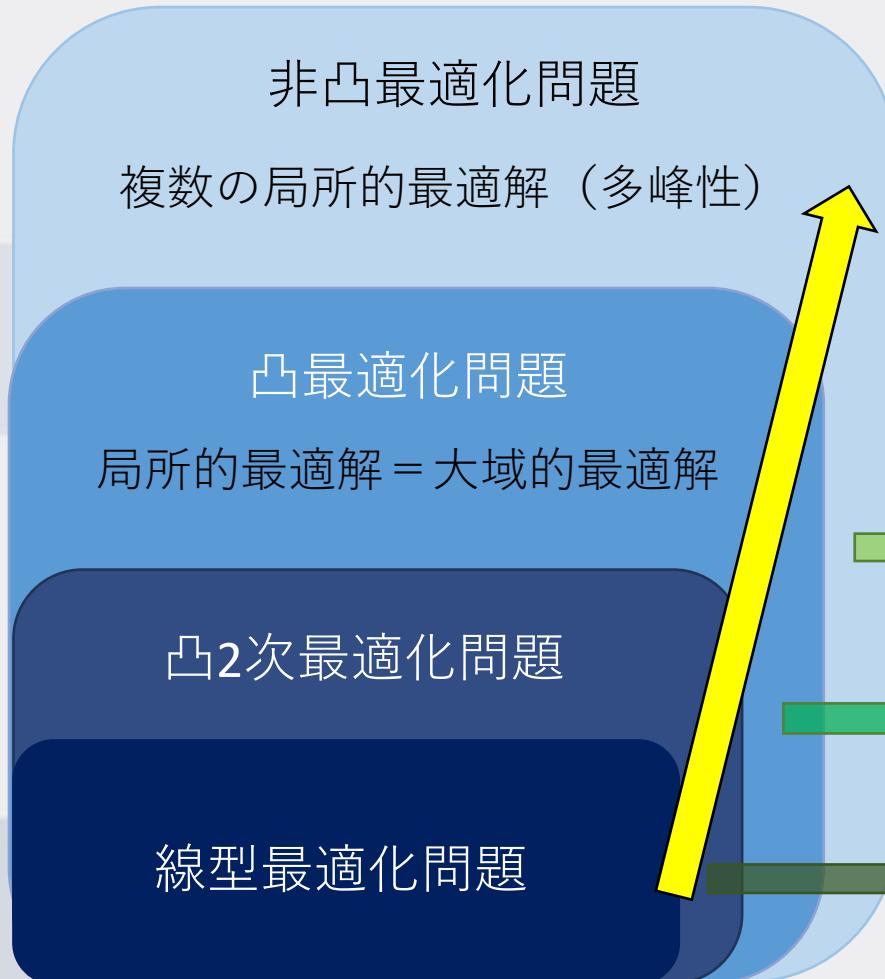
最適化理論・アルゴリズム

周囲技術・環境変化に対応可能な最適化技術が求められる

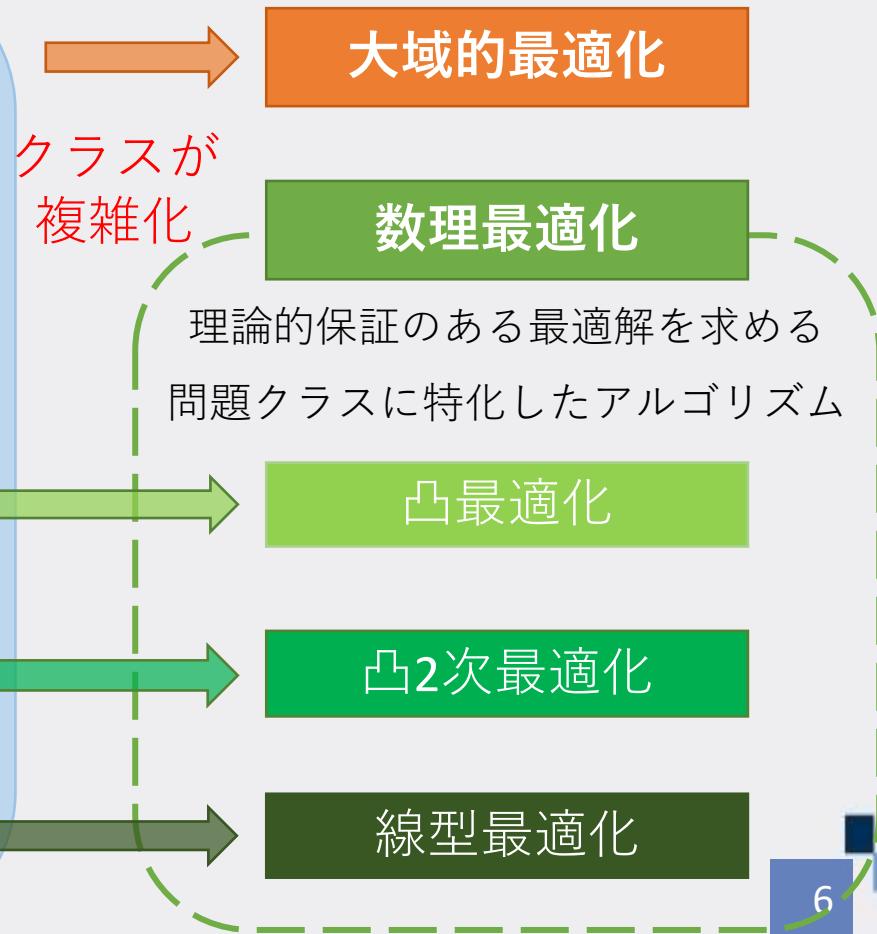
最適化問題クラスとアルゴリズム

問題クラスは複雑化が進むが、大域的最適化は発展途上である。

最適化問題のクラス



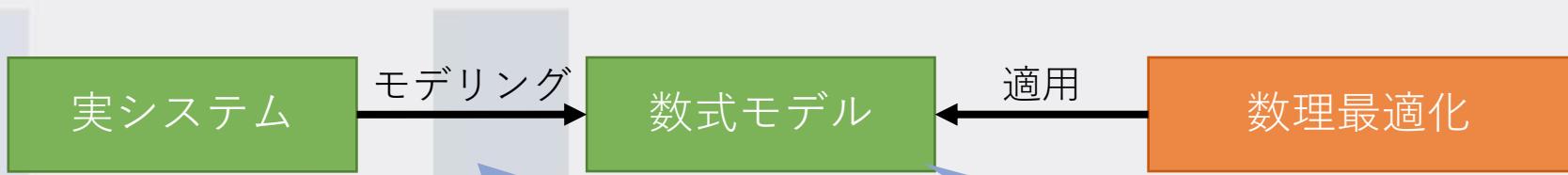
最適化アルゴリズムのパラダイム



数理最適化の課題

数理最適化を用いる場合、その適用範囲が最適化の形態に制約を課すという課題がある。

- ◆ 得た解の実行可能性や有用性が不十分となる
- ◆ 非凸に凸最適化を無理に適用しても、局所解に収束



モデルが数理最適化に対応するクラスになるため、実システムとの間に乖離が生じる

アルゴリズムに特化したクラスにモデリングする必要がある

→ モデルをアルゴリズム側に寄せすぎると、実態と合わない

→ モデルを実システム側に寄せすぎると、適用できない

ブラックボックス最適化 (BBO)

ブラックボックス関数の入出力情報のみを用いて、良好な解を自動的に獲得する問題。

ユーザは目的関数の景観
を知ることができない

目的関数値情報

$$f(x)$$

実システム

数式モデル

最適化アルゴリズム

決定変数値情報

$$x$$

陽な定式化が不可能、
困難なブラックボックス関数

直接探索法 ○



数理最適化 ✗



- 大規模・複雑なシミュレータからの性能評価データを使用
- 実システムからの直接的な計測データを使用

※逐次近似最適化（逐次サンプリング＆モデリング）のアプローチもあるが、代理モデルが多峰性となるため、大域的最適化が必要

BBOの例

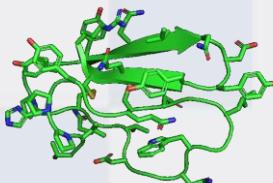
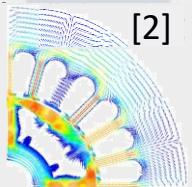
BBOは、モデリング・シミュレーション技術と連携するケースで頻出する。

製品設計案、MLの上位パラメータ、タンパク質構造、素材、実験条件



設計案

電磁界・空力特性、MLアルゴリズム性能、タンパク質の構造的安定性・結合性



上位の最適化



下位のシミュレート

- シミュレーションと連動した製品設計[2]
- 機械学習のハイパープラメータ探索[3]
- **Material/Bio Informatics**による新素材創出・創薬支援[4,5]

コンピュータ、モデリング・シミュレーションなどの周辺技術の発展を背景に、BBO形態はますます増えていくと予想

→ 超スマート社会・Bio-Economyの実現

メタヒューリスティクス

- 数学的な保証がないが、生物の進化過程や群知能などの経験則をアナロジーとして構築される発見的手法。
- 大域的最適化やBBOの有力なアルゴリズム群として考えられる。

直接探索法

発見的手法

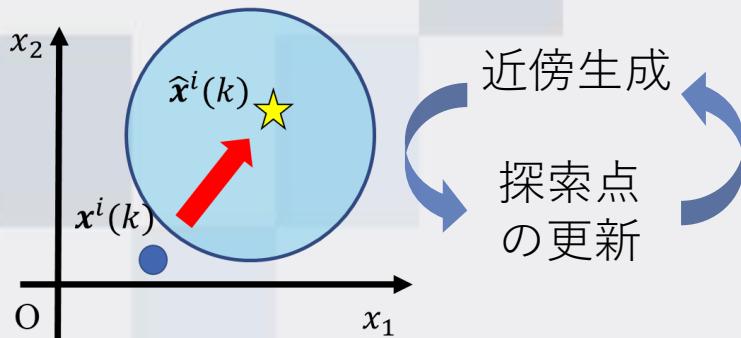
近似解法

Bio-Inspired

確率的解法

多点探索

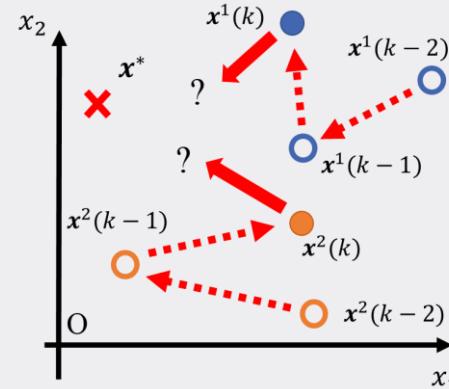
解空間内で探索点が移動する過程に確率的操作が含まれる



メリット

- 多峰性でも局所的最適解に陥りにくい
- 並列処理が適しており、その進化の恩恵を受けやすい

相互作用を及ぼす複数の探索点が群として移動していく

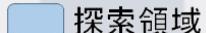


メタヒューリスティクスの最適化能力

近接最適性原理(POP)と多様化・集中化が最適化能力を生み出しており、BBOにおける探索戦略となる。

POP (近接最適性原理)

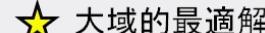
- 「良い解同士には類似構造がある」という経験則
- 良い解を活用すれば問題の有望な領域を探索できる

 探索領域

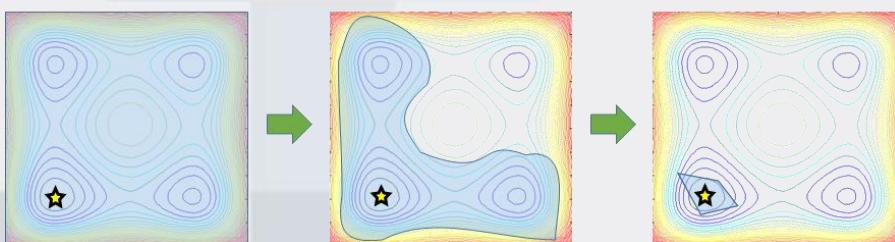
探索序盤

 等高線

探索中盤

 大域的最適解

探索終盤



多様化・集中化

- 多様化：広い範囲を探索し、解空間の偏り構造を把握する
- 集中化：探索範囲を有望な領域に狭め、より優れた解を探索する
- 「探索序盤では多様化、探索終盤では集中化」という戦略が効率的な探索を可能にしている

BBOにおけるメタヒューリスティクスの課題

BBOではユーザがアルゴリズムやパラメータ値を適切に選択・設定できない。

- ◆ 基本性能やパラメータ値は問題の性質から影響を受ける



ユーザが最適化問題として定式化



BBO

事前に問題の情報
が得られない

開発者がアルゴリズムを開発

ユーザが問題に適したアルゴリズムを選択

ユーザが問題に適するようにアルゴリズム
のパラメータを設定

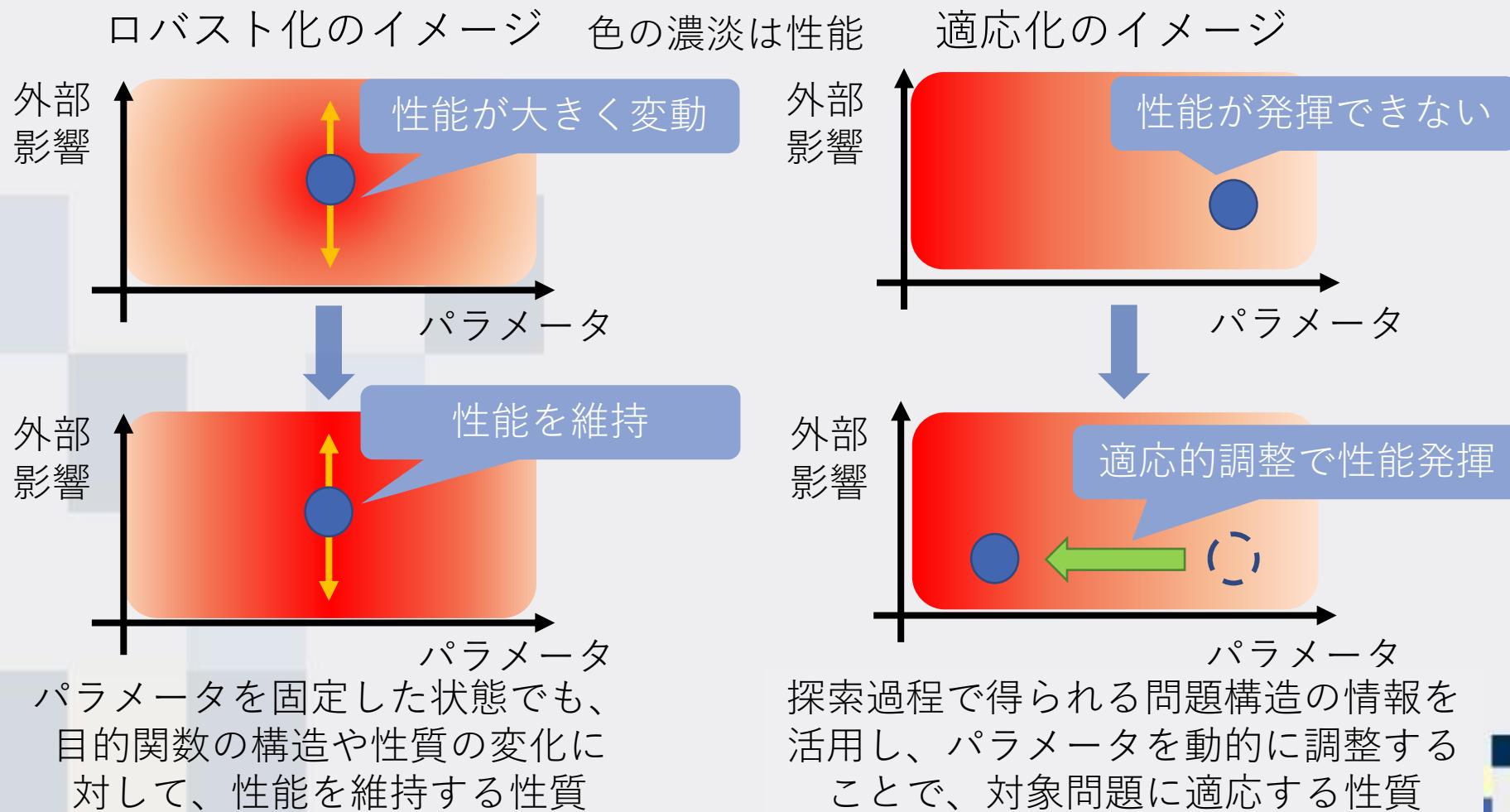
ユーザがアルゴリズムを適用し、得た解を吟味



手戻りが
発生しやすい
(非効率な試行錯誤)

最適化アルゴリズムのロバスト性・適応性

BBOを解くアルゴリズムは、高いロバスト性・適応性を具備すれば、少ない負荷で、満足な解が得られる。



本研究の目的とアプローチ

本研究の目的

BBOのための、高いロバスト性・適応性を具備する最適化アルゴリズムの設計



アプローチ

変換不変性とパラメータ調整に基づく設計指針を構築し、メタヒューリスティクスの設計を行う。

変換不変性

- メタヒューリスティクスに何らかの**理論的な保証**を持たせる
- 目的関数の性質変化に対する**ロバスト性**を向上

(適応的)パラメータ調整

- パラメータ設定に関する**負担軽減**
- 問題構造や次元数などの条件に対する**適応性(平均性能)**を向上

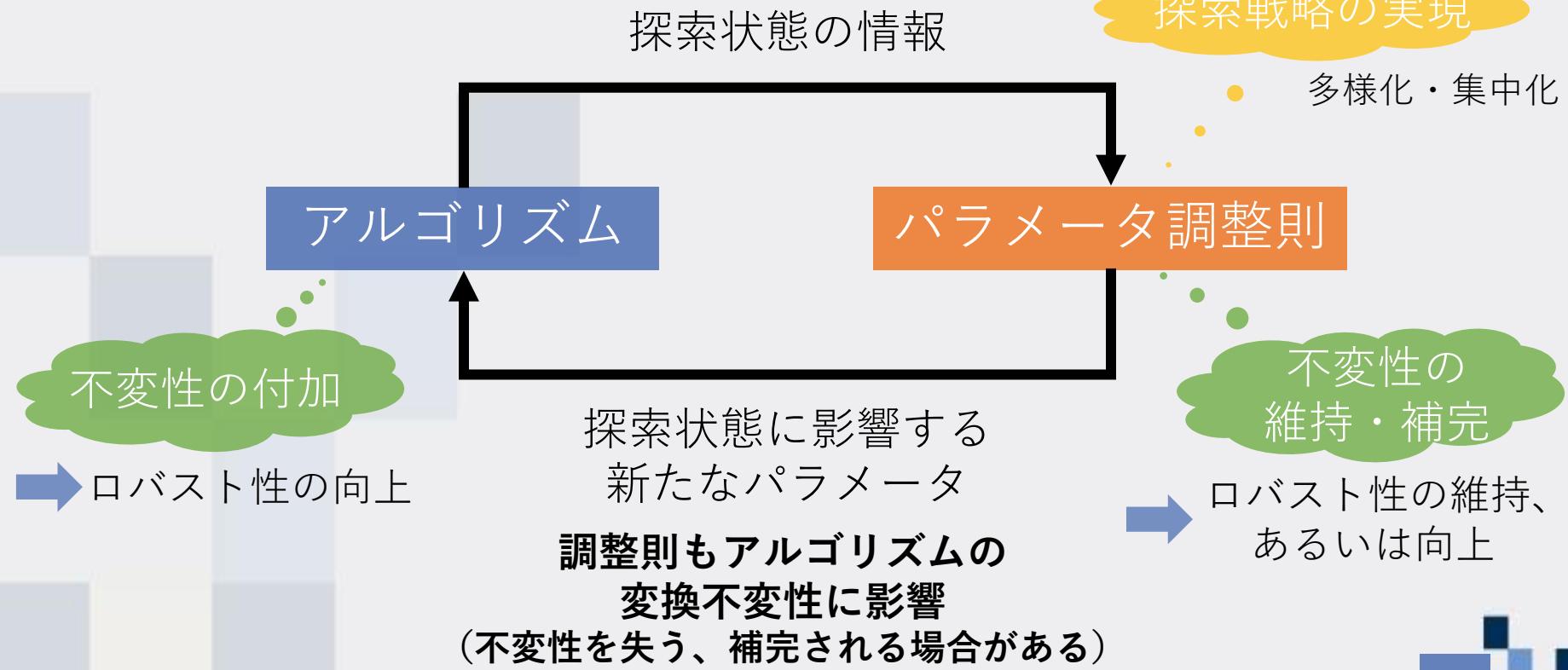
本研究のポイント

- 発見的手法に対して理論的な枠組みによるアルゴリズム設計。
- 変換不变性とパラメータ調整則の相互関係を考慮・活用したアルゴリズム全体の設計。

→ 適応性の向上

探索戦略の実現

多様化・集中化



本研究の構成

本発表2.1章 (学位論文 第4章)

不变性の解析方法確立

適用

不变性の付加と課題提起

課題解決策検討

本発表2.2章

適応的パラメータ調整

(学位論文 第5章)

本発表2.3章

不变性フレームワーク

(学位論文 第6章)

→ アルゴリズム新規設計

本研究の位置付け

網羅的・俯瞰した視点に基づくアプローチ

■ 変換不变性の汎用な解析方法を具体的に与え、既存手法の変換不变性の有無を網羅的に明らかにする。

- ◆ 定義は与えられているが、適用例は少数[6]
- ◆ 方法は与えられているが、特定の不变性に限定[7]
- ◆ 一部のアルゴリズムで明らかにされているが、特定の不变性に限定[7]

■ 不变性の付加アプローチを系統的に分類し、適用。

- ◆ 一部のアルゴリズムに対する不变性を付加するアプローチはあるが、俯瞰した視点で整理されておらず、その事例に限定[7,8,9]

■ 不变性に基づくパラメータ調整則とアルゴリズム設計。

- ◆ 一部のアルゴリズム・調整則は、不变性の有無が指摘されているが、その事例に限定 [6,10,11]

■ 不变性のフレームワークに従うアルゴリズム設計。

- ◆ 特定のアルゴリズムに対する不变性を付加する事例しか見当たらない [7,8,9]

変換不変性の定義

※パラメータや乱数などは省略

定義1：変換不変性の定義

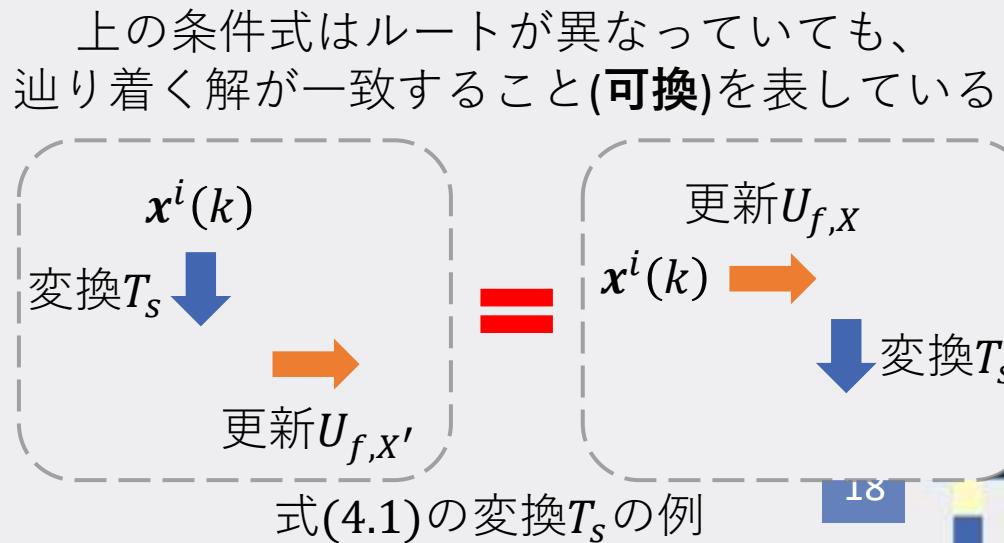
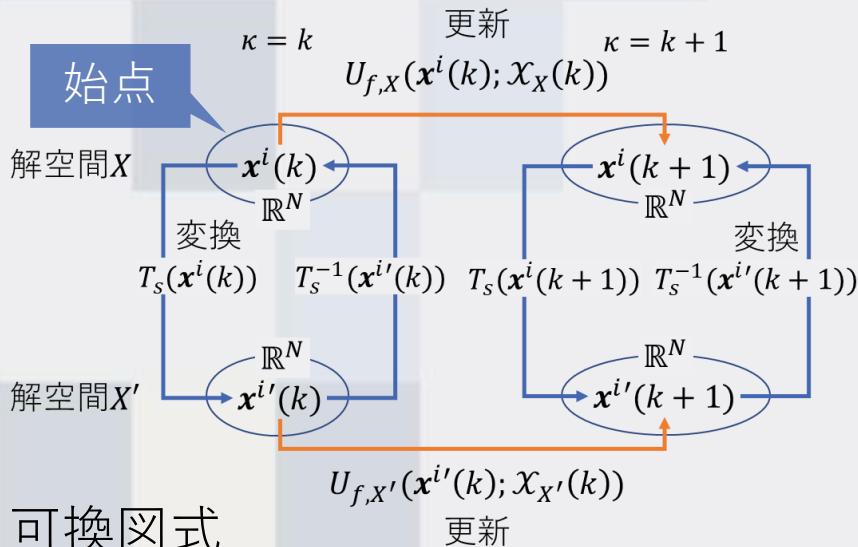
最適化アルゴリズムの更新則 $U_{f,X}$ がある変換 $T_s: \mathbb{R}^N \rightarrow \mathbb{R}^N$, $T_f: \mathbb{R} \rightarrow \mathbb{R}$ に
対して不变であるとは、式(4.1)が成立することだと定義する。

$$U_{T_f \circ f, X'}(T_s(\mathbf{x}^i(k)); \mathcal{X}_{X'}(k)) = T_s \circ U_{f, X}(\mathbf{x}^i(k); \mathcal{X}_X(k)), \forall k \in \mathbb{N} \quad \text{式(4.1)}$$

ただし、 $\mathcal{X}_X(k)$ は反復回数 k までの探索点群の履歴である。ある空間や
量 h がこれらの変換の影響を受けるとき、 h' と表す。

逆変換 T_s^{-1} の存在を仮定すると、式(4.1)は式(4.1)'と等価である。

$$U_{f, X}(\mathbf{x}^i(k); \mathcal{X}_X(k)) = T_s^{-1} \circ U_{T_f \circ f, X'}(T_s(\mathbf{x}^i(k)); \mathcal{X}_{X'}(k)), \forall k \in \mathbb{N} \quad \text{式(4.1)'}$$



解空間のアフィン変換不変性の意義

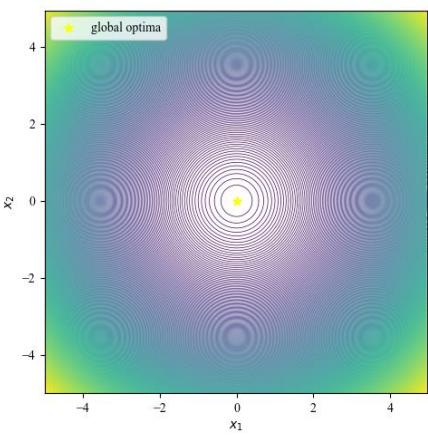
アフィン変換不変性を具備することは、下記を同等の効率で探索可能（＝ロバスト性が高い）。

アフィン変換 = 線形変換 + 平行移動



回転やスケール変換を包含

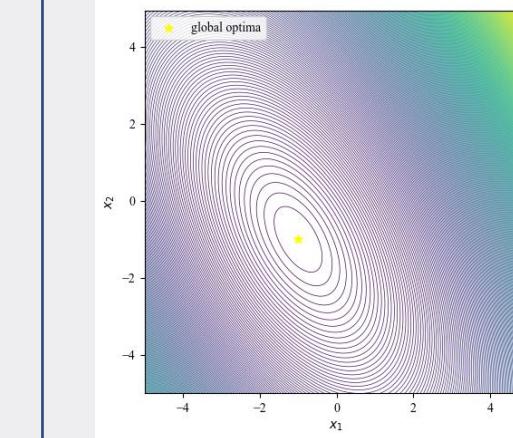
等高線が等方・直交



変数間独立性
良スケール性
原点中心(原点依存)

↓
最適化で容易な性質

等高線が異方・斜交



変数間依存性
悪スケール性
原点非中心

↓
最適化で困難・必須な性質

BBOではこれらの性質に対して、性能が依存しないことが望ましい

※最急降下法は線型変換不変性を有さないため、これらの性質に強く依存する

変換不変性の解析方法

下記の汎用な手順によって、不変性の有無を証明可能。

- 1 更新則 $U_{f,X}$ に含まれる量 h は探索履歴 $\mathcal{X}_X(k)$ に依存するため、特定の変換 T_s, T_f を与えたときに影響を受けた h' を導出。

Particle Swarm Optimization (PSO) の例

パラメータ $w, c_1, c_2 \in C; \mathbf{R}_1, \mathbf{R}_2 \in \mathcal{R}$

$$U_{f,X}(\mathbf{x}^i(k); \mathcal{X}_X(k)) = \mathbf{x}^i(k) + w(\mathbf{x}^i(k) - \mathbf{x}^i(k-1)) + c_1 \mathbf{R}_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{R}_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k))$$

$$\mathbf{x}^i(k), \mathbf{p}^i(k), \mathbf{p}^g(k) \in \mathcal{X}_X(k), \quad \mathbf{x}^{i'}(k) = T_s(\mathbf{x}^i(k)), \mathbf{p}^{i'}(k) = T_s(\mathbf{p}^i(k)), \mathbf{p}^{g'}(k) = T_s(\mathbf{p}^g(k))$$

- 2 不変性の条件式(4.1)に代入して変形し、等式が成立するかを確認。成立すれば与えた変換に対して不変性を有する。

$$U_{T_f \circ f, X'}(\mathbf{x}^{i'}(k); \mathcal{X}_{X'}(k)) = \text{変換した後、更新}$$

= ...

$$= T_s \circ U_{f,X}(\mathbf{x}^i(k); \mathcal{X}_X(k))$$

更新した後、変換

代入・変形

PSOの回転不変性への適用例

定理4

PSOの更新則 U_f は回転不変性を有さない。

証明

回転変換 $T_s: \mathbf{x} \mapsto \mathbf{Bx}$ に対して、

$\mathbf{x}^{i'}(k) = \mathbf{Bx}^i(k), \mathbf{p}^{i'}(k) = \mathbf{Bp}^i(k), \mathbf{p}^{g'}(k) = \mathbf{Bp}^g(k)$ となる。

不变性の条件式(4.1)は、

$$U_{T_f \circ f, X'}(\mathbf{x}^{i'}(k); \chi_{X'}(k)) \quad \text{変換した後、更新}$$

$$= \mathbf{Bx}^i(k) + w(\mathbf{Bx}^i(k) - \mathbf{Bx}^i(k-1)) + c_1 \mathbf{R}_1 (\mathbf{Bp}^i(k) - \mathbf{Bx}^i(k)) + c_2 \mathbf{R}_2 (\mathbf{Bp}^g(k) - \mathbf{Bx}^i(k))$$

$$= \mathbf{B}(\mathbf{x}^i(k) + w(\mathbf{x}^i(k) - \mathbf{x}^i(k-1))) + c_1 \mathbf{R}_1 \mathbf{B}(\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{R}_2 \mathbf{B}(\mathbf{p}^g(k) - \mathbf{x}^i(k))$$

$$\neq \mathbf{B}(\mathbf{x}^i(k) + w(\mathbf{x}^i(k) - \mathbf{x}^i(k-1))) + c_1 \mathbf{B} \mathbf{R}_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{B} \mathbf{R}_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k))$$

$$= \mathbf{B}\left(\mathbf{x}^i(k) + w(\mathbf{x}^i(k) - \mathbf{x}^i(k-1)) + c_1 \mathbf{R}_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{R}_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k))\right)$$

$$= \mathbf{Bx}^i(k+1)$$

$$= T_s \circ U_{f, X}(\mathbf{x}^i(k); \chi_X(k)) \quad \text{更新した後、変換}$$

となり成立しない。計算過程で $\mathbf{R}_\ell \mathbf{B} \neq \mathbf{B} \mathbf{R}_\ell (\ell = 1, 2)$ となることを利用している。
したがって、題意が示された。

各種アルゴリズムの変換不変性の有無

回転不変性か、スケール不変性を欠如することが多い。

アルゴリズム	解空間のアフィン変換不変性			目的関数の 単調変換不変性 ※3
	平行移動不变性	回転不变性	スケール不变性 (相似変換不变性)	
ES	○	○※1	×	○
実数値GA	○	○※1	相似○／スケール×	○(ランクの場合)
実数値SA	○	○※1	×	×
PSO	○	×	○	○
DE	○	×	○	○
ABC	○	×	○	×
Firefly Algorithm	○	○※1	×	○
Cuckoo Search	○	×	×	○

※1：標準正規分布が等方性のためである。

※2：赤字は本論文で新たに指摘した不变性。

※3：単調増加不变性は具備することが多く、具備させることが容易なため、重視しない

変換不変性の付加アプローチの分類

演算形式に合わせて付加アプローチを適切に選択すればよい。

演算形式	付加したい不変性	アプローチ
(1) 基本ベクトルと 差分ベクトルの線形結合	回転不変性	超球の導入 回転行列を係数に作用
(2) 解同士の組合せ	アフィン変換不変性	係数を単位行列とする※1
<hr/>		
(3) ランダムウォーク	回転不変性 スケール不変性 アフィン変換不変性	等方的な確率分布 探索履歴に依存する 回転行列を係数に作用 探索履歴に依存する 対角行列を係数に作用 探索履歴に依存する 回転行列と対角行列を作用

※1：「解の組合せ」で単位行列とすると、線分上に拘束され、多様性が極端に低下
 ※参考にした先行研究は付録に記載

本章のアプローチ

更新則の演算形式の分類と付加したい不変性の種類に従い、先行研究のアプローチを適切に流用することで、不変性を付加し、ロバスト性の向上を検討。

- ◆ 他のアルゴリズムのアプローチを流用しているが、対象のアルゴリズムでは新規性がある

本発表

Particle Swarm Optimization (PSO)

回転不変性を付加

Artificial Bee Colony Algorithm (ABC)

回転不変性を付加

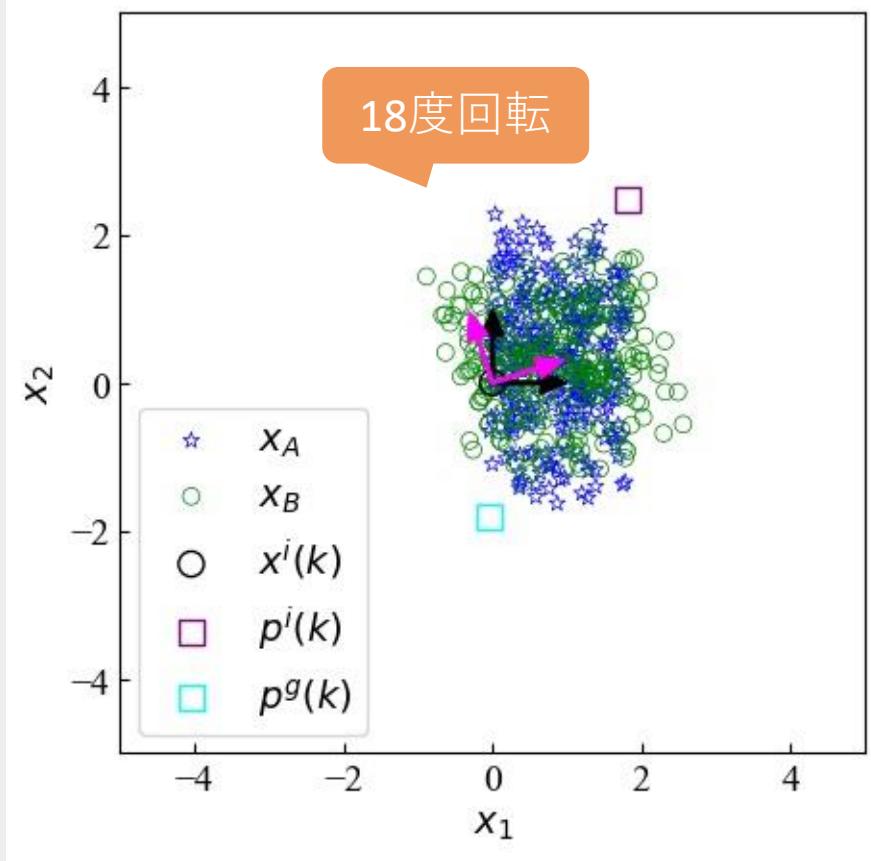
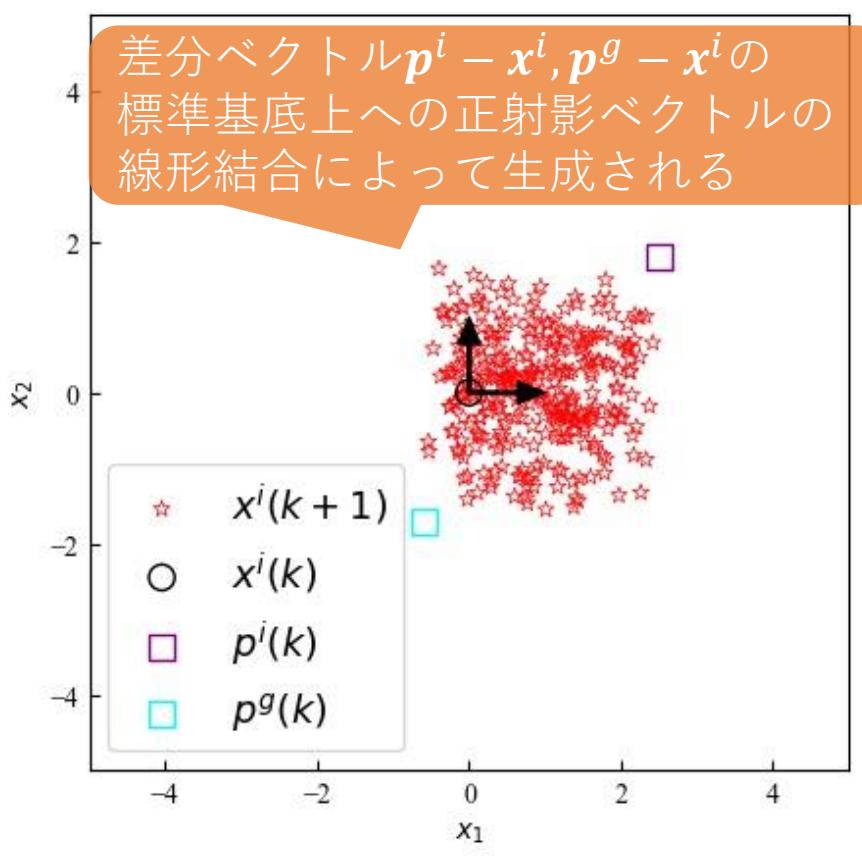
Cuckoo Search (CS)

アフィン変換不変性
を付加

PSOの回転不変性の欠如

300回生成

$x_A \neq x_B$ となるため、回転不変性を欠如している。



※1：慣性項は回転不変性を有するため、 $w = 0$ としている。

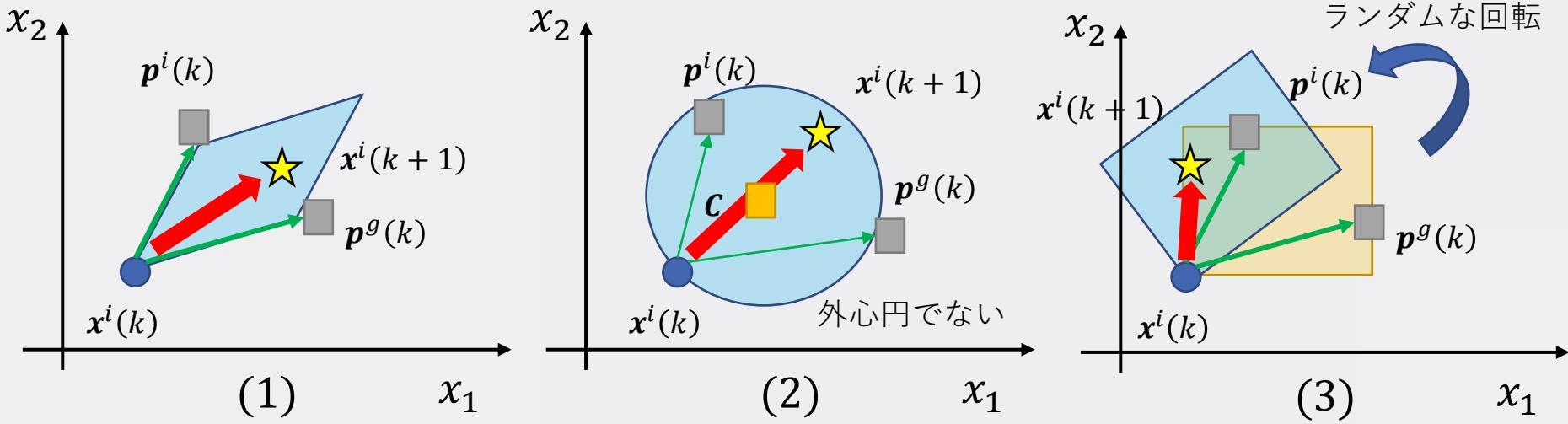
※2：上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

PSOの変換不变性に関する先行研究

形式毎にアプローチは存在するが、(3)はシンプルな事例のみ。

#	アプローチ	特徴	文献
(1)	線形結合の係数をスカラとする	アフィン変換不变性を獲得できるが、 摂動が長方形から平行四辺形となり、多様性が低下	[8]
(2)	超球の導入	探索点、p-best、g-bestの三点の重心を中心とした 超球内部に探索点を移動する（慣性項も加わる）	[9]
(3)	回転行列を線型 結合の係数に作用	ランダムな回転角度を与える回転行列により、 近傍を回転させることで、標準基底に依存しない	[7]

※簡略化のために、 $w = 0, c_1 = c_2 = 1$ として省いている



相関性を用いた回転不変性を有するPSO (CRI-PSO)

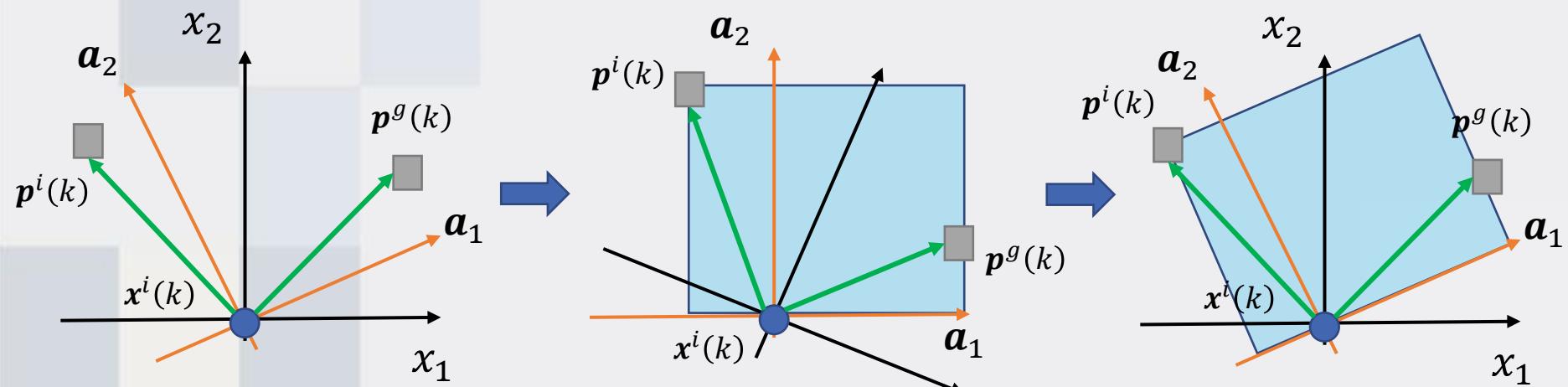
■ 探索履歴の解集合 \mathcal{Z} の共分散行列 Σ (相関)を利用[12,13]。

- ◆ 摂動を回転させることで、回転不変性を獲得
- ◆ 探索で得た解分布を利用してすることで、探索方向が有用

■ 探索点の更新は下記の通りである。 Σの固有ベクトル

- ◆ 基底変換行列 $P = [\mathbf{a}_1, \dots, \mathbf{a}_N]$ を生成し、基底変換する
- ◆ その座標系で近傍生成する
- ◆ 元の座標系へ変換し、その近傍内に探索点を移動させる

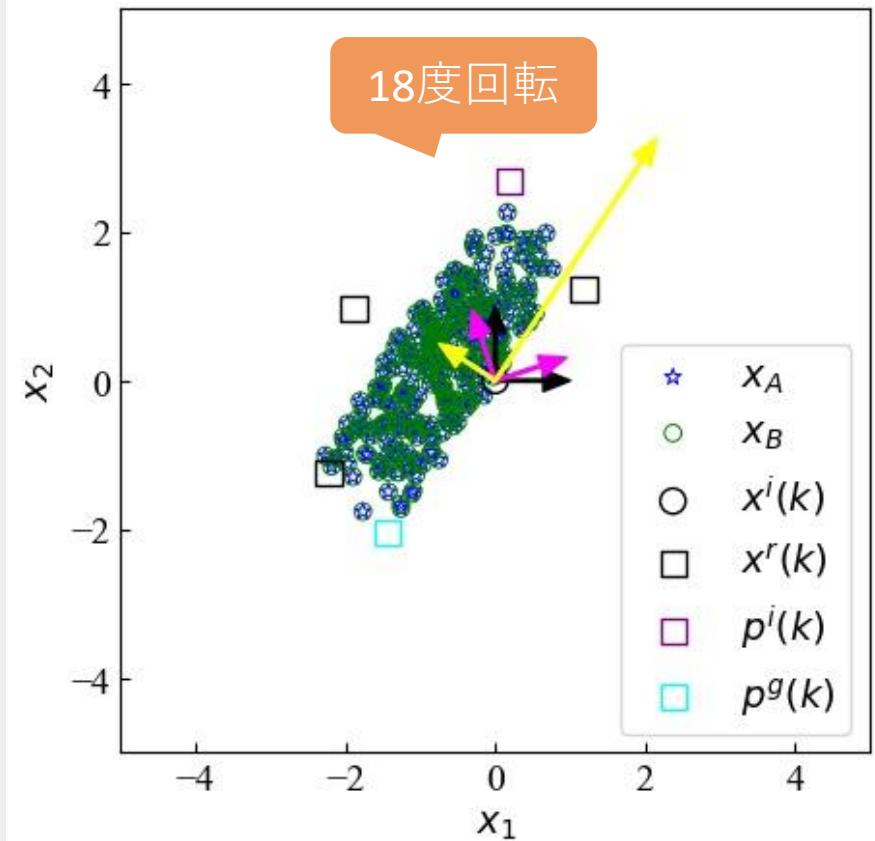
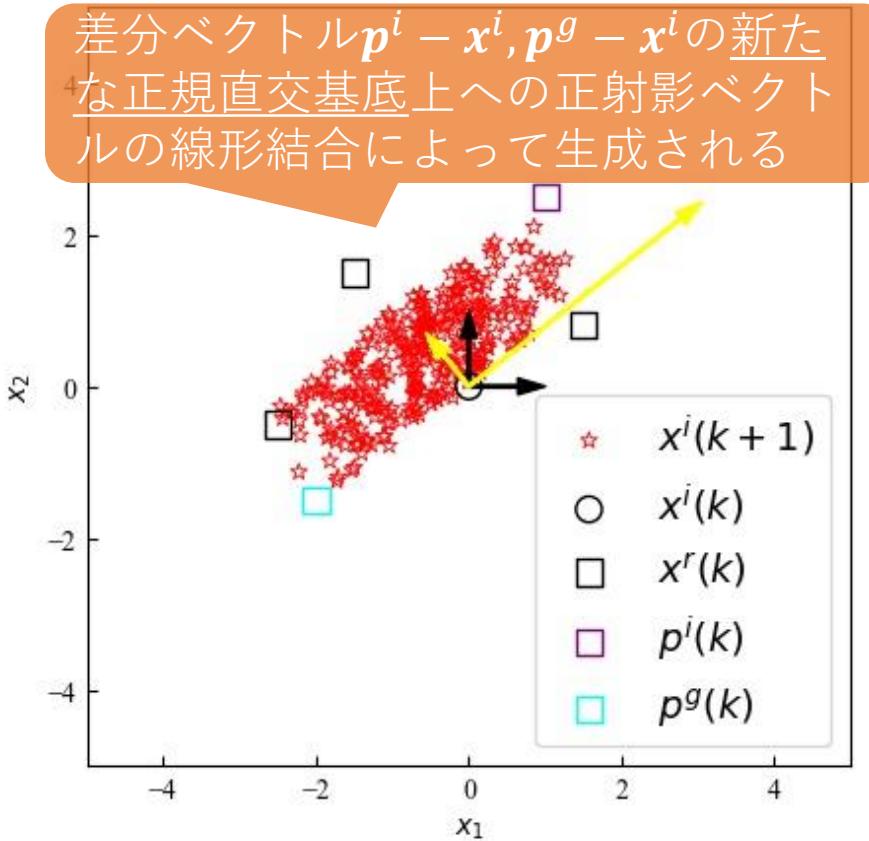
基底変換行列 P は標準基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ を別の正規直交基底 $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ に変換



CRI-PSOの回転不变性の具備

300回生成

$x_A = x_B$ となるため、回転不变性を具備する。



※1：慣性項は回転不变性を有するため、 $w = 0$ としている。

※2：上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

数値実験

■ 数値実験を通じて、CRI-PSOの性能を検証する。

- ◆ 実験1：回転に対するロバスト性に関する検証
- ◆ 実験2：全般的な性能に関する検証

■ 比較手法

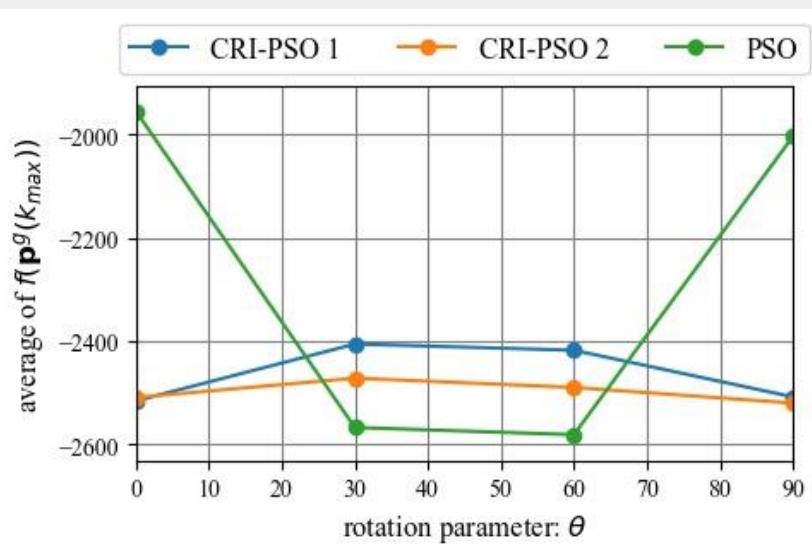
- ◆ PSO：回転不变性を欠如、Constriction Method (CM) [14]
- ◆ CRI-PSO 1：解集合 $Z = \mathcal{X}(k)$ (探索点群)
- ◆ CRI-PSO 2：解集合 $Z = \mathcal{P}(k)$ (p -best集合)

■ 実験1の検証方法

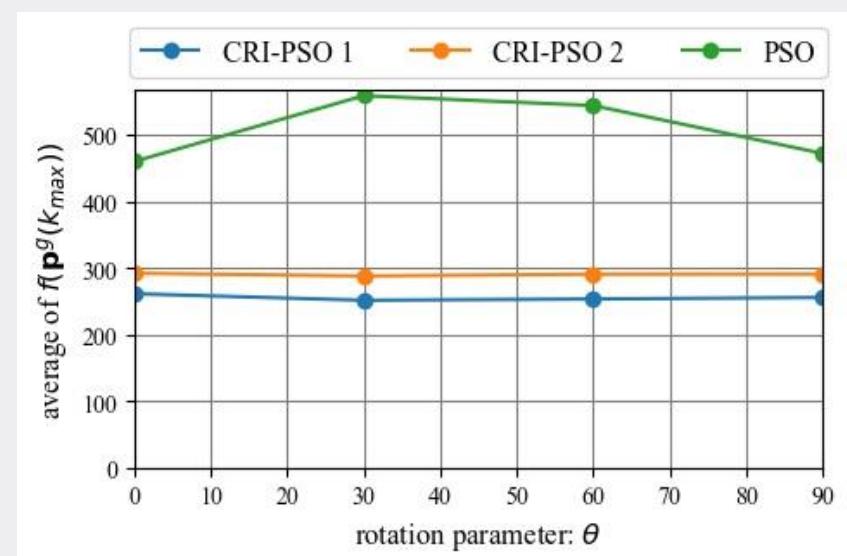
- ◆ 解空間に変換 T_s を与えるには、初期探索点を $T_s^{-1}(\mathbf{x}^i(1))$ とし、 $f(\mathbf{x}^i(k))$ を $f(T_s(\mathbf{x}^i(k)))$ とする。
- ◆ 解空間を θ deg 回転することで、変数間依存性を発生させ、性能の変動を確認する。
 - $\theta = 0, 30, 60, 90$ deg (関数の対称性により、90degは0degと等価)

実験1：回転に対するロバスト性検証結果

回転に対して、PSOはロバストでないが、CRI-PSOはロバストであることが確認できる。



2^N minima 関数



Rastrigin 関数

※ $N = 50, k = 1000$ の結果だが、他の条件でも同様の傾向が確認された

※異なる初期点から50試行、g-bestの目的関数値の平均値を性能とみなす

実験2：全般的性能に関する検証結果

■ 様々な関数や条件に対する全般的性能を比較する。

- ◆ 9種類の関数(单／多峰、変数間独立／依存、良／悪スケール)
- ◆ 次元数 $N = 10, 50, 100, 300$ 、反復回数 $k = 100, 1000$
- ◆ 72個の条件において、3種類手法の中で1位の数をカウント

■ 全般的にCRI-PSO 2が最も優れることを確認。

- ◆ 低次元かつ長期では、PSOが優れており、それ以外では、CRI-PSO 2が優れる (PSOのCMは高次元で性能が劣化[15])
- ◆ CRI-PSO 1の全般的な性能は最も劣る

手法	次元数 N				反復回数 k		All
	10	50	100	300	100	1000	
PSO	14	9	0	0	5	18	23
CRI-PSO 1	0	2	3	1	5	1	6
CRI-PSO 2	4	7	8	17	26	10	36

※異なる初期点から50試行、g-bestの目的関数値の平均値を性能とみなす

まとめと課題

■ PSOは回転変換に対してロバストでないが、CRI-PSOはロバストであることを確認。

- ◆ 他のアプローチを流用することで、回転不变性を付加でき、性能のロバスト性に影響を与えた

■ CRI-PSO 2がPSOとCRI-PSO 1よりも全般的に優れることを確認。

- ◆ PSOの探索点群は激しく動き回るが、p-best集団は有望領域に分布するため、より有望な探索方向を生成できた

■ ただし低次元かつ長期ではPSOが優れる[15]。

- ◆ CRI-PSOのパラメータ設定が短期に特化している(集中化が早い)と同時に、**設定の難しさを認識**

■ CRI-PSOはスケール不变性を失っている。

- ◆ 回転不变性獲得の代わりに、スケール不变性を犠牲
- ◆ **アフィン変換不变性の獲得の難しさを認識**

各種アルゴリズムの変換不变性の有無

アルゴリズム	解空間のアフィン変換不变性			目的関数の 単調変換不变性
	平行移動不变性	回転不变性	スケール不变性 (相似変換不变性)	
ES	○	○	×	○
実数値GA	○	○	×	○
実数値SA	○	○	×	×
PSO	○	✗	○	○
CRI-PSO*	○	○	相似○／スケール×	○
DE	○	✗	○	○
ABC	○	✗	○	×
HBC-ABC*	○	○	相似○／スケール×	×
Firefly Algorithm	○	○	×	○
Cuckoo Search	○	✗	✗	○
CMD-CS*	○	○	○	○

※1：橙背景は、本研究の対象となったアルゴリズム

※2：赤字は、本研究によって影響を受けた不变性

課題解決に資する本論文のアプローチ

以降の章で、各課題に対する解決策を検討する。

2.1章：不变性の付加

学位論文第4章

課題1

アフィン変換不变性獲得の難しさ

課題2

他の条件への対応の必要性

課題3

ユーザにとってのアプローチの難しさ

解決策検討

2.2章：適応的パラメータ調整

学位論文第5章

2.3章：不变性フレームワーク

学位論文第6章

本章のアプローチ

口バスト性・適応性の具備を目指し、
下記指針に基づき、パラメータ調整則を設計する。

不变性に基づく指針

指針1：不变性補完

指針2：不变性維持

→ 本章の新規性

適応性に基づく指針

AFPSOの活性度制御

→ 先行研究[16]を参考

CSの適応化(博士論文5.3節)

不变性補完による口バスト性向上

多様化・集中化の探索戦略による
適応性向上

本発表

CRI-PSOの適応化(博士論文5.4節)

不变性維持による口バスト性維持

多様化・集中化の探索戦略による
適応性向上

Activity Feedback PSO (先行研究[16])

- 多様化・集中化の評価指標として活性度 $P(k)$ を定義し、活性度に基づく PSO の安定性解析を行った。

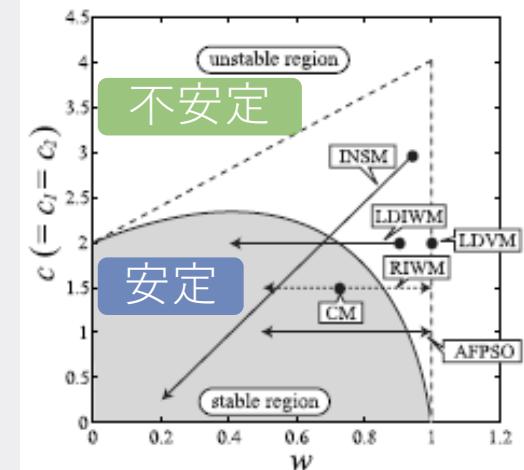
摂動量の L^2 ノルム

$$P(k) = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|x^i(k) - x^i(k-1)\|_2 \quad \text{式(5.40)}$$

- 探索過程で P が徐々に漸減するとき、良好な探索を行うことを指摘 (探索戦略)
- 数値実験的にパラメータ空間内で PSO の安定・不安定領域を求めた
 - w が高い : $P(k)$ が発散 (多様化)
 - w が低い : $P(k)$ がすぐ収束 (集中化)

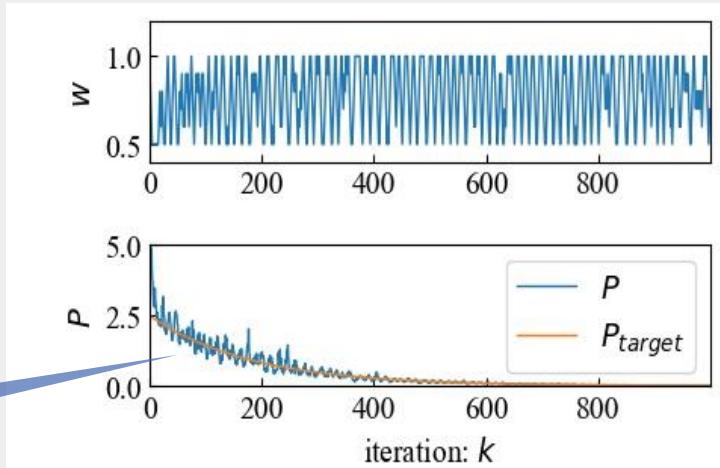
- $P(k)$ を目標値 $P_{\text{target}}(k)$ に追従するように w を調整する AFPSO を提案し、有効性を確認した。

P_{target} を漸減させることで、探索戦略を実現



安定・不安定領域

$$P_{\text{target}}(k) = P_{\text{start}}(P_{\text{end}}/P_{\text{start}})^{k/k_{\max}}$$

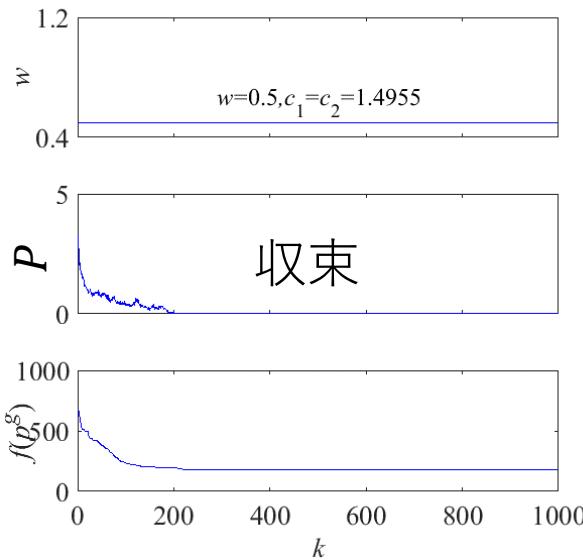


活性度制御による探索戦略実現

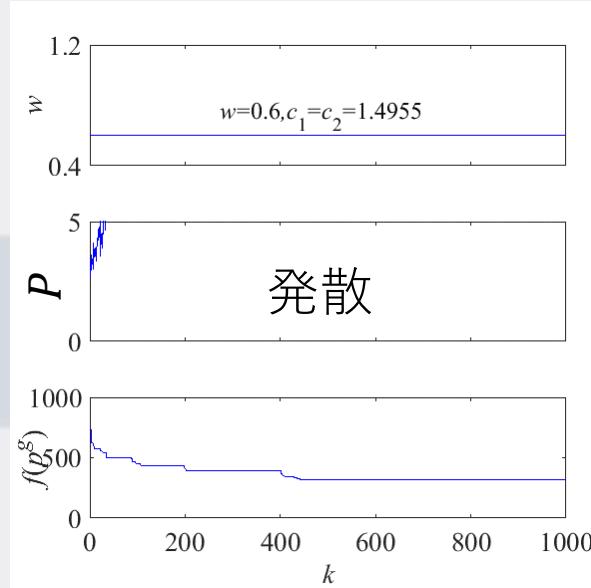
安定性解析に基づくCRI-PSOのパラメータ解析

活性度を用いて、CRI-PSOの安定・不安定領域の境界線を特定。

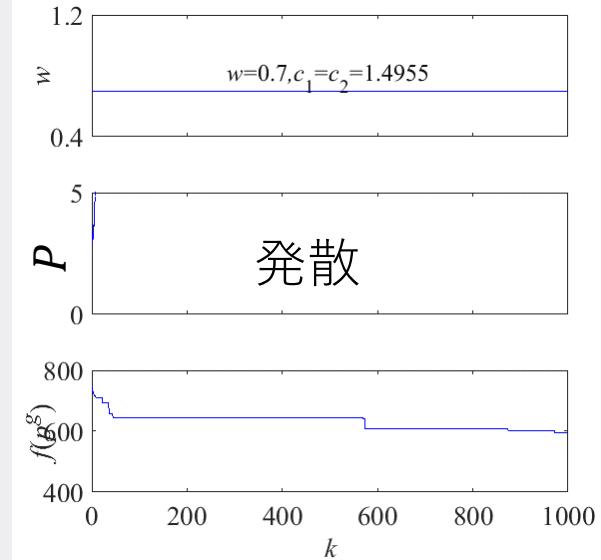
Ackley関数、 $N = 50$



$$w = 0.5$$



$$w = 0.6$$

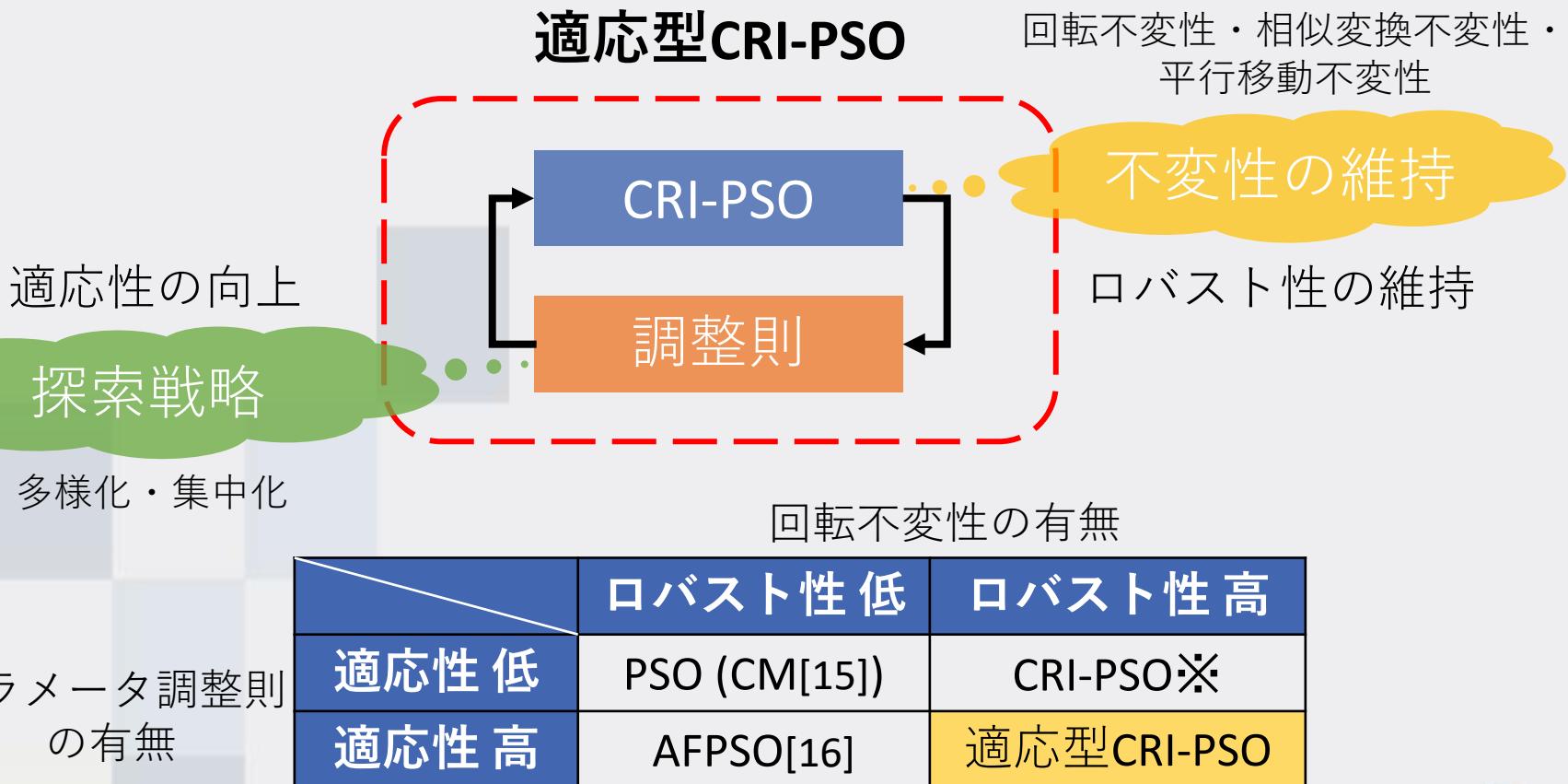


$$w = 0.7$$

- ※他のベンチマーク関数においても、同様の傾向を確認した
- ※PSOと安定・不安定領域の境界線が異なる[15]
- ※ $m = 20, c_1 = c_2 = 1.4955$ に設定

適応型CRI-PSO

下記の観点に基づき、適応型CRI-PSOを設計する。



※スケール不变性の観点では、CRI-PSOはロバスト性が低い

探索戦略に基づくパラメータ調整則

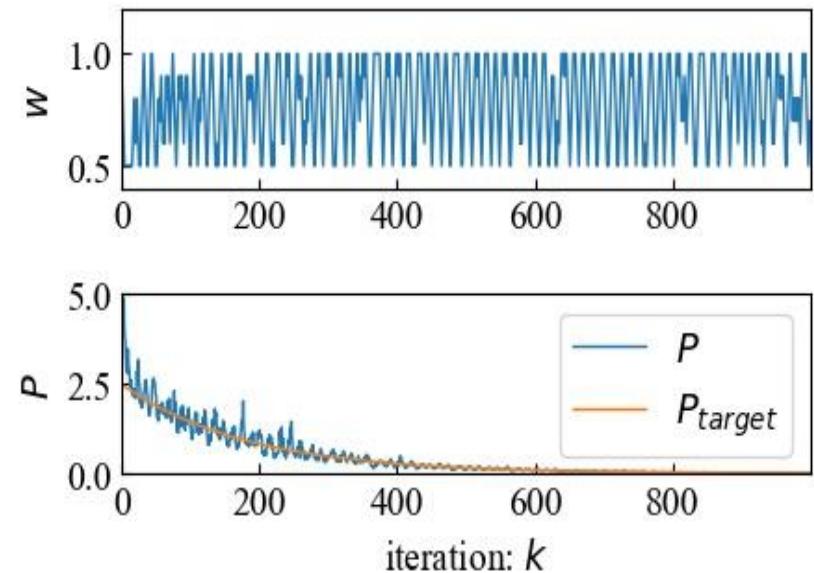
AFPSOと同様の適応的パラメータ調整則を付加する。

- ◆ 活性度 P による探索状態（多様化・集中化）の評価
- ◆ 探索戦略に従って活性度を制御する適応的調整則

$$w(k+1) = \begin{cases} \min\{w(k) + \Delta w, w_{\max}\}, & P(k) \leq P_{\text{target}}(k) \\ \max\{w(k) - \Delta w, w_{\min}\}, & \text{otherwise} \end{cases} \quad \text{活性度制御[16]}$$

$$P(k) = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \left\| \mathbf{x}^i(k) - \mathbf{x}^i(k-1) \right\|_2$$

$$P_{\text{target}}(k) = P_{\text{start}} \left(P_{\text{end}} / P_{\text{start}} \right)^{k/k_{\max}}$$



変換不变性に基づく適応型CRI-PSOの解析

目標値の設定を工夫すれば、CRI-PSOの平行移動不变性・相似変換不变性・回転不变性を維持できる。

初期配置領域 $S_0 = [\gamma_{\min,n}, \gamma_{\max,n}]^N$ とする

$$P_{\text{target}}(k) = P_{\text{start}} \left(\frac{P_{\text{end}}}{P_{\text{start}}} \right)^{\frac{k}{k_{\max}}} = \frac{\varepsilon_{\text{start}}}{\sqrt{N}} \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 \left(\frac{\varepsilon_{\text{end}}}{\varepsilon_{\text{start}}} \right)^{\frac{k}{k_{\max}}} = \varepsilon_1 \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2$$

$$P(k) = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|\boldsymbol{x}^i(k) - \boldsymbol{x}^i(k-1)\|_2 = \varepsilon_2 \sum_{i=1}^m \|\boldsymbol{x}^i(k) - \boldsymbol{x}^i(k-1)\|_2$$

変換によって影響を受ける

例えば、回転変換を与えると、 $P_{\text{target}}(k) > P(k)$ ならば、

$$\begin{aligned} P'_{\text{target}}(k) - P'(k) &= \varepsilon_1 \|\boldsymbol{B}(\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min})\|_2 - \varepsilon_2 \sum_{i=1}^m \|\boldsymbol{B}(\boldsymbol{x}^i(k) - \boldsymbol{x}^i(k-1))\|_2 \\ &= \varepsilon_1 \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 - \varepsilon_2 \sum_{i=1}^m \|\boldsymbol{x}^i(k) - \boldsymbol{x}^i(k-1)\|_2 = P_{\text{target}}(k) - P(k) > 0 \end{aligned}$$

したがって、 $w'(k) = w(k)$ となり、パラメータ調整則の不变性を満たす。
更新則の不变性の有無にも影響を与えない。

数値実験

■ 数値実験を通じて、適応型CRI-PSOの性能を検証する。

- ◆ 実験1：回転変換に対するロバスト性に関する検証
- ◆ 実験2：全般的な性能に関する検証

比較手法

	回転不变性 無	回転不变性 有	
パラメータ調整 無	PSO (CM[14])	CRI-PSO	$Z = \mathcal{P}(k)$
パラメータ調整 有	AFPSO[15]	適応型CRI-PSO	$Z = \mathcal{P}(k)$

wの設定

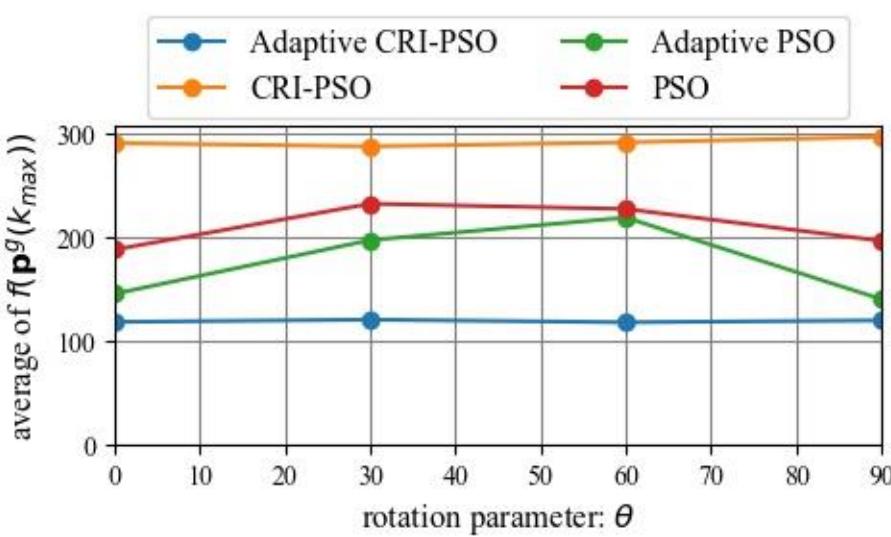
- ◆ PSO : $w = 0.729$ (Constriction Method : CM) [15]
- ◆ CRI-PSO : $w = 0.5$ (安定状態を実現するパラメータ)
- ◆ AFPSO : $w(k) \in [0.5, 1.0]$ [15]
- ◆ 適応型CRI-PSO : $w(k) \in [0.5, 1.0]$

弱い安定領域

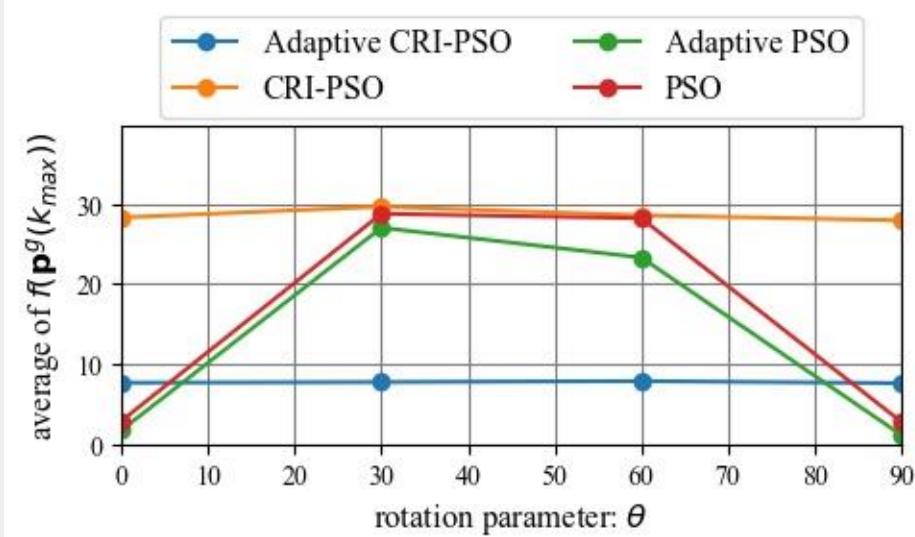
安定・不安定
境界線を跨ぐ

実験1：回転に対するロバスト性検証結果

適応型CRI-PSOはCRI-PSOのロバスト性を維持したまま、性能を向上していることが確認できる。



Rastrigin関数



Alpine関数

※ $N = 50, k = 1000$ の結果だが、他の条件でも同様の傾向が確認された

※異なる初期点から50試行、g-bestの目的関数値の平均値を性能とみなす

実験2：全般的性能に関する検証結果

■ 様々な関数や条件に対する全般的性能を比較する。

- ◆ 9種類の関数(单／多峰、変数間独立／依存、良／悪スケール)
- ◆ 72個の条件において、4種類手法の中で1位の数をカウント

■ 全般的に適応型CRI-PSOが最も優れることを確認。

- ◆ 低次元はAFPSOが優れ、それ以外は全て提案手法が優れる
- ◆ AFPSOは代表的なパラメータ調整則を有するPSOよりも優れている[15]ため、適応型CRI-PSOも同様だと期待できる

アルゴリズム	次元数N				反復回数k		All
	10	50	100	300	100	1000	
PSO (CM[14])	7	0	0	0	3	4	7
AFPSO[15]	8	5	2	0	6	9	15
CRI-PSO	1	0	1	1	3	0	3
適応型CRI-PSO	2	15	15	17	24	25	47

※異なる初期点から50試行、g-bestの目的関数値の平均値を性能とみなす

まとめと課題

- 適応型CRI-PSOは回転・相似変換・平行移動変換に対するロバスト性を維持していることを確認。
 - ◆ 不変性に基づく調整則と更新則の関係を明らかにした
- 適応型CRI-PSOが全般的に優れることを確認。
 - ◆ PSOよりもロバスト性と適応性が向上
 - ◆ 様々な調整則よりも優れていたAFPSOよりも優れている
 - 探索戦略と変数間依存性への対処の効果が現れている
- ただし低次元ではAFPSOが優れる。
 - ◆ 元々CRI-PSOは低次元ではPSOよりも劣っていた
 - ◆ w 範囲に向上の余地がある（より精緻な安定性解析[16]）
- 適応型CRI-PSOはスケール不变性を失っている。
 - ◆ アフィン変換不变性の獲得の難しさは同様であることを認識

変換不变性の付加アプローチの課題

一部未解決であったため、他のアプローチにより、他の課題解決策を検討。

一部未解決

2.1章：不变性の付加

完全未解決

課題1

アフィン変換不变性獲得の難しさ

課題2

他の条件への対応の必要性

課題3

ユーザにとってのアプローチの難しさ

解決策検討

2.2章：適応的パラメータ調整

2.3章：不变性フレームワーク

課題2は解決できるが、

課題1は一部しか解決できない

(補完可能な演算形式・不变性は限定的)

課題1と課題3の解決策を検討

本章のアプローチ

最適化能力の設計自由度を残しながらも、変換不変性を具備するアルゴリズムの枠組みを確立する。

付加アプローチ

2.2章や先行研究

現象A

現象B

現象C

ここで最適化能
力・数理的構造
が決定される

Nature-
Inspired

手法A

手法B

手法C

不変性の
付加

手法A'

手法B'

手法C'

手法毎に
異なる

本章のアプローチ

不变性を具備する
ための条件

埋め込み

高い自由度

メタヒューリスティクスの
設計フレームワーク

最適化能力
の設計

手法A

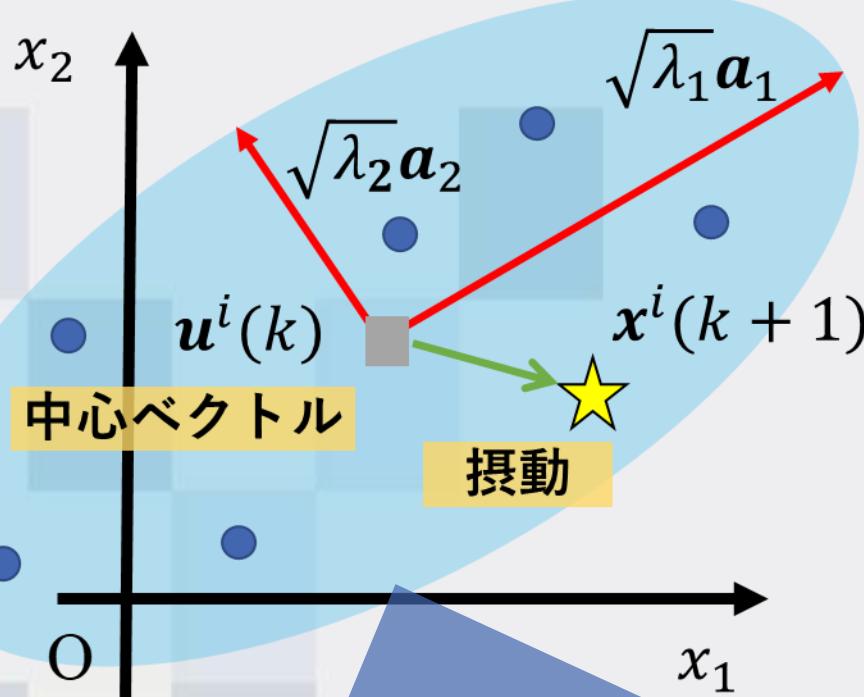
手法B

手法C

共分散行列の固有値分解に基づく直接探索法(CMD-DS)

中心ベクトル \mathbf{u}^i に、共分散行列に基づき基底変換された乱数ベクトル \mathbf{s}^i による摂動を与えることで近傍解を生成する。

近傍のイメージ



摂動の範囲や偏りは確率分布 \mathcal{P} と解集団 \mathcal{Z} に依存

$\Sigma(k)$ の非対称平方根

$$\begin{aligned} \text{中心ベクトル} & \quad \text{乱数ベクトル} \\ \mathbf{x}^i(k+1) &= \mathbf{u}^i(k) + \alpha \underline{\Sigma_*^{1/2}(k)} \mathbf{s}^i(k) \\ \mathbf{s}^i(k) &\sim \mathcal{P} \quad (\text{確率分布 } \mathcal{P}) \end{aligned} \quad \text{式(6.1)}$$

ただし、 α はパラメータ、 $\Sigma(k)$ は解集団 \mathcal{Z} の共分散行列である。

固有ベクトル : \mathbf{a}_n

固有値 : λ_n

アフィン変換不変性を具備するための条件

CMD-DSでは、下記の条件1、条件2が課される。

◆ 条件1

- 中心ベクトル $\mathbf{u}^i(k)$ はアフィン変換不変性を有する

◆ 条件2

- 乱数ベクトル $\mathbf{s}^i(k)$ の確率変数同士が独立であり、
 $\mathbf{s}^i(k)$ が従う分布 \mathcal{P} のパラメータが設計変数とは独立



条件を満たす範囲で設計されたアルゴリズム群は、
アフィン変換不変性の具備が保証される

※定理31は補足に記載

CMD-DSの特徴

- 変換不変性の具備を解析する必要がない。
 - ◆ CMD-DSのアルゴリズム群は不变性の具備が保証されるため、個々のアルゴリズムについて解析する必要がない
- 最適化能力の設計自由度が高い。
 - ◆ 多点探索・確率的探索は共通しているが、経験則に基づいていないため、最適化能力を有していない
 - ◆ 規定範囲で探索戦略を適切に反映するように詳細構造を定めることで、最適化能力を有するアルゴリズムを構成できる
- 共分散行列に基づく基底変換を活用して、アフィン変換不変性を保証する。
 - ◆ $\Sigma_*^{1/2}(k)$ を乱数ベクトル \mathbf{s}^i に適用することで、線形変換不変性を補完する
 - ◆ 最適化能力を決定づけるアルゴリズムの詳細構造において設計自由度が低い点で、先行研究とは異なる

変換不変性定理の成立条件・例

条件の成立条件を定理として与えた。

- ◆ 反復法のアルゴリズムに立脚しているため、多くの手法のアフィン変換不変性の有無を指摘することが可能

条件 1 の成立条件・例

\mathbf{u}^i がアフィン変換不変性を具備

中心ベクトル $\mathbf{u}^i(k)$ は代表的な形式の場合の変換不変性定理を導出。
(証明は補足資料に記載)

#	形式例	条件
①	$\mathbf{u}^i = \mathbf{A}\mathbf{v}$	$\mathbf{A} = \mathbf{I}$
②	$\mathbf{u}^i = \sum_j \mathbf{A}_j(\mathbf{x}_j - \mathbf{y}_j) + \mathbf{v}$	$\mathbf{A}_j = \delta_j \mathbf{I}$

$$\mathbf{A}, \mathbf{A}_j \in \mathbb{R}^{N \times N}, \mathbf{x}_j, \mathbf{y}_j, \mathbf{v} \in \mathbb{R}^N, \delta_j > 0$$

条件 2 の成立条件・例

分布 \mathcal{P} が探索とは独立、要素間が独立

乱数ベクトル $\mathbf{s}^i(k)$ の確率分布が $\mathcal{P}(x) = \prod_{n=1}^N \mathcal{P}(x_n)$ と表される。

- ✓ $s_n^i \sim \mathcal{P}(x_n)$ (各要素が独立)
- \mathcal{P} の例
 - ✓ 標準正規分布
 - ✓ 一様分布
 - ✓ 標準Cauchy分布

CMD-DSに基づくPSO着想アルゴリズム \mathcal{A}_{PSO}

- \mathcal{A}_{PSO} の更新則における \mathbf{u}^i と \mathbf{s}^i を下記のように設計する。

$$\mathbf{u}^i(k) = \mathbf{x}^i(k) + c_1 r_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 r_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k)) \quad \text{式(6.5)}$$

$$\mathbf{s}^i(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad \text{式(6.6)}$$

ただし、 $Z = \mathcal{P}(k)$ 、 $\alpha = \beta(1 - c_1 r_1 - c_2 r_2)$ 、 $r_1, r_2 \sim \mathcal{U}_{\mathbb{R}}(0,1)$ である。

- この更新則は、条件1、条件2を満たすため、アフィン変換不変性を有する。(定理33)

◆ 条件1の成立

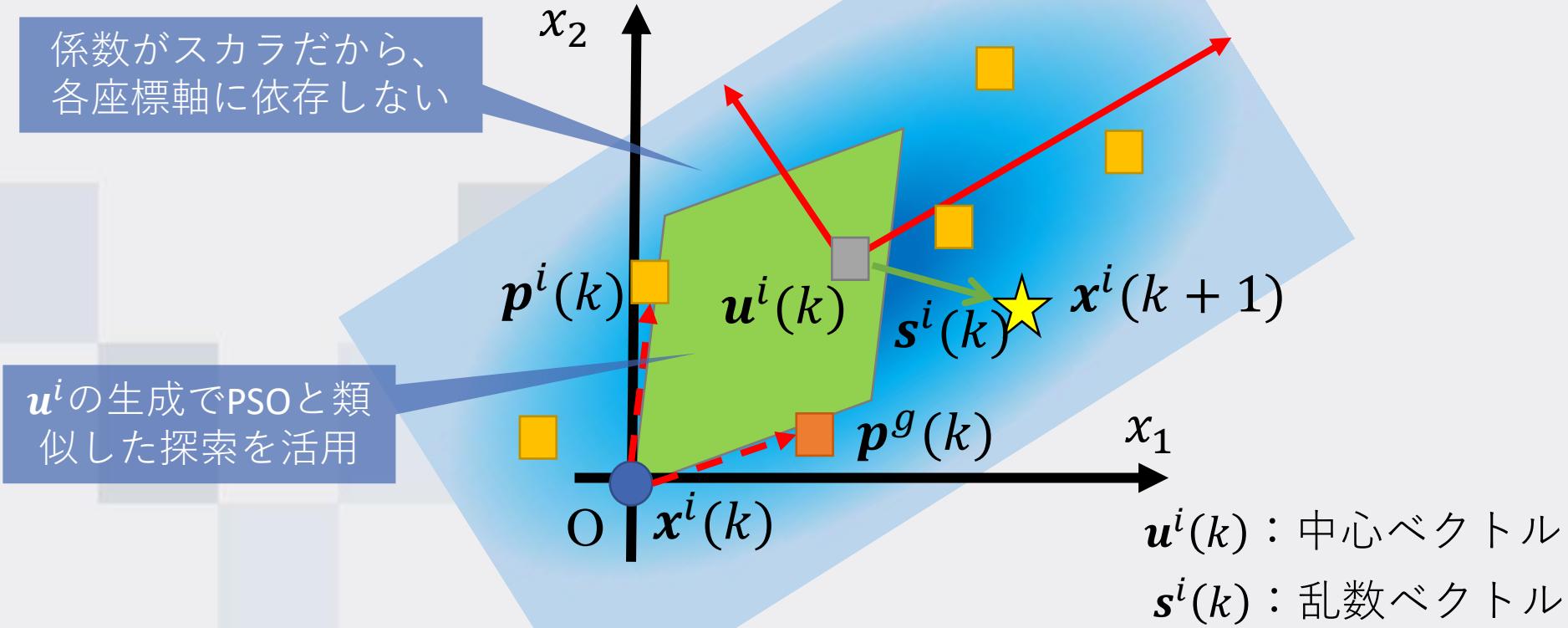
\mathbf{u}^i は定理33の形式 ($\mathbf{u}^i = \sum_j \mathbf{A}_j(\mathbf{x}_j - \mathbf{y}_j) + \mathbf{v}$) に該当するが、この場合の \mathbf{u}^i がアフィン変換不変性を具備するための条件は、 $\mathbf{A}_j = \delta_j \mathbf{I}$ が成立することである。式(6.5)は明らかにこれを満たす。

◆ 条件2の成立

\mathcal{P} は標準正規分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ に従うが、明らかに標準正規分布は確率変数同士で独立であり、パラメータは設計変数とは独立している。

\mathcal{A}_{PSO} の近傍のイメージ

$$\begin{aligned} \boldsymbol{x}^i(k+1) &= \boldsymbol{u}^i(k) + \beta(1 - c_1 r_1 - c_2 r_2) \boldsymbol{\Sigma}_*^{1/2}(k) \boldsymbol{s}^i(k) \\ \boldsymbol{u}^i(k) &= \boldsymbol{x}^i(k) + c_1 r_1 (\boldsymbol{p}^i(k) - \boldsymbol{x}^i(k)) + c_2 r_2 (\boldsymbol{p}^g(k) - \boldsymbol{x}^i(k)) \end{aligned}$$



- ※ $u^i(k)$ の生成はアフィン変換不変性を有するLinear PSO[8]と類似
- ※ 実質的に、 $\boldsymbol{x}^i(k+1) \sim \mathcal{N}(\boldsymbol{u}^i(k), \alpha^2 \boldsymbol{\Sigma}(k))$ と等価

\mathcal{A}_{PSO} の特徴

■ \mathcal{A}_{PSO} の最適化能力

- ◆ PSOの「良い解が存在する有望領域に向かう」というPOPの思想が導入されているため、PSOと同等の最適化能力が期待される

■ \mathcal{A}_{PSO} の変換不変性

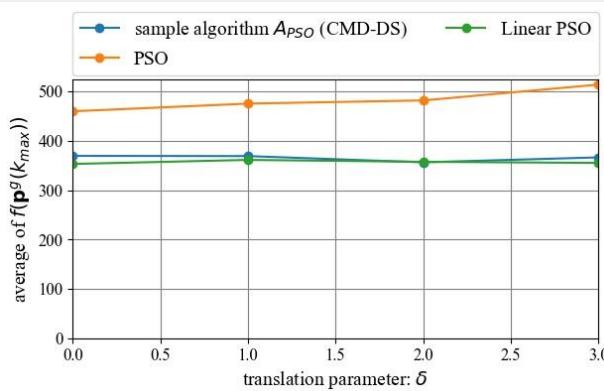
- ◆ CMD-DSが課す条件を満たすため、アフィン変換不変性を有する
- ◆ アフィン変換に対する探索性能のロバスト性が高いことが期待される

アフィン変換不変性が保証される枠組みのもとで、
PSOの最適化能力を埋め込んだ

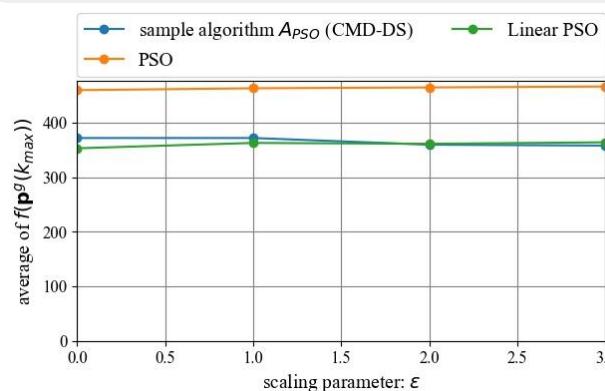
アフィン変換に対するロバスト性検証結果

この条件では、最適化能力がPSO、Linear PSO[8]と同等以上で、高いロバスト性を確認できた。

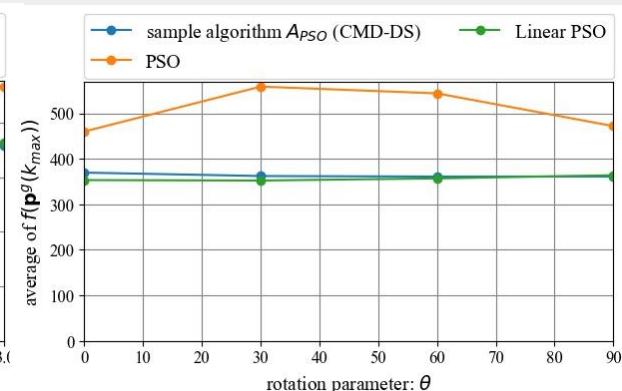
Rastrigin 実験A（平行移動）



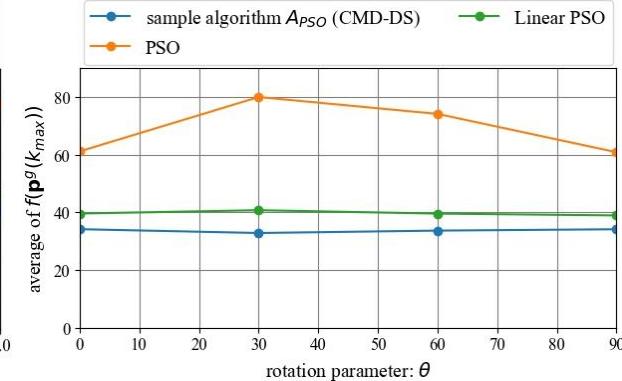
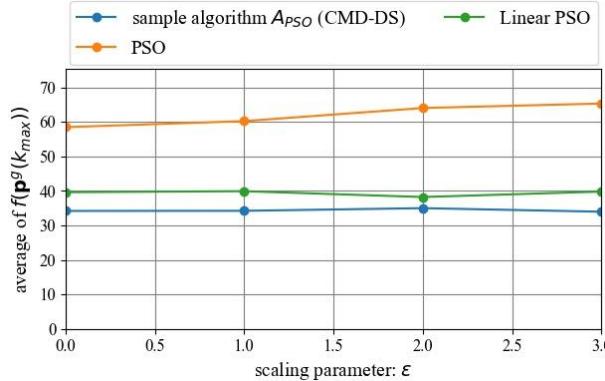
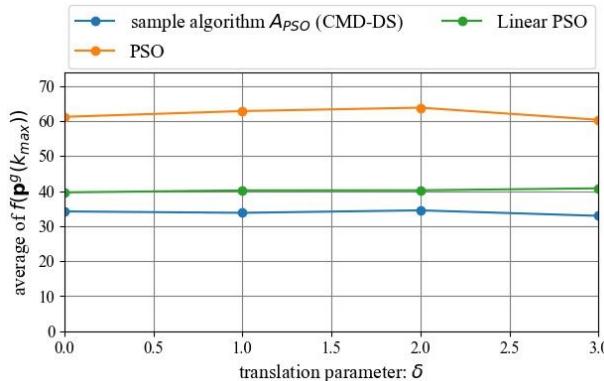
実験B（スケール）



実験C（回転）



Alpine



* $N = 50, k_{\max} = 100$ 、全般的な性能検証は割愛

まとめと課題

- アルゴリズム設計において、変換不変性の証明が不要なフレームワークとして、CMD-DSを提案した。
- CMD-DSに従って既存のアルゴリズムの最適化性能を埋め込むことで、アフィン変換不変性の具備が保証されたサンプルアルゴリズム \mathcal{A}_{PSO} を設計した。
- \mathcal{A}_{PSO} はPSOの基本的な性能を引き継ぐと同時に、アフィン変換のロバスト性を具備していることを確認。
 - ◆ CMD-DSはアフィン変換不変性の具備が保証されながら、優秀な最適化能力を発揮するポテンシャルを有する
(極めて高い汎用性を確認)
- 不変性の指針2に従って調整則を付加することで、ロバスト性の維持と適応性の向上も期待される。

全般的な性能に関する総合比較

下記12種類アルゴリズムを総合比較する。

アルゴリズム	ロバスト性 (変換不变性)	適応性 (パラメータ調整)
PSO (CM)[15]	× (回転不变性)	×
AFPSO[16]	× (回転不变性)	○
CRI-PSO*	△ (スケール不变性)	×
適応型CRI-PSO*	△ (スケール不变性)	○
ABC	× (回転不变性)	×
HBC-ABC*	△ (スケール不变性)	×
Cuckoo Search (CS)	× (回転・相似変換不变性)	×
CMD-CS*	○ (−)	×
適応型CS*	△ (回転・スケール不变性)	○
Linear PSO[8]	○ (−)	×
$\mathcal{A}_{\text{PSO}}^*$	○ (−)	×
(μ_w, λ) -CMA-ES[17]	○ (−)	○

※ *は本研究で設計したアルゴリズム（合計6種類）

全般的な性能に関する総合比較結果

- 様々な条件における全般的な性能を比較する。
 - ◆ 72個の条件において、12種類手法の中で4位以上の数をカウント（合計数の1位赤、2位青、3位緑、4位黄）
- 全般的に適応型CRI-PSOが最も優れることを確認。
 - ◆ 特に高次元では本研究のアルゴリズムが優位

条件	優劣関係（上位6位のみ）
全般	適応型CRI-PSO* > HBC-ABC* > ABC > AFPSO = CMA-ES > 適応型CS*
$T = 2 \times 10^3$ (短期)	適応型CRI-PSO* > HBC-ABC* > $\mathcal{A}_{\text{PSO}}^*$ > AFPSO > Linear PSO = CMD-CS*
$T = 2 \times 10^4$ (長期)	ABC > 適応型CRI-PSO* > CMA-ES > 適応型CS* = HBC-ABC* > CS
$N = 10$	CMA-ES > ABC = AFPSO > PSO > HBC-ABC* > 適応型CRI-PSO*
$N = 50$	適応型CRI-PSO* > AFPSO = HBC-ABC* > ABC > CMA-ES > CS
$N = 100$	適応型CRI-PSO* > HBC-ABC* > $\mathcal{A}_{\text{PSO}}^*$ > ABC > 適応型CS* = CMD-CS*
$N = 300$	HBC-ABC* > 適応型CRI-PSO* > $\mathcal{A}_{\text{PSO}}^*$ > Linear PSO = CS > 適応型CS*

※ *は本研究で設計したアルゴリズム（合計6種類）

全般的な性能に関する総合比較結果

■ 適応型CRI-PSO*が全般的に最も優れている。

- ◆ 高いロバスト性・適応性の具備、p-best集団の活用

■ HBC-ABC*・ABCが次に優れている。

- ◆ ABCは元々最適化能力が優れているが、高次元では劣化
 - 緩やかな集中化と、再配置による多様化が可能
- ◆ HBC-ABCは超球化することで高次元でも性能を維持した
 - 変数間依存性と高次元化へのロバスト性を具備

■ CMA-ES・AFPSO・適応型CS*が次に優れている。

- ◆ 適応性（多様化・集中化）によって平均性能の向上

■ それ以外のアルゴリズムは、適応性が低い、ロバスト性が低い、元々の最適化能力が低い、に該当。

適応型CRI-PSO・HBC-ABC・適応型CSはパラメータ設定を試行錯誤せずに、平均的に優れた性能を発揮した

本研究のまとめ

組み合わせることで、高いロバスト性・適応性を有するメタヒューリスティクスの設計が可能であることを示唆した。

2.1章：不变性の付加

課題1

アフィン変換不变性獲得の難しさ

課題2

他の条件への対応の必要性

課題3

ユーザにとってのアプローチの難しさ

2.2章：適応的パラメータ調整

2.3章：不变性フレームワーク

課題2は解決できる

(補完可能な不变性は限定的)

課題1と課題3が解決できる

本研究の成果

■ 本研究の新規性・有用性

- ◆ 汎用な不变性の解析フレームワークを確立し、代表的な手法の不变性を理論的に明らかにした。
- ◆ 理論に基づき、不变性の付加アプローチを適切に流用することで、口バスト性の向上が可能であることを確認。
- ◆ 不変性に基づくパラメータ調整則の設計指針に従うことで、口バスト性・適応性の高いアルゴリズムの設計が可能であることを確認。State of the Artな手法よりも優れることも確認。
- ◆ 不変性が保証されるもとで自由に最適化能力を埋め込まれるアルゴリズムのフレームワークを提案した。

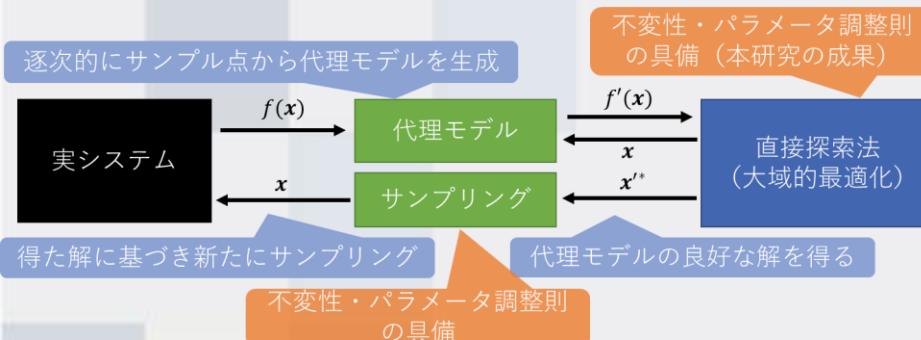
■ 本研究の発展性

- ◆ 新たなアルゴリズムに対しても同様に、理論的な方法で、不变性の有無を指摘し、付加することが可能。
- ◆ 有用なパラメータ調整則の導入により、より口バスト性・適応性に優れたアルゴリズム設計が期待できる。

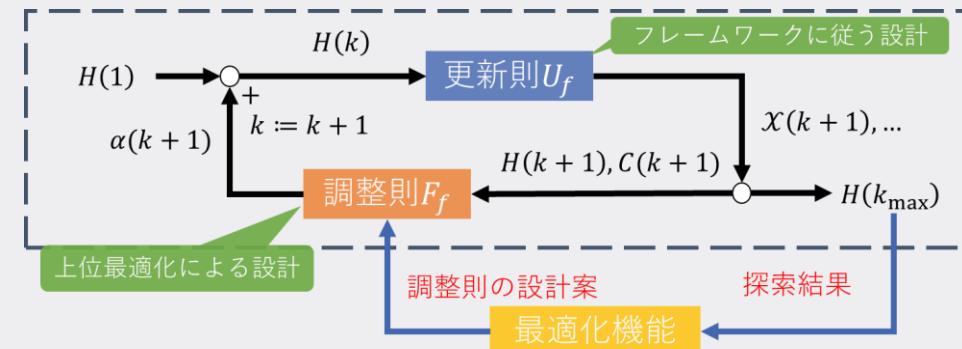
今後の展望

- 逐次近似最適化においても導入することで、高いロバスト性・適応性を具備する総合的なBBOの開発が期待できる。
 - ◆ AI技術やコンピュータパワーなどの外部環境をフルに活用可能
- 変換不変性のフレームワークとパラメータ調整則の融合により、より優れたアルゴリズム設計が期待できる。
 - ◆ CMA-ESの調整則[17]、安定性が保証されたフレームワーク[18]など
 - ◆ 上位最適化機能により、調整則の設計案を自動的に獲得[11]

逐次近似最適化



調整則の自動設計[11]



BBOの有力なアプローチの一種

参考文献①（本発表資料）

- [1] 「メタヒューリスティク手法の電力・エネルギー分野への適用例」, 電気学会B部門誌, Vol.124, No.5, pp.679-682 (2004)
- [2] 「回転機のトポロジー最適化および3次元形状最適化に関する研究」, 北海道大学博士論文 (2017)
- [3] "Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, Vol.13, pp.281-305 (2012)
- [4] 「マテリアルズインフォーマティクスの現状と課題」, 情報知識学会誌, Vol.27, No.4, pp.297-304 (2017)
- [5] 「バイオインフォマティクスと人工知能の相互作用」, 人工知能学会誌, Vol.22, No.1, pp.42-48 (2007)
- [6] "Principled Design of Continuous Stochastic Search: From Theory to Practice," *Theory and Principled Methods for the Design of Metaheuristics*, pp 145-180 (2013)
- [7] "Comparison of Linear and Classical Velocity Update Rules in Particle Swarm Optimization: Notes on Scale and Frame Invariance," *Int. Journal for Numerical Methods in Engineering*, Vol.70, pp.985-1008 (2007)
- [8] "A New Particle Swarm Optimiser for Linearly Constrained Optimisation," *Proc. of the 2003 IEEE Cong. on Evo. Comp.*, pp.1839-1844 (2003)
- [9] "Standard Particle Swarm Optimisation," *HAL Open Access Archive* (2012)

参考文献② (本発表資料)

- [10] 「適応的実数値交叉AREX の提案と評価」, 人工知能学会論文誌, Vol.24, No.6A, pp.446-458 (2009)
- [11] 「メタヒューリスティクスに対する遺伝的プログラミングによる創発的パラメータ調整則の自動設計」, 慶應義塾大学博士論文 (2016)
- [12] "Completely Derandomized Self-Adaptation in Evolution Strategies," MIT Press, Evol. Comp., Vol.9, No.2, pp.159-195 (2001)
- [13] "Artificial Bee Colony Algorithm with Principal Component Analysis," Journal of Electronics and Communications in Japan, Vol.99, No.9, pp.47-61 (2016)
- [14] "Swarm Intelligence," Morgen Kaufmann Publishers (2001)
- [15] 「安定性解析に基づく多点型確率的最適化アルゴリズムの設計に関する研究」, 慶應義塾大学博士論文 (2015)
- [16] "Particle Swarm Optimization: A Numerical Stability Analysis and Parameter Adjustment Based on Swarm Activity," IEEJ Trans. on Electrical and Electronic Engineering, Vol.3, No.6, pp.642-659 (2008)
- [17] "The CMA Evolution Strategy: a Comparing Review," Towards a New Evolutionary Computation, Advances on Estimation of Distribution Algorithms, pp.75-102, Springer (2006)
- [18] "Regulated Evolution Strategies: A Framework of Evolutionary Algorithms with Stability Analysis Result," Trans. on Electrical and Electronic Engineering (2020)

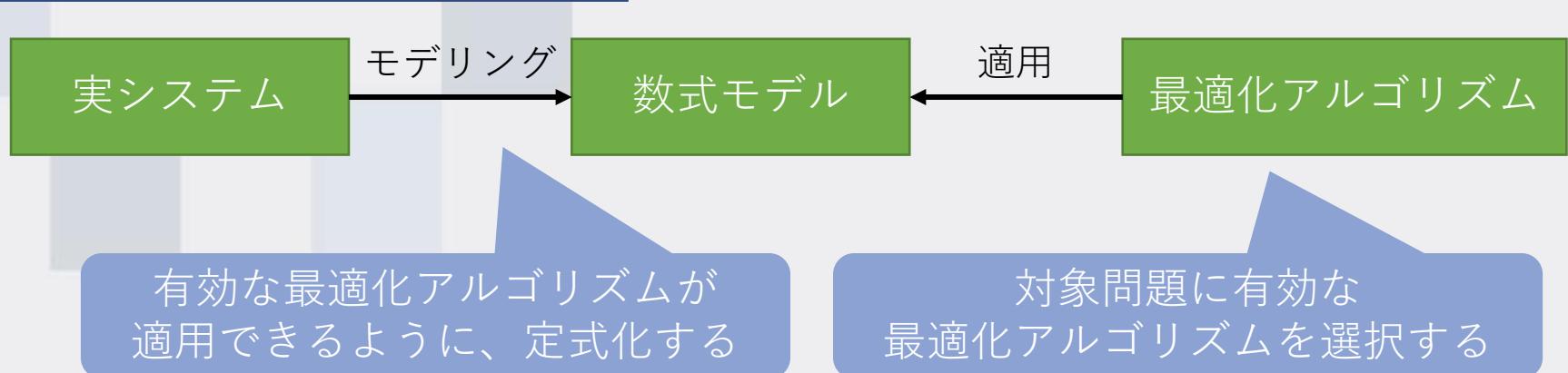
発表は以上です、ありがとうございました。

付録

工学における最適化の形態

一般に工学における最適化では、「アルゴリズムのためのモデリング」と「モデルのためのアルゴリズム選択」を行うことが常道である。

工学における最適化の形態



工学における最適化プロセス

ステップ	内容
1	ユーザが原問題のモデリングを行い、最適化問題として定式化する
2	開発者がある問題クラスに適したアルゴリズムを開発する
3	ユーザが問題に適したアルゴリズムを選択する
4	ユーザが問題に適するようにアルゴリズムのパラメータを設定する
5	ユーザがアルゴリズムを適用し、得た解を吟味する

モデルとアルゴリズムは密接な関係にあるため、
ユーザには各プロセスで適切な作業が求められる

タンパク質の新規設計

- Protein Data Bankはタンパク質と核酸(DNA&RNA)の世界的公共構造データベースで、簡単にBig Dataを入手可能。
- 酵素は触媒となる生体内高分子化合物(タンパク質)で、今後の生活の基盤となりうる。
 - バイオリファイナリの中核となる
 - バイオマスを原料として燃料や樹脂を製造する技術
⇒バイオマスは環境負荷が少なく、化石原料の代替が期待される
 - バイオエタノール生産には、アミラーゼによるデンプン分解が必要
 - 一方、自然界(菌や微生物)から採取するため、非効率かつ不明なことが多い上に、人工設計は未踏の技術
 - 名古屋議定書(2014)により遺伝資源の産業利用に大きな制限
⇒遺伝資源の公平分配が目的（途上国(保有)と先進国(利用)間の対立を防ぐ）
⇒全て分配や契約が必須となるなど、利用に対する制限が強く、過去の資源に対してまで分配を要求されるなどのリスクがある
- タンパク質のDNA⇒構造予測は困難だが、コンペティションで初参加のAlphaFold（2018）がトップとなった。

コンピュータの計算能力

■ ムーアの法則は限界を迎えていたが、計算能力の増加は可能である。

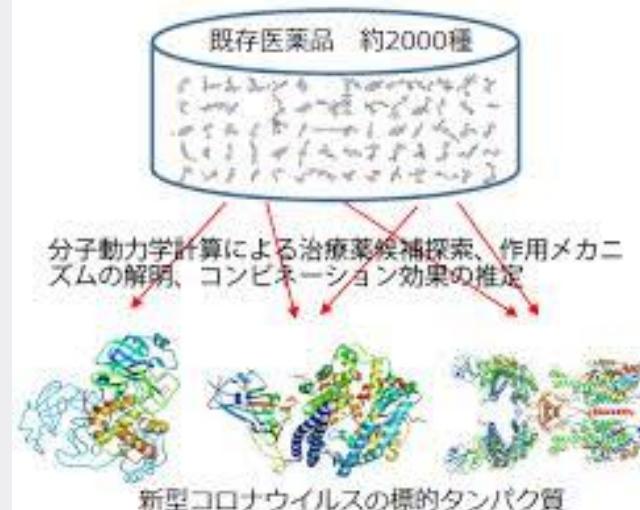
◆ CPUの進化は限界を迎えていたと言われているが、あくまでCPUに限定された議論で、GPGPUやFPGAのようなアクセラレータユニットによって更なる加速は可能である

■ 「Top500」(スパコン性能ランキング)の2020年版の上位では、GPGPUやスーパースカラが採用されている。

◆ MicrosoftやIntelなどもスパコン大手と組み、開発に注力
◆ 富岳（理研・富士通）が1位

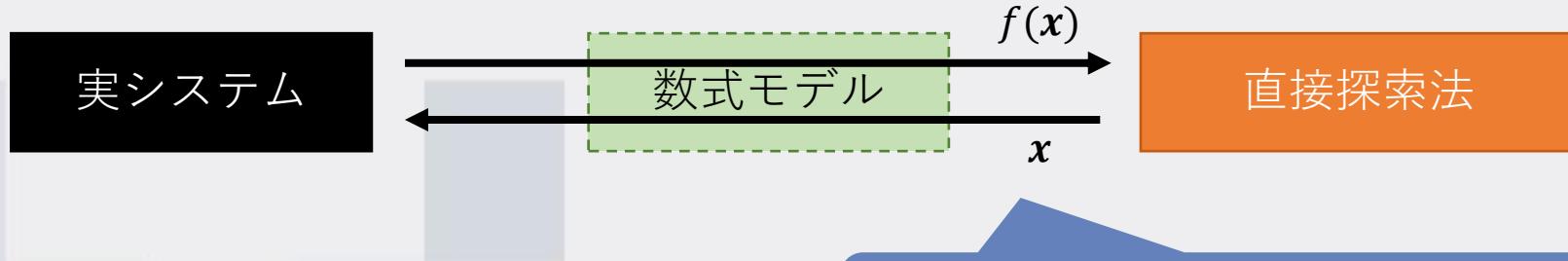
- スーパースカラ方式
- 新型コロナウイルスの治療薬候補を、10日で約2000種類の医薬品から選別（先月発表、京なら1年以上）

※理研：<https://www.rccs.riken.jp/jp/fugaku/corona/projects/okuno.html>



直接探索法

- BBOでは、システムの入出力情報しか得られず、目的関数の景観が不明なため、**直接探索法**の適用が必要。
 - ◆ 直接探索法は大域的最適化の側面も同時に有する



(1960年頃)

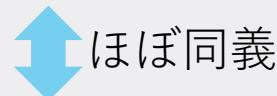
- 古典的な直接探索法は実問題で有効とは言い難い。
 - ◆ Hooke-Jeeves法（座標軸に依存した探索）
 - ◆ 滑降シンプレックス法（低次元のみ有効）
 - ◆ ランダムサーチ(無限に計算時間を与えた場合のみ有効)

メタヒューリスティクス

- 直接探索法の新たな枠組みとして、「メタヒューリスティクス」が知られている。
- 数学的な保証がないが、生物の進化過程や群知能などの経験則をアナロジーとして構築される発見的手法。

“Meta-heuristics”

複数の**発見的操作**が**上位概念**により束ねられ、全体の秩序が形成されたアルゴリズムの総称



上位概念

発見的操作

発見的操作

“Nature-Inspired”

自然現象に着想

進化計算

生物の進化過程

群知能

群の知的行動

※“Meta-” = 個別の要素を構造的に司る上位概念を表す接頭辞

※“Heuristics” = 発見的操作

メタヒューリスティクスの歴史

1960年代

Evolution Strategy(ES)・Genetic Algorithm(GA)（進化計算が登場）

1980年代

Simulated Annealing(SA)・Tabu Search(TS)(メタヒューリスティクスが登場)

→ GAやSAが組合せ最適化専用、ESやSAが単点探索であるなど、
枠組みとして体系化されていなかった

1990年代

実数値最適化への拡張・多点化などの進化計算が発展すると同時に、
PSOやDEなどの群知能が登場

→ 進化計算や群知能がメタヒューリスティクスとして融合され
始めた

2000年代

IEEEやACMでこの分野の大規模な国際会議が開催されると同時に、
ABC、FA、CSなどの新たなアルゴリズムの開発が盛んになる

2010年代

コウモリ、花粉媒介、キノコ、オオカミ、ペンギンなど、多様な
アナロジーに基づくアルゴリズム開発が進んでいる

BBOにおけるメタヒューリスティクスの成果

メタヒューリスティクスは、BBOで成果を挙げており、有力な最適化アルゴリズムとして地位を確立している。

シミュレータを活用した製品設計

- 電磁アクチュエータ[2]
- エンジン制御パラメータ[45]
- 新幹線先頭車両
(N-700系・N-700S)[46]



音声認識システム

言語モデルを人工知能でモデリングし、ハイパープラメータ探索[48]



実質的に右下と同じ

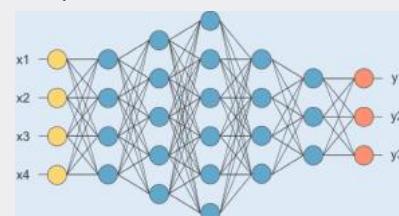
プラント設備運用

- エネルギー設備 [47]
- 生産設備



機械学習

- ハイパープラメータ探索[49]
- Neural Architecture Search[50]
- ✓ 深層学習のネットワーク構造を含めて、全て最適化する



現状トップ性能のアルゴリズム

■ 分野トップ国際会議のワークショップでは、BBOのコンペティションが開催されている。

- ◆ CEC : Competition on Real Parameter Single Objective Optimization (2005年～)
- ◆ GECCO : Black-Box-Optimization-Benchmarking (BBOB) (2009年～)

■ 総合ランキング上位には、適応型DEやCovariance Matrix Adaptation – ES (CMA-ES)の亜種が位置する。

- ◆ CMA-ESは、多くの変換不变性と適応的調整型パラメータ調整則の両方を具備するアルゴリズムとして代表的手法

※メタヒューリстиクスの性能比較は非常に困難だが、様々な条件・問題で平均性能を測定しているため、一つの指標となる

本研究の目的(1)

■ 数理的証明ベースの、不变性の汎用的な解析方法を確立する。

- ◆ 先行研究では、特定の不变性を明らかにする方法である、あるいは定義を示しただけで具体的な導出例がESのみに限定されている

■ 各種アルゴリズムの不变性の有無を明らかにする。

- ◆ 先行研究では、ES、PSO、DE、実数値GA、FAの不变性を解析した例があるが、一部の不变性のみ

■ 更新則の数理的構造と不变性の具備の関係について明らかにする。

- ◆ 更新則の数理的構造毎に不变性の付加アプローチが異なるため、系統的に分類・整理し、後の設計に活かす

本研究の目的(2)

■ PSO、ABC、CSに対して不变性の付加を検討する。

- ◆ 先行研究において、他のアルゴリズムでは適用されているが、対象のアルゴリズムでは事例がないアプローチを流用し、不变性を獲得する

■ 不変性の付加アプローチの課題を指摘する。

- ◆ 単に流用しただけでは、特定の不变性を付加できるが、他の不变性を犠牲にし、一般のアフィン変換不变性を獲得するのは困難であることを指摘する

本研究の目的(3)

■ 高いロバスト性・適応性を有するメタヒューリスティクスの設計指針を提示する。

- ◆ 有効な探索戦略である多様化・集中化に基づき、パラメータ調整則を設計する
- ◆ 不変性に基づき、パラメータ調整則を設計する

■ 上記指針に従い、適応型メタヒューリスティクスの設計を通じて、その可能性・有効性を明らかにする。

- ◆ CSに対して、相似変換不变性を補完するように、パラメータ調整則を付加する
- ◆ 目的(2)で提案した回転不变性を有するPSOに対して、不变性を維持しながら、パラメータ調整則を付加する

本研究の目的(4)

■ (2)、(3)を通じて明らかとなつた、不变性を付加するアプローチの課題解決を検討する。

- ◆ 各アルゴリズム特有のアナロジーに基づき更新則が設計されているため、付加アプローチはそのバリエーションまで考慮されていない

■ アフィン変換不变性を具備するメタヒューリスティクスのフレームワークを構築する。

- ◆ 不变性定理を与えることで、不变性の具備が保証される
- ◆ 様々な最適化能力を自由に付与でき、パラメータ調整則を付加することも可能

非線型最適化で直面する目的関数の性質

非線型の目的関数には、下記の性質がしばしば現れ、解の探索の難易度に影響を与える。

易

单峰性

変数分離可能
(変数間独立)

良スケール性
(良条件)

難

多峰性 (非凸)

変数分離不可能
(変数間依存)

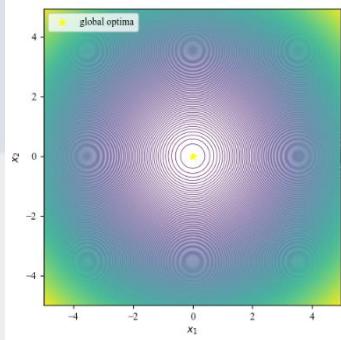
スケール性

BBOでは、目的関数が左右の性質を有しているのかが不明

変数分離可能性（変数間依存性）

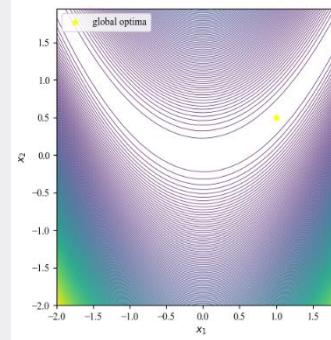
- 目的関数を変数毎に分解することで、多次元の最適化問題を1次元最適化問題へ分離可能な性質。
 - ◆ 座標軸に沿わない方向の探索が困難となりやすい
 - ◆ 変数間独立の問題を回転させても、依存性が発生

変数分離可能（変数間独立）

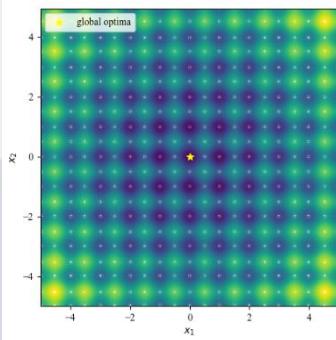


Sphere 関数

変数分離不可能（変数間依存）

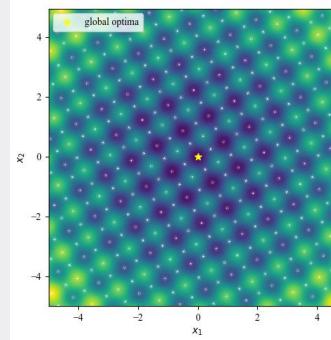


Rosenbrock 関数



Rastrigin 関数

回転

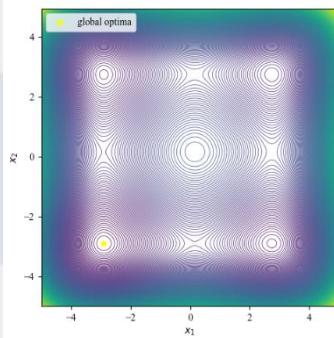
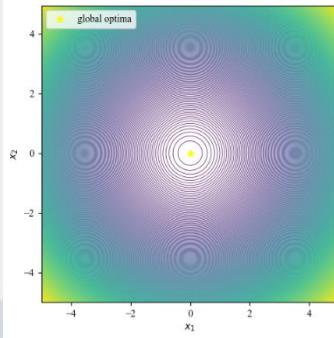


Rastrigin 関数
(30度回転)

悪スケール性（悪条件数）

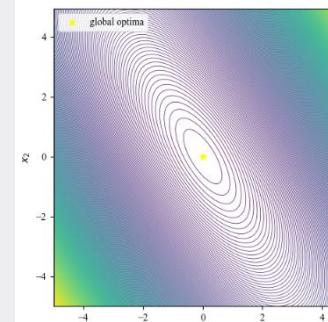
- 変数毎に目的関数への貢献度合いが大きく異なる性質。
- ◆ 全変数に異なるスケールに合わせた探索が困難となりやすい
- ◆ 良スケールの問題をスケーリングすると、悪スケール性が発生

良スケール性（良条件）

 2^N minima 関数

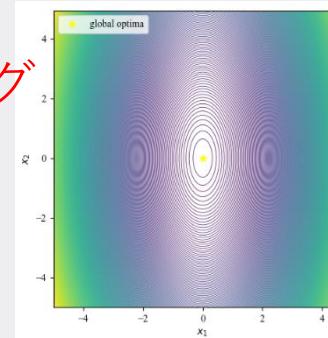
Sphere 関数

悪スケール性（悪条件）



Schwefel 関数

スケーリング

Sphere 関数
($x_2 \mapsto 2x_2$:
条件数を1から4に)

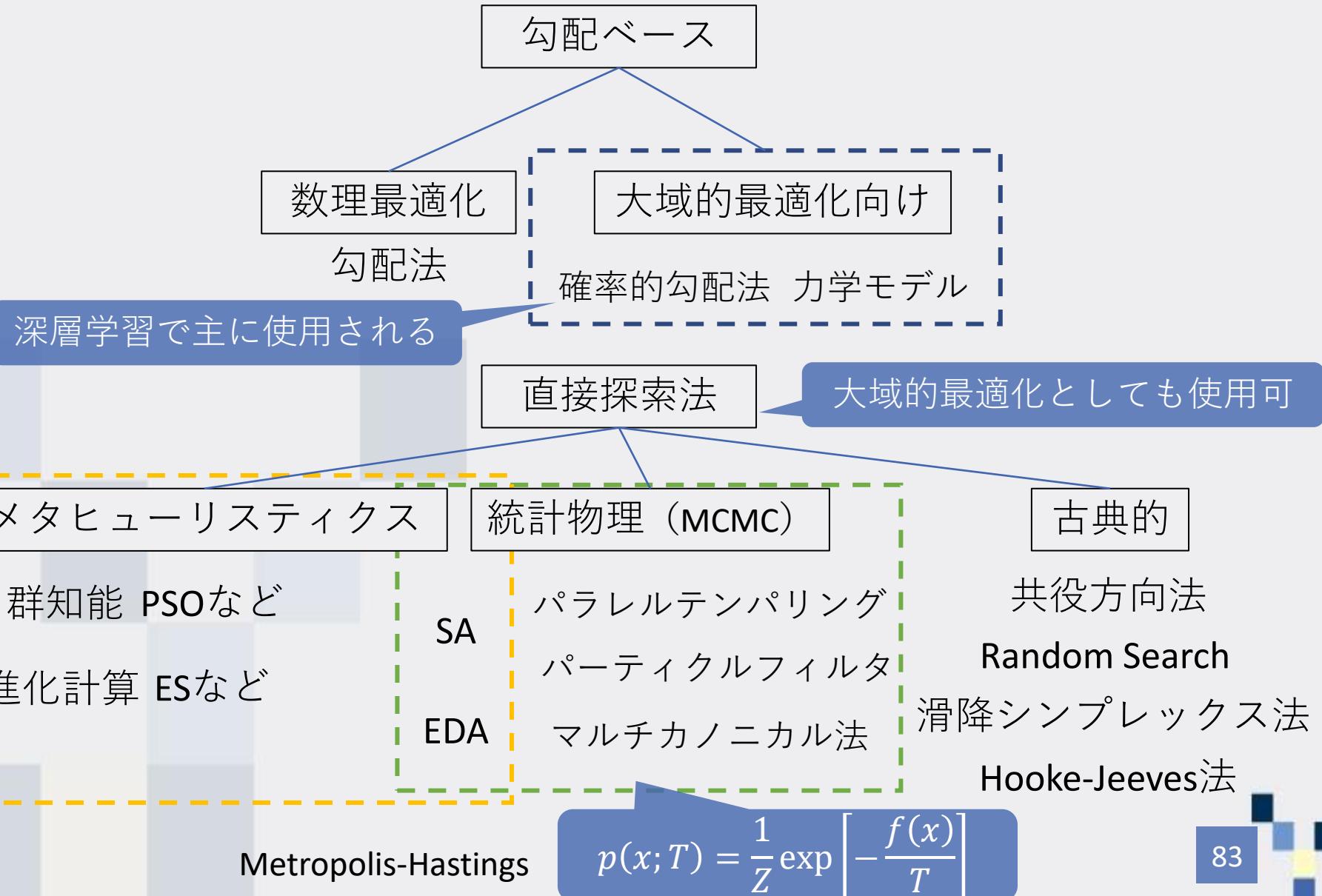
※条件数：目的関数のHesse行列の条件数（最大固有値と最小固有値の比）

数理最適化

- 線型や凸など、一部の問題クラスに特化した最適化アルゴリズム。
- 目的関数の解析的情報*を使用することで、解への収束性や最適性に対して数学的な保証を与える。
- 古典的方法は18世紀以前から存在していたといわれている。
 - ◆ 最急降下法、Newton-Raphson法
- 1960年代以降、軍事用途でのニーズが高まる、コンピュータの開発が進むなどを背景にして、発展系が構築された。
 - ◆ 共役勾配法、準Newton法

*勾配ベクトルやHesse行列などの微分情報を指す

最適化アルゴリズムの分類

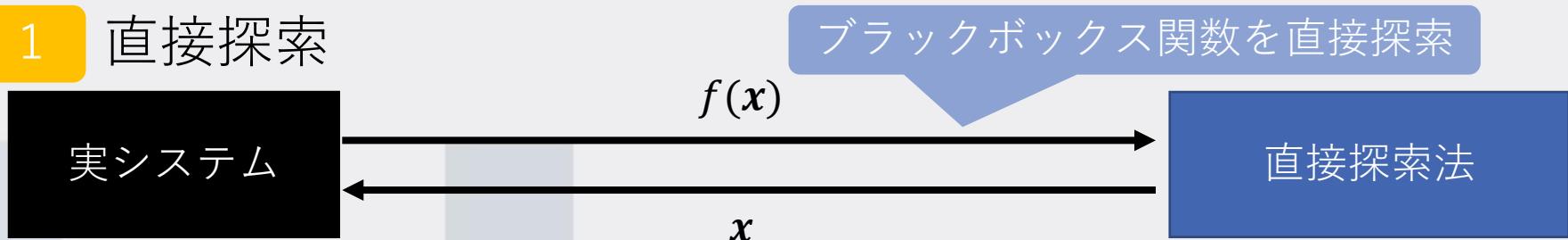


直接探索法を用いたBBOの方法論

直接探索法とモデリング技術を併用すれば、BBOに
対して様々な方法が適用可能

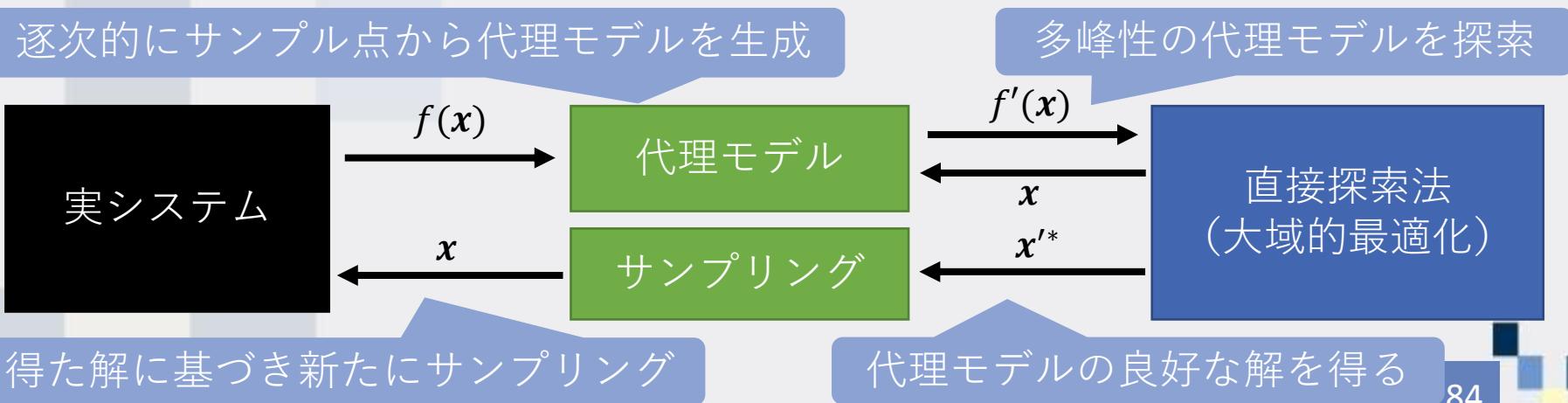
- ◆ これらの発展・融合により、更なる有効な方法論が期待

1 直接探索



2 逐次近似最適化（実験計画法、応答曲面法、Bayes最適化）

逐次的にサンプル点から代理モデルを生成



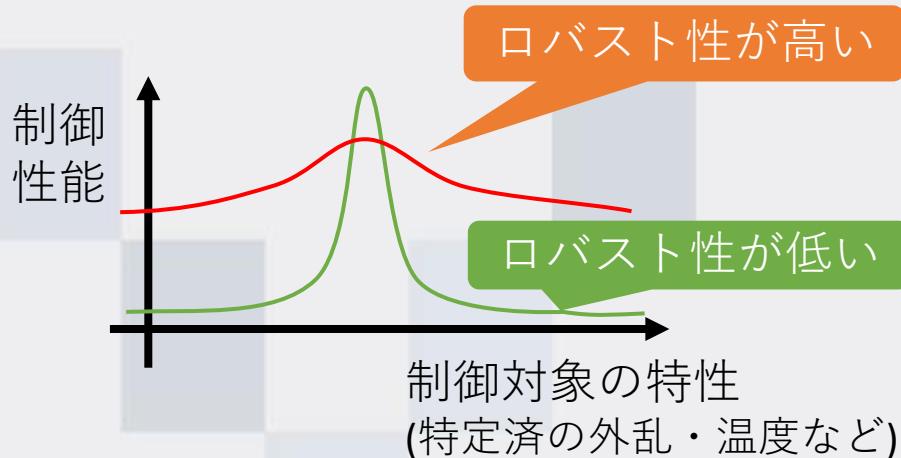
変換不变性によるロバスト化

- **変換不变性**とは、解空間や目的関数の変換に対して、アルゴリズムの性能が再現性を有する理論的な性質。
 - ◆ 同等あるいは類似の結果が得られる
- メタヒューリスティクスが有する変換不变性は、アルゴリズムによって様々である。
 - ◆ 実際、ある変換に対する不变性を欠如するアルゴリズムは、その変換の有無によって、性能が大幅に変化することがよく見られる
- 非線型最適化で直面する目的関数の性質の変化に対応した不变性を具備することは重要。
 - ◆ 解空間の変換：アフィン変換（線型変換 + 平行移動）
 - ◆ 目的関数の変換：単調増加変換

BBOにおけるロバスト性

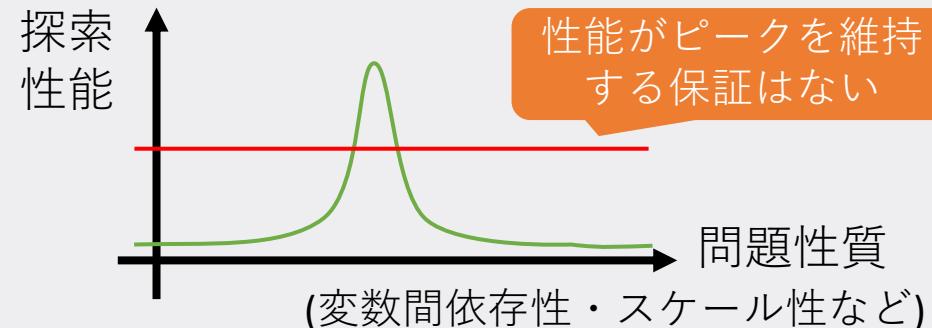
ロバスト性は、パラメータを固定した状態でも、目的関数の構造や性質の変化に対して性能を維持する性質。

ロバスト制御におけるロバスト性
(H_∞ 制御理論など)

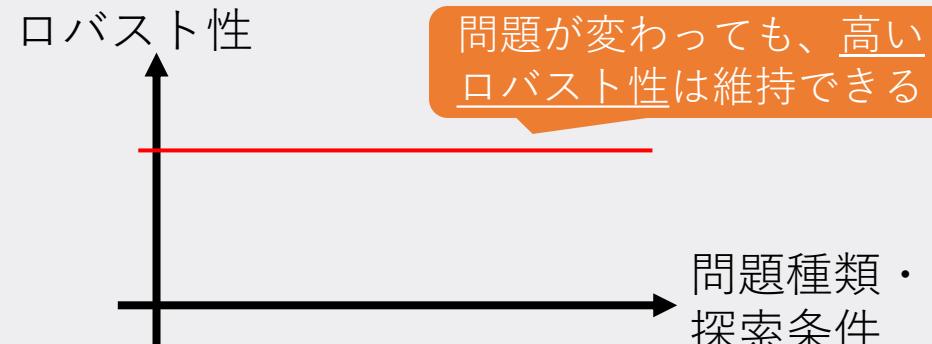


- ※コントローラのパラメータを固定
- ※ロバスト制御では、制御対象・特性が特定済のため、ピーク性能をほぼ維持
- ※制御性能：安定性など

最適化アルゴリズムのロバスト性



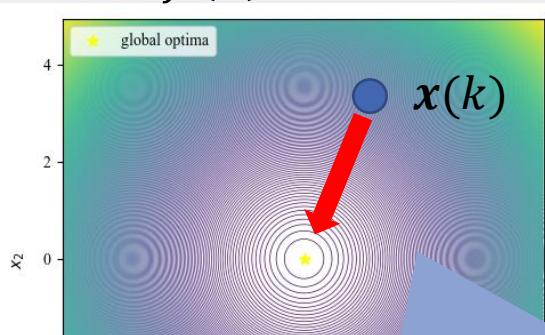
ロバスト性



- ※BBOでは問題種類の汎用性も一定以上考慮すべき

線型変換不変性とロバスト性の関係

標準型 $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$



最急降下法

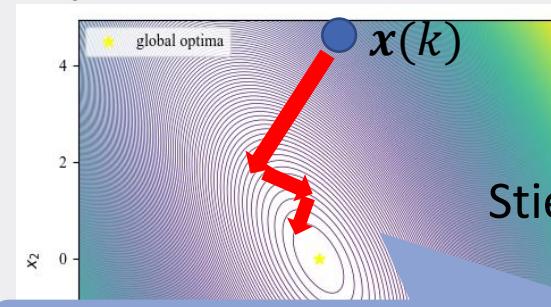
線型変換不
変性を欠如



線型変換に
対してロバ
ストでない

等高線の法線方向に進むため、
理論的に一度で最適解に到達

二次形式 $f(\mathbf{x}) = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{D}^2 \mathbf{B} \mathbf{x}$



線型変換
 $\mathbf{x} \mapsto \mathbf{B}\mathbf{D}\mathbf{x}$



Stiefelの鳥籠

線型変換不変性を有さないため、
最適解到達には複数回必要

$$\mathbf{B}(30^\circ), \mathbf{D} = \text{diag}(1, 2)$$

共役方向法

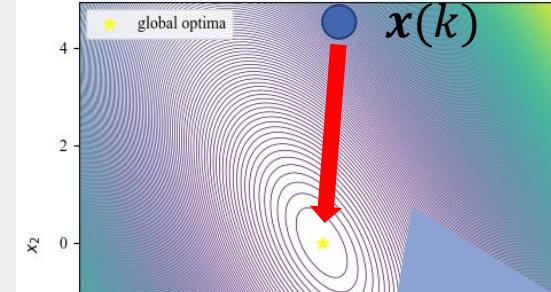
線型変換不
変性を具備



線型変換に
対してロバ
スト

等高線の法線方向に探索したとする
(勾配を使用しないため、厳密には異なる)

線型変換
 $\mathbf{x} \mapsto \mathbf{B}\mathbf{D}\mathbf{x}$



左の共役方向に進むため、
左と同等の効率で探索

線型変換不变性とロバスト性の関係

変換不变性とロバスト性の関係は、数理最適化における結果から推測される。

アルゴリズム	線型変換不变性	2次関数における ロバスト性	より複雑な関数における 性能・ロバスト性
最急降下法	欠如	理論的に線型変換 前後で性能が変動 (ロバストでない)	数値実験的に性能は悪い (ロバストでないから)
共役方向法、 Newton-Raphson法、 共役勾配法	具備	理論的に線型変換 前後で性能を維持 (ロバスト)	数値実験的に性能は良い (ロバストだから)
メタヒューリスティクス	欠如(仮定)	数値実験的に線型 変換前後で性能を ほぼ維持 (ロバスト)	数値実験的にロバスト性 が高く、性能も良いこと を期待

乱数系列が同一であれば、
理論的にロバストである

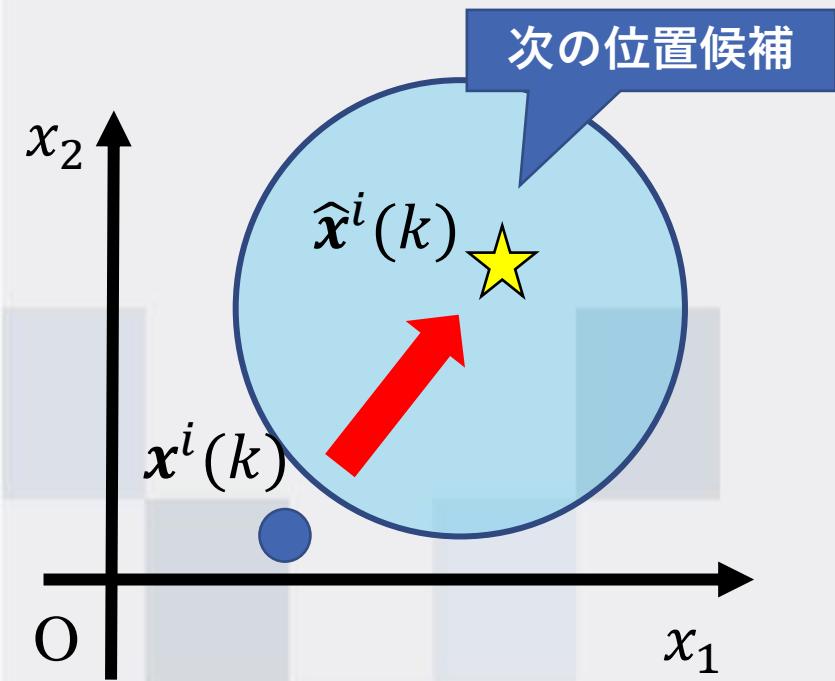
本研究の興味

パラメータ調整による適応化

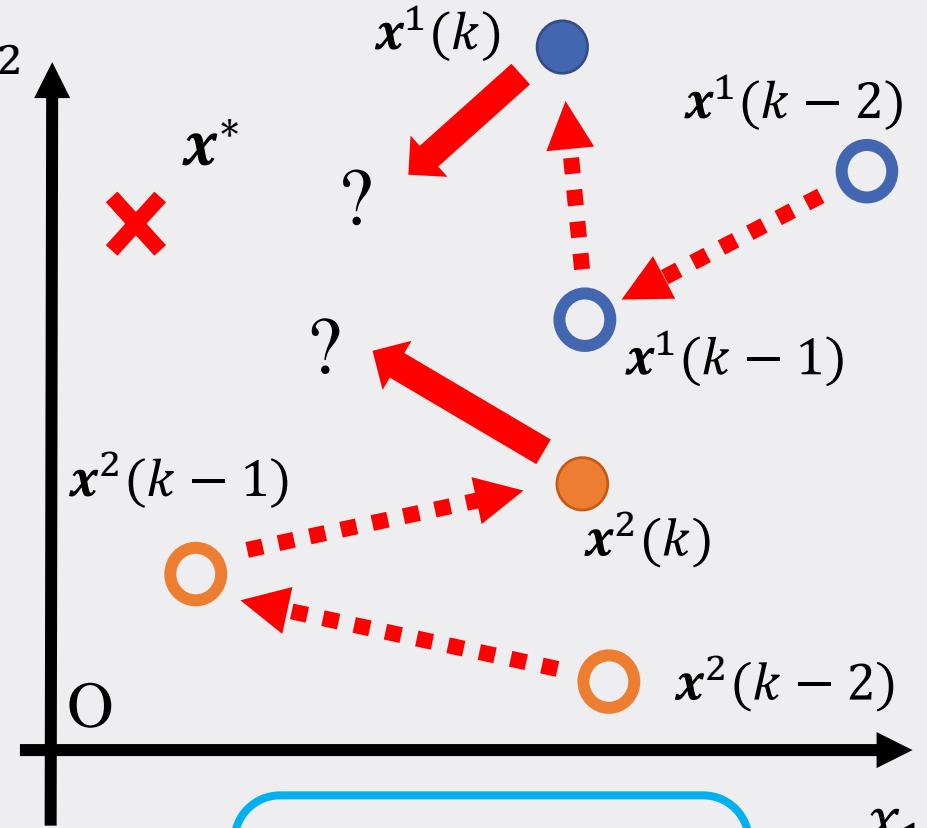
- 適応的調整型のパラメータ調整則は、探索の時間経過に限らず、探索過程で得られた情報を活用して、パラメータを適応的に調整する機構。
 - ◆ 時系列パラメータを逐次生成する「フィードバック型」
- 1980年代から現在まで、ES、PSO、DEなどの多くのアルゴリズムにおいて、多くのパラメータ調整則が精力的に研究してきた。
- パラメータ設定・調整に対する有効な知見を抽出し、系統的に分類・整理することや、パラメータの自律的調整の可能性を検討することは重要。

基本的な探索構造

近傍生成と探索点の移動を繰り返して、探索する。



-  : 探索点 $x^i(k)$
-  : 近傍解 $\hat{x}^i(k)$
-  : 移動ベクトル
-  : 近傍 (乱数に基づく撮動)



-  : 探索点 $x^i(k)$
- k : 反復回数
-  : 最適解 x^*

更新則の定義

定義3.1：更新則の定義

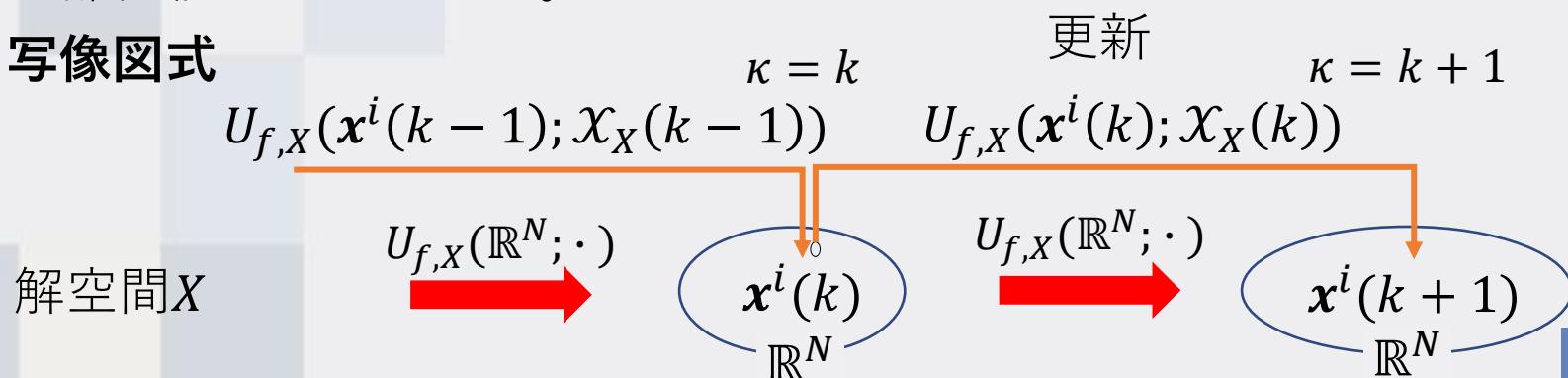
探索点 $\mathbf{x}^i(k)$ の更新則 $U_{f,X}$ は式(3.4)'で定義される。

$$\mathbf{x}^i(k+1) = U_{f,X}(\mathbf{x}^i(k); \mathcal{X}_X(k), \mathcal{C}(k), \mathcal{R}(k)) = U_{f,X}(\mathbf{x}^i(k); \mathcal{X}_X(k)) \text{ 式(3.4)'} \quad \text{簡潔な表記}$$

ただし、 $\mathcal{X}_X(k)$ は探索履歴、 $\mathcal{C}(k)$ は更新用パラメータ集合、 $\mathcal{R}(k)$ は確率変数用パラメータ集合である。 $r_j(k) \in \mathcal{R}(k)$ は $r_j(k) \sim \mathcal{P}_j$ 、 \mathcal{P}_j は任意の確率分布である。

※写像の引数ではパラメータとして扱うが、実際の更新則で使用する疑似乱数を表しており、確率分布 \mathcal{P}_j によって、スカラ・ベクトル、実数値・離散値などが定まる。

写像図式



近傍生成則と探索点の移動則の定義

定義3.2：近傍生成則と探索点の移動則の定義

探索点 $\mathbf{x}^i(k)$ の更新則 $U_{f,X}$ は近傍生成則式(3.5)と探索点の移動式(3.6)を用いて式(3.4)'で書き換えられる。

$$\mathbf{x}^i(k+1) = U_{f,\text{mov}} \circ U_{f,\text{nei}}(\mathbf{x}^i(k); \mathcal{X}_X(k)) \quad \text{式(3.4)'}$$

探索点の移動

$$\mathbf{x}^i(k+1) = U_{f,\text{mov}}(\hat{\mathbf{x}}^i(k); \mathcal{X}_X(k)) \quad \text{式(3.5)}$$

近傍生成

$$\hat{\mathbf{x}}^i(k) = U_{f,\text{nei}}(\mathbf{x}^i(k); \mathcal{X}_X(k)) \quad \text{式(3.6)}$$

探索点の移動 $U_{f,\text{mov}}$

近傍生成 $U_{f,\text{nei}}$

絶対移動

$$\mathbf{x}^i(k+1) = \hat{\mathbf{x}}^i(k)$$

アルゴリズム毎にバリエーションがある

改善移動

$$\mathbf{x}^i(k+1) = \begin{cases} \hat{\mathbf{x}}^i(k), f(\hat{\mathbf{x}}^i(k)) < f(\mathbf{x}^i(k)) \\ \mathbf{x}^i(k), \text{otherwise} \end{cases}$$

確率的移動

$$\mathbf{x}^i(k+1) = \begin{cases} \hat{\mathbf{x}}^i(k), f(\hat{\mathbf{x}}^i(k)) < f(\mathbf{x}^i(k)) \vee P > \theta \\ \mathbf{x}^i(k), \text{otherwise} \end{cases}, \theta \sim \mathcal{P}$$

各種メタヒューリスティクス

アルゴリズム	探索構造の特徴
Evolution Strategy (ES)	正規分布に基づくランダムウォーク
実数値Genetic Algorithm (GA)	探索点分布内へのサンプリング
実数値Simulated Annealing (SA)	正規分布に基づくランダムウォーク・確率的移動
Particle Swarm Optimization (PSO)	良い解に向かう差分ベクトルと慣性項による摂動の線型結合
Differential Evolution (DE)	探索点間の差分ベクトル・解の組合せ
Artificial Bee Colony Algorithm (ABC)	探索点間の差分ベクトル・解の組合せ・停滯探索点のランダム配置
Firefly Algorithm (FA)	良い探索点に向かう差分ベクトルと乱数ベクトルの線型結合
Cuckoo Search (CS)	安定分布に基づくランダムウォーク・探索点間の差分ベクトル

多様化・集中化の定義

メタヒューリスティクスの探索構造の解析を経て、多様化・集中化の操作と効果を定義する。

項目	多様化	集中化
目的	長期的な解の改善	短期的な解の改善
効果・現象	探索領域／探索点分布の拡大	探索領域／探索点分布の縮小
操作①	特定領域／方向への指向性抑制	特定領域／方向への指向性促進
操作②	摂動／分布の拡張	摂動／分布の縮小・限定

更新則の構成要素

更新則の構成要素として、下記が挙げられる。

1 探索点などの位置ベクトル

探索点群の各位置ベクトル、探索点の位置履歴（探索履歴）

2 スカラ乱数、乱数ベクトルなど

ある確率分布に従って生成されるスカラ乱数・乱数ベクトル

3 目的関数值

探索点群の目的関数值、探索過程で得た目的関数值

4 上記要素を持つ行列

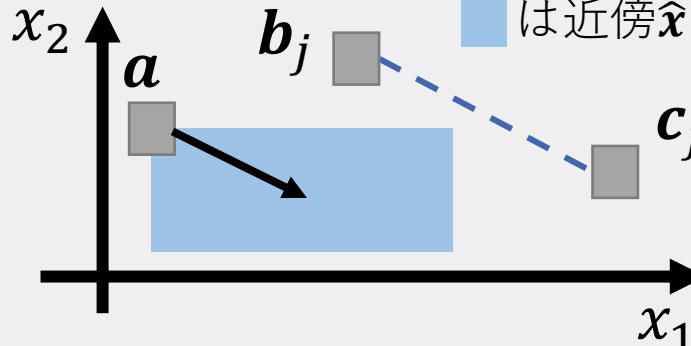
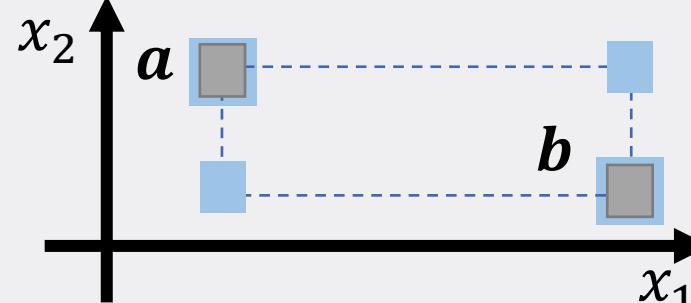
5 単項演算子の集合（ベクトル演算のため）

算術演算子（四則演算子） $\{+, -, \times, \div\}$

単純な非線型関数： $\{\sqrt{\cdot}, (\cdot)^n, \exp(\cdot), \log(\cdot)\}$ など

更新則の分類

更新則の演算形式は、下記の3種類に大別できる。

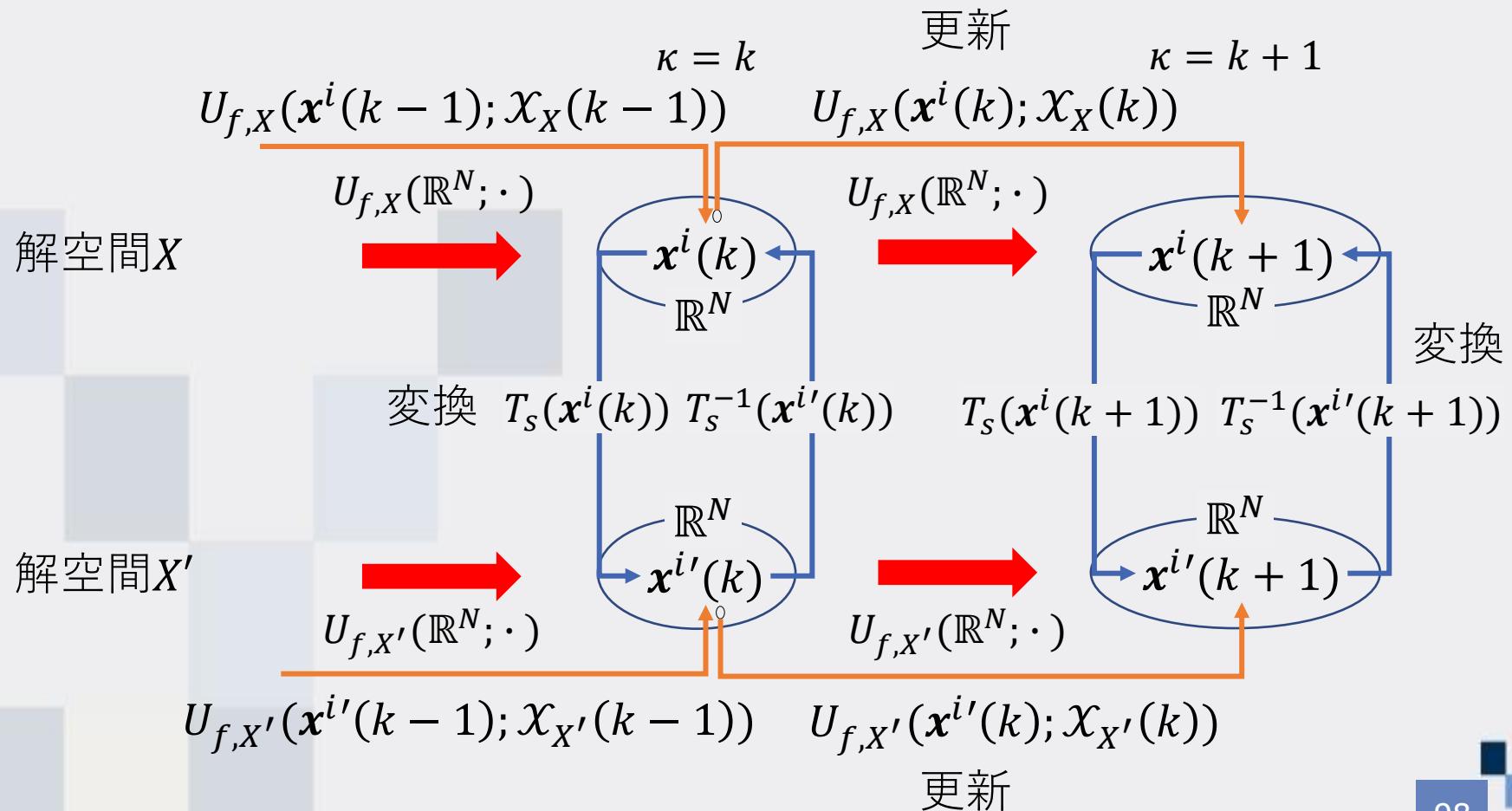
項目	イメージ	演算形式
(1) 基本ベクトルと 差分ベクトルの 線形結合		$\hat{x} = a + \sum_j A_j(b_j - c_j)$ 式(3.48) a, b_j, c_j は位置ベクトル
(2) 解同士の 組合せ		$\hat{x} = M_1 a + M_2 b$ 式(3.49) $M_1 + M_2 = I$ 式(3.50) M_1, M_2 は対角行列 $M_{1,n} \in \{0,1\}$
(3) ランダム ウォーク		$\hat{x} = a + As$ 式(3.51) s は乱数ベクトル

各種アルゴリズムの多様化・集中化

手法	多様化の操作	集中化の操作	探索戦略の実現方法
(1+1)-ES	乱数ベクトルによる摂動	改善移動	改善頻度に基づく摂動スケールの増減
実数値GA	統計量の遺伝・世代交代モデル	探索点分布内への近傍生成・世代交代モデル	探索過程で分布を縮小
実数値SA	乱数ベクトルによる摂動・確率的な悪化の許容	基本的に改善移動	経過時間に基づく悪化確率の減少
PSO	慣性項・差分ベクトルの係数で乱数使用	最良解に向かう差分ベクトル	弱い安定状態
DE	解の組合せで乱数の使用・探索点のランダム選出	探索点間の差分ベクトル・改善移動	差分ベクトルによる探索状態の反映
ABC	解の組合せで乱数の使用・探索点のランダム選出 ・有望でない探索点の再配置	探索点間の差分ベクトル・改善移動・優れた探索点に重みづけした選択	差分ベクトルによる探索状態の反映
FA	乱数ベクトルによる摂動・探索点間の距離による係数の減衰	より優れた探索点に向かう差分ベクトル	多様化・集中化の同時実現・経過時間に基づく摂動スケールの減少
CS	乱数ベクトルによる摂動・探索点のランダム選出・差分ベクトルの係数で乱数使用	探索点間の差分ベクトル・改善移動	多様化・集中化の同時実現

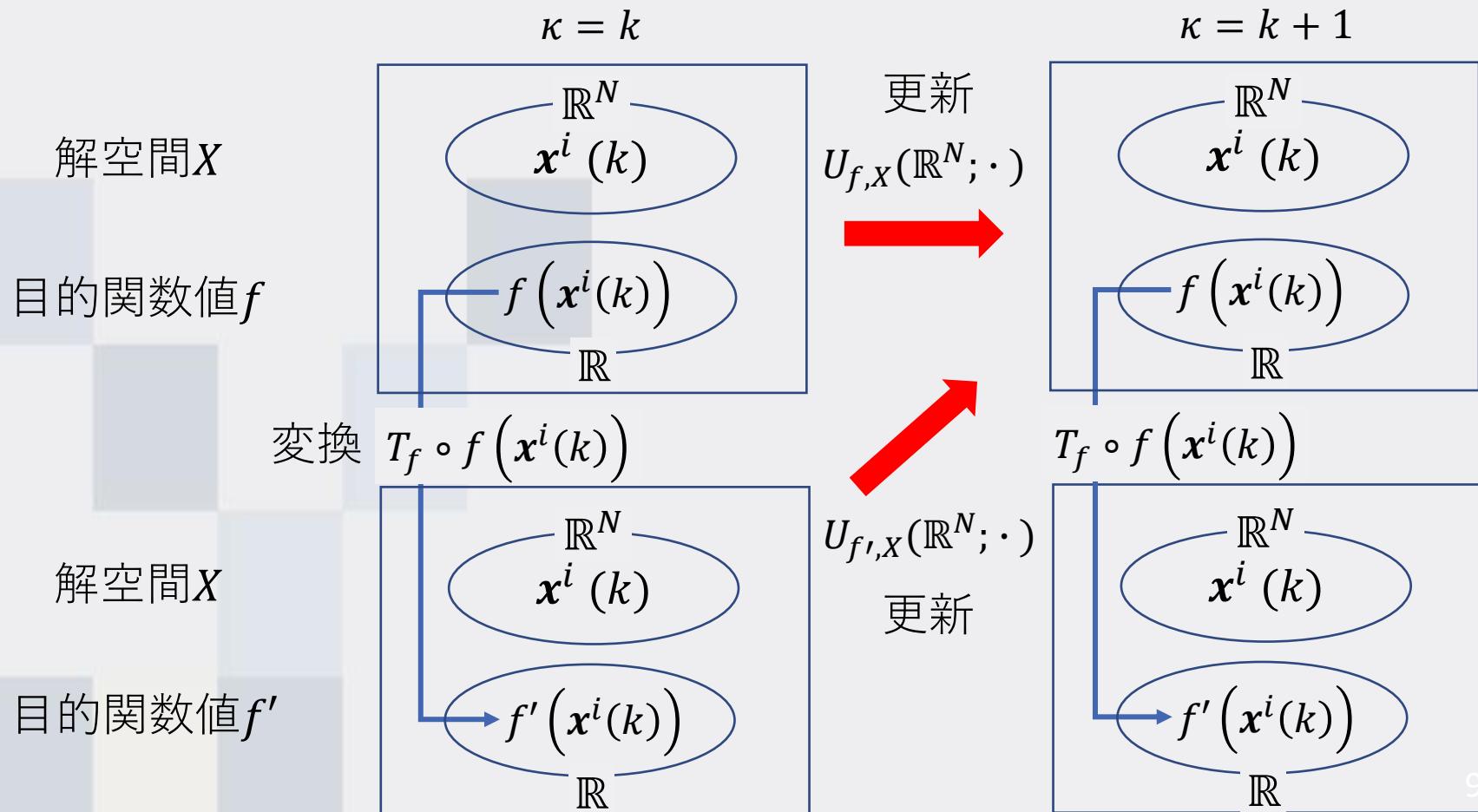
解空間の変換不変性の定義

更新則 $U_{f,X}$ が解空間 T_s, T_s^{-1} に対して不変であるとは、下記の可換図式が $\forall k \in \mathbb{N}$ について成立することである。



目的関数値の変換不変性の定義

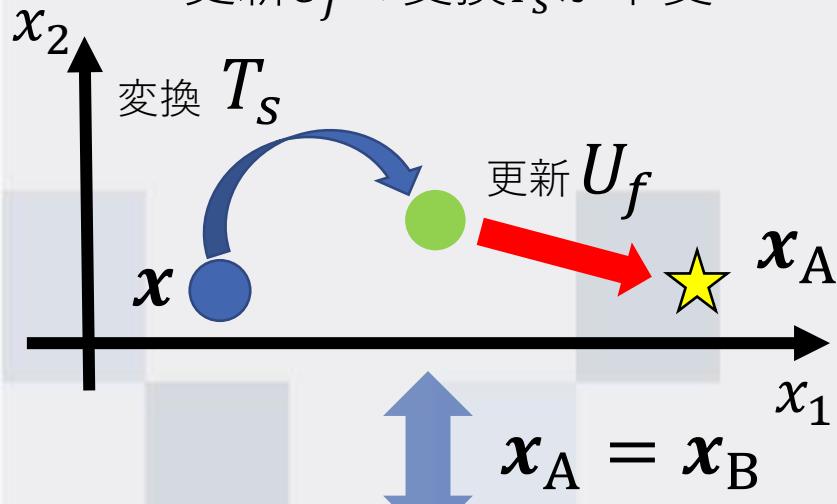
更新則 $U_{f,X}$ が空間 T_f に対して不変であるとは、下記の可換図式が $\forall k \in \mathbb{N}$ について成立することである。



変換不変性の定義の解釈

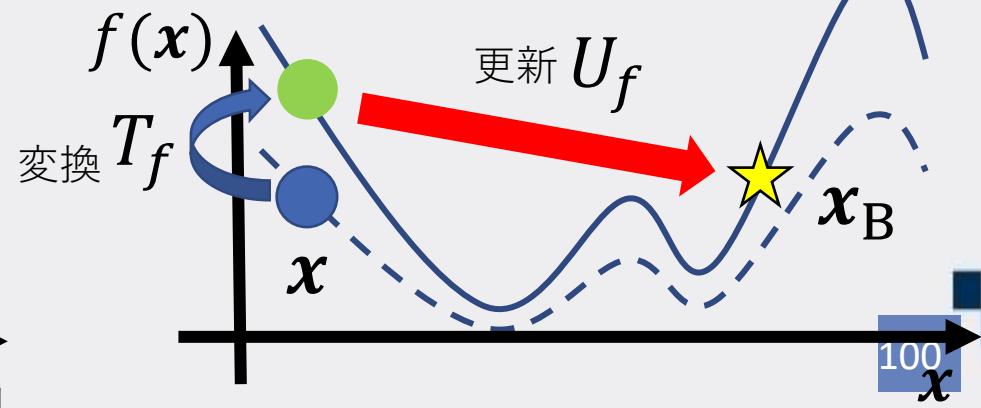
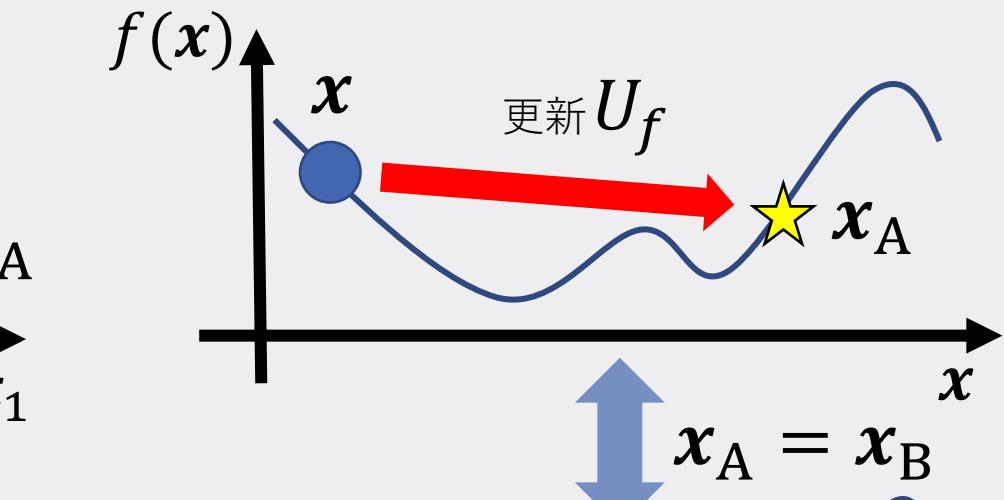
解空間の変換 T_s

変換 $T_s \Rightarrow$ 更新 U_f と、
更新 $U_f \Rightarrow$ 変換 T_s が不变



目的関数値の変換 T_f

更新 U_f と、変換 $T_f \Rightarrow$ 更新 U_f が不变



各種不变性

下記の各変換に対して、不变性を定義できる。

種類	変換	補足
相似変換不变性	$T_s: \mathbf{x} \mapsto \delta \mathbf{x}, \forall \delta > 0$	各座標軸に等しいスケーリング
スケール不变性	$T_s: \mathbf{x} \mapsto \mathbf{D}\mathbf{x}, \forall \mathbf{D}$ は対角行列	各座標軸に異なるスケーリング 悪スケール性 を再現する 相似変換を包含
平行移動不变性	$T_s: \mathbf{x} \mapsto \mathbf{x} - \mathbf{t}, \forall \mathbf{t}$	原点依存性 を再現する
回転不变性	$T_s: \mathbf{x} \mapsto \mathbf{B}\mathbf{x}, \forall \mathbf{B}$ は直交行列	変数間依存性 を再現する
線型変換不变性	$T_s: \mathbf{x} \mapsto \mathbf{C}\mathbf{x}, \forall \mathbf{C}$ は正則行列	スケール変換と回転変換を包含
アフィン変換不变性	$T_s: \mathbf{x} \mapsto \mathbf{C}\mathbf{x} - \mathbf{t},$ $\forall \mathbf{C}$ は正則行列, $\forall \mathbf{t}$	線型変換と平行移動の合成変換 (<u>上記変換を全て包含</u>)
単調増加変換不变性 ※	$T_f: f \mapsto g \circ f,$ $\forall g$ は単調増加変換	単調増加変換を再現する

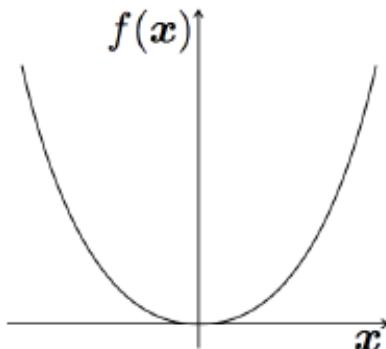
※単調増加変換は目的関数 f に対する変換

目的関数の単調増加変換不変性の意義

目的関数の単調増加変換不変性を具備することは、下記を同等効率で探索可能（ロバスト性が高い）。

目的関数の単調増加変換

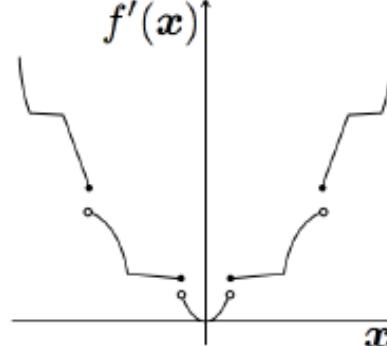
凸・連続



(a) f

最適化で容易
な性質

非凸（单峰）・不連續



(b) $f' = h \circ f$

最適化で困難・
必須な性質

BBOではこれらの性質に対して、性能が依存しないことが望ましい

※勾配法やBayes最適化はこれを有さないため、この性質変化に強く依存する

変換不変性の証明に関する先行研究

本研究は先行研究に比べて、汎用性が高い、網羅的に適用した点で異なる。

文献	類似点	相違点
Hansen[6]	不変性の定義と アルゴリズムへの適用	<ul style="list-style-type: none"> ➤ ES系への適用例しかない ➤ 不変性の定義が曖昧 ➤ 回転・相似変換・平行移動不変性しか定義していない ➤ PSOへの適用例しかない
Wilke[7]	不変性の定義と アルゴリズムへの適用	<ul style="list-style-type: none"> ➤ ES系への適用例しかない ➤ 不変性の定義が曖昧 ➤ 回転・相似変換・平行移動不変性しか定義していない ➤ PSOへの適用例しかない

各手法の変換不变性の有無に関する先行研究

先行研究では一部の不变性・手法に対して、実験的あるいは理論的に不变性の有無を確認している。

手法	指摘した不变性	実験的	理論的
ES	相似変換不变性の欠如	—	Hansen[6], 金政[11]
CMA-ES	アフィン変換不变性の具備	—	Hansen[6]
実数値SA	相似変換不变性の欠如	—	本研究
実数値GA UNDX,BLX- α	回転不变性・スケール不变性の欠如	Tsutsui[19], 坂井[20]	Ras[21]
PSO	回転不变性の欠如	Hansen[22]	Wilke[7]
DE	回転不变性の欠如	Sutton[23]	金政[11], Sakai[24]
ABC	回転不变性の欠如	—	Mori[13], Hariya[25]
FA	相似変換不变性の欠如	—	Cleghorn[26]
CS	回転不变性・相似変換不变性の欠如	—	本研究

※他にもあるが、学位論文を参考されたい。

※本研究は実数値GA以外理論的に示している。

変換不变性の付加に関する先行研究

先行研究では一部の手法に対して、一部の不变性を付加し、実験的あるいは理論的に確認している。

手法	付加した不变性	実験的	理論的
ES	アフィン変換不变性 相似変換不变性	— 金政[11](多点型ESに調整則)	Hansen[6](CMA-ES) —
実数値 GA	アフィン変換不变性 回転不变性	Tsutsui[19](SPX), 坂井[20](BLX- α) —	— Ras[21]
PSO	アフィン変換不变性 回転不变性	Raquet[8] Clerk[9]	— Wilke[7]、本研究
DE	回転不变性	Guo[27]	金政[11], Sakai[24]
ABC	回転不变性	—	Mori[13]、本研究
CS	アフィン変換不变性	—	本研究

※他にもあるが、学位論文を参考されたい。

不变性の付加に関する先行研究

先行研究では一部の手法に対して付加しているが、他の手法ではまだなされていない事例がある。
乱数ベクトル無

手法	係数をスカラ	超球	係数に回転行列
実数値GA	Tsutsui[19](SPX)	Ras[21](BLX- α)	坂井[20](BLX- α)
PSO	Raquet[8]	Clerk[9]	Wilke[7]、本研究
DE	—※1	金政[11]	Sakai[24], Guo[27]
ABC	—※1	本研究	Mori[13]

ランダムウォーク

手法	回転行列と対角行列
ES	Hansen[6](CMA-ES)
CS	本研究

※1：解の組合せをスカラにすると、多様性が極端に低下する。

※2：他にもあるが、学位論文を参考されたい。

不变性に関する重要な定理

更新則に探索とは独立な対角行列やベクトルが含まれると、何らかの不变性を失うケースが多い。

- ◆ 探索とは独立した量は変換に対しても独立だから

PSO 亂数から構成される対角行列が係数に作用

$$\boldsymbol{x}^i(k+1) = \boldsymbol{x}^i(k) + w(\boldsymbol{x}^i(k) - \boldsymbol{x}^i(k-1)) + c_1 \mathbf{R}_1 (\boldsymbol{p}^i(k) - \boldsymbol{x}^i(k)) + c_2 \mathbf{R}_2 (\boldsymbol{p}^g(k) - \boldsymbol{x}^i(k))$$

DE バイナリ値から構成される対角行列が係数に作用

$$\hat{\boldsymbol{x}}^i(k) = \mathbf{M}^1 \boldsymbol{u}^i(k) + \mathbf{M}^2 \boldsymbol{x}^i(k)$$

ES 亂数ベクトルを単独で使用

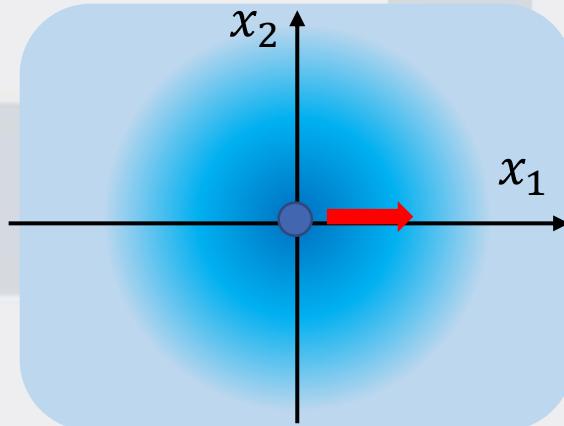
$$\hat{\boldsymbol{x}}(k+1) = \boldsymbol{x}(k) + \sigma \boldsymbol{s}(k)$$

不变性に関する重要な定理

正規乱数は標準正規乱数をアフィン変換すれば得られる（定理1）。共分散行列を適切に与えれば、アフィン変換された正規乱数が得られる。

◆CMA-ESなどで見られるテクニック

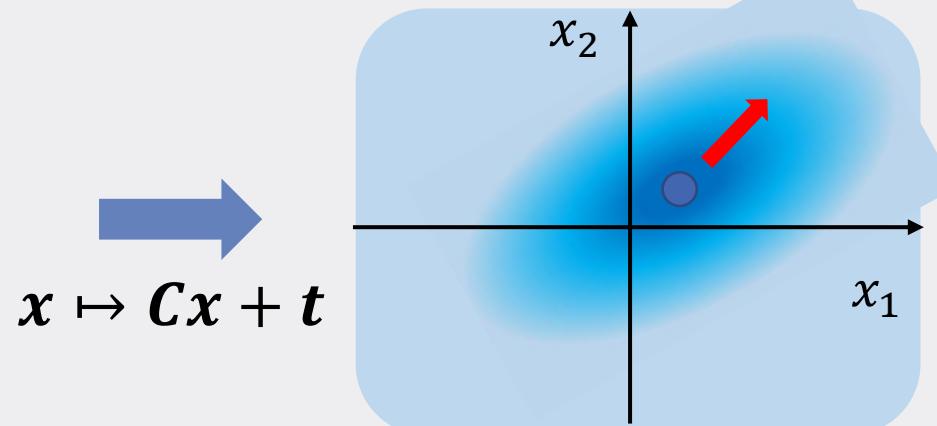
同時確率分布は等方・直交
(等方性なら回転不变性を持つ)



標準正規乱数

$$\mathbf{s}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

スケーリングかつ回転により、
異方性・斜交となる



正規乱数

$$\mathbf{s} \sim \mathcal{N}(\mathbf{t}, \mathbf{CC}^T)$$

不变性に関する重要な定理

探索とは独立な量に作用させれば変換不变性を付加できる

補助定理3

探索履歴に含まれる解集合を対象とした共分散行列 Σ の非対称平方根は式(4.2)と表される。

$$\Sigma_*^{1/2} = P \Lambda^{1/2} \quad \text{式(4.2)}$$

$P = [a_1, \dots, a_N]^T$ は基底変換行列、 Λ は固有値行列、
 $\{(a_1, \lambda_1), \dots, (a_N, \lambda_N)\}$ は Σ の固有ベクトルと固有値である。

※ A の非対称平方根とは、 $A = A^{1/2}(A^{1/2})^T$ を満たす $A^{1/2}$ のこと

補助定理4

$\Sigma_*^{1/2}$ はアフィン変換に対して、式(4.3)と変換される。

$$\Sigma_*^{1/2'} = C \Sigma_*^{1/2} \quad \text{式(4.3)}$$

アフィン変換は $x \mapsto Cx - t$ である。

共分散行列の非対称平方根

補助定理3

共分散行列の非対称平方根は $\Sigma_*^{1/2} = \mathbf{P}\Lambda^{1/2}$ と表せる。ただし、 $\mathbf{P} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N]$ は基底変換行列、 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ は固有値行列、 $\{(\mathbf{a}_1, \lambda_1), (\mathbf{a}_2, \lambda_2), \dots, (\mathbf{a}_N, \lambda_N)\}$ は $\Sigma(k)$ の固有ベクトルと固有値である。

証明

共分散行列の非対称平方根の定義は、式(A1)を満たす $\Sigma_*^{1/2}$ である。

$$\Sigma = \Sigma_*^{1/2} (\Sigma_*^{1/2})^T \quad \text{式(A.35)}$$

また、 Σ は実対称行列であるため、 $\Sigma = \mathbf{P}\Lambda\mathbf{P}^T$ と対角化できる。
これを変形すると、式(A.36)となる。

$$\Sigma = \mathbf{P}\Lambda^{1/2}\Lambda^{1/2}\mathbf{P}^T = \mathbf{P}\Lambda^{1/2}(\Lambda^{1/2})^T\mathbf{P}^T = \mathbf{P}\Lambda^{1/2}(\mathbf{P}\Lambda^{1/2})^T \quad \text{式(A.36)}$$

式(A.35)と式(A.36)を恒等的に比較すると、 $\Sigma_*^{1/2} = \mathbf{P}\Lambda^{1/2}$ が得られる。 ■

回転不变性の付加アプローチ

本研究では、先行研究のアプローチと対象手法の組合せが新しい。

探索にあまり
有効でない

手法	係数を スカラ	超球	係数に回転行列		
			ランダム	Gram-Schmidtによる固有ベクトル	共分散行列の固有ベクトル
実数値GA	Tsutsui[19]	Ras[21]	—	—	坂井[20](BLX- α)
PSO	Raquet[8]	Clerk[9]	Wilke[7]	—	本研究
DE	—※1	金政[11]	—	Sakai[24]	Guo[27]
ABC	—※1	本研究	—	—	Mori[13]
ES	—※2	—※2	Schwefel[28]	—	Hansen[6](アフィン)
CS	—※2	—※2	—	—	本研究(アフィン)

※1：解の組合せをスカラにすると、多様性が極端に低下する。

※2：ランダムウォークなので不可能。

ABCの回転不変性欠如の証明

定理10

ABCの更新則 U_f は回転不変性を有さない。

証明

回転変換 $T_s: \mathbf{x} \mapsto \mathbf{Bx}$ に対して、 $\mathbf{x}^{i'} = \mathbf{Bx}^i, \mathbf{x}^{r'} = \mathbf{Bx}^r$ となる。
このとき、 \mathbf{u}^i は、

$$\begin{aligned}\mathbf{u}^{i'} &= \mathbf{x}^{i'} + \phi(\mathbf{x}^{r'} - \mathbf{x}^{i'}) = \mathbf{Bx}^i + \phi(\mathbf{Bx}^r - \mathbf{Bx}^i) = \mathbf{Bx}^i + \phi\mathbf{B}(\mathbf{x}^r - \mathbf{x}^i) \\ &= \mathbf{B}(\mathbf{x}^i + \phi(\mathbf{x}^r - \mathbf{x}^i)) = \mathbf{Bu}^i\end{aligned}$$

となるから、 $\mathbf{u}^{i'} = \mathbf{Bu}^i$ となる。このとき、不变性の条件式は、

$$\begin{aligned}(U_f \circ T_s)(H; C) &= \mathbf{M}^1 \mathbf{x}^{i'} + \mathbf{M}^2 \mathbf{u}^{i'} \\ &= \mathbf{M}^1 \mathbf{Bx}^i + \mathbf{M}^2 \mathbf{Bu}^i \\ &\neq \mathbf{B} \mathbf{M}^1 \mathbf{x}^i + \mathbf{B} \mathbf{M}^2 \mathbf{u}^i \\ &= \mathbf{B}(\mathbf{M}^1 \mathbf{x}^i + \mathbf{M}^2 \mathbf{u}^i) \\ &= \mathbf{B} \hat{\mathbf{x}}^i = (T_s \circ U_f)(H; C)\end{aligned}$$

となる。計算過程で $\mathbf{M}^\ell \mathbf{B} \neq \mathbf{B} \mathbf{M}^\ell$ となることを利用している。
したがって、題意が示された。

CSの回転不変性欠如の証明

定理16

CSの更新則 U_f は回転不変性を有さない。

証明

回転変換 $T_s: \mathbf{x} \mapsto \mathbf{Bx}$ に対して、 $\mathbf{x}^{i'} = \mathbf{Bx}^i$ となる。

このとき、 \mathbf{s}^i は安定分布 \mathcal{L} に従う乱数ベクトルであるため、

回転しても同一のベクトルでない。よって、 $\mathbf{s}^{i'} \neq \mathbf{Bs}^i$ である。

このとき、Global Searchにおける不变性の条件式は、

$$\begin{aligned}
 (U_f \circ T_s)(H; C) &= \mathbf{x}^{i'} + \alpha \mathbf{s}^{i'} \\
 &= \mathbf{Bx}^i + \alpha \mathbf{s}^{i'} \\
 &\neq \mathbf{Bx}^i + \alpha \mathbf{Bs}^i \\
 &= \mathbf{B}(\mathbf{x}^i + \alpha \mathbf{s}^i) \\
 &= \mathbf{B}\hat{\mathbf{x}}^i = (T_s \circ U_f)(H; C)
 \end{aligned}$$

となる。したがって、題意が示された。

CSの相似変換不変性欠如の証明

定理17

CSの更新則 U_f は相似変換不変性・スケール不変性を有さない。

証明

相似変換 $T_s: \mathbf{x} \mapsto \delta\mathbf{x}$ に対して、 $\mathbf{x}^{i'} = \delta\mathbf{x}^i$ となる。

このとき、 \mathbf{s}^i は安定分布 \mathcal{L} に従う乱数ベクトルであるため、

相似変換しても同一のベクトルでない。よって、 $\mathbf{s}^{i'} \neq \delta\mathbf{s}^i$ である。

このとき、Global Searchにおける不变性の条件式は、

$$\begin{aligned}
 (U_f \circ T_s)(H; C) &= \mathbf{x}^{i'} + \alpha\mathbf{s}^{i'} \\
 &= \delta\mathbf{x}^i + \alpha\mathbf{s}^{i'} \\
 &\neq \delta\mathbf{x}^i + \alpha\delta\mathbf{s}^i \\
 &= \delta(\mathbf{x}^i + \alpha\mathbf{s}^i) \\
 &= \delta\hat{\mathbf{x}}^i = (T_s \circ U_f)(H; C)
 \end{aligned}$$

となる。よって、相似変換不変性を有さないことが示された。

さらに、相似変換はスケール変換 $T_s: \mathbf{x} \mapsto \mathbf{Cx}$ の対角行列 \mathbf{C} が $\mathbf{C} = \delta\mathbf{I}$ とした場合である。このため、相似変換不変性を有さないならば、スケール不変性も有さない。したがって、題意が示された。

定理19

CRI-PSOの更新則 U_f は回転不変性・相似変換不変性・平行移動不変性を有する。

証明

合成変換 $T_s: \mathbf{x} \mapsto \delta\mathbf{B}\mathbf{x} - \mathbf{t}$ に対して、 $\mathbf{x}^{i'}(k) = \delta\mathbf{B}\mathbf{x}^i(k) - \mathbf{t}, \mathbf{p}^{i'}(k) = \delta\mathbf{B}\mathbf{p}^i(k) - \mathbf{t}, \mathbf{p}^{g'}(k) = \delta\mathbf{B}\mathbf{p}^g(k) - \mathbf{t}$ となる。

また、 $\mathbf{v}^{i'}(k) = \mathbf{x}^{i'}(k) - \mathbf{x}^{i'}(k-1) = \delta\mathbf{B}\mathbf{v}^i(k)$ である。さらに、 $\mathbf{P}' = \mathbf{BP}$ である。

このとき、不变性の条件式は、

$$\begin{aligned}
 & (U_f \circ T_s)(H(k); C(k)) \\
 &= \delta\mathbf{B}(\mathbf{x}^i(k) + w\mathbf{v}^i(k)) + c_1 \mathbf{B} \mathbf{P} \mathbf{R}_1 (\mathbf{B} \mathbf{P})^T (\delta\mathbf{B}\mathbf{p}^i(k) - \delta\mathbf{B}\mathbf{x}^i(k)) \\
 &\quad + c_2 \mathbf{B} \mathbf{P} \mathbf{R}_2 (\mathbf{B} \mathbf{P})^T (\delta\mathbf{B}\mathbf{p}^g(k) - \delta\mathbf{B}\mathbf{x}^i(k)) - \mathbf{t} \\
 &= \delta\mathbf{B}(\mathbf{x}^i(k) + w\mathbf{v}^i(k) + c_1 \mathbf{P} \mathbf{R}_1 \mathbf{P}^T \mathbf{B}^T \mathbf{B}(\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{P} \mathbf{R}_2 \mathbf{P}^T \mathbf{B}^T \mathbf{B}(\mathbf{p}^g(k) - \mathbf{x}^i(k))) - \mathbf{t} \\
 &= \delta\mathbf{B}(\mathbf{x}^i(k) + w\mathbf{v}^i(k) + c_1 \mathbf{P} \mathbf{R}_1 \mathbf{P}^T (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 \mathbf{P} \mathbf{R}_2 \mathbf{P}^T (\mathbf{p}^g(k) - \mathbf{x}^i(k))) - \mathbf{t} \\
 &= \mathbf{B}\mathbf{x}^i(k+1) = (T_s \circ U_f)(H(k); C(k)) \quad \mathbf{B}^{-1} = \mathbf{B}^T \text{なることを利用している。}
 \end{aligned}$$

となる。したがって、題意が示された。

CRI-PSOのスケール不变性欠如の証明

定理20

CRI-PSOの更新則 U_f はスケール不变性を有さない。

証明

変換 $T_s: \mathbf{x} \mapsto \mathbf{Dx}$ に対して、 $\mathbf{x}^{i'}(k) = \mathbf{Dx}^i(k), \mathbf{p}^{i'}(k) = \mathbf{Dp}^i(k), \mathbf{p}^{g'}(k) = \mathbf{Dp}^g(k)$ となる。

また、 $\mathbf{v}^{i'}(k) = \mathbf{x}^{i'}(k) - \mathbf{x}^{i'}(k-1) = \mathbf{Dv}^i(k)$ である。さらに、 $\mathbf{P}' = \mathbf{P}$ である。

このとき、不变性の条件式は、

$$\begin{aligned}
 & (U_f \circ T_s)(H(k); C(k)) \\
 &= \mathbf{D} \left(\mathbf{x}^i(k) + w\mathbf{v}^i(k) \right) + c_1 \mathbf{PR}_1 \mathbf{P}^T \left(\mathbf{Dp}^i(k) - \mathbf{Dx}^i(k) \right) + c_2 \mathbf{PR}_2 \mathbf{P}^T \left(\mathbf{Dp}^g(k) - \mathbf{Dx}^i(k) \right) \\
 &= \mathbf{D} \left(\mathbf{x}^i(k) + w\mathbf{v}^i(k) \right) + c_1 \mathbf{PR}_1 \mathbf{P}^T \mathbf{D} \left(\mathbf{p}^i(k) - \mathbf{x}^i(k) \right) + c_2 \mathbf{PR}_2 \mathbf{P}^T \mathbf{D} \left(\mathbf{p}^g(k) - \mathbf{x}^i(k) \right) \\
 &\neq \mathbf{D} \left(\mathbf{x}^i(k) + w\mathbf{v}^i(k) \right) + c_1 \mathbf{DPR}_1 \mathbf{P}^T \left(\mathbf{p}^i(k) - \mathbf{x}^i(k) \right) + c_2 \mathbf{DPR}_2 \mathbf{P}^T \left(\mathbf{p}^g(k) - \mathbf{x}^i(k) \right) \\
 &= \mathbf{D} \left(\mathbf{x}^i(k) + w\mathbf{v}^i(k) + c_1 \mathbf{PR}_1 \mathbf{P}^T \left(\mathbf{p}^i(k) - \mathbf{x}^i(k) \right) + c_2 \mathbf{PR}_2 \mathbf{P}^T \left(\mathbf{p}^g(k) - \mathbf{x}^i(k) \right) \right) \\
 &= \mathbf{Dx}^i(k+1) = (T_s \circ U_f)(H(k); C(k)) \quad \mathbf{DP} \neq \mathbf{PD} \text{となることを利用している。}
 \end{aligned}$$

となる。したがって、題意が示された。

数値実験条件

■ 探索性能

- ◆ 初期配置領域内に初期探索点をランダムに配置した後、終了条件を満たすまで探索を進める試行を50回行い、目的関数値 $f(\mathbf{p}^g)$ の平均値を探索性能とする。

■ パラメータ設定

- ◆ 探索点数 : $m = 20$
- ◆ PSO : $w = 0.729, c_1 = c_2 = 1.4955$ (Constriction Method[4])
- ◆ CRI-PSO : $w = 0.5, c_1 = c_2 = 1.4955$

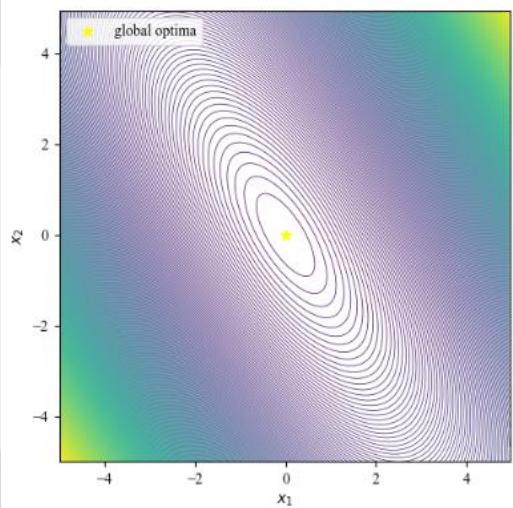
■ CRI-PSOは解集合 \mathcal{Z} の設定により、2種類使用する。

- ◆ CRI-PSO 1 : $\mathcal{Z} = \mathcal{X}(k)$ (探索点群)
- ◆ CRI-PSO 2 : $\mathcal{Z} = \mathcal{P}(k)$ (p-best集合)

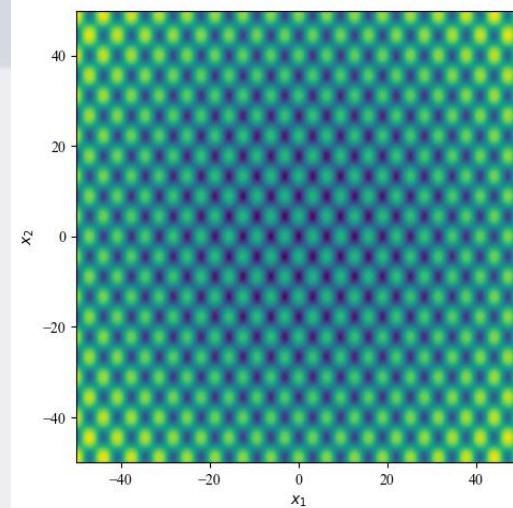
補足資料 > 学位論文4.4節 相関性を用いた回転不变性を有するPSO
ベンチマーク関数

関数	单／多峰性	变数間独立／依存性	良／悪スケール性
f_1	单峰性	变数間独立性	良スケール性
f_2, f_6, f_8	多峰性	变数間依存性	良スケール性
f_3, f_4, f_7, f_9	多峰性	变数間独立性	良スケール性
f_5	单峰性	变数間依存性	悪スケール性

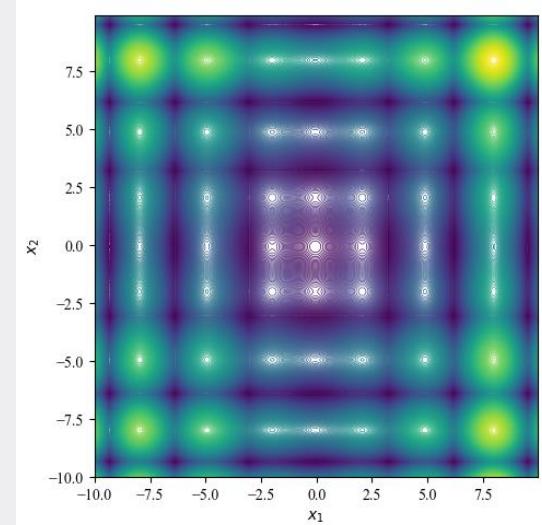
f_5 : Schwefel関数



f_8 : Griewank関数



f_9 : Alpine関数



実験2：実験結果

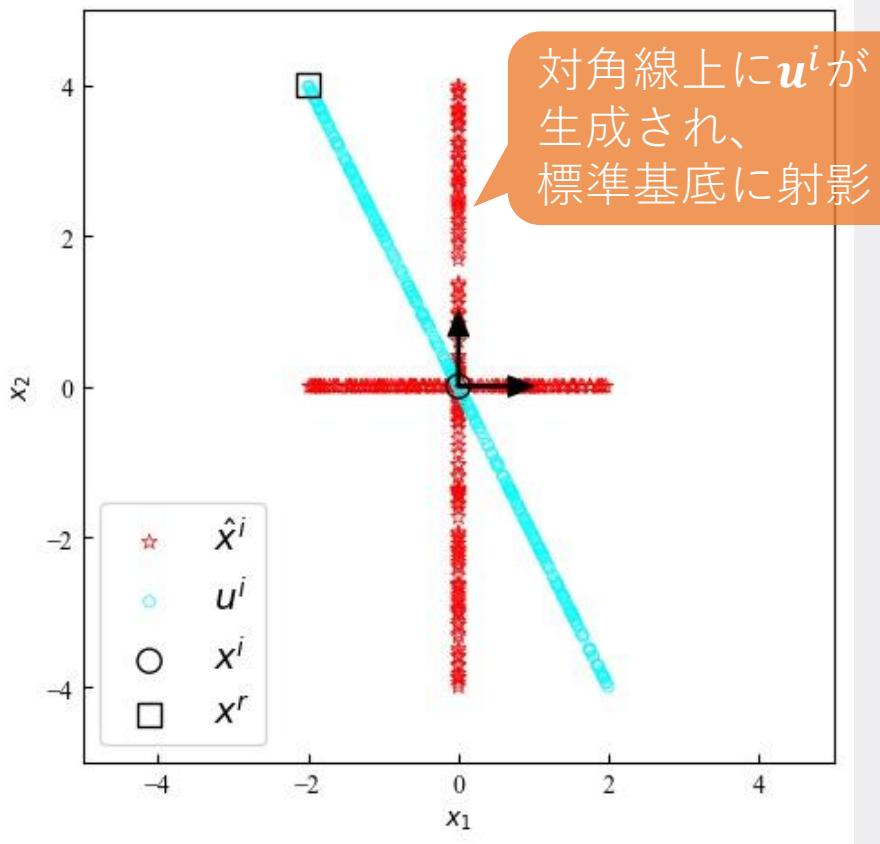
■ 変数間依存性の関数で提案手法が優れることが多い。

関数	N	$k = 100$			$k = 1000$		
		PSO	CRI-PSO 1	CRI-PSO 2	PSO	CRI-PSO 1	CRI-PSO 2
f_2	50	2236	805	606*	99.6*	966	624
	100	29456	2871	1970*	1831*	2627	2176
	300	111619	12399	9342*	111776	12776	8829*
f_3	50	-1955	-2516*	-2509	-3344	-2519	-2533*
	100	-2285	-4257*	-4257	-2480	-4290	-4249*
	300	-4931	-10266*	-10197	-4898	-10146	-10313*
f_5	50	297	85.2	70.7*	34.1*	84.1	69.2
	100	1469	379	314*	363	353	324*
	300	15518	3420	2878*	4968	3584	2886*

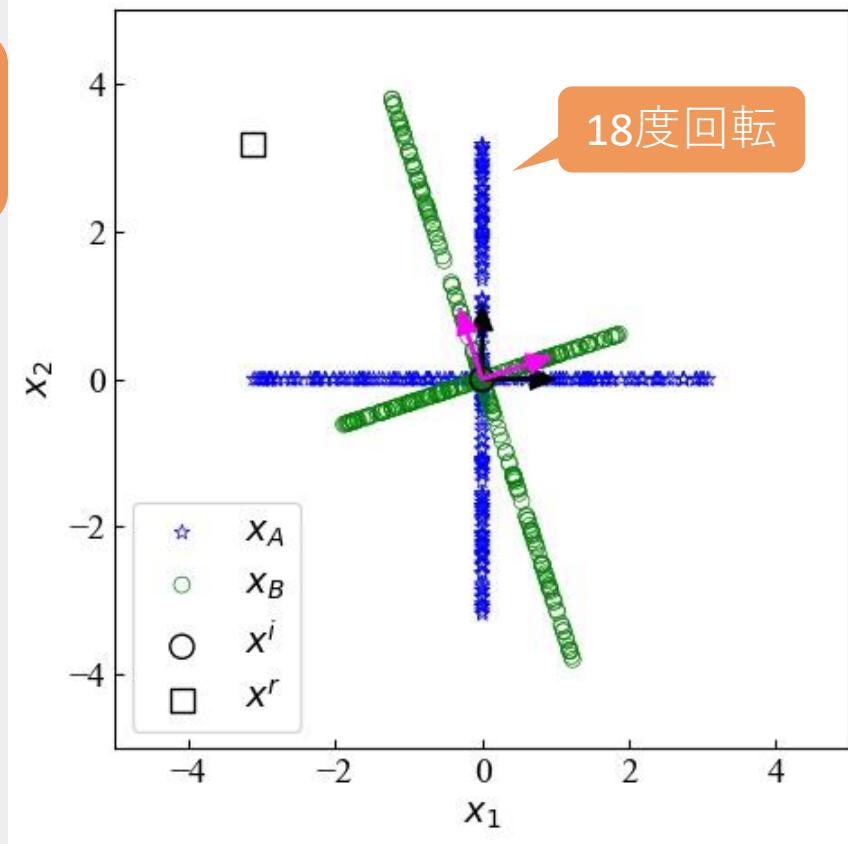
※一部抜粋

ABCの回転不变性の欠如

■ $x_A \neq x_B$ となるため、回転不变性を欠如している。



近傍生成

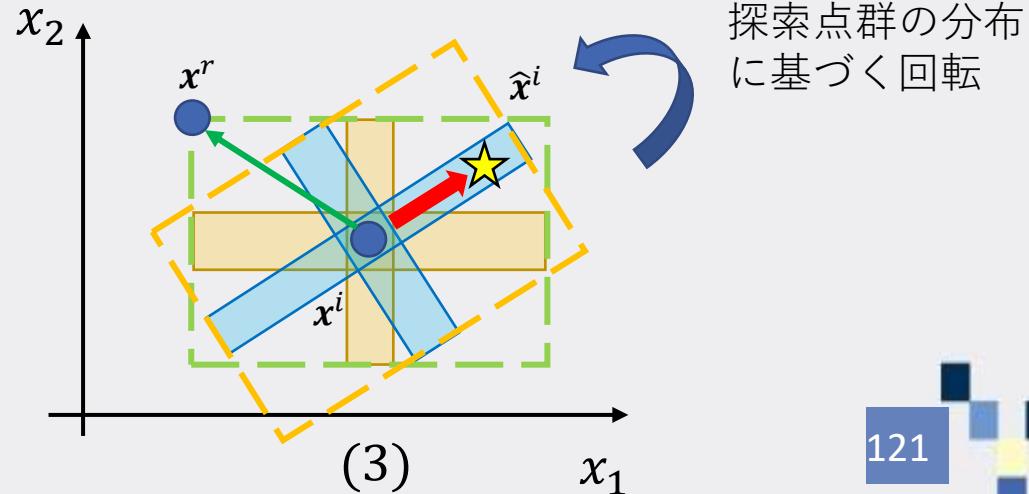
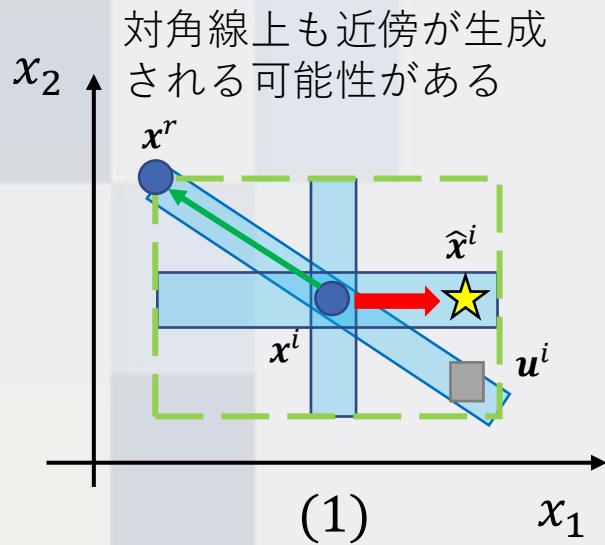
 x_A, x_B の生成

※上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

ABCの変換不变性に関する先行研究

■ 他手法におけるアプローチ(2)の事例はない。

#	アプローチ	特徴	文献
(1)	線形結合の係数をスカラとする	完全なスカラではないが、複数要素を組合せる (完全にスカラにすると、解の組合せではなく、線分上に生成され、多様性が極端に低下する)	[5]
(2)	超球の導入	—	—
(3)	回転行列を線型結合の係数に作用	探索点群の共分散行列に基づく回転行列により、近傍を回転させることで、標準基底に依存しない	[6]



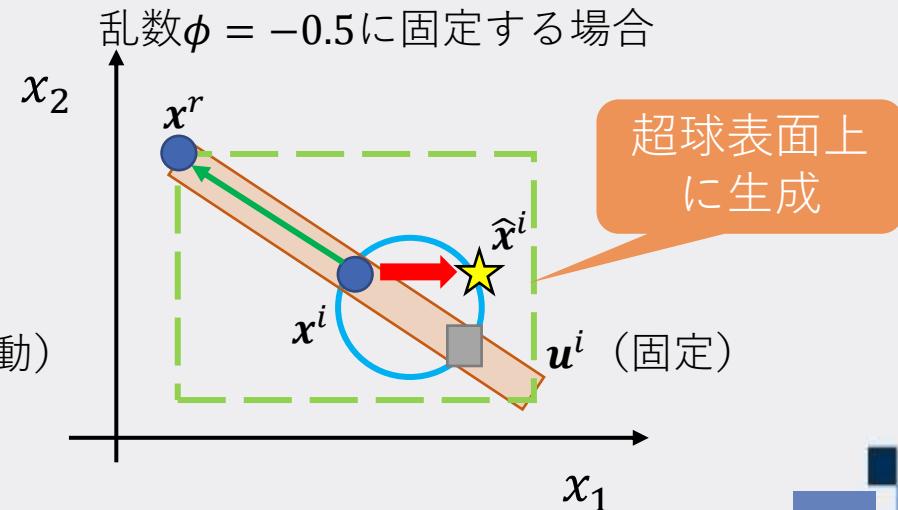
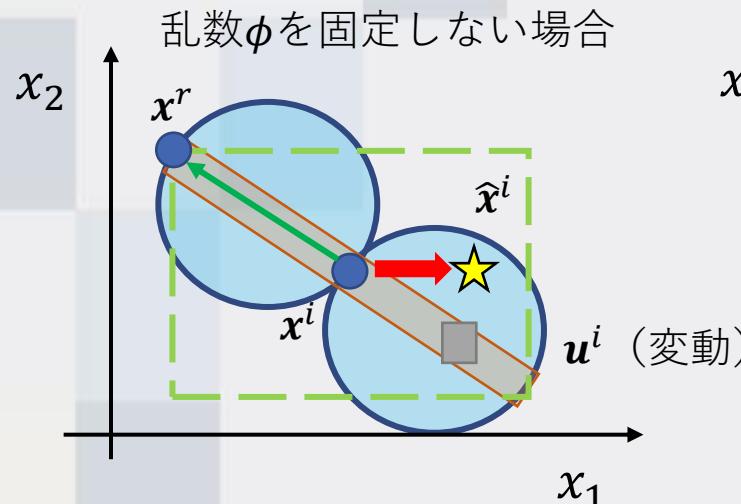
超球を用いた回転不变性を有するABC

■ HBC-ABCは超球表面上に近傍を生成する。

- ◆ 射影先を標準基底ではなく超球表面上に変更することで、回転不变性を獲得

■ 近傍生成は下記の通りである。

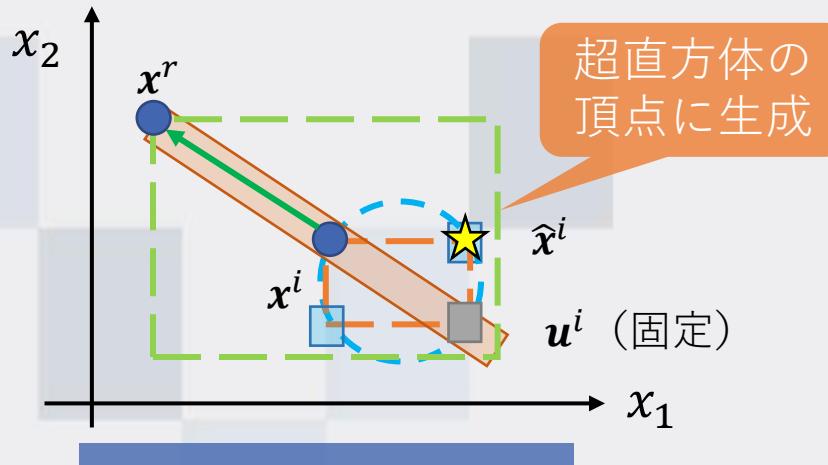
- ◆ $\mathbf{u}^i = \mathbf{x}^i + \phi(\mathbf{x}^i - \mathbf{x}^r)$, $\phi \sim \mathcal{U}_{\mathbb{R}}(-1,1)$ を生成する
- ◆ 中心 $(\mathbf{x}^i + \mathbf{u}^i)/2$ で、半径 $\|(\mathbf{x}^i - \mathbf{u}^i)/2\|_2$ の超球表面上に近傍生成する



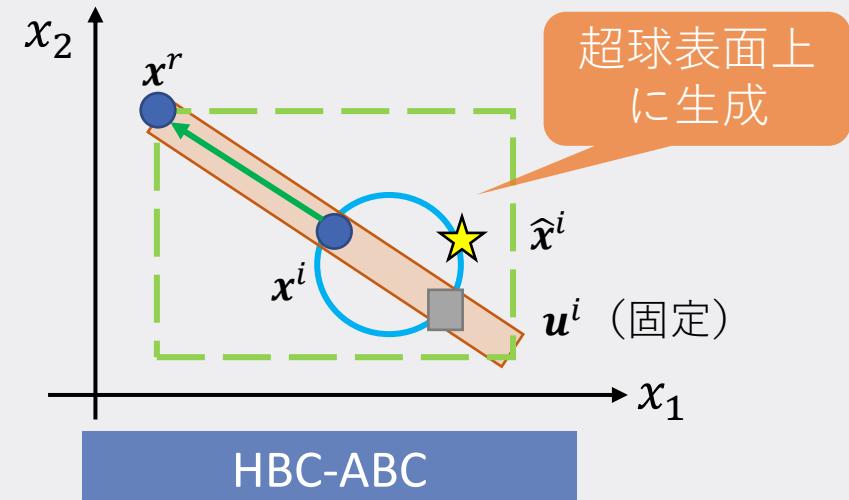
HBC-ABCとABCの比較

- ABCの近傍も同様の超球表面上の一部に生成される
とみなせるため、HBC-ABCはABCを包含している
 - ◆ 超球に内接する超直方体の頂点に生成されるから

対角線上に \mathbf{u}^i が生成され、
それを標準基底に射影



対角線上に \mathbf{u}^i が生成され、 \mathbf{x}^i と
 \mathbf{u}^i を直径とする超球表面に射影



※乱数 $\phi = -0.5$ に固定する場合

HBC-ABCの回転・相似変換・平行移動不变性具備の証明

定理22

HBC-ABCの更新則 U_f は回転不变性・相似変換不变性・平行移動不变性を有する。

証明

合成変換 $T_s: \mathbf{x} \mapsto \delta \mathbf{Bx} - \mathbf{t}$ に対して、 $\mathbf{x}^{i'} = \delta \mathbf{Bx}^i - \mathbf{t}, \mathbf{x}^{r'} = \delta \mathbf{Bx}^r - \mathbf{t}$ となる。このとき、 \mathbf{u}^i は、

$$\begin{aligned}\mathbf{u}^{i'} &= \mathbf{x}^{i'} + \phi(\mathbf{x}^{r'} - \mathbf{x}^{i'}) = \delta \mathbf{Bx}^i - \mathbf{t} + \phi(\delta \mathbf{Bx}^r - \delta \mathbf{Bx}^i) = \delta \mathbf{Bx}^i + \phi \delta \mathbf{B}(\mathbf{x}^r - \mathbf{x}^i) - \mathbf{t} \\ &= \delta \mathbf{B}(\mathbf{x}^i + \phi(\mathbf{x}^r - \mathbf{x}^i)) - \mathbf{t} = \delta \mathbf{B}\mathbf{u}^i - \mathbf{t}\end{aligned}$$

このとき、 \mathbf{z}^i は標準正規分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ に従う乱数ベクトルであるため、 $\mathbf{z}^{i'} = \mathbf{Bz}^i = \mathbf{z}^i$ である。よって、 $\mathbf{s}^{i'} = \mathbf{z}^{i'}/\|\mathbf{z}^{i'}\|_2 = \mathbf{z}^i/\|\mathbf{z}^i\|_2 = \mathbf{s}^i = \mathbf{Bs}^i$ となる。

このとき、不变性の条件式は、

$$\begin{aligned}(U_f \circ T_s)(H; C) &= (\mathbf{x}^{i'} + \mathbf{u}^{i'})/2 + \|(\mathbf{x}^{i'} - \mathbf{u}^{i'})/2\|_2 \mathbf{s}^{i'} \\ &= \delta \mathbf{B}(\mathbf{x}^i + \mathbf{u}^i)/2 + \|\delta \mathbf{B}(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{Bs}^i - \mathbf{t} \\ &= \delta \mathbf{B}(\mathbf{x}^i + \mathbf{u}^i)/2 + \delta \|(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{Bs}^i - \mathbf{t} \\ &= \delta \mathbf{B}\left((\mathbf{x}^i + \mathbf{u}^i)/2 + \|(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{s}^i\right) - \mathbf{t} \\ &= \delta \mathbf{B}\hat{\mathbf{x}}^i - \mathbf{t} = (T_s \circ U_f)(H; C)\end{aligned}$$

※目的関数值に無関係なので、探索点の移動則についても成立する。

$\|\mathbf{Bx}\|_2 = \|\mathbf{x}\|_2$ なることを利用している。

となる。したがって、題意が示された。

HBC-ABCのスケール不变性欠如の証明

定理23

HBC-ABCの更新則 U_f はスケール不变性を有さない。

証明

変換 $T_s: \mathbf{x} \mapsto \mathbf{D}\mathbf{x}$ に対して、 $\mathbf{x}^{i'} = \mathbf{D}\mathbf{x}^i, \mathbf{x}^{r'} = \mathbf{D}\mathbf{x}^r$ となる。
このとき、 \mathbf{u}^i は、

$$\begin{aligned}\mathbf{u}^{i'} &= \mathbf{x}^{i'} + \phi(\mathbf{x}^{r'} - \mathbf{x}^{i'}) = \mathbf{D}\mathbf{x}^i + \phi(\mathbf{D}\mathbf{x}^r - \mathbf{D}\mathbf{x}^i) = \mathbf{D}\mathbf{x}^i + \phi\mathbf{D}(\mathbf{x}^r - \mathbf{x}^i) \\ &= \mathbf{D}\left(\mathbf{x}^i + \phi(\mathbf{x}^r - \mathbf{x}^i)\right) = \mathbf{D}\mathbf{u}^i\end{aligned}$$

このとき、 \mathbf{z}^i は標準正規分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ に従う乱数ベクトルであるため、 $\mathbf{z}^{i'} = \mathbf{z}^i$ である。よって、 $\mathbf{s}^{i'} = \mathbf{z}^{i'}/\|\mathbf{z}^{i'}\|_2 = \mathbf{z}^i/\|\mathbf{z}^i\|_2 = \mathbf{s}^i$ となる。

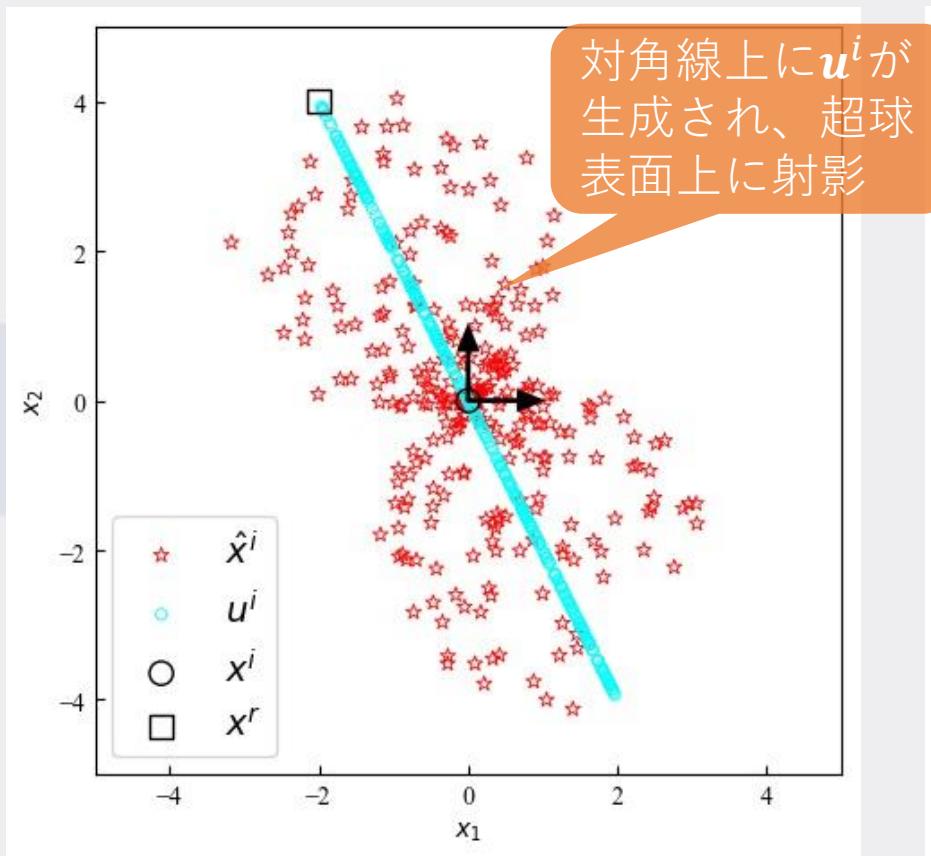
このとき、不变性の条件式は、

$$\begin{aligned}(U_f \circ T_s)(H; C) &= (\mathbf{x}^{i'} + \mathbf{u}^{i'})/2 + \|(\mathbf{x}^{i'} - \mathbf{u}^{i'})/2\|_2 \mathbf{s}^{i'} \\ &= \mathbf{D}(\mathbf{x}^i + \mathbf{u}^i)/2 + \|\mathbf{D}(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{s}^i \\ &\neq \mathbf{D}(\mathbf{x}^i + \mathbf{u}^i)/2 + \mathbf{D}\|(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{s}^i \\ &= \mathbf{D}\left((\mathbf{x}^i + \mathbf{u}^i)/2 + \|(\mathbf{x}^i - \mathbf{u}^i)/2\|_2 \mathbf{s}^i\right) \\ &= \mathbf{D}\hat{\mathbf{x}}^i = (T_s \circ U_f)(H; C)\end{aligned} \quad \|\mathbf{D}\mathbf{x}\|_2 \neq \mathbf{D}\|\mathbf{x}\|_2 \text{なることを利用している。}$$

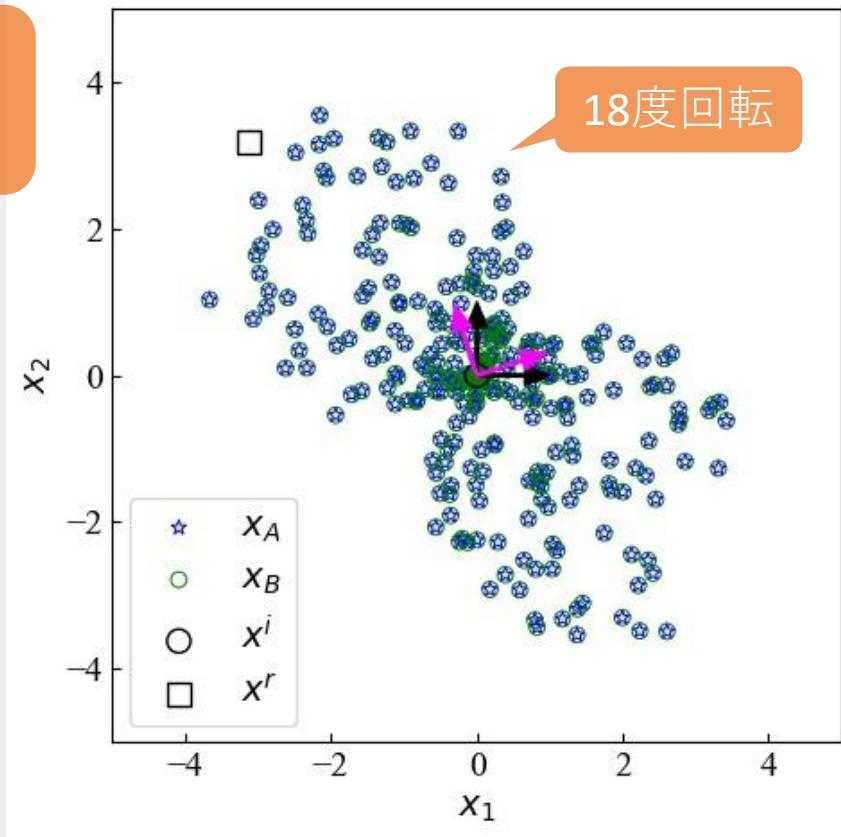
となる。したがって、題意が示された。

HBC-ABCの回転不变性の具備

■ $\mathbf{x}_A = \mathbf{x}_B$ となるため、回転不变性を具备している。



近傍生成

 $\mathbf{x}_A, \mathbf{x}_B$ の生成

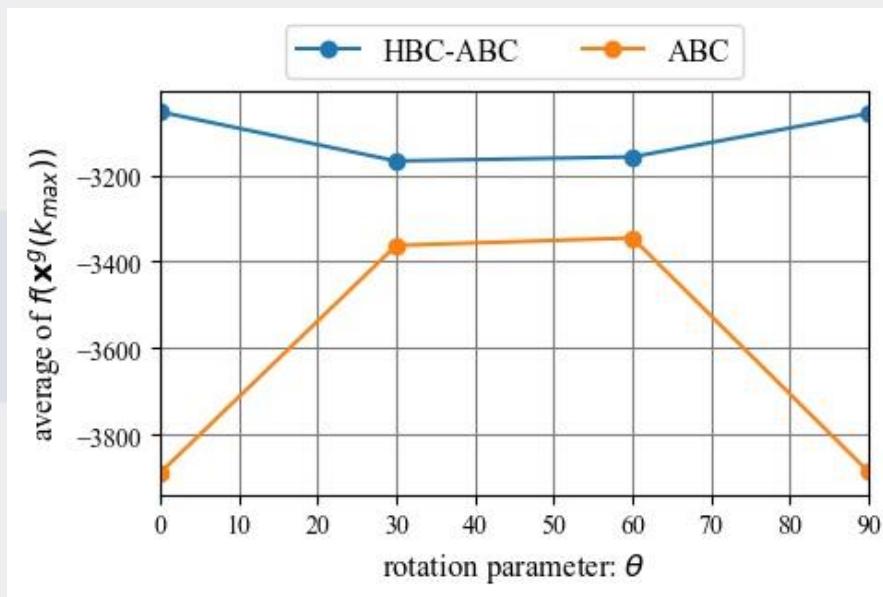
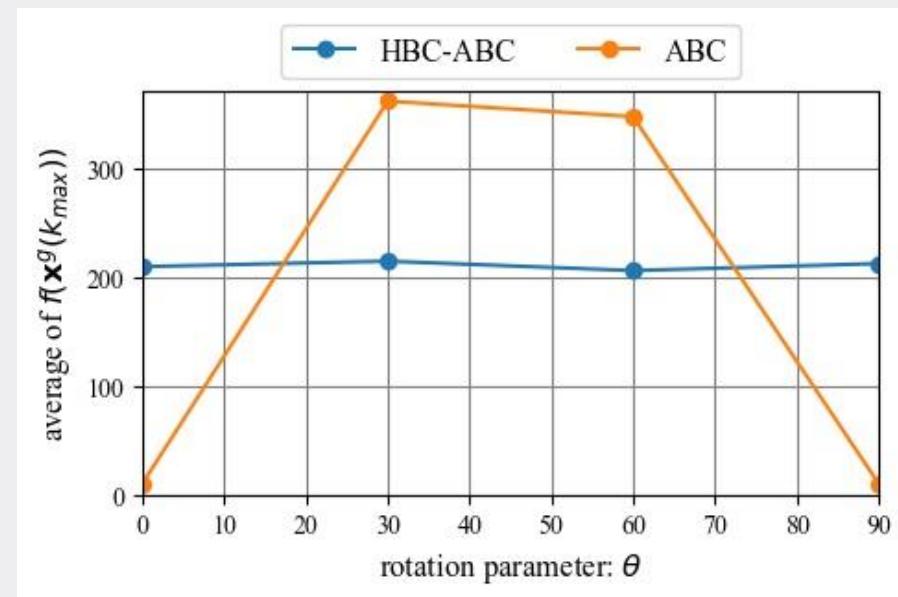
※上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

数値実験

- 数値実験を通じて、HBC-ABCの下記を検証する。
 - ◆ 実験1：回転に対するロバスト性に関する検証
 - ◆ 実験2：全般的な性能に関する検証
- 回転不变性を欠如するABCと比較する。
- 初期配置領域内に初期探索点をランダムに配置した後、終了条件を満たすまで探索を進める試行を50回行い、最良解の目的関数値 $f(\mathbf{x}^b)$ の平均値を探索性能とする。
- 終了条件は評価回数が $T \geq T_{\max}$ を満たすことである。
- パラメータ設定
 - ◆ 探索点数： $m_e = m_o = 10$
 - ◆ 停滞回数の閾値： $Limit = 5N$

実験1：実験結果

■ 回転に対して、ABCは口バストでないが、HBC-ABCは口バストであることが確認できる。

 2^N minima 関数

Rastrigin 関数

※ $N = 50, T = 20000$ の結果だが、他の条件でも同様の傾向が確認された

実験2：実験結果

■ 変数間依存性の関数で提案手法が優れることが多い。

関数	N	$T = 2000$		$T = 20000$	
		ABC	HBC-ABC	ABC	HBC-ABC
f_2	50	2736	198*	102	48.6*
	100	15367	920*	408	102*
	300	93025	6620*	9028	743*
f_3	50	-2835*	-2306	-3889*	-3054
	100	-4145*	-3989	-7437*	-5539
	300	-7104	-9932*	-18749*	-13038
f_5	50	317	38.3*	126	2.90*
	100	1246	193*	541	34.1*
	300	11541	1767*	4627	452*

※一部抜粋

実験2：実験結果

■ 各条件において、2種類の手法の性能の中で1位の数をカウントした結果、下記が確認できる。

- ◆ 全般的には、HBC-ABCが優れる
- ◆ 低次元では、ABCが優れており、それ以外では、HBC-ABCが優れる

N	ABC			HBC-ABC		
	T = 2000	T = 20000	All	T = 2000	T = 20000	All
10	6	5	11	3	4	7
50	2	5	7	7	4	11
100	1	5	6	8	4	12
300	0	4	4	9	5	14
All	9	19	28	27	17	44

まとめと課題

■ ABCは回転変換に対してロバストでないが、HBC-ABCはロバストであることを確認。

- ◆ 他のアプローチを流用することで、回転不变性を付加でき、性能のロバスト性に影響を与えた

■ HBC-ABCがABCよりも全般的に優れることを確認。

- ◆ 特に変数間依存性を有する関数では、顕著な差が出た
- ◆ HBC-ABCが**変数間独立／依存性の影響を受けにくい**

■ ただし低次元ではABCが優れる。(詳細な理由は不明)

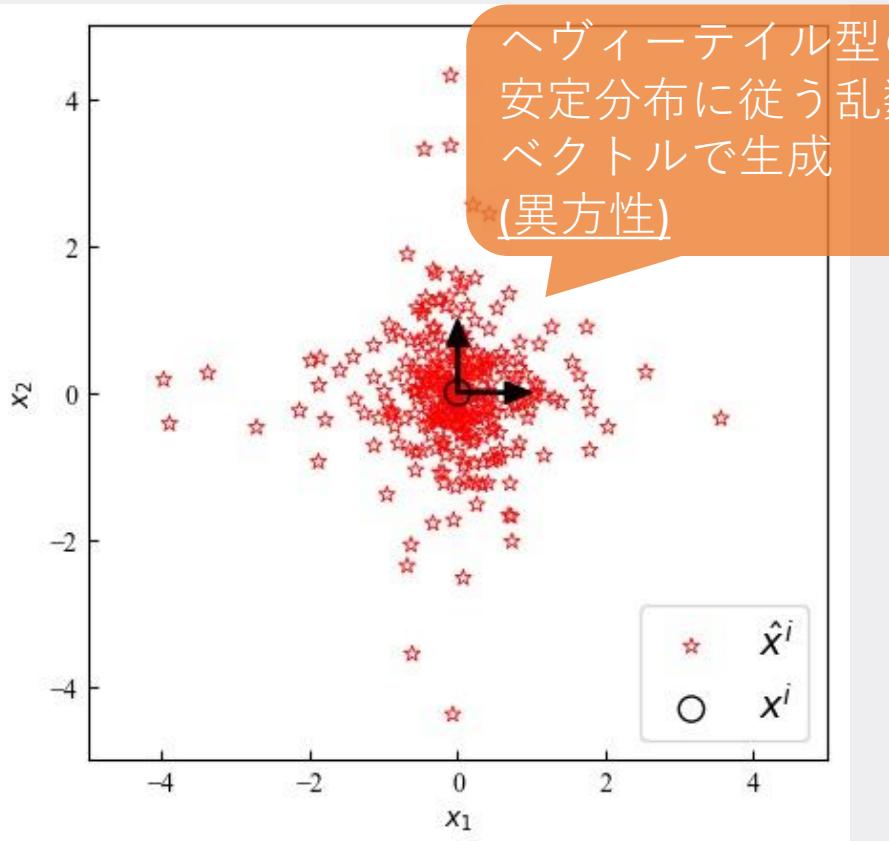
- ◆ 射影先を単一の標準基底から超球表面に変更したことでの多様性を獲得する一方、収束性に差が出た
- ◆ パラメータ設定が比較的短期に特化している(集中化が早い)と推測すると同時に、**設定の難しさを認識**

■ HBC-ABCはスケール不变性を失っている。

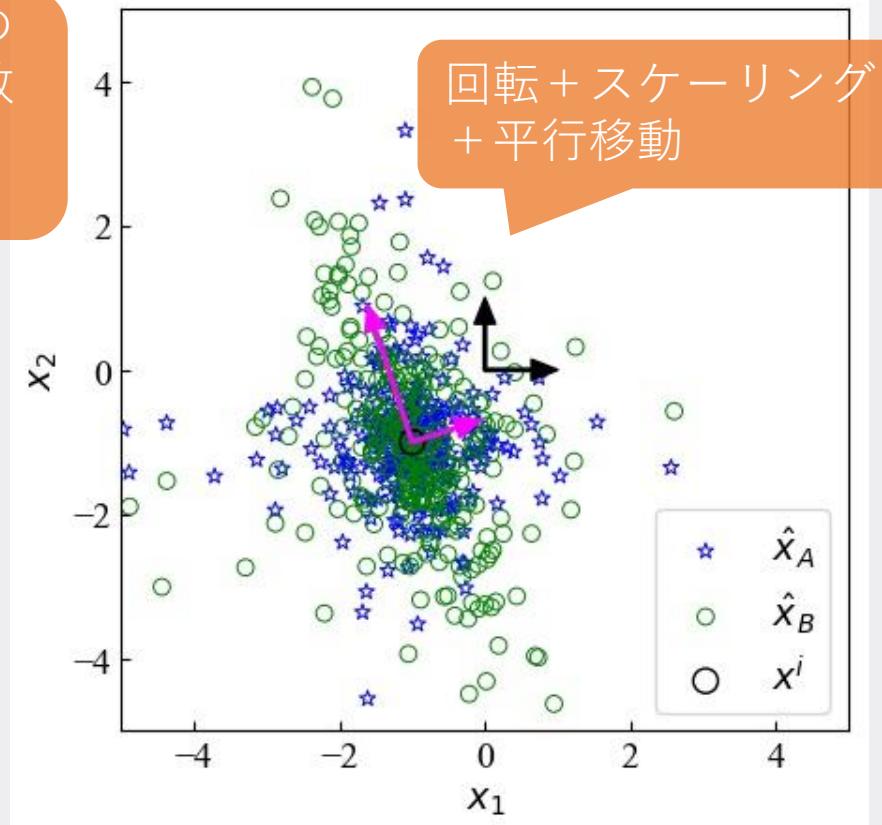
- ◆ 回転不变性獲得の代わりに、スケール不变性を犠牲
- ◆ アフィン変換不变性の獲得の難しさを認識

CSのアフィン変換不変性の欠如

■ $x_A \neq x_B$ なため、アフィン変換不変性を欠如している。



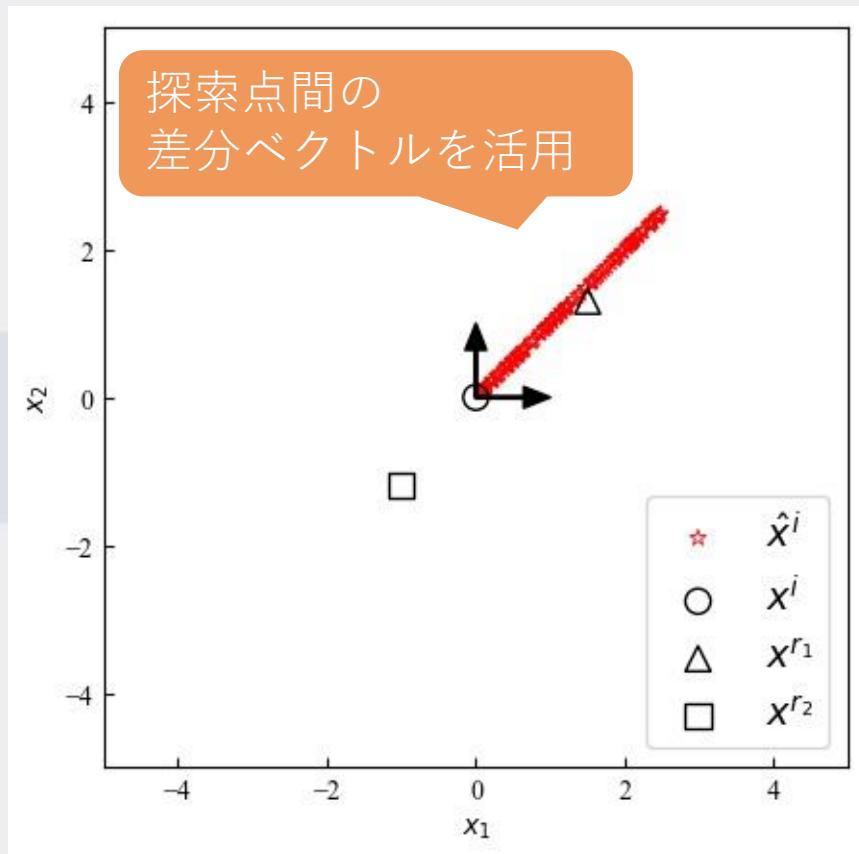
近傍生成

 x_A, x_B の生成

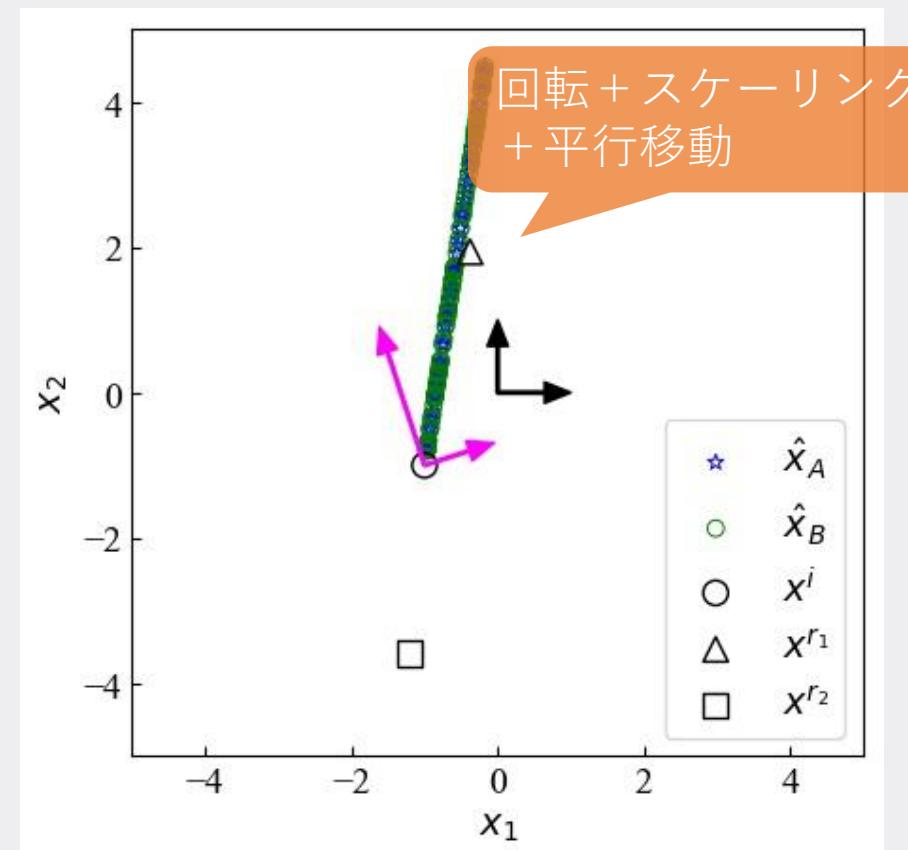
※上記はGlobal Searchの近傍生成で、Local Searchの近傍生成は不変性を有する。
※上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

CSのLSのアフィン変換不変性の具備

■ $x_A = x_B$ なため、アフィン変換不変性を具備している。



近傍の生成

 x_A, x_B の生成

※上記はLocal Searchの近傍生成で、Global Searchの近傍生成は不変性を有さない。
※上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

CSの変換不变性に関する先行研究

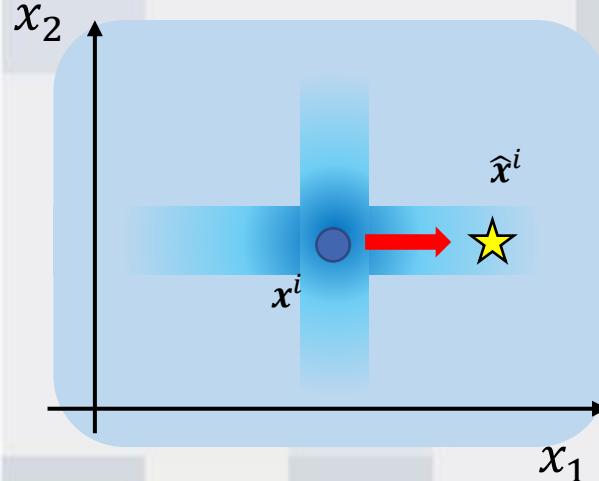
■ 他手法におけるアプローチの事例は、CSにはない。

アプローチ	特徴	文献
探索履歴に依存する回転行列と対角行列を乱数ベクトルの係数に作用	探索点群の共分散行列に基づく回転行列・対角行列により、近傍を回転・スケーリングさせることで、標準基底に依存しない	ES [7]

遠くにも生成する確率がある
同時確率分布は異方性となる

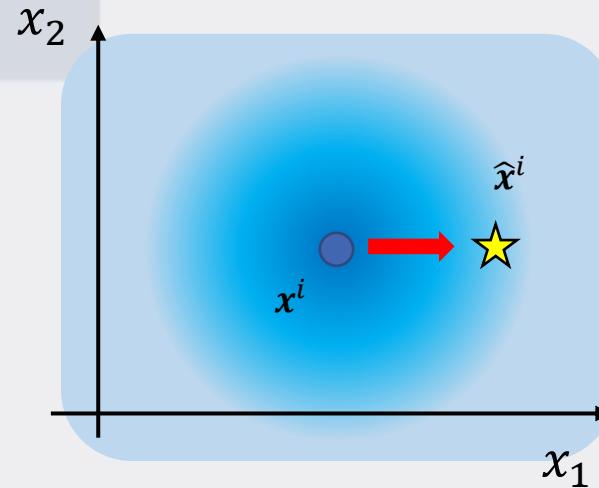
同時確率分布は等方性となる
(等方性なら回転不变性を持つ)

スケーリングかつ回転により、
楕円状の異方性となる

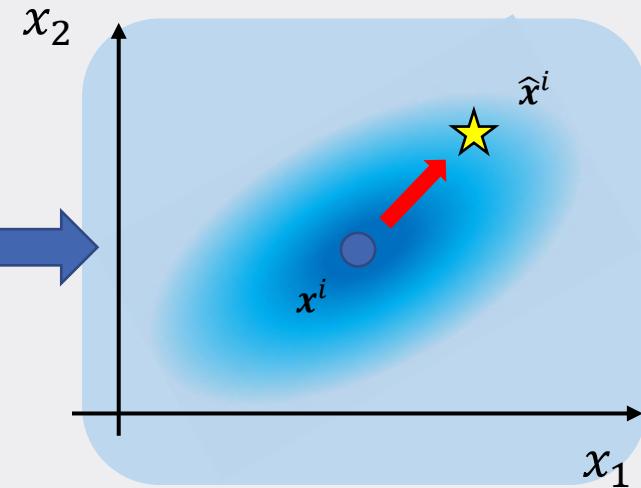


CS (Global Search)

※濃淡は確率密度を表す



ES (正規分布)



ESでのアプローチ

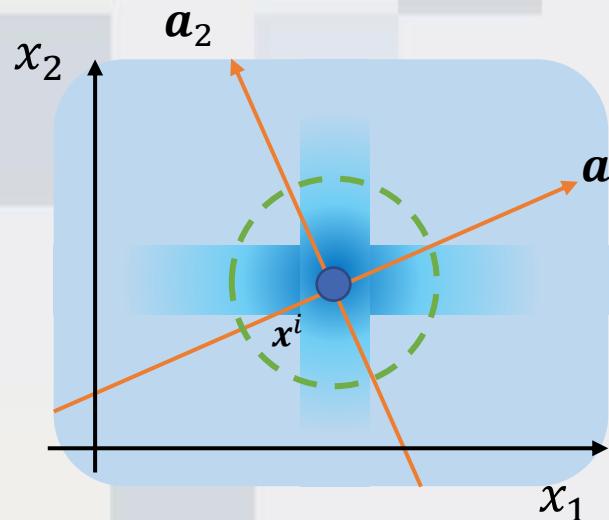
基底変換を用いたアフィン変換不変性を有するCS

■ CMD-CSは探索点群 χ の共分散行列 $\Sigma(\chi)$ を利用し、近傍に基底変換を作用する。

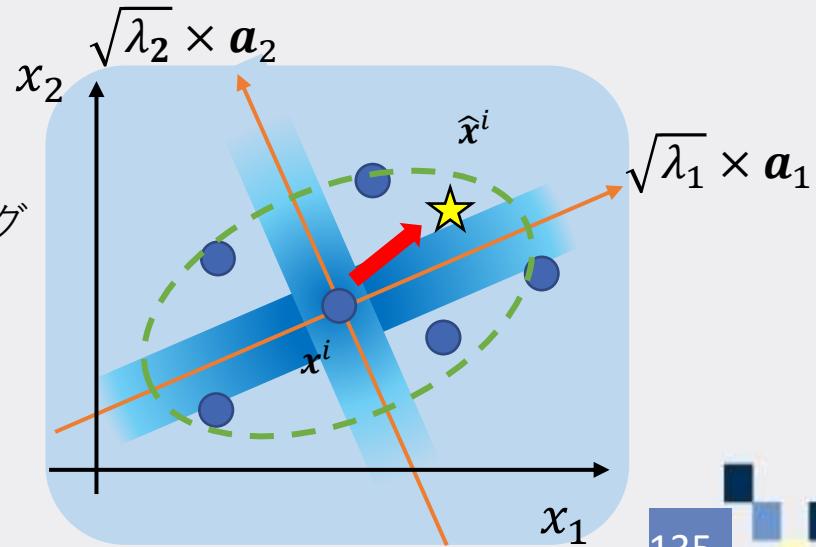
- ◆ 近傍を標準基底ではなく、探索過程で得た基底に依存することで、アフィン変換不変性を獲得

■ 近傍生成は下記の通りである。 $\Sigma(\chi)$ の固有ベクトルと固有値

- ◆ 基底変換行列 $P = [a_1, \dots, a_N]$ 、スケーリング $\Lambda = [\lambda_1, \dots, \lambda_N]$ を生成し、乱数ベクトル s を基底変換する

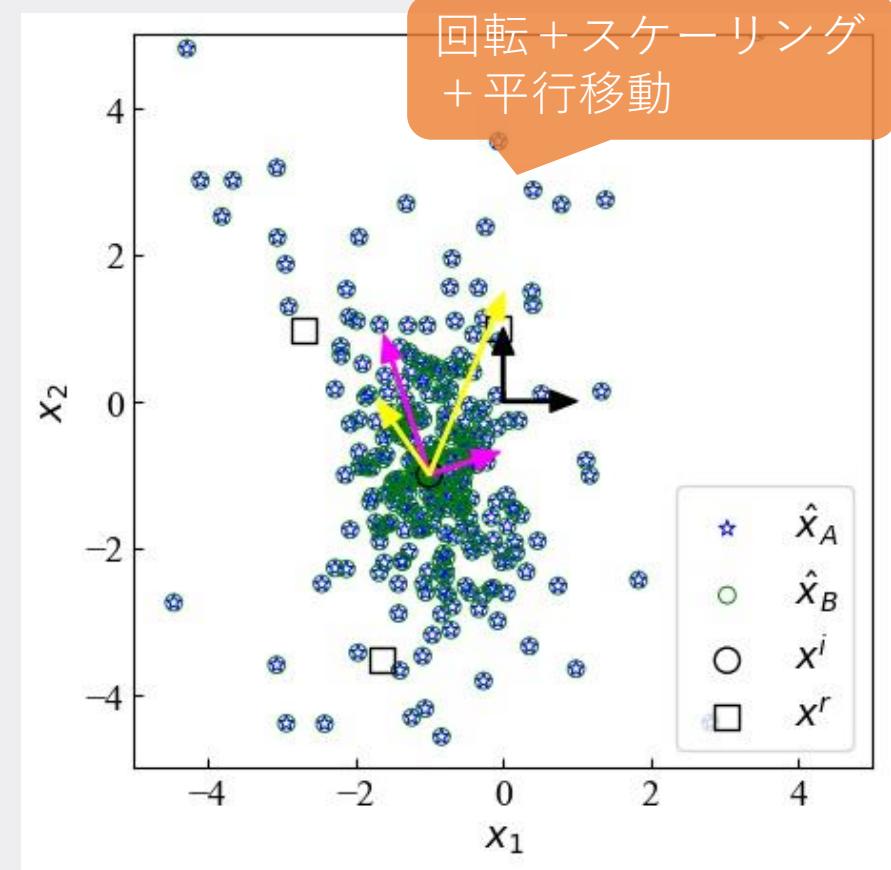
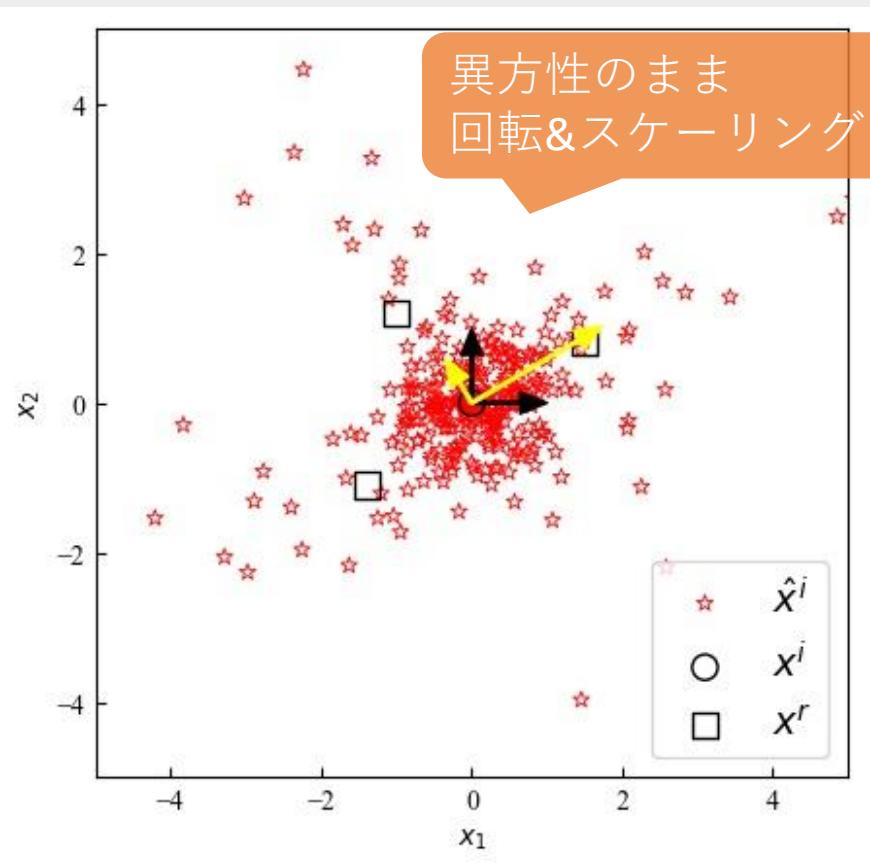


回転&
スケーリング



CMD-CSのアフィン変換不变性の具備

■ $x_A = x_B$ のため、アフィン変換不变性を具备する。



※上記はGlobal Searchの近傍生成で、Local Searchの近傍生成も不变性を有する。
※上記は数値実験的な検証だが、本論文では一般的な条件で証明している。

CMD-CSのアフィン変換不変性具備の証明

定理25

CMD-CSの更新則 U_f はアフィン変換不変性を有する。

証明

変換 $T_s: \mathbf{x} \mapsto \mathbf{Cx} - \mathbf{t}$ に対して、 $\mathbf{x}^{i'} = \mathbf{Cx}^i - \mathbf{t}$ となる。

このとき、 \mathbf{s}^i は安定分布 \mathcal{L} に従う乱数ベクトルであるため、

$\mathbf{s}^{i'} = \mathbf{s}^i$ である。また、 $\Sigma_*^{1/2}$ は $\Sigma_*^{1/2'} = \mathbf{C}\Sigma_*^{1/2}$ となる。

このとき、Global Search (GS) における不変性の条件式は、

$$(U_f \circ T_s)(H; C) = \mathbf{x}^{i'} + \alpha \Sigma_*^{1/2'} \mathbf{s}^{i'} = \mathbf{Cx}^i - \mathbf{t} + \alpha \mathbf{C}\Sigma_*^{1/2} \mathbf{s}^i = \mathbf{C}(\mathbf{x}^i + \alpha \Sigma_*^{1/2} \mathbf{s}^i) - \mathbf{t} = \mathbf{C}\hat{\mathbf{x}}^i \\ = (T_s \circ U_f)(H; C)$$

※目的関数値に無関係なので、探

となる。よって、GSでは条件式を満たす。 索点の移動則についても成立する。

このとき、Local Search (LS) における不変性の条件式は、

$$(U_f \circ T_s)(H; C) = \mathbf{x}^{i'} + r(\mathbf{x}^{i'} - \mathbf{x}^{r'}) = \mathbf{Cx}^i - \mathbf{t} + r\mathbf{C}(\mathbf{x}^i - \mathbf{x}^r) = \mathbf{C}(\mathbf{x}^i + r(\mathbf{x}^i - \mathbf{x}^r)) - \mathbf{t} = \mathbf{C}\hat{\mathbf{x}}^i \\ = (T_s \circ U_f)(H; C)$$

となる。よって、LSでは条件式を満たす。

したがって、題意が示された。

数値実験

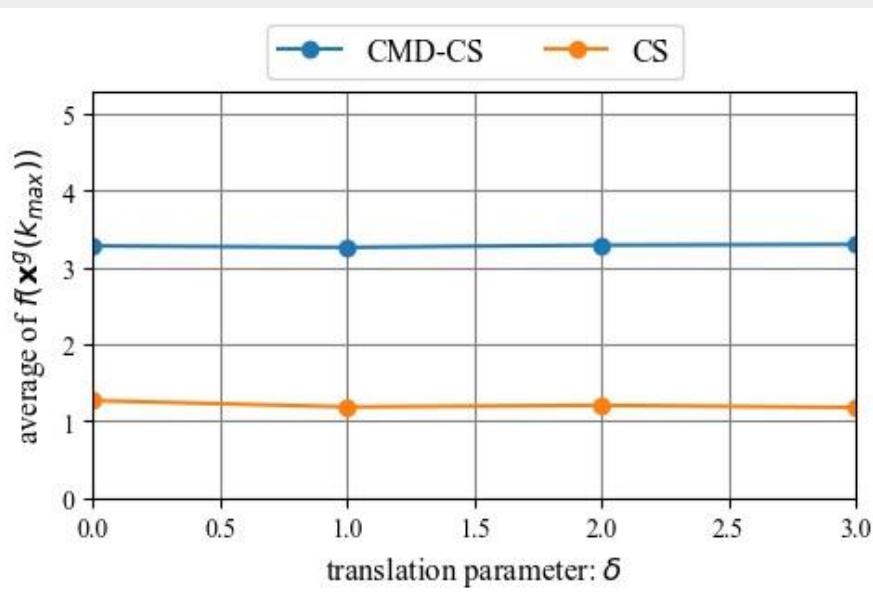
- 数値実験を通じて、CMD-CSの下記を検証する。
 - ◆ 実験1：回転に対するロバスト性に関する検証
 - ◆ 実験2：全般的な性能に関する検証
- アフィン変換不変性を欠如するCSと比較する。
- 初期配置領域内に初期探索点をランダムに配置した後、終了条件を満たすまで探索を進める試行を50回行い、最良解の目的関数値 $f(\mathbf{x}^b)$ の平均値を探索性能とする。
- 終了条件は評価回数が $T \geq T_{\max}$ を満たすことである。
- パラメータ設定
 - ◆ 探索点数 : $m = 20$
 - ◆ パラメータ : $\alpha = 0.1, \beta = 1.9$
 - ◆ 排斥確率 : $P_a = 0.25$

実験1：アフィン変換に対する口バスト性の検証

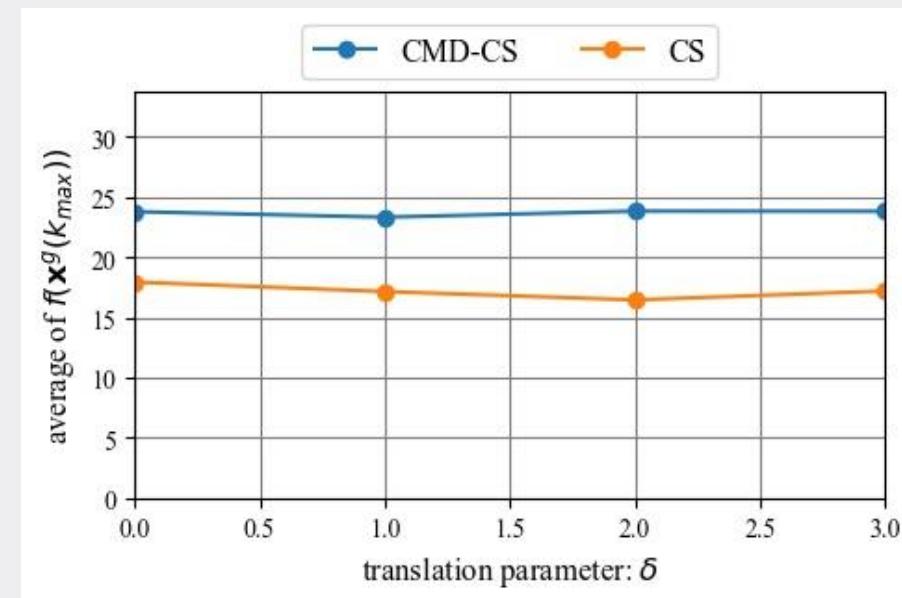
- 解空間に変換 T_s を与えるには、初期探索点を $T_s^{-1}(\mathbf{x}^i)$ とし、 $f(\mathbf{x}^i)$ を $f(T_s(\mathbf{x}^i))$ とする。
- 実験1-1：解空間を $\delta\mathbf{1}$ で平行移動し、性能の変動を確認する。
 - ◆ $\delta = 0, 1, 2, 3$ ($[1, \dots, 1]^T, [2, \dots, 2]^T, [3, \dots, 3]^T$)
 - ◆ 平行移動することで、原点依存性を発生する
- 実験1-2：解空間を \mathbf{D}^ε でスケーリングし、性能の変動を確認する。
 - ◆ $\varepsilon = 0, 1, 2, 3$ ($\text{diag}[1, \dots, N], \text{diag}[1^2, \dots, N^2], \text{diag}[1^3, \dots, N^3]$)
 - ◆ スケーリングすることで、目的関数に悪スケール性を発生する
- 実験1-3：解空間を θdeg 回転し、性能の変動を確認する。
 - ◆ $\theta = 0, 30, 60, 90\text{deg}$
 - ◆ 関数の対称性により、 90deg は 0deg と等価
 - ◆ 回転させることで、目的関数に変数間依存性を発生する
- 変数間独立・多峰性・良スケール性の関数を対象とする。

実験1-1：実験結果

- 平行移動に対して、CSとCMD-CSは口バストであることが確認できる（両方不变性を有する）。



Ackley関数

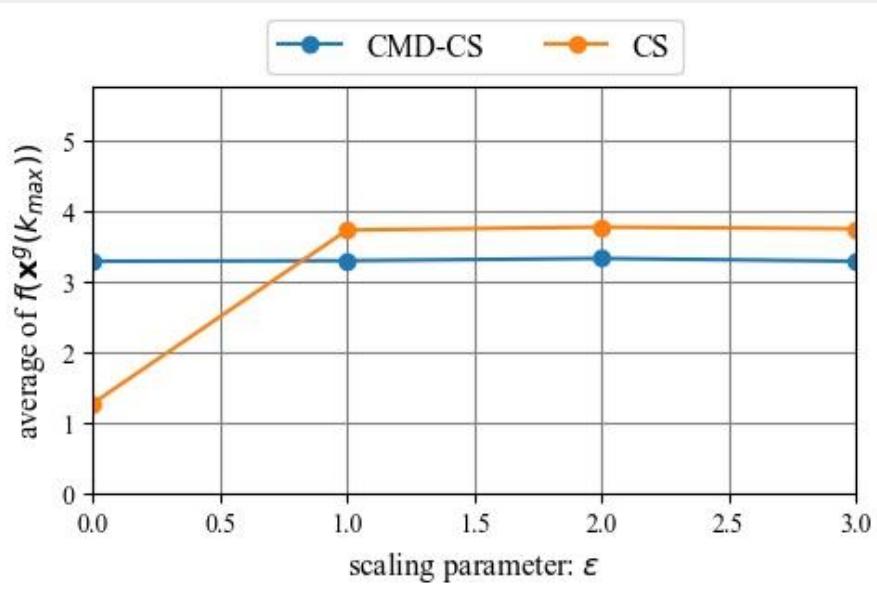


Alpine関数

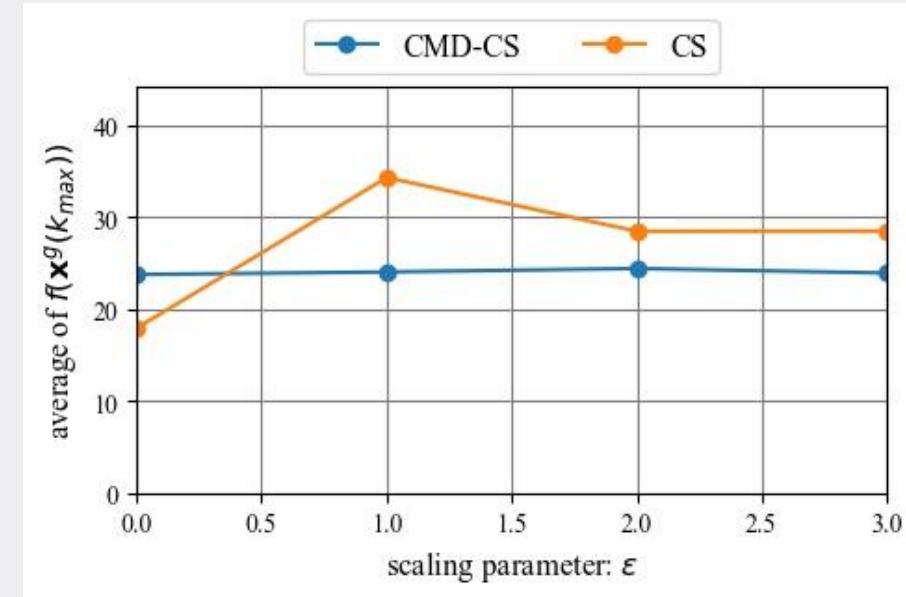
※ $N = 50, T = 20000$ の結果だが、他の条件でも同様の傾向が確認された

実験1-2：実験結果

■ スケーリングに対して、CSはロバストでないが、CMD-CSはロバストであることが確認できる。



Ackley関数

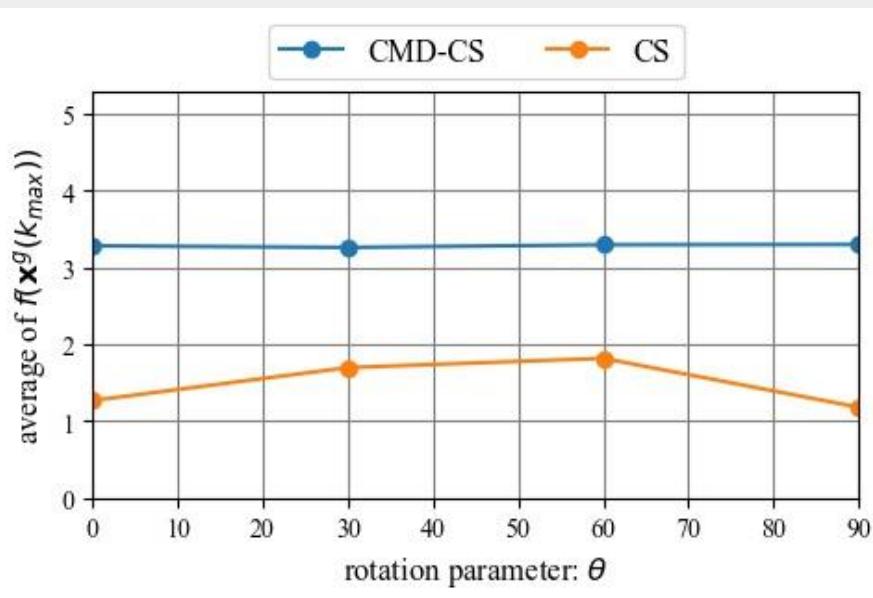


Alpine関数

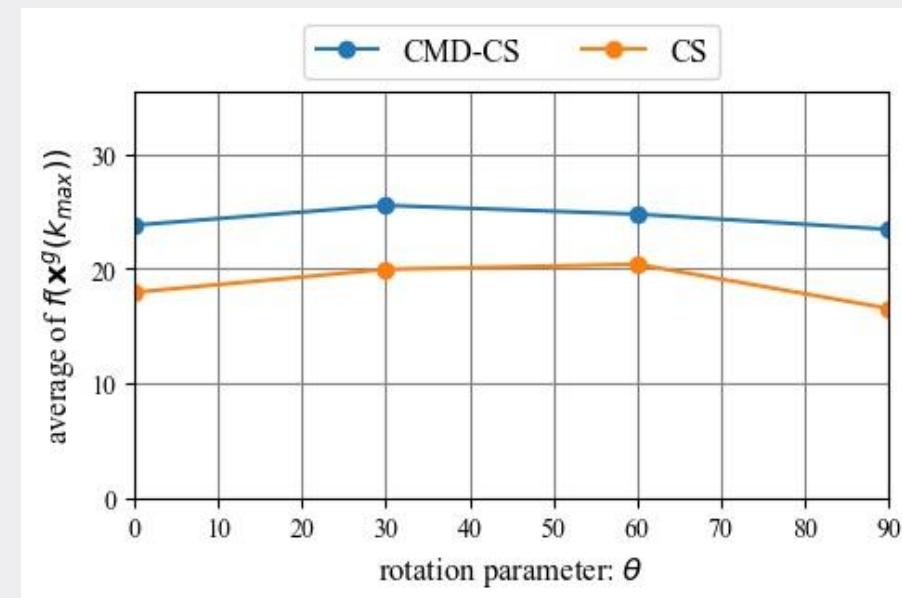
※ $N = 50, T = 20000$ の結果だが、他の条件でも同様の傾向が確認された

実験1-3：実験結果

■ 回転に対して、CSとCMD-CSは口バストだが、CSは(わずかに)比較的変動が大きいことが確認できる。



Ackley関数



Alpine関数

CSは回転不变性を有さないが、探索点付近の近傍は等方性であるため

※ $N = 50, T = 20000$ の結果だが、他の条件でも同様の傾向が確認された

実験2：実験結果

- 短期で提案手法が優れることが多い。
- ◆ 必ずしも想定している変数間依存性や悪スケール性に
対して有効となるとは限らない

関数	N	$T = 2000$		$T = 20000$	
		CS	CMD-CS	CS	CMD-CS
f_2	50	459018	7697*	79.0*	278
	100	1215876	19129*	301*	719
	300	4371190	70585*	2427*	2599
f_3	50	-554	-2330*	-3371*	-2578
	100	-1170	-4150*	-6158*	-4417
	300	-10619	-11185*	-14714*	-11611
f_5	50	850	72.0*	6.15*	14.7
	100	2928	292*	36.4*	63.6
	300	27265	2903*	375*	627

※一部抜粋

実験2：実験結果

■ 各条件において、2種類の手法の性能の中で1位の数をカウントした結果、下記が確認できる。

- ◆ 全般的には、CMD-CSが優れる
- ◆ 長期では、CSが優れており、短期では、CMD-CSが優れる

N	CS			CMD-CS		
	$T = 2000$	$T = 20000$	All	$T = 2000$	$T = 20000$	All
10	0	6	6	9	3	12
50	0	9	9	9	0	9
100	0	9	6	9	0	9
300	0	7	7	9	2	11
All	0	31	31	36	5	41

まとめと課題

- CSは回転とスケーリングに対してロバストでないが、CMD-CSはロバストであることを確認。
 - ◆ 他のアプローチを流用することで、アフィン変換不变性を付加でき、性能のロバスト性に影響を与えた
- CMD-CSがCSよりも全般的に優れることを確認。
- ただし長期ではCSが優れる。
 - ◆ CMD-CSは固有値のスケーリングが加わるため、収束性に差が出た (CSはパラメータでしかスケールを調整できない)
 - ◆ スケールパラメータの値を統一したが、上記理由からCMD-CSが短期に特化している(集中化が早い)と推測すると同時に、**設定の難しさを確認**
- CMD-CSはアフィン変換不变性を獲得するが、全般的な性能改善には、ロバスト性では不十分。
 - ◆ **適応性の必要性を認識**

不変性の付加アプローチの課題

1. 完全にアフィン変換不変性を付加することが困難。
 - ◆ アプローチによっては他の不変性を犠牲にする
2. 不変性の獲得は、想定した性質に対するロバスト性向上に貢献するが、他の条件（探索時間や多峰性などの問題構造）に貢献するとは限らない。
 - ◆ パラメータ設定が探索(多様化・集中化)に与える効果がオリジナルと同様でない
3. 各種アルゴリズムの数理的構造と不変性の関係に着目し、アプローチを適切に選択する必要があるが、理解が薄いユーザにとっては困難

パラメータ調整則の意義と分類

■ メタヒューリスティクスのパラメータ設定は、多様化・集中化に影響を与え、基本性能を左右する。

- ◆ オフラインチューニング(固定型)による性能向上の難しさ

■ オンラインチューニングは探索過程でパラメータ値を動的に変動させることで、多様化・集中化の「多様性」を拡張できるため、適切に与えれば、ユーザの労力削減と同時に性能向上が期待できる。

- ◆ **適応的調整型**は最も適応性の発揮が期待できる

スケジューリング的調整型

事前に定めたスケジュールに従って、パラメータ値を調整

→ フィードフォワード型

適応的調整型

探索過程で逐次得る情報を時系列パラメータの生成に活用・調整

→ **フィードバック型**

先行研究のパラメータ調整則の分類

■ 既存の調整則も多く提案されており、分類できる。

手法	固定型	スケジューリング的調整型	適応的調整型
ES(σ)	初期領域に依存	多点 : (μ, λ) -ES[28]	<ul style="list-style-type: none"> ➤ 単点 : 1/5ルール [28] ➤ 多点 : CMA-ES[6]
PSO(w)	CM[14]	<ul style="list-style-type: none"> ➤ RIWM[29] : ランダムに生成 ➤ LDIWM[29] : 時間経過に応じて減少 	Activity Feedback PSO[16]
DE (F, CR)	—※3	SaDE[30]、BiDE[31] : 確率的に生成	<ul style="list-style-type: none"> ➤ jDE[32]、JADE[33]、SHADE[34] : 確率的生成だが、生成確率の偏りを移動成功によって調整 ➤ FCDE[35] : 探索状態の評価と制御
FA(α)	初期領域に依存	時間経過に応じて減少[36]	—
CS(α)	初期領域に依存	時間経過に応じて減少[37]	探索点群内の相対的な良さや時間経過に応じて調整[38,39]

※1 : 橙背景は、本研究の対象とするアルゴリズム

※2 : パラメータを限定しているが、他のパラメータを調整する方法もある

※3 : DEはロバストな固定型の設定はないといわれている

適応型CSの先行研究

- α を探索点群内の相対的良さに基づき決定 [38,39]
 - ◆ 良ければ最小、平均より悪ければ探索時間に応じて減少

$$\alpha^i = \begin{cases} \alpha_{\min} + (\alpha_{\max} - \alpha_{\min})F^i, & f(\mathbf{x}^i) \leq f_{\text{ave}} \\ \frac{\alpha_{\max}}{\sqrt{T}}, & \text{otherwise} \end{cases}$$

$$F^i = \frac{f(\mathbf{x}^i) - f(\mathbf{x}^b)}{f_{\text{ave}} - f(\mathbf{x}^b)} \quad f_{\text{ave}} = \sum_{j=1}^m f(\mathbf{x}^j)$$

- ◆ 探索時間に応じて減少、その指數は良ければ小さくする

$$\alpha^i = \left(\frac{1}{T}\right)^{|F^i|} \quad F^i = \frac{f(\mathbf{x}^i) - f(\mathbf{x}^b)}{f(\mathbf{x}^w) - f(\mathbf{x}^b)}$$

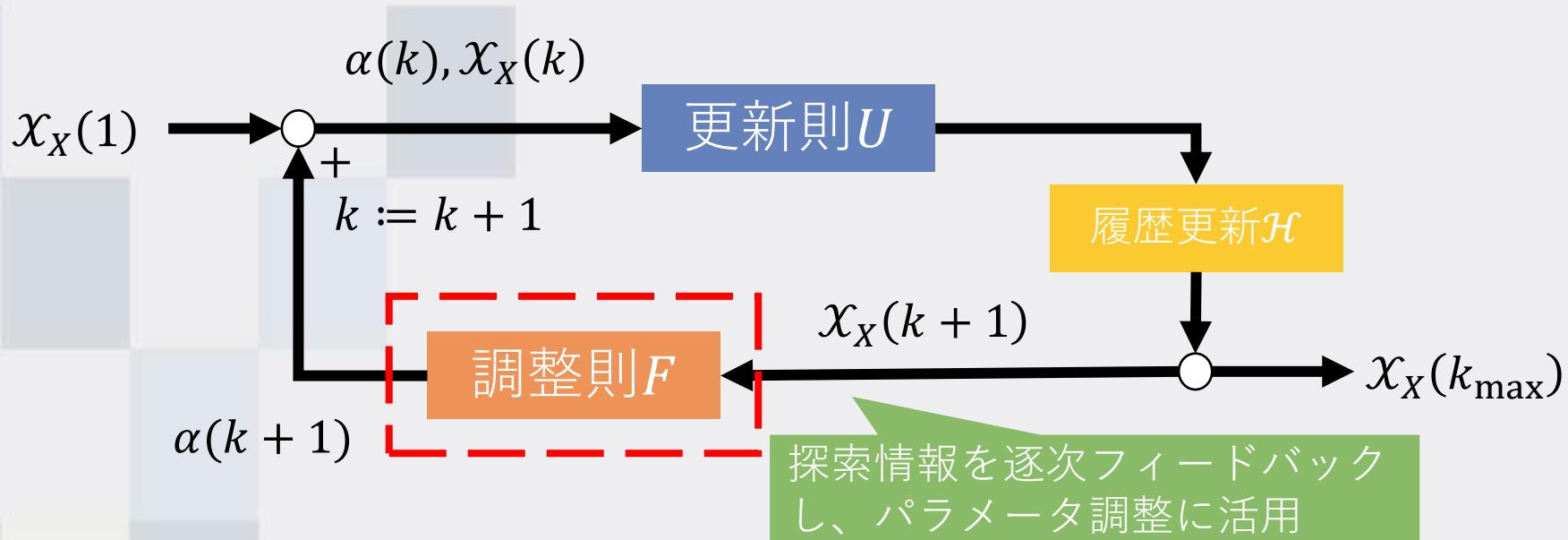
\mathbf{x}^b は最良点、 \mathbf{x}^w は最悪点

パラメータ調整則の一般形

定義5.1

更新則 $U_{f,X}$ 内のパラメータ $\alpha \in \mathbb{R}$ がパラメータ調整則 $F_{f,X}$ によって時系列で調整されるとき、 F_f は式(5.1)で表される。

$$\alpha(k+1) = F_{f,X}(\alpha(k); \mathcal{X}_X(k)) \quad \text{式(5.1)}$$



パラメータ調整則の変換不变性

定義5.2

パラメータ調整則 F_f が変換不变性を有するとは、ある変換 T に対して、パラメータ調整則 F_f が式(5.40)を満たすことである。

$$F_f(\alpha(k), T(H(k)); C(k)) = F_f(\alpha(k), H(k); C(k)) \quad \text{式(5.40)}$$



$\alpha'(k + 1) = F_f(\alpha(k), T(H(k)); C(k))$ とおく

$$\alpha'(k) = \alpha(k) \quad \text{式(5.40)'}$$



変換 T に対して α が不变であることを表している

明らかに式(5.41)は式(5.40)の十分条件である※

$$T(H(k)) = H(k) \Leftrightarrow H'(k) = H(k) \quad \text{式(5.41)}$$

※式(5.41)が成立しなくても演算形式によっては式(5.40)を満たすこともある

パラメータ調整則の情報と不变性の関係

- 調整則の不变性は用いる情報に依存する。
- ◆ ノルムの種類や演算形式によっても不变性の有無が変わる

情報	形式	具体例	十分条件を満たす変換不变性
探索点群の状態	差分ベクトルの L^2 ノルム	$\ \mathbf{a} - \mathbf{b}\ _2$	解空間の回転・平行移動・ 目的関数の単調増加変換
	Mahalanobis距離	$d_M(\mathbf{a}, \mathbf{b})$	解空間のアフィン変換・ 目的関数の単調増加変換
探索の経過時間	経過時間の割合	k/k_{\max}	解空間のアフィン変換・ 目的関数の単調増加変換
目的関数値	値	$f(\mathbf{a}) - f(\mathbf{b})$	解空間のアフィン変換
	全順序関係	$f(\mathbf{a}) \leq f(\mathbf{b})$	解空間のアフィン変換・ 目的関数の単調増加変換

指針1に従う場合、上記の要素を適切に選択することで、
調整則の不变性を獲得でき、更新則の不变性維持が可能

不变性に基づくパラメータ調整則の解析

- 調整則の不变性の具備は、更新則の場合と同様に示すことが可能。
 - ◆ 調整則が含む量が変換によってどのように変化するかを確認
 - ◆ それらを用いて、不变性の条件式を満たすかどうかを導出
- よって、不变性に基づく設計指針に従い、適切に調整則の選択・設計が可能である。

指針1：不变性維持

ロバスト性維持・適応性向上

指針2：不变性補完

ロバスト性向上・適応性向上

多様化・集中化の評価指標

先行研究の調整則では、下記情報が使用されている。

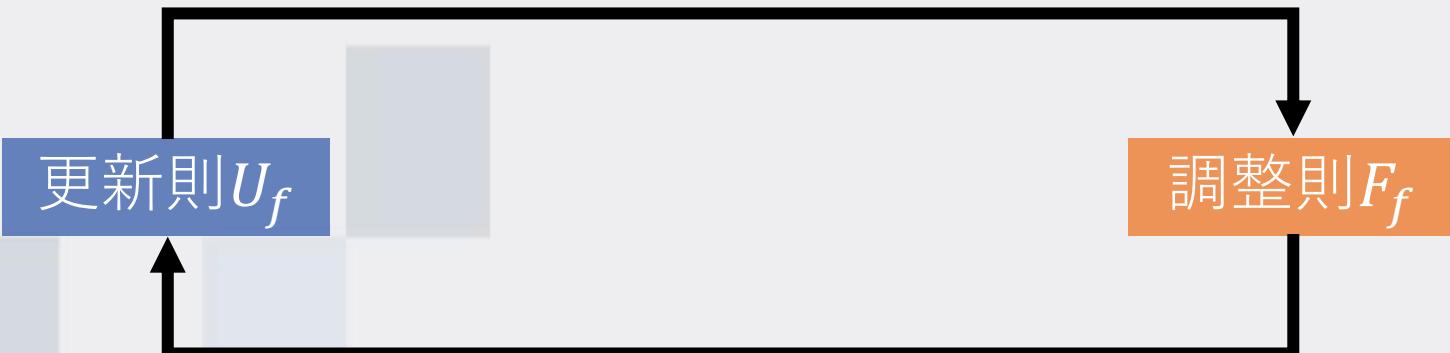
情報	形式	具体例	先行研究の調整則
探索点群の状態	探索点などの摂動の L^2 ノルム	$\frac{1}{m\sqrt{N}} \sum_{i=1}^m \ \hat{\mathbf{x}}^i(k) - \mathbf{x}^i(k)\ _2$	<ul style="list-style-type: none"> ➤ CMA-ES (平均位置) ➤ AFPSO (各探索点)[1]
	探索点から最良解への差分ベクトルの L^1 ノルム	$\frac{1}{mN} \sum_{i=1}^m \ \mathbf{p}^g(k) - \mathbf{x}^i(k)\ _1$	<ul style="list-style-type: none"> ➤ FCDE
探索の経過時間	経過時間	k (反復回数), T (評価回数)	<ul style="list-style-type: none"> ➤ 適応型CS
	経過時間の割合	$k/k_{\max}, T/T_{\max}$	<ul style="list-style-type: none"> ➤ AFPSO[1] ➤ FCDE
目的関数値(良さ)	探索点群内の目的関数値の相対値	$\frac{f(\mathbf{x}^i(k)) - f(\mathbf{x}^b(k))}{f(\mathbf{x}^w(k)) - f(\mathbf{x}^i(k))}$	<ul style="list-style-type: none"> ➤ 適応型CS
	目的関数値の全順序関係(改善頻度など)	$f(\hat{\mathbf{x}}^i(k)) \leq f(\mathbf{x}^i(k))$	<ul style="list-style-type: none"> ➤ 1/5ルール, CMA-ES ➤ jDE, JADE, SHADE

パラメータ調整則と変換不变性の関係

■ パラメータ調整則を付加すると、更新則の変換不变性へ影響を与える場合がある。

◆ パラメータ固定を前提とした不变性であったため

あるパラメータ値における探索状態を与える



探索状態に応じて、新たなパラメータ値を与える

→ **更新則の不变性を失う・補完することがある**

パラメータ調整則にも変換不变性の定義を適用可能

本章の指針

不变性に基づくパラメータ調整則の設計指針に基づき、高いロバスト性・適応性を具備するメタヒューリスティクスを開発する。

指針1：不变性維持

更新則 U_f 不変性を具備



調整則 F_f 不変性を維持

ロバスト性維持・適応性向上

指針2：不变性補完

更新則 U_f 不変性を欠如



調整則 F_f 不変性を補完

ロバスト性向上・適応性向上

指針1に従った適応型CSの設計

CSは相似変換不変性（スケール不変性）を欠如しているため、パラメータ調整則によって補完する。

不変性に基づく指針

指針1：不変性補完

指針2：不変性維持

1.2 : CSの適応化

相似変換不変性補完による
ロバスト性向上

多様化・集中化の探索戦略による
適応性向上

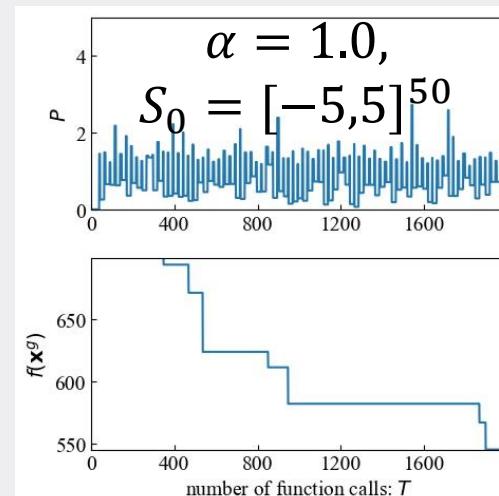
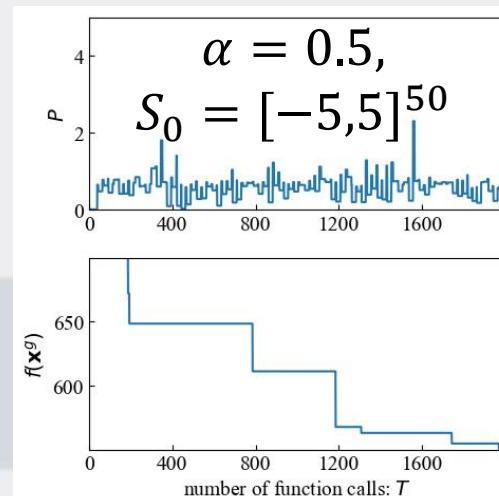
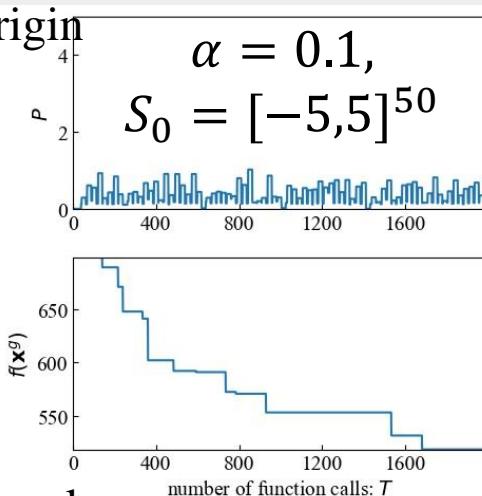
適応性に基づく指針

AFPSOの活性度制御

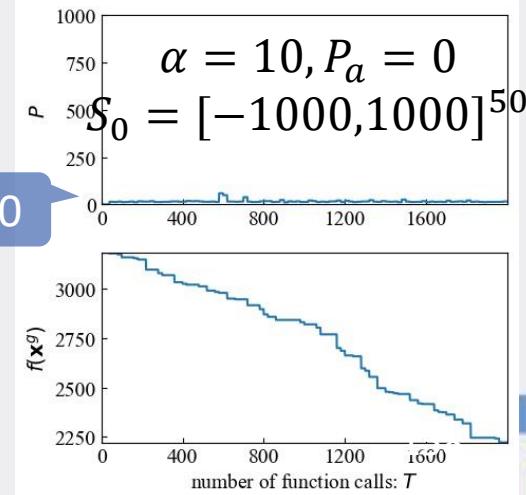
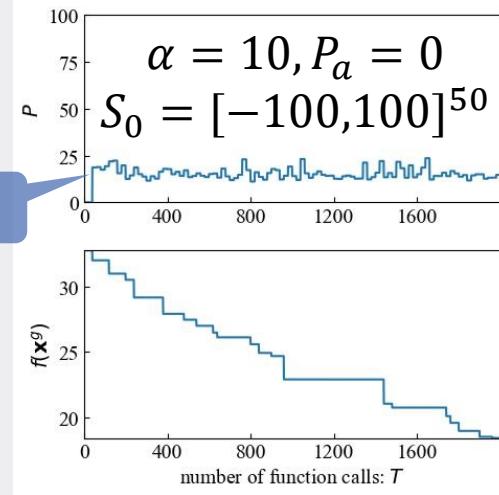
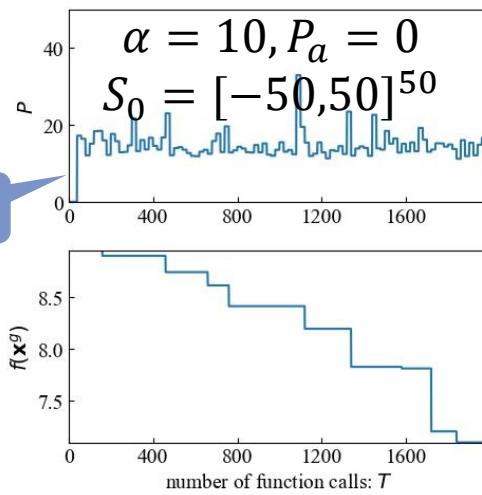
多様化・集中化の評価指標に基づくCSの解析

- パラメータ α の値で、活性度（摂動）の様子が変わる。
- 配置領域のサイズを変えても、活性度は変わらない。

Rastrigin



Griewank



相似変換不变性を補完した適応型CS

- 活性度を制御するように、 α を適応的に調整する。
- 探索状態を評価しながら、スケールパラメータ α を適応させることで、相似変換不变性を補完する。

✓ P を目標値 P_{target} に追従

(Global Searchでは)

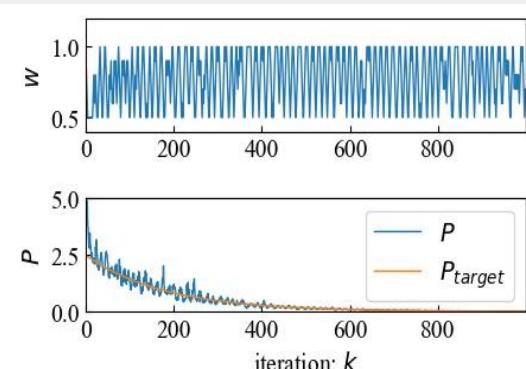
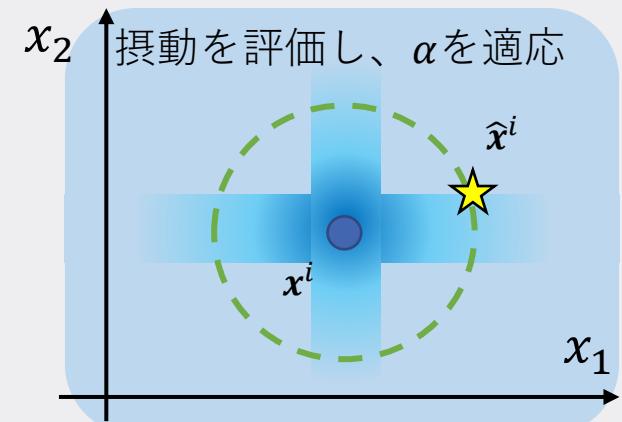
$$P = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|\hat{\mathbf{x}}^i - \mathbf{x}^i\|_2 = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|\alpha \mathbf{s}^i\|_2$$

(local Searchでは)

$$P = \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|r(\mathbf{x}^i - \mathbf{x}^j)\|_2 \quad (\alpha \text{とは無関係})$$

✓ P_{target} を漸減させることで、探索戦略を実現

$$P_{\text{target}}(T) = P_{\text{start}} (P_{\text{end}}/P_{\text{start}})^{T/T_{\text{max}}}$$



AFPSOと同様の調整則

パラメータ調整則による相似変換不変性の補完

不変性を補完するためにパラメータ調整則を工夫。

初期配置領域 $S_0 = [\gamma_{\min,n}, \gamma_{\max,n}]^N$ とする

$$\begin{aligned}
 P_{\text{target}}(T) &= P_{\text{start}} \left(\frac{P_{\text{end}}}{P_{\text{start}}} \right)^{\frac{T}{T_{\max}}} = \varepsilon_{\text{start}} \chi_{\text{width}} \left(\frac{\varepsilon_{\text{end}}}{\varepsilon_{\text{start}}} \right)^{\frac{T}{T_{\max}}} \\
 &= \frac{\varepsilon_{\text{start}}}{\sqrt{N}} \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 \left(\frac{\varepsilon_{\text{end}}}{\varepsilon_{\text{start}}} \right)^{\frac{T}{T_{\max}}} = \varepsilon_1 \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 \quad \text{定数部分を置いた} \\
 P &= \frac{1}{m\sqrt{N}} \sum_{i=1}^m \|\alpha \mathbf{s}^i\|_2 = \varepsilon_2 \sum_{i=1}^m \|\alpha \mathbf{s}^i\|_2 \quad \text{定数部分を置いた}
 \end{aligned}$$

ここで、 P を P_{target} に十分追従したと仮定すると、 $P \sim P_{\text{target}}$ である。

$$\begin{aligned}
 \varepsilon_1 \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 &= \varepsilon_2 \sum_{i=1}^m \|\alpha \mathbf{s}^i\|_2 \quad \text{定数部分を置いた} \\
 \Leftrightarrow \alpha &= \frac{\varepsilon_1}{\varepsilon_2 \sum_{i=1}^m \|\mathbf{s}^i\|_2} \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2 = \varepsilon \|\boldsymbol{\gamma}_{\max} - \boldsymbol{\gamma}_{\min}\|_2
 \end{aligned}$$

パラメータ調整則による相似変換不変性の補完

相似変換不変性を補完していることがわかる。

相似変換 $T_s: \mathbf{x} \mapsto \delta\mathbf{x}$ を与えると、

$$\alpha' = \varepsilon \|\gamma_{\max}' - \gamma_{\min}'\|_2 = \varepsilon \|\delta\gamma_{\max} - \delta\gamma_{\min}\|_2 = \delta\varepsilon \|\gamma_{\max} - \gamma_{\min}\|_2 = \delta\alpha$$

また、 $\mathbf{x}^{i'} = \delta\mathbf{x}^i, \mathbf{s}^{i'} = \mathbf{s}^i$ である。

よって、(Global Searchの) 近傍生成則全体は、

$$(U_f \circ T_s)(H; C) = \mathbf{x}^{i'} + \alpha' \mathbf{s}^{i'} = \delta\mathbf{x}^i + \delta\alpha \mathbf{s}^i = \delta(\mathbf{x}^i + \alpha \mathbf{s}^i) = \delta\hat{\mathbf{x}}^i = (T_s \circ U_f)(H; C)$$

となる。

よって、パラメータ調整則の不変性は有さないが、近傍生成則全体の不変性を補完していることが示された。

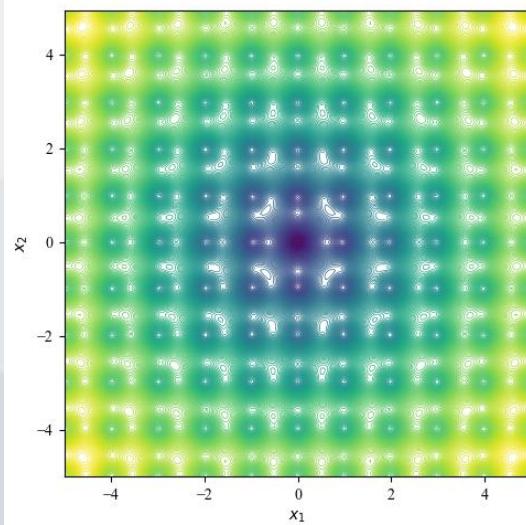
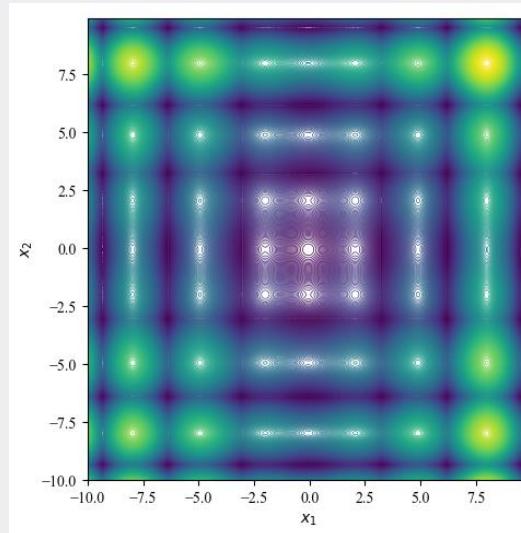
※Local Searchは元々アフィン変換不変性を満たしており、 α とは無関係

数値実験

- 数値実験を通じて、適応型CSの下記を検証する。
 - ◆ 実験1：スケール変換に対するロバスト性に関する検証
 - ◆ 実験2：全般的な性能に関する検証
- 相似変換不変性を欠如するCSと比較する。
- 初期配置領域内に初期探索点をランダムに配置した後、終了条件を満たすまで探索を進める試行を50回行い、最良解の目的関数値 $f(\mathbf{x}^b)$ の平均値を探索性能とする。
- 終了条件は評価回数が $T \geq T_{\max}$ を満たすことである。
- パラメータ設定
 - ◆ 探索点数 $m = 20$ 、排斥確率 $P_a = 0.25$ 、 $\beta = 1.9$
 - ◆ CS : α の設定は後述
 - ◆ 適応型CS : $\alpha_{\max} = 1.0, \alpha_{\min} = 0.01, \Delta\alpha = 0.1$

実験1：スケール変換に対する口バスト性の検証

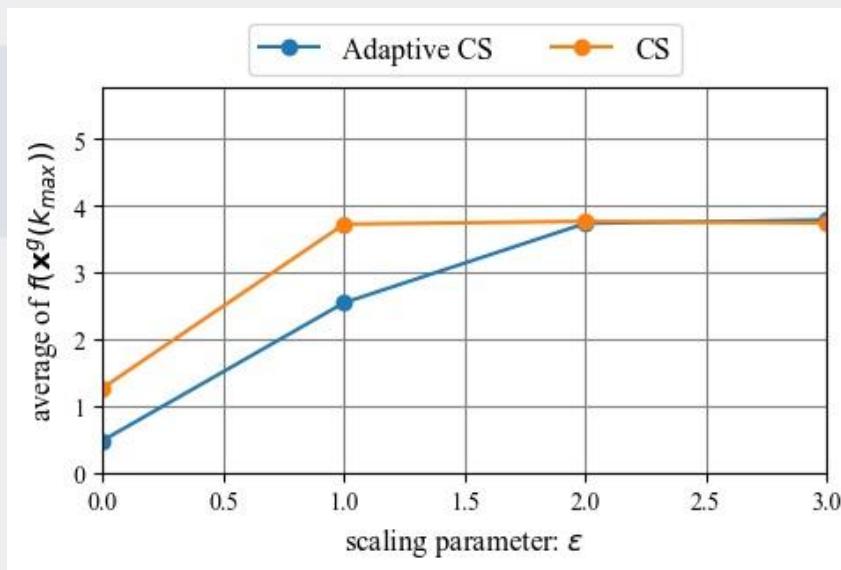
- 解空間に変換 T_s を与えるには、初期探索点を $T_s^{-1}(\mathbf{x}^i)$ とし、 $f(\mathbf{x}^i)$ を $f(T_s(\mathbf{x}^i))$ とする。
- 解空間を \mathbf{D}^ε でスケーリングし、性能の変動を確認する。
 - ◆ $\varepsilon = 0, 1, 2, 3$ ($\text{diag}[1, \dots, N], \text{diag}[1^2, \dots, N^2], \text{diag}[1^3, \dots, N^3]$)
- スケーリングすることで、目的関数に悪スケール性を発生すると同時に、解空間のサイズが変わる。
- 変数間独立・多峰性・良スケール性の関数を対象とする。

 f_7 : Ackley関数 f_9 : Alpine関数

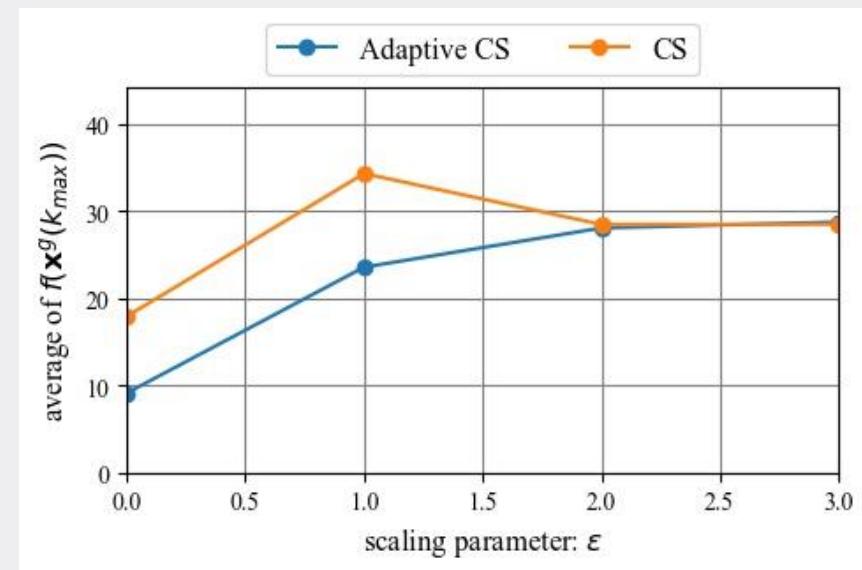
実験1：実験結果

■ スケーリングに対して、CSは口バストでないが、適応型CSは口バスト性がわずかに向上していることが確認できる。

- ◆ 強い悪スケール性(D^2, D^3)だと、相似変換不变性でも対処できない



Ackley関数



Alpine関数

※ $N = 50, T = 20000$ の結果だが、他の条件でも同様の傾向が確認された

実験2：全般的な性能に関する検証

- CSの平均的性能と比較するために、各条件で11通りの α の値における性能を導出しておく。
 - ◆ $\alpha = 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$
- このうち6位（中央値）は平均的性能とみなせる。
 - ◆ 知識・経験のないユーザは、ランダムに設定するため
- 提案手法を含めた12通りの中の順位を確認する。
 - ◆ 6位以上であれば、平均的に優れていると考えられる

結果例	最良値	中央値	最悪値
-----	-----	-----	-----

#	1位	3位	5位	6位	7位	9位	11位
条件1	$\alpha = 0.2$	$\alpha = 0.5$	$\alpha = 0.1$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 0.6$	$\alpha = 1.0$
条件2	$\alpha = 0.7$	$\alpha = 0.3$	$\alpha = 0.8$	$\alpha = 0.2$	$\alpha = 0.6$	$\alpha = 0.1$	$\alpha = 0.5$
条件3	$\alpha = 1.0$	$\alpha = 0.7$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.3$

実験2：実験結果 ($T = 2000$)

適応型CSはCSの平均的性能よりも優れている。

関数	N	CS (11種類の α)						適応型CS	
		1位	α	3位	α	6位	α	性能	順位
f_2	50	1204	0.1	1907	0.3	2415	0.6	1638■	3
	100	4359	0.1	4888	0.01	5388	1.0	3596*	1
	300	17798	0.1	18170	0.8	18623	1.0	24957	11
f_4	50	444	0.1	505	0.01	528	0.5	446■	2
	100	1020	0.1	1079	0.01	1147	0.4	1004*	1
	300	3380	0.1	3505	0.2	3596	0.9	3224*	1
f_8	50	2.47	1.0	2.70	0.8	2.79	0.7	2.56■	2
	100	4.73	0.9	4.91	1.0	5.05	0.7	4.58*	1
	300	14.4	0.8	14.8	0.4	15.2	0.5	16.0	11

※初期配置領域 S_0 は f_2 が $[-2,2]^N$ 、 f_4 が $[-5,5]^N$ 、 f_8 が $[-50,50]^N$

※一部抜粋

※*は1位、 ■は6位以上

実験2：実験結果 ($T = 20000$)

適応型CSはCSの平均的性能よりも優れている。

関数	N	CS (11種類の α)						適応型CS	
		1位	α	3位	α	6位	α	性能	順位
f_2	50	49.5	0.01	145	0.2	280	0.5	50.5 ■	2
	100	135	0.01	563	0.2	841	0.6	143 ■	2
	300	931	0.01	3062	0.9	3067	1.0	1010 ■	2
f_4	50	101	0.01	339	0.2	397	0.5	141 ■	2
	100	287	0.01	818	0.2	886	0.9	350 ■	2
	300	1593	0.01	2776	1.0	2855	0.2	1960 ■	2
f_8	50	0.27	0.3	0.30	0.2	0.35	0.7	0.15*	1
	100	0.83	0.4	0.84	0.5	0.97	0.8	0.71*	1
	300	1.77	0.4	1.82	0.3	2.04	0.2	2.12	8

※初期配置領域 S_0 は f_2 が $[-2,2]^N$ 、 f_4 が $[-5,5]^N$ 、 f_8 が $[-50,50]^N$

※一部抜粋

※*は1位、 ■は6位以上

実験2：実験結果

■ 各条件において、CSの1位、3位、6位と適応型CSの2つの性能の中で1位の数をカウントした結果、下記が確認できる。

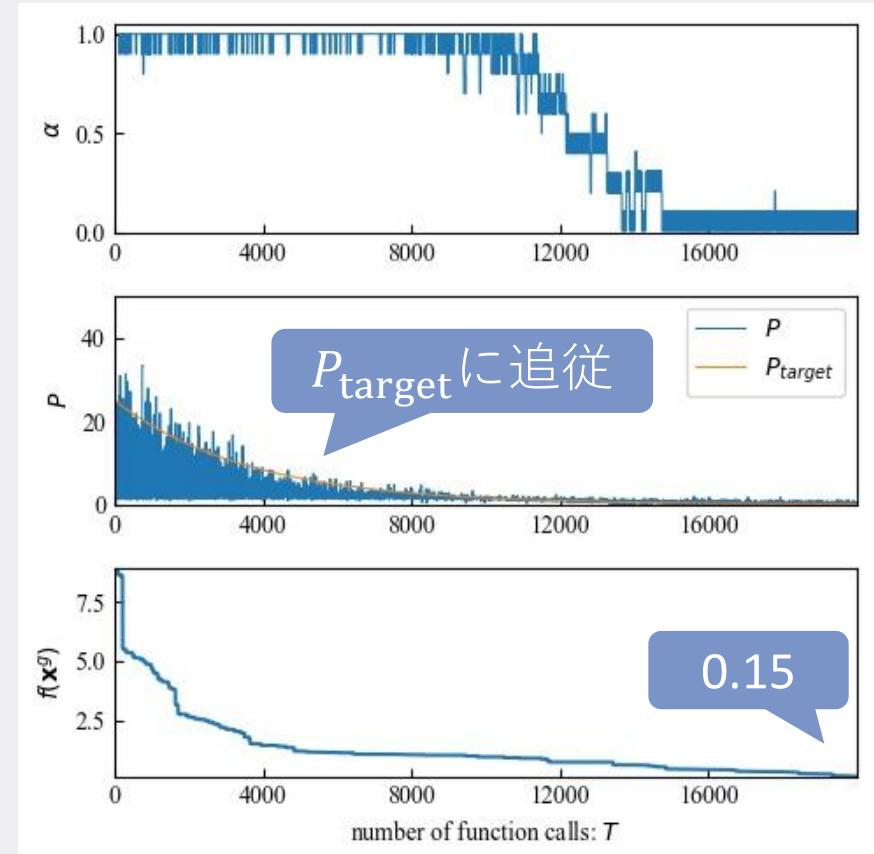
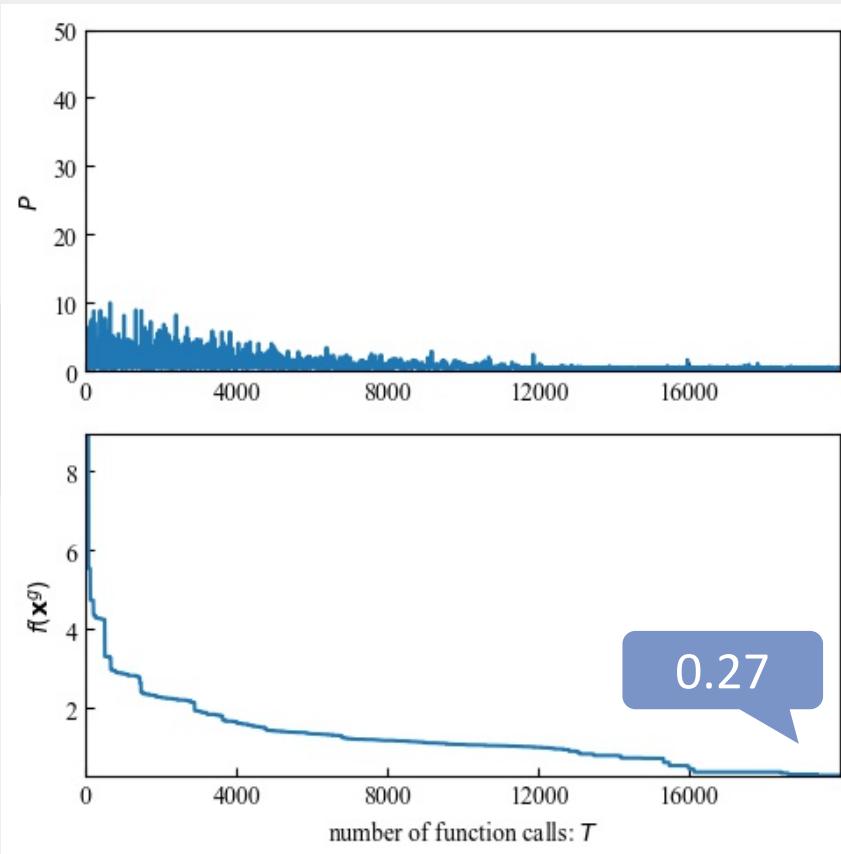
- ◆ 全般的に、best CSは適応型CSよりも優れる
- ◆ 全般的に、適応型CSはCSの3位よりも優れる

N	best CS (vs 適応型CS)			3rd CS (vs 適応型CS)			6th CS (vs 適応型CS)		
	2000	20000	All	2000	20000	All	2000	20000	All
10	7 (2)	2 (7)	9 (9)	6 (3)	2 (7)	8 (10)	4 (5)	1 (8)	5 (13)
50	7 (2)	4 (5)	11 (7)	4 (5)	0 (9)	4 (14)	3 (6)	0 (9)	3 (15)
100	5 (4)	8 (1)	13 (5)	3 (6)	1 (8)	4 (14)	3 (6)	0 (9)	3 (15)
300	6 (3)	9 (0)	15 (3)	6 (3)	4 (5)	10 (8)	6 (3)	1 (8)	7 (11)
All	25 (11)	23 (13)	48 (24)	19 (17)	7 (29)	26 (46)	16 (20)	2 (34)	18 (54)

考察

Griewank関数、 $T = 20000, N = 50$

探索戦略を実現するように α を適応的に調整。



まとめと課題

■ 適応型CSはスケール変換に対してロバストであることを確認。

- ◆ パラメータ調整則を付加することで、相似変換不变性を補完した

■ 適応型CSがCSの平均的性能よりも全般的に優れることを確認。

- ◆ パラメータ調整則を付加することで、適応性が向上

■ ただし、平均的性能にも劣っている場合もいくつか見られる。

- ◆ パラメータ調整則自身の改良の余地がある (ES系を参考)

■ パラメータ調整則による不变性の補完が可能な更新則は、ランダムウォーク系しかないと考えられる。

- ◆ 「調整則の設計指針1」の適用範囲の限界を認識

実験2：実験結果

変数間依存性の関数で提案手法が優れることが多い。

関数	N	$k = 100$				$k = 1000$			
		PSO	AFPSO	CRI-PSO	適応型 CRI-PSO	PSO	AFPSO	CRI-PSO	適応型 CRI-PSO
f_2	50	2236	606	574**	380*	99.6	624	84.8**	47.6*
	100	29456	1970**	3759	1570*	1831	2176	531**	104*
	300	111619	9342**	35519	7877*	111776	8829**	9170	1287*
f_4	50	460	292	265**	254*	186	291	145**	118*
	100	1364	766**	784	653*	592	752	458**	250*
	300	5029	2809**	3462	2660*	5015	2835	2591**	1112*
f_5	50	297	70.7**	153	68.3*	34.1	69.2	24.9**	1.02*
	100	1469	314*	700	320**	363	324	265**	54.4*
	300	15518	2878*	7204	2891**	4968	2886**	3333	928*

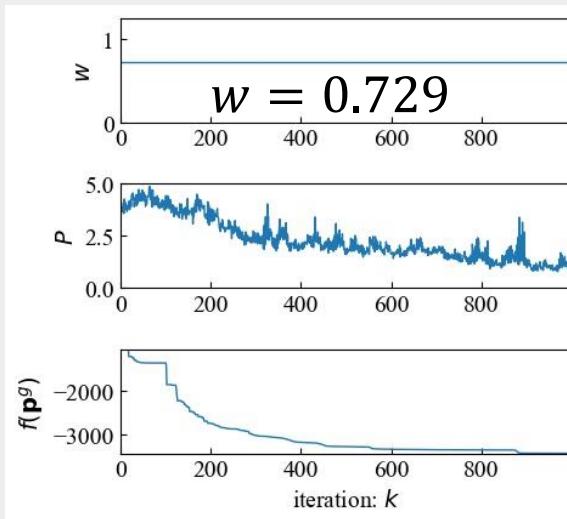
※一部抜粋

※ * は1位、 ** は2位

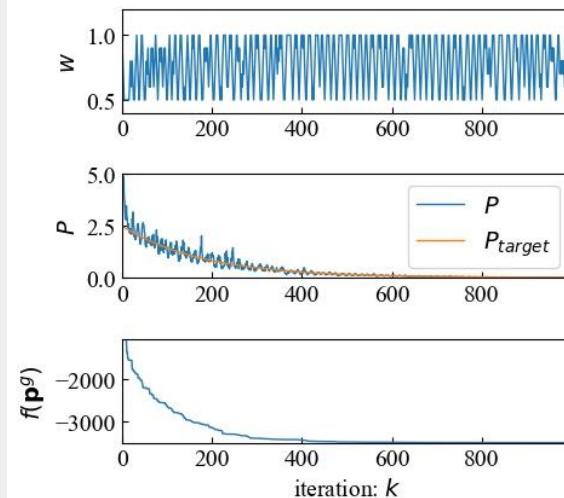
考察

 2^N minima関数、 $T = 20000, N = 50$ 探索戦略を実現するように w を適忯的に調整。

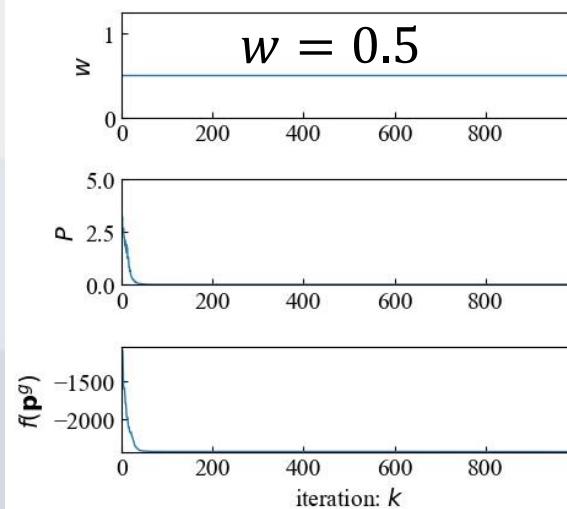
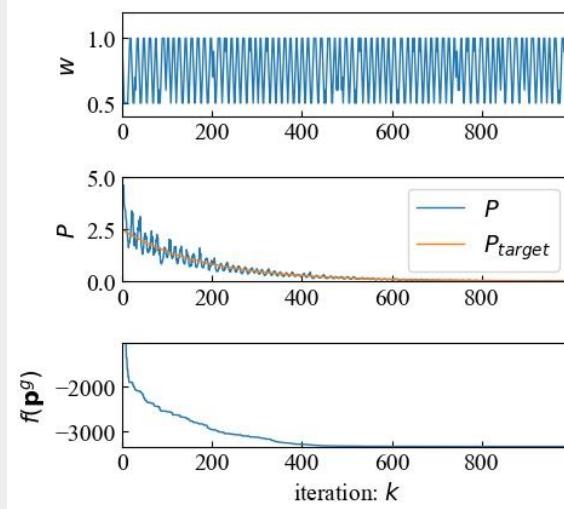
PSO



AFPSO



CRI-PSO

適忯型
CRI-PSO

CMD-DSの変換不变性定理

定理31

CMD-DSの更新則が条件1と条件2を満たすとき、アフィン変換不变性を有する。

証明

変換 T_s は解空間に属する任意の量（ベクトル、あるいはそれに基づく行列）に影響を及ぼす。このとき、更新則 U_f に含まれる量 $h \in H(k)$ が T_s の影響によってどのように変換されるのかを考える。 T_s によって影響を受けた任意の量 h を h' と表すことにする。

また、アフィン変換 T_s に対して、 $\mathbf{u}^i(k)$ 、 $\mathbf{s}^i(k)$ 、 $\boldsymbol{\Sigma}_*^{1/2}(k)$ は下記のように表せる。

$$\mathbf{u}^{i'}(k) = \mathbf{C}\mathbf{u}^i(k) - \mathbf{t}$$

$$\mathbf{s}^{i'}(k) = \mathbf{s}^i(k)$$

$$\boldsymbol{\Sigma}_*^{1/2'}(k) = \mathbf{C}\boldsymbol{\Sigma}_*^{1/2}(k) \quad (\text{補題4})$$

(次のページ)

CMD-DSの変換不变性定理

よって、次が成り立つ。

$$\begin{aligned} & (U_f \circ T_s)(H(k); C(k)) \\ &= \mathbf{u}^{i'}(k) + \alpha \boldsymbol{\Sigma}_*^{1/2'}(k) \mathbf{s}^{i'}(k) \\ &= \mathbf{C}\mathbf{u}^i(k) - \mathbf{t} + \alpha \mathbf{C}\boldsymbol{\Sigma}_*^{1/2}(k) \mathbf{s}^i(k) \\ &= \mathbf{C} \left(\mathbf{u}^i(k) + \alpha \boldsymbol{\Sigma}_*^{1/2}(k) \mathbf{s}^i(k) \right) - \mathbf{t} \\ &= \mathbf{C}\mathbf{x}^i(k+1) - \mathbf{t} \\ &= (T_s \circ U_f)(H(k); C(k)) \end{aligned}$$

これは式(4.1)を満たすことである。

以上から、題意は満たされた。



中心ベクトル \mathbf{u}^i に関する不变性定理①

定理32

CMD-DSの更新則において、 \mathbf{u}^i の形式が $\mathbf{u}^i = \mathbf{A}\mathbf{v}$ であるとき、 \mathbf{u}^i がアフィン変換不变性を有するための条件は $\mathbf{A} = \mathbf{I}$ が成立することである。ただし、 \mathbf{I} は単位行列である。

証明

アフィン変換 T_s に対して、 $\mathbf{v}' = \mathbf{C}\mathbf{v} - \mathbf{t}$ となる。
また、 $\mathbf{A}' = \mathbf{I} = \mathbf{A}$ である。よって、

$$\begin{aligned}\mathbf{u}^{i'}(k) &= \mathbf{A}'\mathbf{v}' = \mathbf{I}(\mathbf{C}\mathbf{v} - \mathbf{t}) = \mathbf{C}\mathbf{v} - \mathbf{t} = \mathbf{C}\mathbf{I}\mathbf{v} - \mathbf{t} = \mathbf{C}\mathbf{A}\mathbf{v} - \mathbf{t} \\ &= \mathbf{C}\mathbf{u}^i(k) - \mathbf{t}\end{aligned}$$

となる。以上から、題意は満たされた。 ■

中心ベクトル \mathbf{u}^i に関する不变性定理②

定理33

CMD-DSの更新則において、 \mathbf{u}^i の形式が $\mathbf{u}^i = \sum_j \mathbf{A}_j(\mathbf{x}_j - \mathbf{y}_j) + \mathbf{v}$ であるとき、 \mathbf{u}^i がアフィン変換不变性を有するための条件は $\mathbf{A}_j = \delta_j \mathbf{I}$ が成立することである。ただし、 $\delta_j > 0$ である。

証明

アフィン変換 T_s に対して、 $\mathbf{x}_j, \mathbf{y}_j, \mathbf{v}$ は $\mathbf{x}' = \mathbf{C}\mathbf{x} - \mathbf{t}$ となる。また、 $\mathbf{A}_j' = \delta_j \mathbf{I} = \mathbf{A}_j$ である。よって、

$$\begin{aligned}
 \mathbf{u}^{i'}(k) &= \sum_j \mathbf{A}_j'(\mathbf{x}_j' - \mathbf{y}_j') + \mathbf{v}' = \sum_j (\delta_j \mathbf{I})((\mathbf{C}\mathbf{x}_j - \mathbf{t}) - (\mathbf{C}\mathbf{y}_j - \mathbf{t})) + \mathbf{C}\mathbf{v} - \mathbf{t} \\
 &= \sum_j (\delta_j \mathbf{I})\mathbf{C}(\mathbf{x}_j - \mathbf{y}_j) + \mathbf{C}\mathbf{v} - \mathbf{t} = \mathbf{C} \sum_j (\delta_j \mathbf{I})(\mathbf{x}_j - \mathbf{y}_j) + \mathbf{C}\mathbf{v} - \mathbf{t} \\
 &= \mathbf{C} \left(\sum_j \mathbf{A}_j (\mathbf{x}_j - \mathbf{y}_j) + \mathbf{v} \right) - \mathbf{t} = \mathbf{C}\mathbf{u}^i(k) - \mathbf{t}
 \end{aligned}$$

となる。以上から、題意は満たされた。

不变性検証用サンプルアルゴリズム $\mathcal{A}_{\text{test}}$

■ $\mathcal{A}_{\text{test}}$ の更新則における \mathbf{u}^i と \mathbf{s}^i を下記のように設計する。

$$\mathbf{u}^i(k) = \mathbf{x}^i(k) \quad \text{式(6.3)}$$

$$\mathbf{s}^i(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad \text{式(6.4)}$$

ただし、 $\mathcal{Z} = \{\mathbf{x}^i(k) | i = 1, 2, \dots, m\}$ 、 $\alpha = 1$ 、 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ は標準正規分布である。

■ この更新則は、条件1、条件2を満たすため、アフィン変換不变性を有する。（定理1）

◆ 条件1の成立

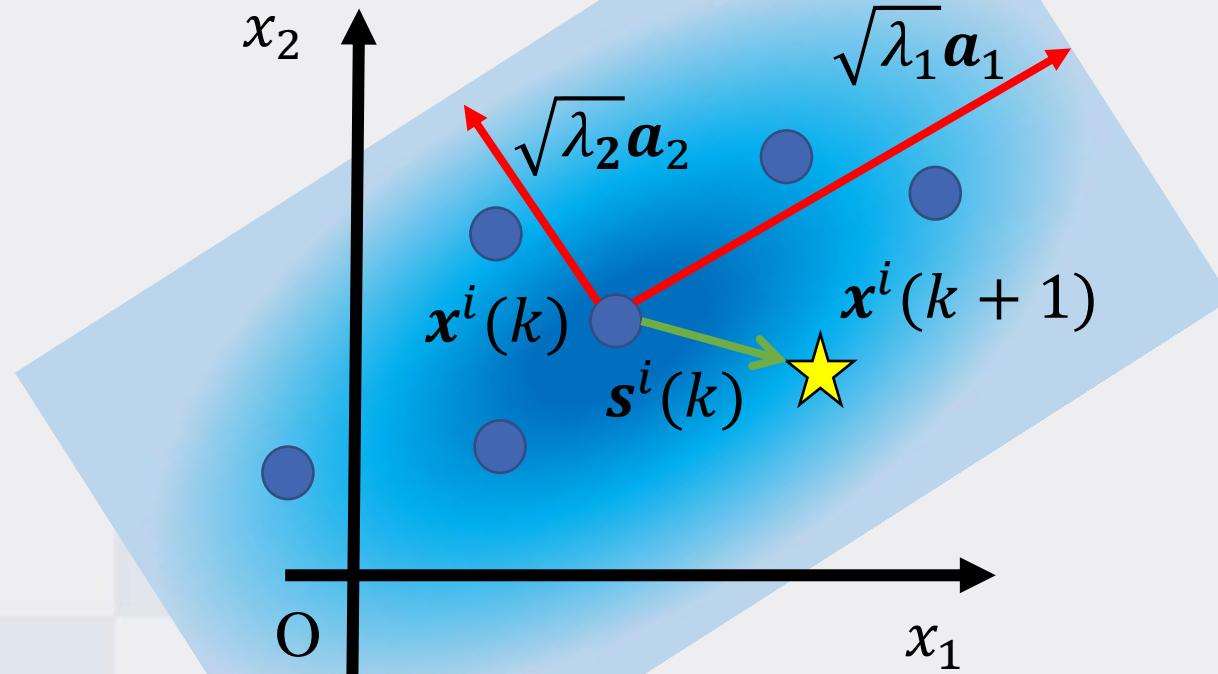
\mathbf{u}^i は定理2の形式 ($\mathbf{u}^i = \mathbf{Av}$) に該当するが、この場合の \mathbf{u}^i がアフィン変換不变性を具備するための条件は、 $\mathbf{A} = \mathbf{I}$ が成立することである。式(6.3)は明らかにこれを満たす。

◆ 条件2の成立

\mathbf{P} は標準正規分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ に従うが、明らかに標準正規分布は確率変数同士で独立であり、パラメータは設計変数とは独立している。

$\mathcal{A}_{\text{test}}$ の近傍のイメージ

$$\boldsymbol{x}^i(k+1) = \boldsymbol{x}^i(k) + \Sigma_*^{1/2}(k) \boldsymbol{s}^i(k)$$

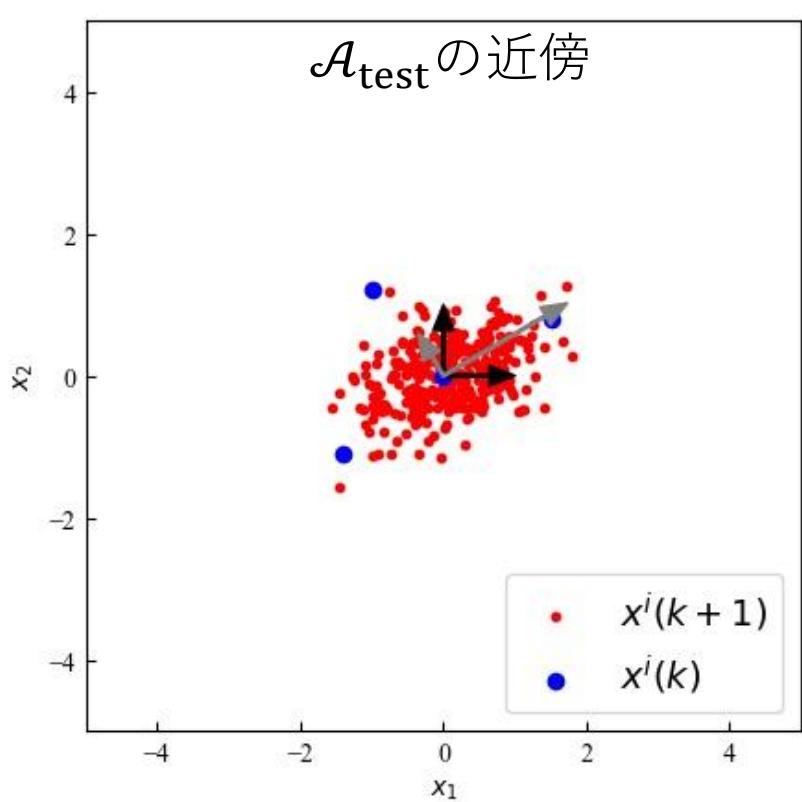


ただし、 $\boldsymbol{s}^i(k)$ は乱数ベクトル、 $\boldsymbol{x}^i(k)$ は探索点である。

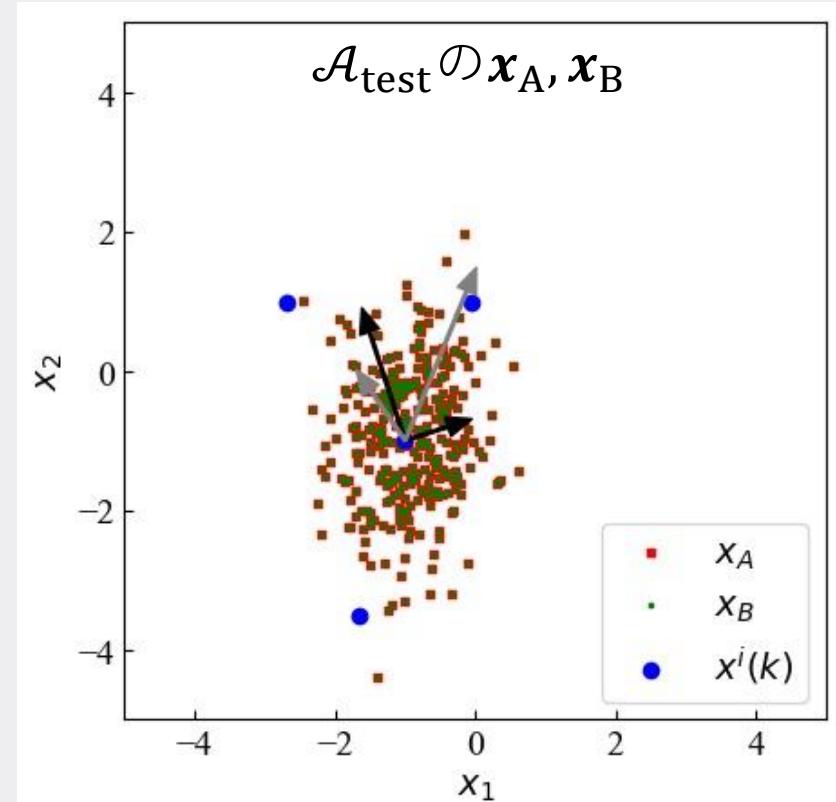
実質的に、 $\boldsymbol{x}^i(k+1) \sim \mathcal{N}(\boldsymbol{x}^i(k), \boldsymbol{\Sigma}(k))$ である。

不変性定理の数値実験的検証結果①

CMD-DSのサンプルアルゴリズムがアフィン変換不変性を具備することを確認



近傍が各固有ベクトルに沿っていると同時に、各固有値によってスケーリングされている。

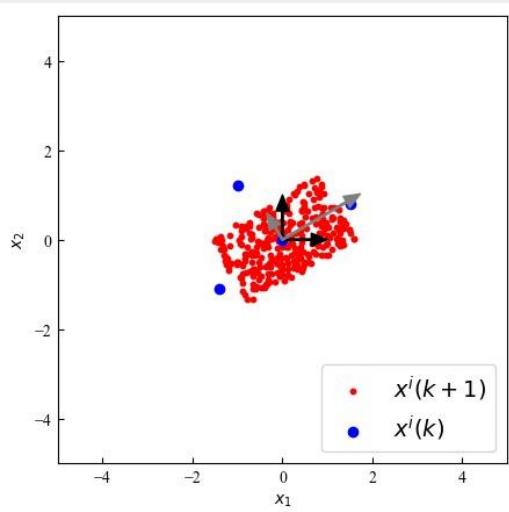


x_A, x_B が一致している。つまり、近傍生成がアフィン変換に依存していない。

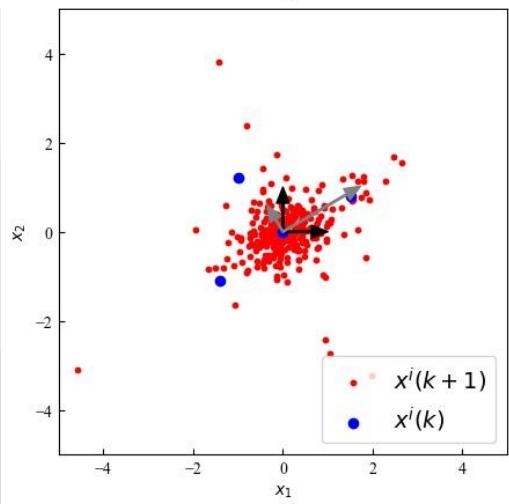
不変性定理の数値実験的検証結果②

正規分布以外も、アフィン変換不変性の具備を確認
通常の近傍

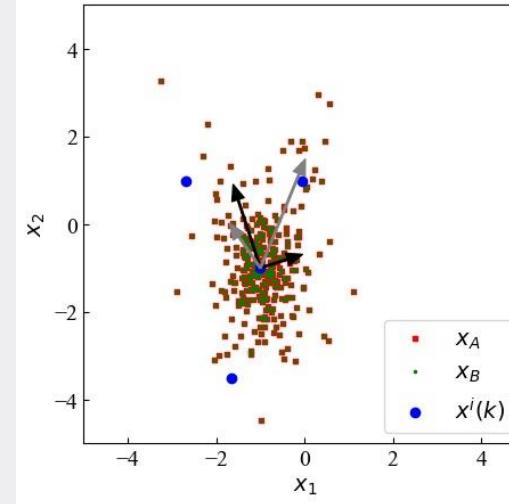
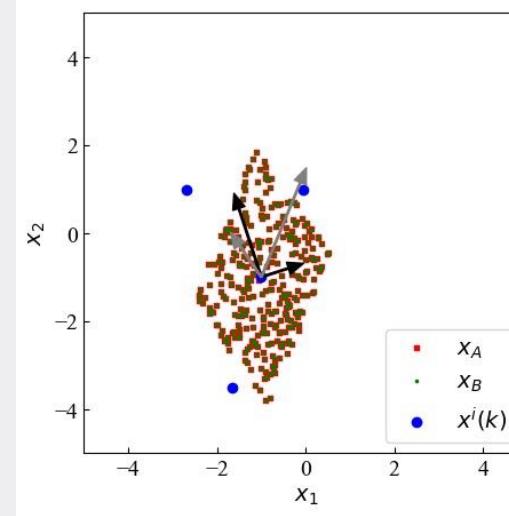
一様分布



Lévy分布



\hat{x}_A, \hat{x}_B



\mathcal{A}_{PSO} と Linear PSOとの類似点

\mathcal{A}_{PSO} は Linear PSO の最適化能力を引き継いでいる。

Linear PSOの更新式は下記の通り。

$$\mathbf{x}^i(k+1) = \mathbf{x}^i(k) + w\mathbf{v}^i(k) + c_1 r_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 r_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k))$$

ここで、 $\mathbf{z}^i(k) = (c_1 r_1 \mathbf{p}^i(k) + c_2 r_2 \mathbf{p}^g(k))/\phi$, $\phi = c_1 r_1 + c_2 r_2$ とおくと、

$$\mathbf{x}^i(k+1) = (1 - \phi)\mathbf{x}^i(k) + \phi\mathbf{z}^i(k) + w\mathbf{v}^i(k) \quad \text{摂動}$$

$\lim_{k \rightarrow \infty} \mathbf{v}^i(k) = \mathbf{0}$ ※であるため、探索点は \mathbf{x}^i と \mathbf{z}^i の内分位置が平衡点となる。つまり、有望領域である p-best と g-best に収束する点が最適化能力である。



\mathcal{A}_{PSO} の更新則は下記の通り。

$$\begin{aligned} \mathbf{x}^i(k+1) &= \mathbf{x}^i(k) + c_1 r_1 (\mathbf{p}^i(k) - \mathbf{x}^i(k)) + c_2 r_2 (\mathbf{p}^g(k) - \mathbf{x}^i(k)) + \beta(1 - c_1 r_1 - c_2 r_2) \boldsymbol{\Sigma}_*^{1/2}(k) \mathbf{s}^i(k) \\ &= (1 - \phi)\mathbf{x}^i(k) + \phi\mathbf{z}^i(k) + \beta(1 - c_1 r_1 - c_2 r_2) \boldsymbol{\Sigma}_*^{1/2}(k) \mathbf{s}^i(k) \quad \text{摂動} \end{aligned}$$

※パラメータを安定領域に設定した場合

数値実験

- 数値実験を通じて、 \mathcal{A}_{PSO} の下記を検証する。
 - ◆ 実験1：アフィン変換に対するロバスト性に関する検証
 - ◆ 実験2：全般的な性能に関する検証
- PSO、Linear PSO（アフィン変換不変性を具備）と比較する。
- 初期配置領域内に初期探索点をランダムに配置した後、終了条件を満たすまで探索を進める試行を50回行い、最良解の目的関数値 $f(\mathbf{x}^g(k_{\max}))$ の平均値を探索性能とする。
- パラメータ設定
 - ◆ 探索点数 $m = 20$ 、 $c_1 = c_2 = 1.4955$
 - ◆ PSO、Linear PSO : $w = 0.729$
 - ◆ \mathcal{A}_{PSO} : $\beta = 2$

実験2：実験結果

多くの条件で提案手法が優れることが多い。

関数	N	$k = 100$			$k = 1000$		
		PSO	Linear PSO	\mathcal{A}_{PSO}	PSO	Linear PSO	\mathcal{A}_{PSO}
f_2	50	2236	602**	577*	99.6*	590	538**
	100	29456	1528**	1286*	1831	1521**	1366*
	300	111619	5426**	4457*	111776	5666**	4185*
f_4	50	460	353*	368**	186*	354	350**
	100	1364	870**	831*	592*	841	823**
	300	5029	2964**	2858*	5015	2956**	2813*
f_5	50	297	60.3**	56.4*	34.1*	60.6**	63.2
	100	1469	248*	285**	363	236*	248**
	300	15518	2090*	2476**	4968	2141*	2455**

※一部抜粋

※*は1位、 **は2位

\mathcal{A}_{PSO} の実験2：実験結果

■ 各条件において、3種類の手法の性能の中で1位の数をカウントした結果、下記が確認できる。

- ◆ 全般的には、PSOが最も優れる
- ◆ 特に低次元では、PSOが優れており、それ以外では、 \mathcal{A}_{PSO} が優れる
- ◆ PSOは1位か3位しかない（高次元では大きく劣化）。

N	PSO			Linear PSO			\mathcal{A}_{PSO}		
	$k = 100$	$k = 1000$	All	$k = 100$	$k = 1000$	All	$k = 100$	$k = 1000$	All
10	9	9	18	0	0	0	0	0	0
50	0	9	9	3	0	3	6	0	6
100	0	5	5	2	1	3	7	3	10
300	0	0	0	2	2	4	7	7	14
All	9	23	32	7	3	10	20	10	30

実験2：実験結果

■ 各条件において、3種類の手法の性能の中で1位、2位の数をカウントした結果、下記が確認できる。

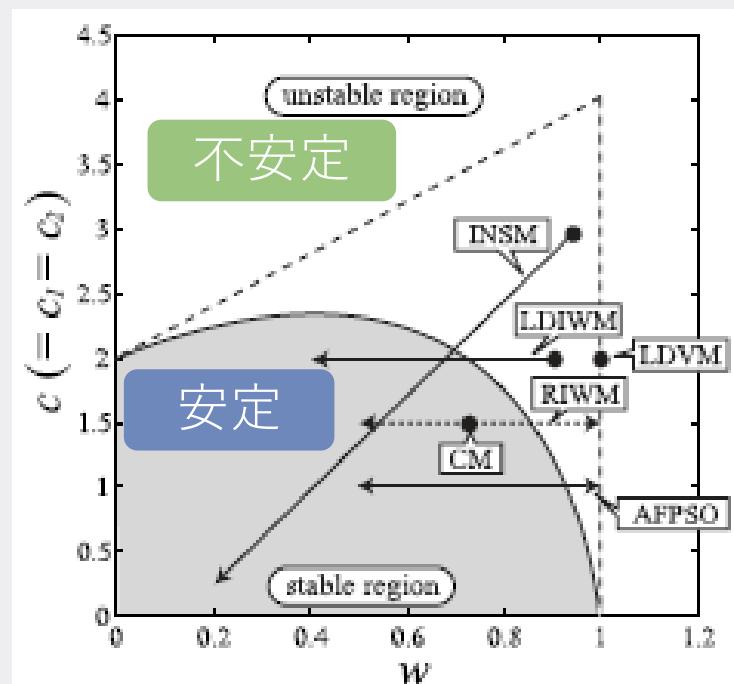
- ◆ 全般： $\mathcal{A}_{\text{PSO}} > \text{Linear PSO} > \text{PSO}$
- ◆ 低次元： $\text{PSO} > \mathcal{A}_{\text{PSO}} > \text{Linear PSO}$
- ◆ 高次元： $\mathcal{A}_{\text{PSO}} > \text{Linear PSO} > \text{PSO}$

➤ 1位の数はPSOが最も多いが、2位がない(高次元は3位のみ)

N	PSO			Linear PSO			\mathcal{A}_{PSO}		
	$k = 100$	$k = 1000$	All	$k = 100$	$k = 1000$	All	$k = 100$	$k = 1000$	All
10	9	9	18	4	5	9	5	4	9
50	0	9	9	9	2	11	9	7	16
100	0	5	5	9	4	13	9	9	18
300	0	0	0	9	4	13	9	14	23
All	9	23	32	31	15	46	32	34	66

高次元における結果の考察

- 通常のPSOの安定領域の境界線は、高次元になると安定領域側に多少小さくなる[1]。
 - ◆ CMは不安定領域に近づくため、高い性能を発揮できない
 - ◆ 係数乱数が行列であるため、全ての要素で安定状態が実現しにくくなる(集中化が弱い)
- 一方、Linear PSOや \mathcal{A}_{PSO} は係数乱数がスカラであるため、高次元でも多様化・集中化が実現できた。
 - ◆ 高次元ではPSOと大差がつく
- Linear PSOは近傍の多様性が低いが、 \mathcal{A}_{PSO} は乱数ベクトルによる摂動による多様性を維持できる。



PSOの安定・不安定領域

各種アルゴリズムの変換不変性の有無

手法	解空間のアフィン変換不変性			目的関数の 単調変換不変性
	平行移動不变性	回転不变性	スケール不变性	
ES	○	○	×	○
実数値GA	○	○	×	○
実数値SA	○	○	×	×
PSO	○	×	○	○
CRI-PSO ・ 適応型CRI-PSO	○	○	相似○／スケール×	○
DE	○	×	○	○
ABC	○	×	○	×
HBC-ABC	○	○	相似○／スケール×	×
FA	○	○	×	○
CS	○	×	×	○
CMD-CS	○	○	○	○
適応型CS	○	×	相似△／スケール△	○
CMD-DS	○	○	○	○

本研究の(個別の)有用性

- 不変性を付加したアルゴリズムは、ロバスト性を向上したこと。
- 適応性を付加したアルゴリズムは、ロバスト性を向上あるいは維持すると同時に、適応性を向上したこと。
- 不変性に基づくフレームワークに従い、PSOの最適化能力を埋め込むことで簡潔に設計したアルゴリズムが、PSOの基本的な最適化能力を引き継ぎながら、高いロバスト性を具備すること。

性質	CRI-PSO	HBC-ABC	CMD-CS	適応型CS	適応型CRI-PSO	$\mathcal{A}_{\text{PSO}}(\text{CMD-DS})$
ロバスト性	○	○	◎	△	○	◎
適応性	×	×	×	○	◎	×

今後の展望①

フレームワーク + 適応的パラメータ調整則による高い
ロバスト性・適応性を具備するアルゴリズム設計。

最適化能力を埋め込む

CMD-DS

アフィン変換不変性の具備
が保証される

$$\boldsymbol{x}^i(k+1) = \boldsymbol{u}^i(k) + \alpha \boldsymbol{\Sigma}_*^{1/2}(k) \boldsymbol{s}^i(k)$$

更新則 U_f

$$\alpha(k+1)$$

探索状態

調整則 F_f

$$\alpha(k+1) = F_f(\alpha(k), H(k); C(k))$$

先行研究で優れた
適応的調整則を採用

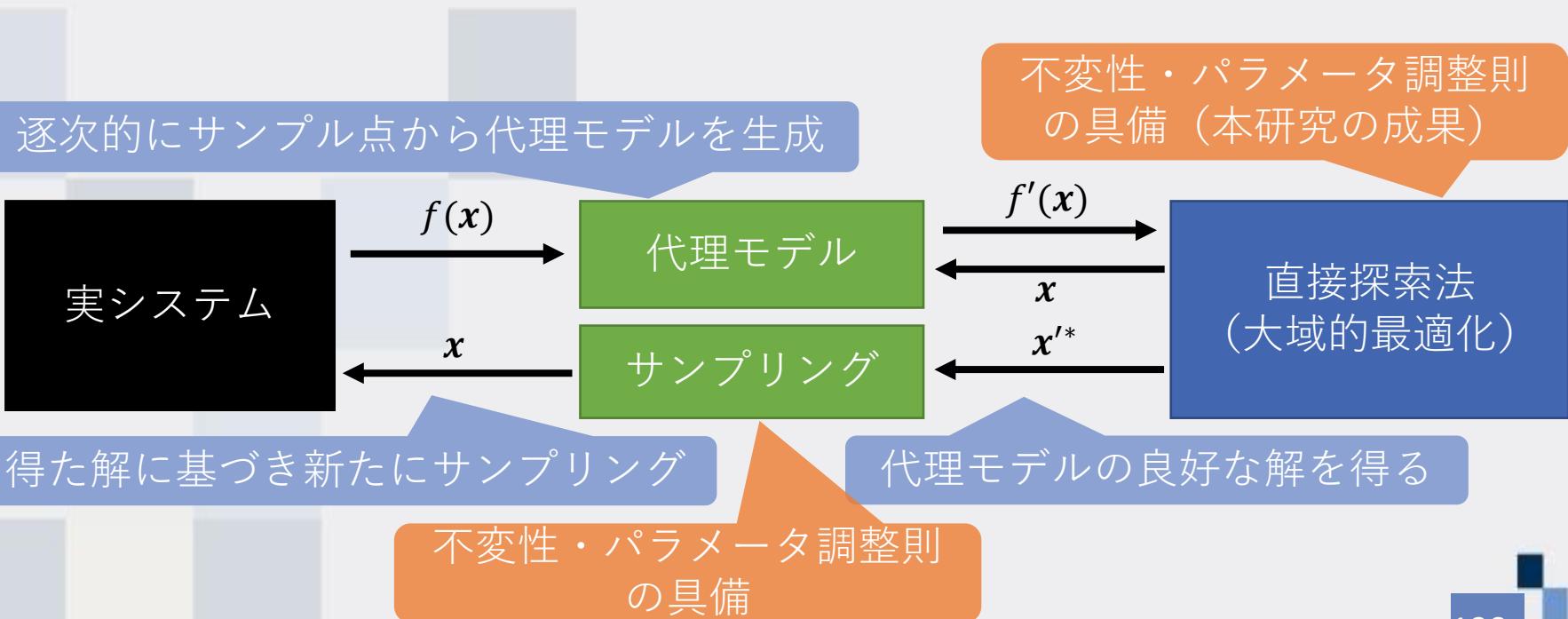
指針2（アフィン変換不変性
の維持）に基づく設計

優れたCMA-ESの調整則[2]、
安定性が保証されたフレームワーク[3]など

今後の展望②

BBOに有効な逐次近似最適化においても、本研究のアプローチを導入することで、高いロバスト性・適応性を具備する総合的な最適化技術の開発が期待される。

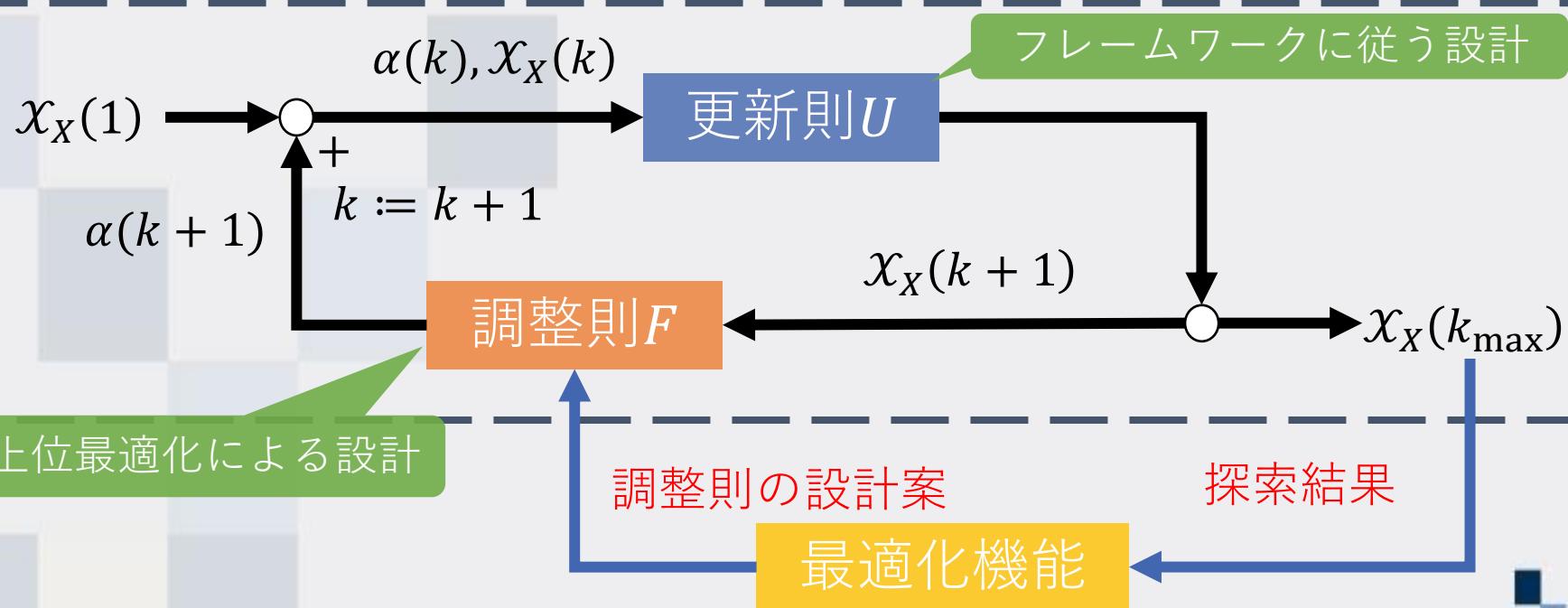
- ◆ AI技術やビッグデータ、コンピュータパワーなどの外部環境を活用可能な枠組み



今後の展望③

変換不变性のフレームワークとパラメータ調整則の融合により、より優れたアルゴリズム設計が期待される。

- ◆ [4,5,6,7]では上位最適化機能により、パラメータ調整則の設計案を自動的に獲得する（BBOに適している）



今後の展望（その他）

- 本研究は連続・単一目的・無制約を扱ったが、適用範囲を拡張しても、変換不变性の枠組みはそのまま適用可能である。
 - ◆ 多目的最適化、有制約最適化など(離散では一部活用可能)
- 近年、メタヒューリスティクスの探索の動特性・安定性・収束性に関する理論的研究も進んでいる。変換不变性の定理とこれらの定理を共存させることで、理論面の成果をさらに充実できる。
 - ◆ ESに関する理論的研究[6,18,40]
 - ◆ Markov連鎖による収束性に関する解析[41,42]

CMA-ES[6]との違い

CMA-ESはアフィン変換不变性を具備するアルゴリズムとして代表的だが、本研究とは下記の点で異なる。

項目	CMA-ES	CMD-DS（本研究）
最適化能力・汎用性	ES限定	限定されないが、最適化能力はない
パラメータ	適応的調整型 最適化能力を左右する適応的調整の具備が前提	固定型 調整則を前提としない
共分散行列 \mathbf{C} の更新	能動的 統計量を徐々に学習させ、統計量から標本を抽出	受動的 標本に基づき統計量を推定

Weight Recombination : 上位の近傍解に重みづけしたものを平均ベクトル \mathbf{m}

CSA : \mathbf{m} の変位がランダムなら一定、変位が変われば σ も調整

Hybrid CMA : \mathbf{C} を \mathbf{m} の変位方向の相関を強めながら、上位の摂動だけから構成される共分散行列に近づける

小熊さんの学位論文[15]との違い

LNRは安定性が保証されるアルゴリズムのフレームワークだが、本研究とは下記の点で異なる。

項目	LNR	CMD-DS（本研究）
最適化能力・汎用性	限定されないが、最適化能力はない	限定されないが、最適化能力はない
安定性・不变性	安定性を保証 不变性は保証されない (結果的に具備しやすいが)	不变性を保証 安定性は保証されない
共分散行列 \mathbf{C} の更新	受動的 探索点群に基づき 統計量を推定	受動的 標本に基づき 統計量を推定

本研究は「安定性」が範囲外であるため、この点では[15]のほうが優れている

金政さんの学位論文[11]との違い

これはGenetic Programmingによる適応的パラメータ調整則の設計だが、本研究とは下記の点で異なる。

アルゴリズムとの親和性
が低い可能性がある

項目	Genetic Programming	本研究（第5章）
設計方法・汎用性	自動・限定されない 手法や問題に合わせて自動設計	手動・限定可能性あり 設計指針は手動設計の支援 手法や問題に合わせて手動設計
不変性	一部に限定・理論でない 調整則と更新則の不变性の関係を指摘したが、例示に限定され、 証明ベースでない	網羅的・理論的 調整則と更新則の不变性の関係について、 網羅的かつ理論的に指摘・付加

本研究は「質の良い調整則を設計する」という観点は重要視していないため、この点では[11]のほうが優れている

田邊さんの学位論文[43]との違い

これは適応型DEの設計だが、本研究とは下記の点で異なる。

アルゴリズムとの親和性
が低い可能性がある

項目	田邊さん	本研究（第5章）
調整則の解析方法	統一的・一般性が高い 理想（Oracleパラメータ）に基づくシミュレーションにより、環境変化へのロバスト性・適応性の検証が可能	発見的・限定可能性あり 探索状態の指標を導入し、その制御を目指すが、その推移を視覚的に確認するか、実際の性能を確認するに過ぎない
調整則の設計	信頼度が高く、非常に性能が高い 上記のシミュレーションに基づき設計するため、調整則の信頼度が高い	発見的・限定可能性あり ある探索状態の指標の制御を目指すが、発見的であり、信頼度が高いとはいえない

本研究は「一般的な調整則の解析・評価」は実施していないため、この点では[11]のほうが優れている

A. Atamnaの学位論文[44]との違い

これは有制約BBOのための適応型アルゴリズムの設計だが、本研究とは下記の点で異なる。

項目	Atamna	本研究（第5章）
最適化能力・汎用性	ESに限定	限定されないが、最適化能力はない
収束性・不变性	不变性・ 収束性の理論 に基づく マルコフ連鎖モデルとして 線型収束性を解析	不变性の理論に基づくが、 収束性は未確認
調整則の設計	ESに限定 CMA-ES系の調整則に注力	ある程度の汎用性 発見的だが、統一的な視点に基づく

本研究は「収束性」は範囲外であるため、この点では[44]のほうが優れている

参考文献③（補足資料）

- [19] "Multi-Parent Recombination with Simplex Crossover in Real Coded Genetic Algorithms," n GECCO, Vol.1, pp.657-664 (1999)
- [20] 「斜交座標系に基づく回転不变なブレンド交叉の提案」, 情報処理学会研究報告, Vol.2017-MPS-113, pp.1-6 (2017)
- [21] "On Rotationally Invariant Continuous-Parameter Genetic Algorithms," Journals of Advances in Engineering Software Vol.78, pp.52-59 (2014)
- [22] "Impact of Invariance in Search: When CMA-ES and PSO Face Ill-Conditioned and Non-Separable Problems," Journal of Applied Soft Computing, pp.5755-5769 (2011)
- [23] "Differential Evolution and Nonseparability: Using Selective Pressure to Focus Search," Proc. of the 9th Annual Conference on Genetic and Evolutionary Computation, pp.1428–1435 (2007)
- [24] "Solving Nonlinear Optimization Problems by Differential Evolution with a Rotation-Invariant Crossover Operation Using Gram-Schmidt Process," Proc. of NABIC 2010 (2010)
- [25] "A Study of Robustness of PSO for Non-Separable Evaluation Functions," Proc. of 2015 International Symposium on Nonlinear Theory and its Applications, pp.724-727 (2015)
- [26] "Firefly Optimization: A Study on Frame Invariance," Proceedings of 2017 IEEE Symposium Series on Computational Intelligence, pp.423-428 (2017)
- [27] "Enhancing Differential Evolution Utilizing Eigenvector-Based Crossover Operator," IEEE Transactions on Evolutionary Computation, Vol.19, No.1, pp.31-49 (2015)
- [28] "Evolution and Optimum Seeking," Wiley (1995)
- [29] "Particle Swarm Optimization," IEEE Connections, Vol.2, No.1, pp.8-13 (2004)

参考文献④ (補足資料)

- [30] "Differential Evolution Algorithm with Strategy Adaptation for Global Numerical Optimization," IEEE Trans. on Evo. Comp., Vol.13, No.2, pp.398-417 (2009)
- [31] "Differential Evolution Based on Covariance Matrix Learning and Bimodal Distribution Parameter Setting," Journal of Applied Soft Computing, Elsevier, Vol.18, pp. 232-247 (2014)
- [32] "Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems," IEEE Trans. Evo. Comp., Vol.10, No.6, pp.646-657 (2006)
- [33] "JADE: Adaptive Differential Evolution With Optional External Archive," IEEE Trans. Evo. Comp., Vol.13, No.5, pp.945-958 (2009)
- [34] "Success-History Based Parameter Adaptation for Differential Evolution," Proc. of the 2013 IEEE Cong. on Evo. Comp., pp.71-78 (2013)
- [35] "Concept of Feedback Controlled Differential Evolution and Its Realization," IEEJ Trans. on Electrical and Electronic Engineering, Vol.10, No.4, pp.423-437 (2015)
- [36] "Firefly Algorithm: Recent Advances and Applications," Int. Journal of Swarm Intelligence, Vol.1, No.1, pp.36–50 (2013)
- [37] "Improved Cuckoo Search Algorithm for Global Optimization," Int. Journal of Communications and Information Technology, Vol.1, No.1, pp.31-44 (2011)
- [38] "Adaptive Cuckoo Search Algorithm for Unconstrained Optimization," the Science World Journal, Vol.2014, pp.1-8 (2014)
- [39] "A New Adaptive Cuckoo Search Algorithm," Proc. of 2015 IEEE 2nd Int. Conf. on Recent Trends in Information Systems, Vol.49, No.5, pp.4677-4683 (2015)

参考文献⑤（補足資料）

- [40] 「Evolution Strategiesによる連続最適化—CMA-ESの設計原理と理論的基盤」，システム/制御/情報，Vol.60，No.7（2016）
- [41] “Linear Convergence of Comparison-Based Step-Size Adaptive Randomized Search via Stability of Markov Chains,” Society for Industrial and Applied Mathematics Journal on Optimization, Vol.26, No3, pp.1589-1624 (2016)
- [42] “Global Convergence Analysis of the Flower Pollination Algorithm: A Discrete-Time Markov Chain Approach,” Proc. of the Int. Conf. on Computational Science, Vol.108, No.4 (2017)
- [43] 「関数最適化問題に対する適応型差分進化法の研究」，東京大学 博士論文(2016)
- [44] “Analysis of Adaptive Randomized Algorithms for Black-Box Continuous Constrained Optimization,” Doctoral Thesis, Paris-Sud University (2017)
- [45] 「エンジン応答局面モデルと制御パラメータ最適化手法の開発」，計測自動制御学会論文集, Vol.47, No.10, pp.501-510 (2011)
- [46] 「N700系新幹線車両の低騒音化技術」，精密工学会誌, Vol.76, No.1 (2010)
- [47] 「 ϵ 制約メタヒューリстиクスを用いたエネルギー原発最適運用」，電気学会研究会資料産業計測制御研究会, IIC-11-49, pp.1-6 (2011)
- [48] 「音声認識とブラックボックス最適化」，日本音響学会誌, Vol.72, No.10, (2016)
- [49] “CMA-ES for Hyperparameter Optimization of Deep Neural Networks,” Proceedings of the 33rd International Conference on Machine Learning, Workshop Track, pp.1-8 (2016)
- [50] “Adaptive Stochastic Natural Gradient Method for One-Shot Neural Architecture Search,” Proc. of the 36th International Conference on Machine Learning, pp.171-180 (2019)