

九州大学 海洋システム工学専攻

システム設計特論（木村）

非線形数値最適化(3)

遺伝的手法(GA)による数値最適化

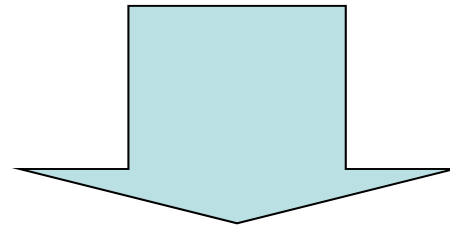
授業の資料等は

<http://sysplan.nams.kyushu-u.ac.jp/gen/index.html>

背景

勾配法 = 勾配を下って極小値を探す方法

多峰性関数に勾配法は使えない!



【解決策】

探索性能の下限

- 1. ランダムに動く → **ランダムサーチ・SA**
- 2. 多点で探す → **滑降シンプレックス法**

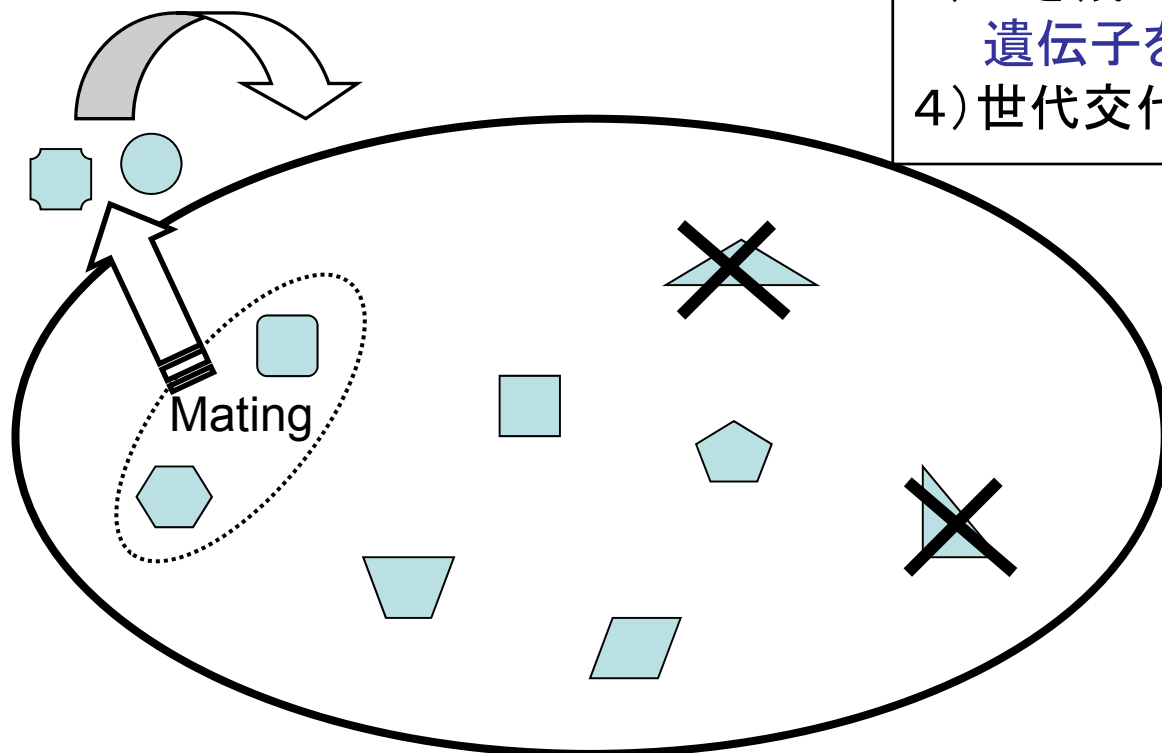
パラメータ数が5~6個程度までならお奨め

両方の特徴を併せ持つ最適化手法: **遺伝的アルゴリズム**

【遺伝的手法による最適化】

生物の進化の仕組みを模倣

生物個体 → 遺伝子によるコード化
優れた個体 → 高い評価値・生存率



【一般的な処理手順】

- 1) 設計したい物をパラメータ化 → 遺伝子
- 2) いろいろな個体を集めて集団を作り、
個体を評価に応じて自然淘汰
- 3) 生き残った個体同士で
遺伝子を交換 → 新しい個体生成
- 4) 世代交代を繰り返し、集団を進化

【探索の特徴】

- ・評価値の悪い個体でも
生き残るチャンスが与えられる
- ・優れた親同士を掛け合わせる
ことで、両親の優れた形質を
受け継いだ子個体が生成される
- ・集団全体が進化していく

たった1つの解候補を改善していくよりも、
より優れた解や多様な評価を持つ複数の解を発見できる

【遺伝的手法による最適化】

Simple GA (SGA) 遺伝的アルゴリズム研究黎明期の方法

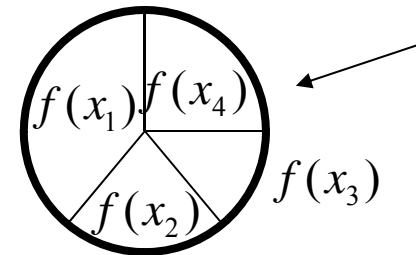
1) 解候補を固定長文字列に表現 → 遺伝子 例) 00101110110000

2) いろいろな個体を集めてN個体集団を作り、
各個体の評価値を計算して非負の「適応度(fitness)」を算出

3) [淘汰(selection)]

現在のN個の個体集合から、以下の確率 $P(x_i)$ に基づき重複を許してN個選択

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^N f(x_j)} \quad \begin{array}{l} \text{ルーレット選択,} \\ f(x) \text{は適応度} \end{array}$$



4) [交叉(crossover)・突然変異(mutation)]

・一定の割合で2個の個体ペアを作り、遺伝子文字列の部分を変換

1110 0011 (1111) ➡ 1110 0011 1010
0000 1110 (1010) ➡ 0000 1110 1111

・一定の割合で個体の遺伝子文字列の一部をランダムに変化

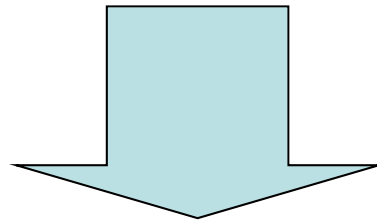
1110 0011 1111 ➡ 1110 1011 1010
 ↑

5) 手順2へ戻って繰り返す

【遺伝的手法による最適化】

Simple GA (SGA)の性能は、しばしばランダムサーチ以下

- ・連続関数最適化において文字列変換して交叉・突然変異操作することに意味があるか？
- ・組合せ最適化問題において交叉・突然変異操作すると、ほとんど致死遺伝子になる！
- ・集団全体が一気に世代交代 → 多様性の喪失／生物現象からの乖離



Simple GA (SGA)は、非実用的

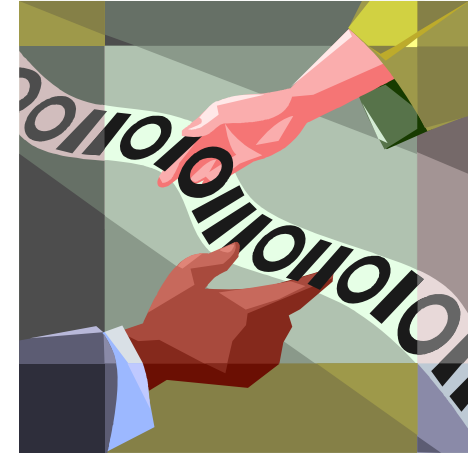
実用的なGAへ

- 連続関数における交叉方法: UNDX, SPX, REX
- TSPにおいて致死遺伝子を生成しない交叉方法: EAX
- 序々に世代交代: MGGモデル, JGGモデル
- 究極のGA: CMA-ES + NESs

【より洗練された遺伝的手法の枠組み1】

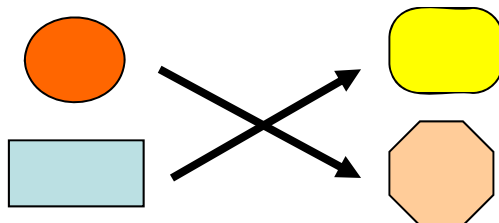
(1) ビットストリングにこだわらない遺伝子表現

遺伝子は解のコード化「表現」であるので、
解が完全に表現可能で、かつ
操作・改変が容易な表現であることがのぞましい



(2) 親の形質(特徴)を受け継ぐ交叉オペレーション

交叉は、親の優れた特徴を受け継ぐことが可能なオペレーション
でなければならない
ビットストリングGAのように親の特徴を破壊しては意味がない



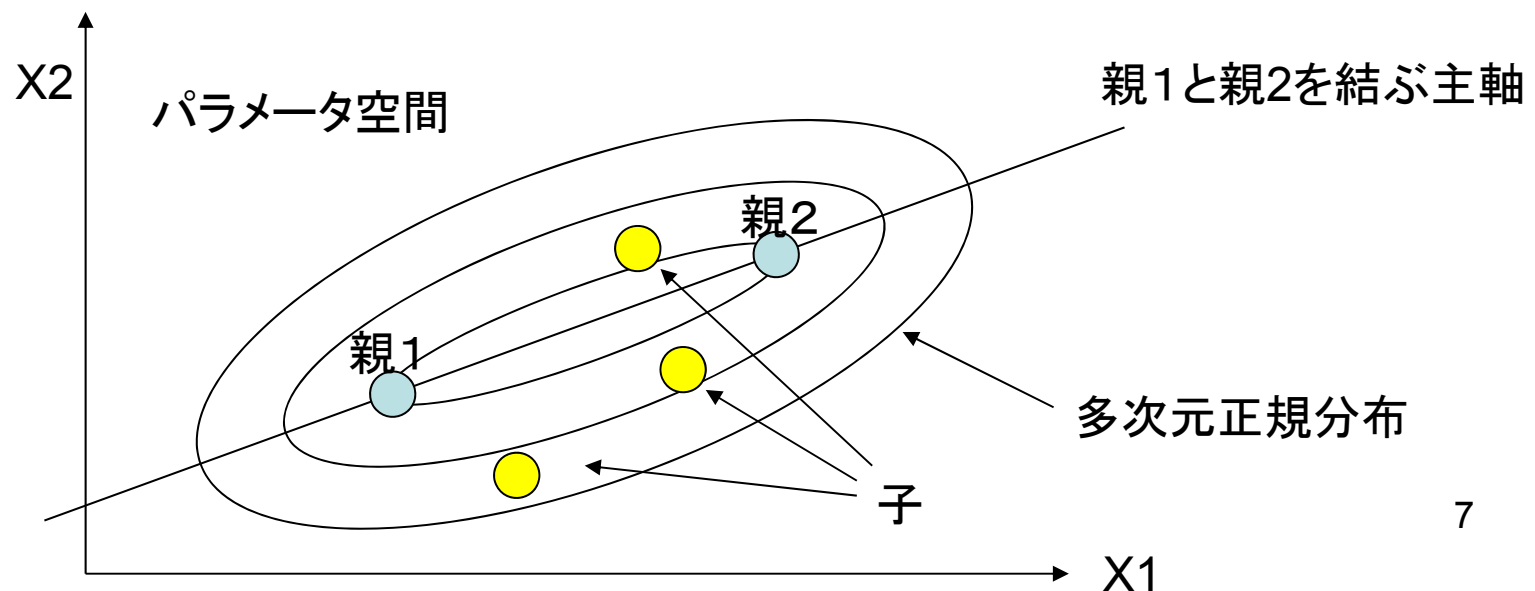
【遺伝的手法による最適化】

実数値 GA (RGA: Real-coded GA) 実用的手法として初めて認知

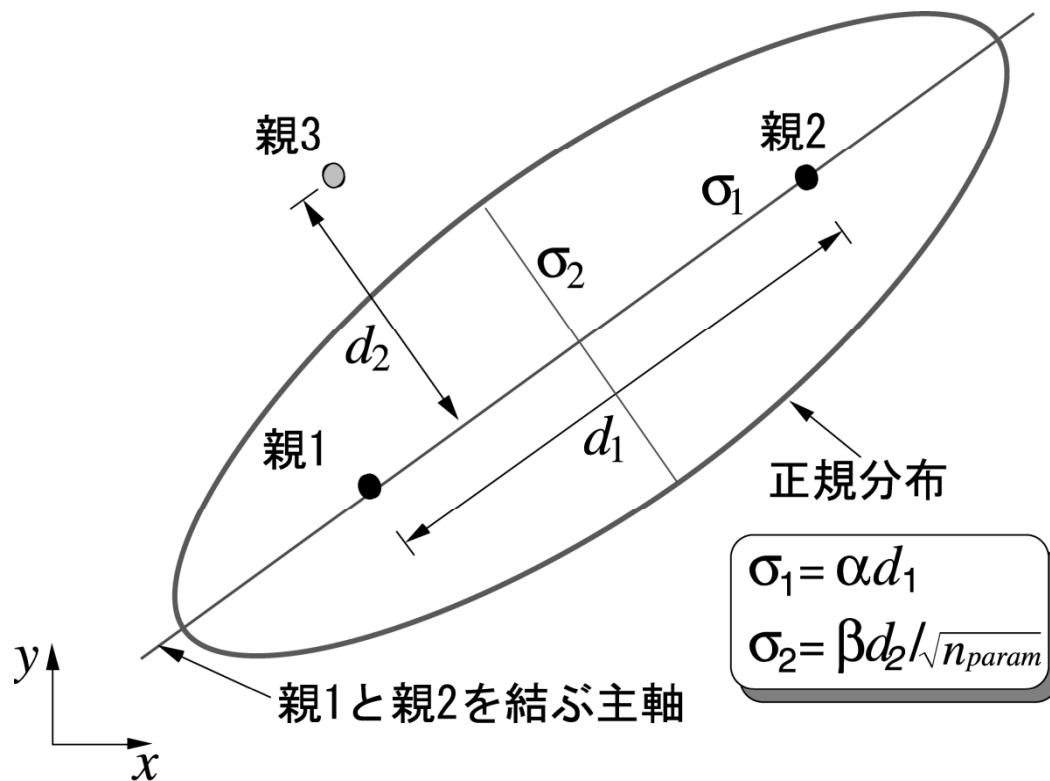
連続関数最適化問題における解候補の「形質」 = パラメータ空間中の座標
解候補の類似性 = パラメータ空間でのユークリッド距離

2つの親個体から、「交叉」オペレーションにより
形質を継承した子個体を生成する → パラメータ空間でユークリッド幾何学的に考える

例) 正規分布交叉: UNDX



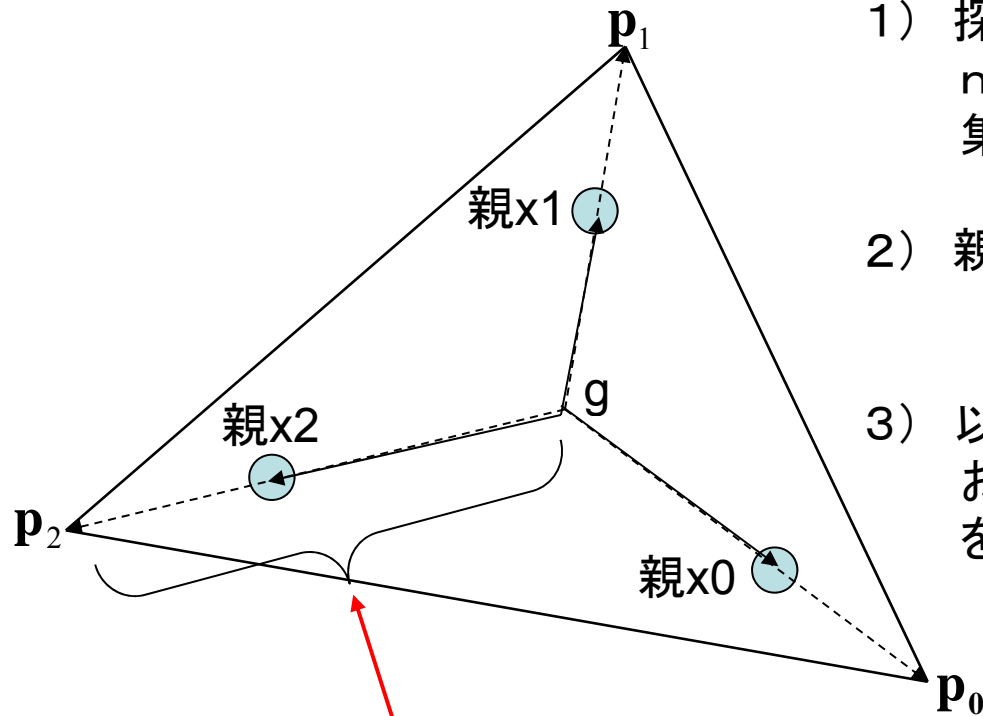
単峰性正規分布交叉UNDX[Ono 97]



設計パラメータの
連続値ベクトル
= 遺伝子

設計パラメータベクトル
空間中において、
親とする個体同士の間
の空間に子個体をラン
ダムに生成

シンプレックス交叉 (SPX) [樋口, 筒井, 山村2001]



重心からそれぞれの親個体までのベクトルを $\sqrt{n+2}$ 倍まで拡張したシンプレックス状多角形の範囲に均一に子個体を生成する

1) 探索するパラメータ空間が n 次元の場合、 $n+1$ 個の親個体 $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ を集団からランダムに選ぶ。

2) 親個体の重心を $\mathbf{g} = \frac{1}{n+1} \sum_{i=0}^n \mathbf{x}_i$ とする

3) 以下の式に従ってベクトル $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n$ およびベクトル $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_n$ を計算する:

$$\mathbf{p}_0 = \mathbf{g} + \sqrt{n+2}(\mathbf{x}_0 - \mathbf{g})$$

$$\mathbf{c}_0 = \mathbf{0}$$

$$\mathbf{p}_k = \mathbf{g} + \sqrt{n+2}(\mathbf{x}_k - \mathbf{g})$$

$$\mathbf{c}_k = r_{k-1}(\mathbf{p}_{k-1} - \mathbf{p}_k + \mathbf{c}_{k-1})$$

ただし r_k は区間 $[0, 1]$ の一様乱数 u から以下の式で変換された乱数である:

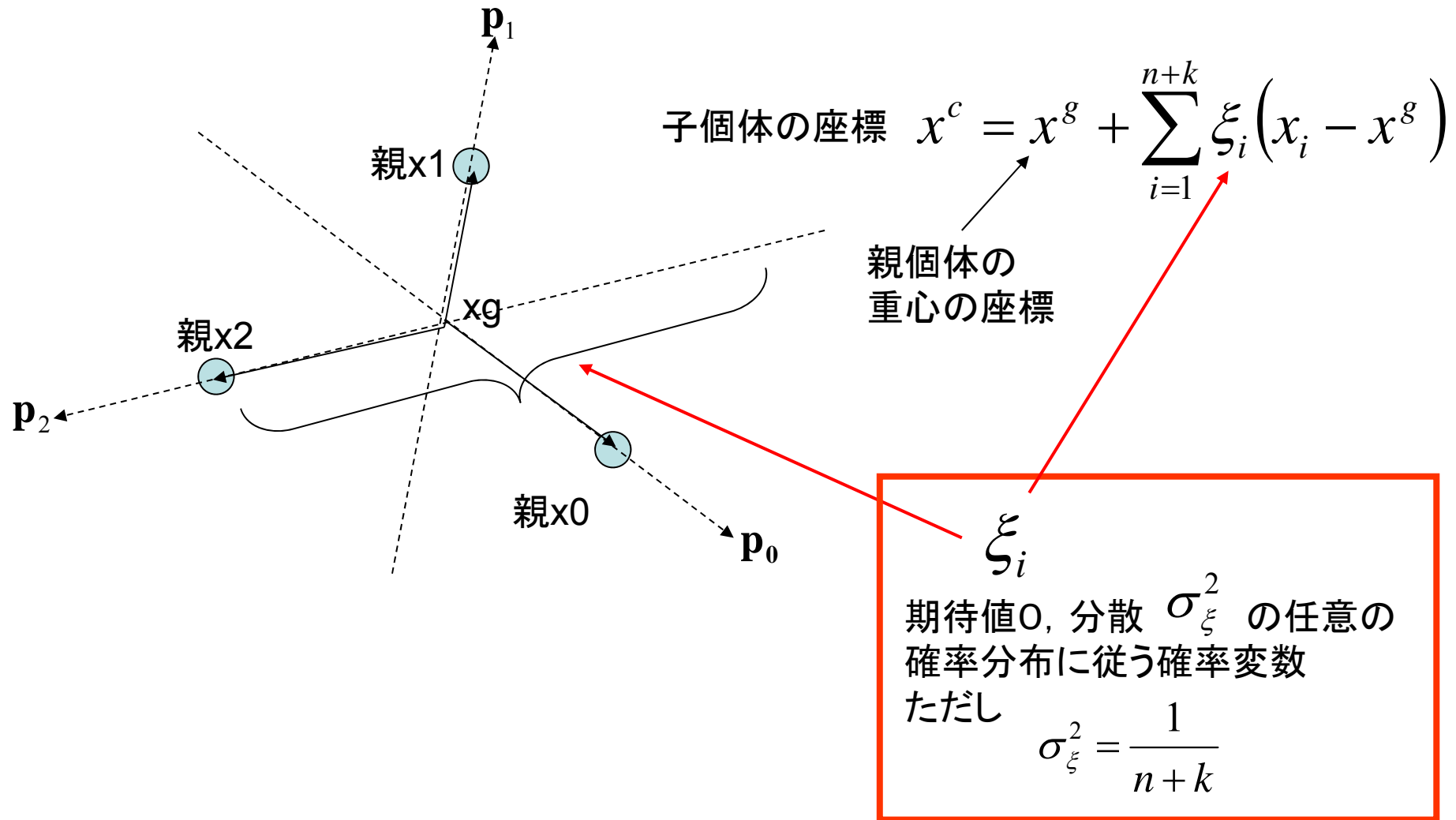
$$r_k = (u)^{\frac{1}{k+1}}$$

4) 子ベクトル \mathbf{c} を次式で得る:

$$\mathbf{c} = \mathbf{p}_n + \mathbf{c}_n$$

多親交叉REX [小林2007]

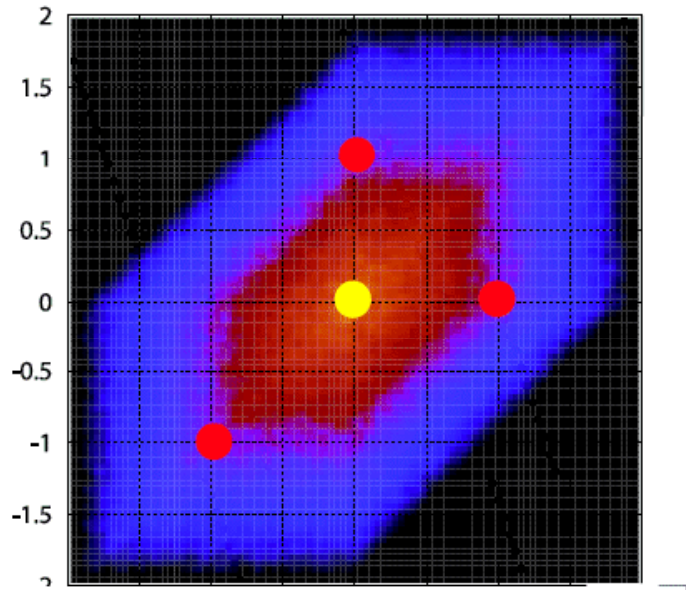
対象とするコスト関数の次元数 n
交叉によって生成する子個体数 $n+k$



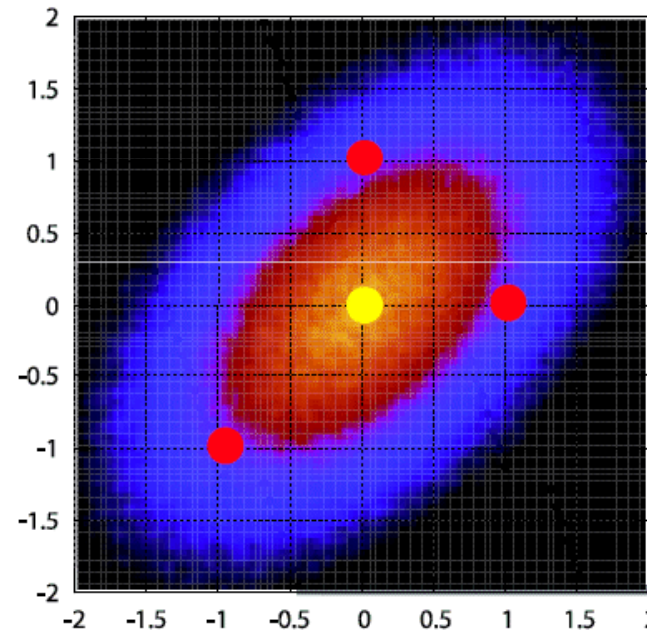
$k=0.5n \sim 2.0n$ を推奨

多親交叉REX [小林2007]

REX($U, n+1$)



REX($N, n+1$)



確率変数 ξ_i が一様分布に従う場合
ただし区間 $[-a, a]$ で一様で、

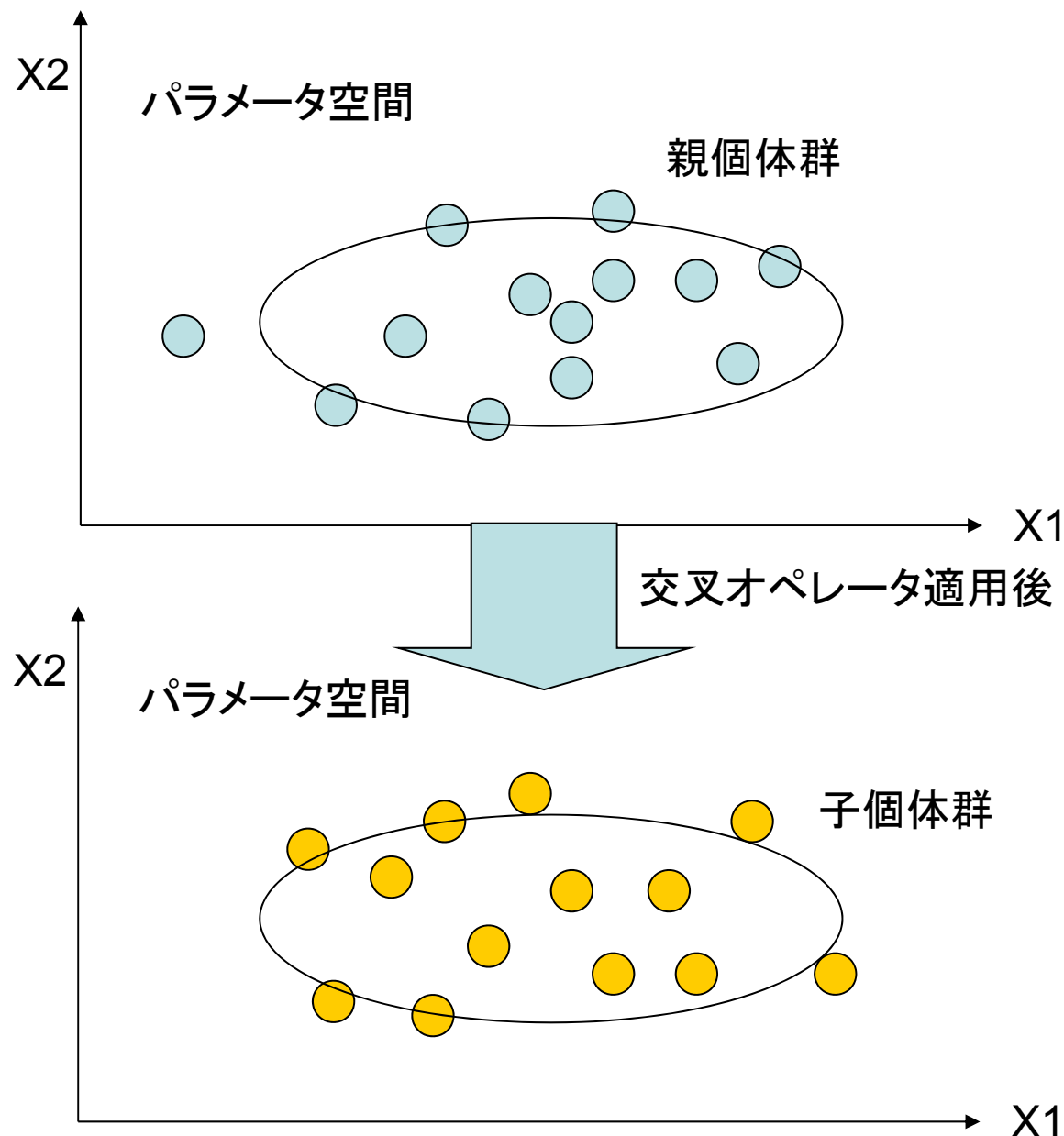
$$a = \sqrt{\frac{3}{n+k}} \text{ とする。}$$

(このとき分散 $\sigma_{\xi}^2 = \frac{1}{n+k}$ となる)

確率変数 ξ_i が正規分布に従う場合
ただし期待値0,

$$\text{分散 } \sigma_{\xi}^2 = \frac{1}{n+k}$$

実数値 GA の交叉オペレータにおいて求められる性質



分布の統計量が
親個体群の分布と
同じになることが望ましい

UNDXもSPXもREXも
この条件を満たしている

機能分担仮説(山村98)

交叉オペレータは
親個体群が張る
部分空間の補間的
探索に徹して、
探索領域の絞込みは
世代交代モデルに
委ねるのが望ましい

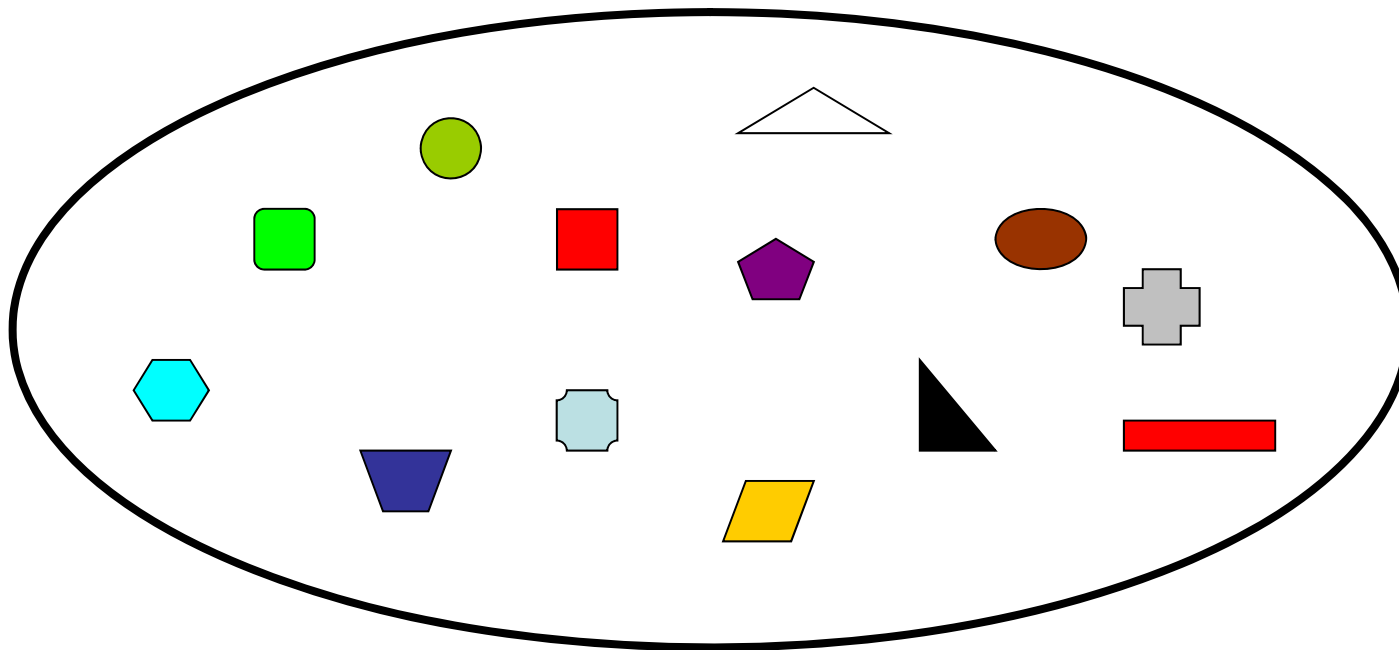
【より洗練された遺伝的手法の枠組み2】

(3) 集団の「多様性」を維持する世代交代方法

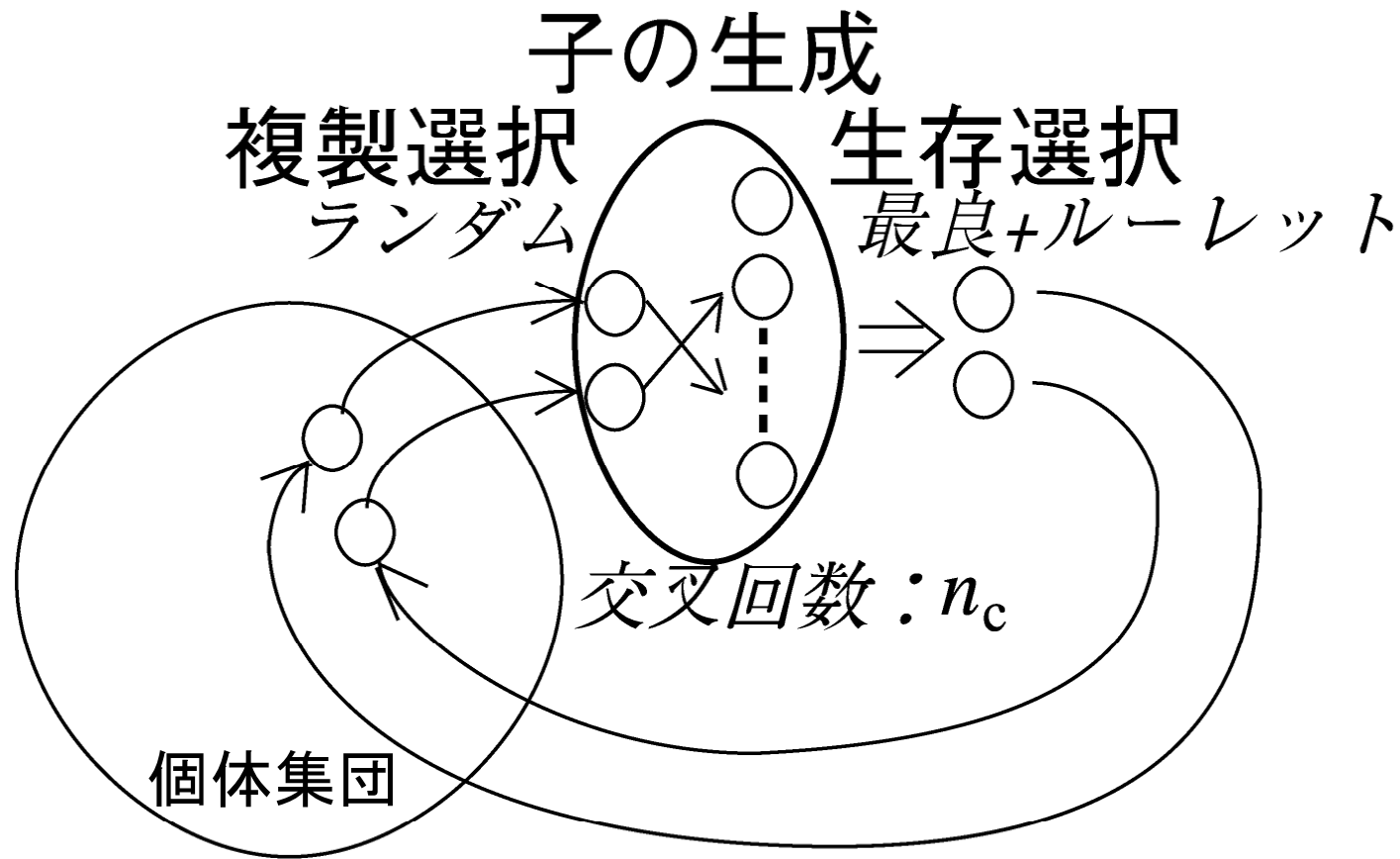
集団の一部しか生き残れないような世代交代では
多様性は維持できない

ビットストリングGAで行う「淘汰」のような操作は不適當

(英語では「選択」も「淘汰」も「selection」と訳される)

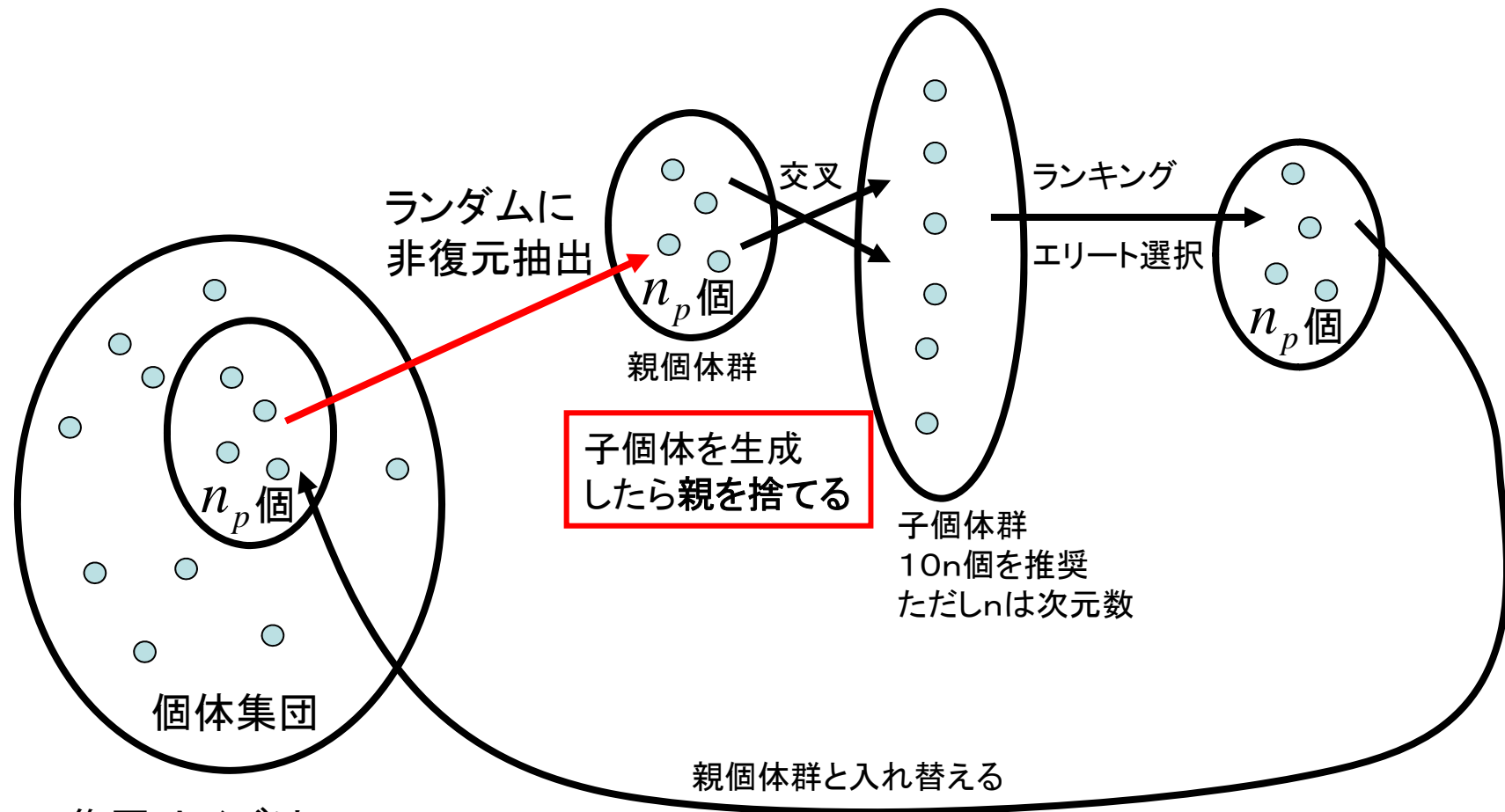


(2親) 世代交代モデルMGG [佐藤1997]



集団からランダムに**2個体**を選択し、何らかの交叉オペレーションによって子個体を n_c 個生成、**両親と生成された全ての子**を合わせた**個体集合**から、1つは最良個体、もう1つは評価値のランクに基づくルーレット選択により選び、両親個体と入れ替える

多親世代交代モデルJGG [小林2007]



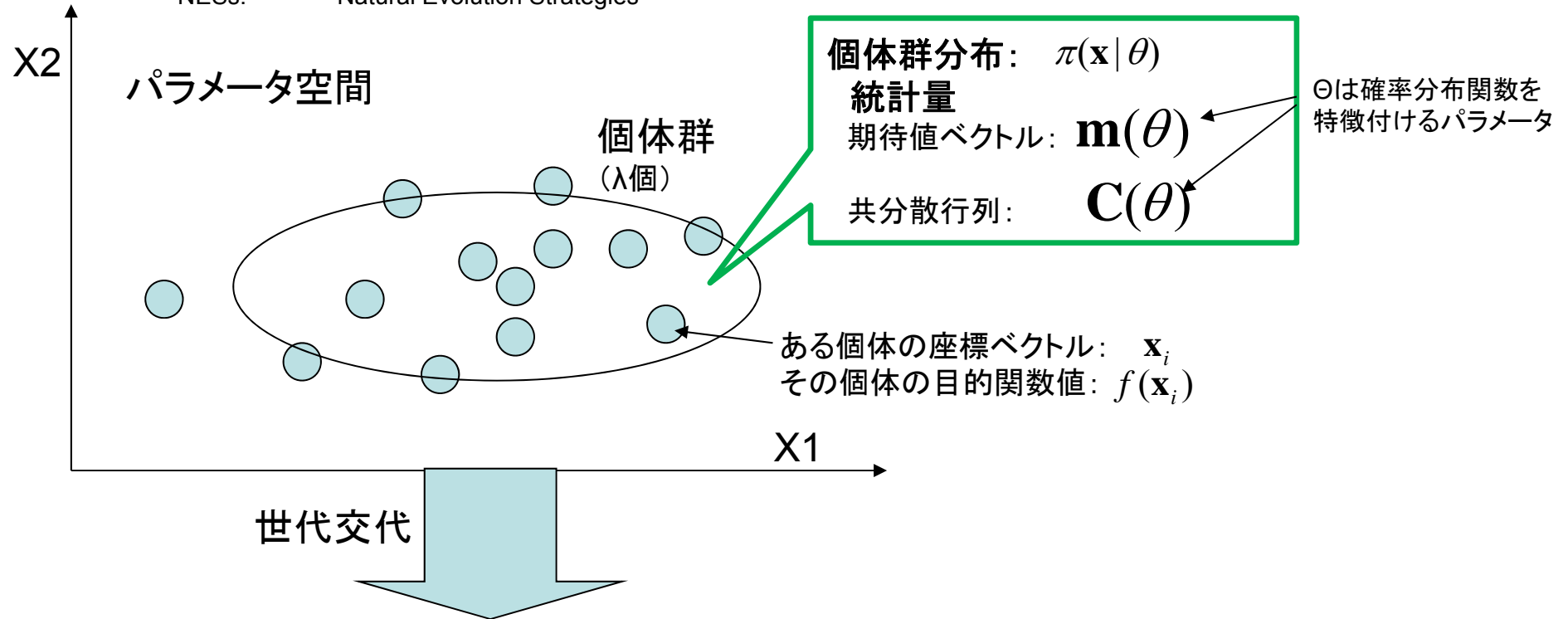
集団サイズは
 $15n \sim 50n$ 個を推奨
ただし n は次元数

MGGと比較すると 2.5~13倍の性能アップ

究極の実数値 GA: CMA-ES + NESSs

CMA-ES の学習率に関する理論的考察
秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).

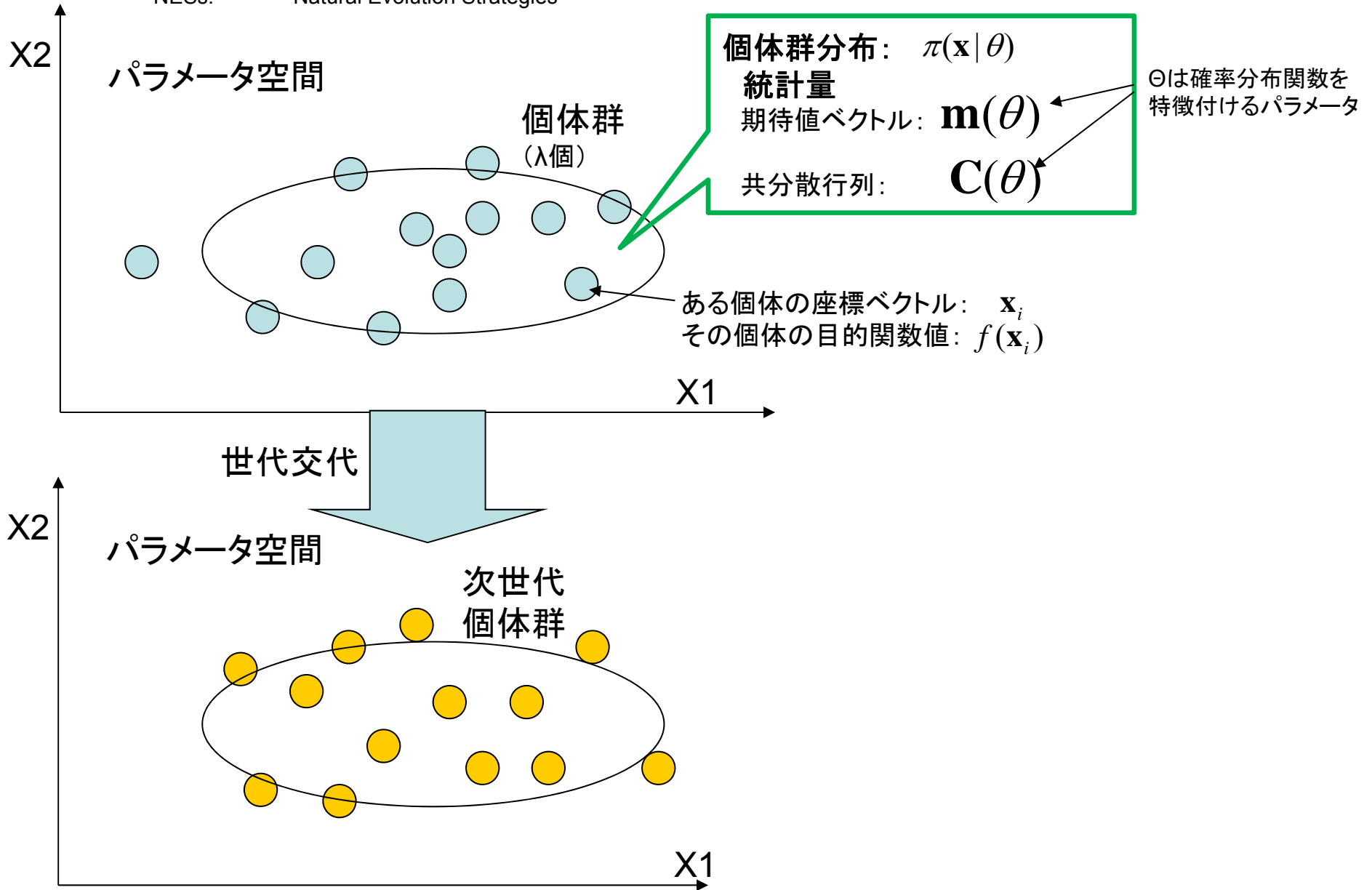
CMA-ES: Covariance Matrix Adaptation Evolution Strategies
NESSs: Natural Evolution Strategies



究極の実数値 GA: CMA-ES + NESSs

CMA-ES の学習率に関する理論的考察
秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).

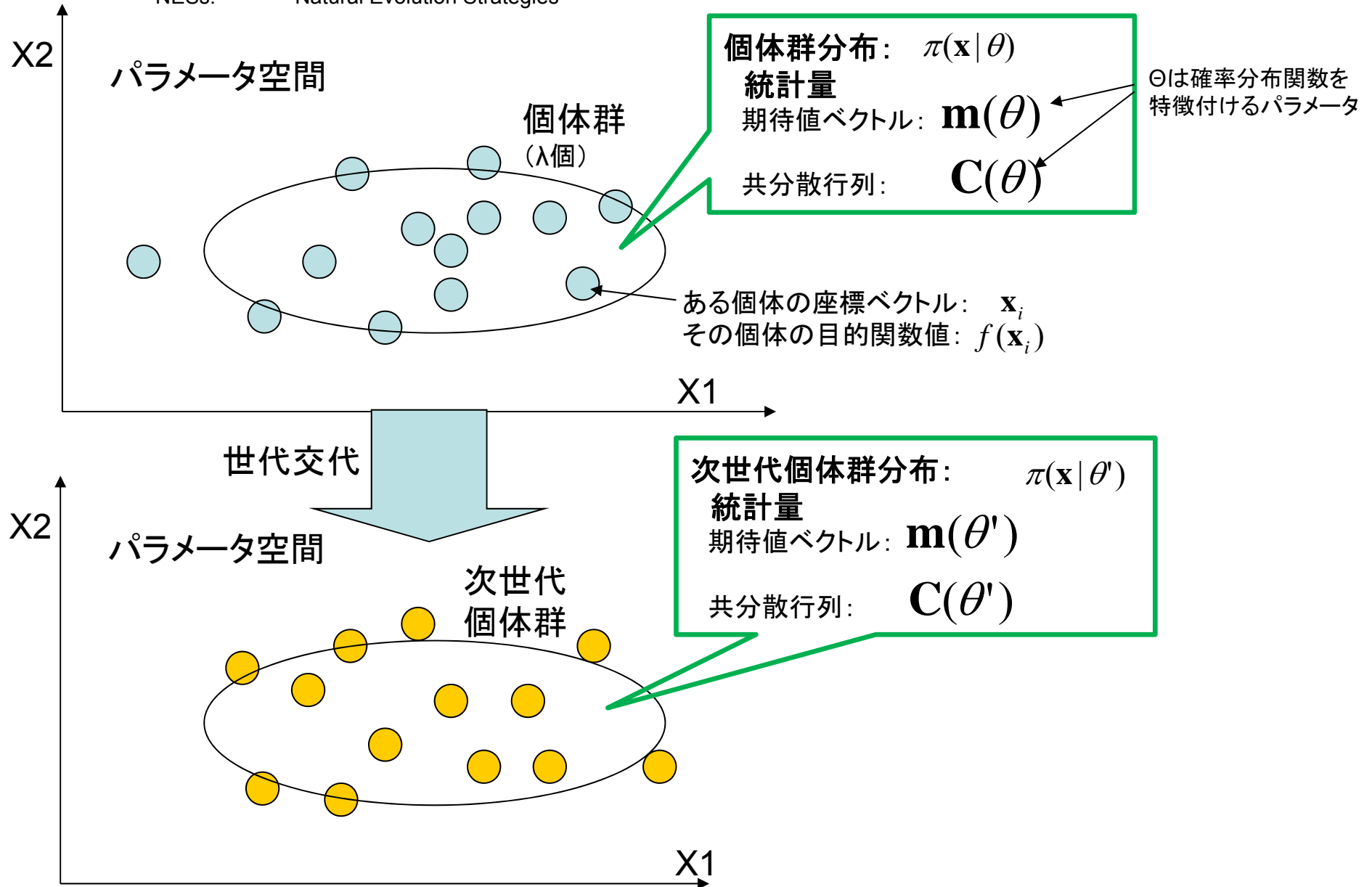
CMA-ES: Covariance Matrix Adaptation Evolution Strategies
NESSs: Natural Evolution Strategies



究極の実数値 GA: CMA-ES + NESSs

CMA-ES の学習率に関する理論的考察
秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).

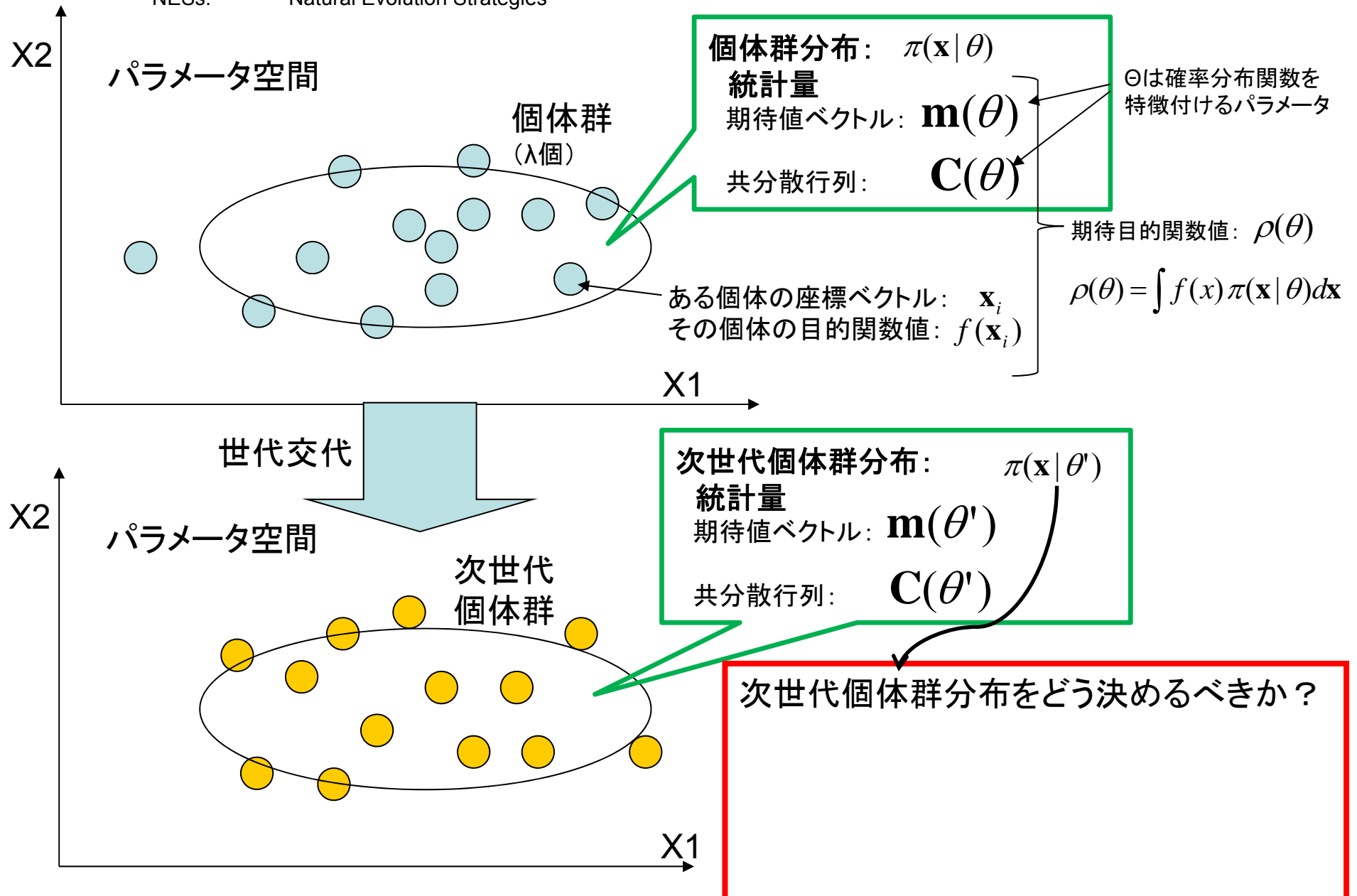
CMA-ES: Covariance Matrix Adaptation Evolution Strategies
NESSs: Natural Evolution Strategies



究極の実数値 GA: CMA-ES + NESSs

CMA-ES の学習率に関する理論的考察
秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).

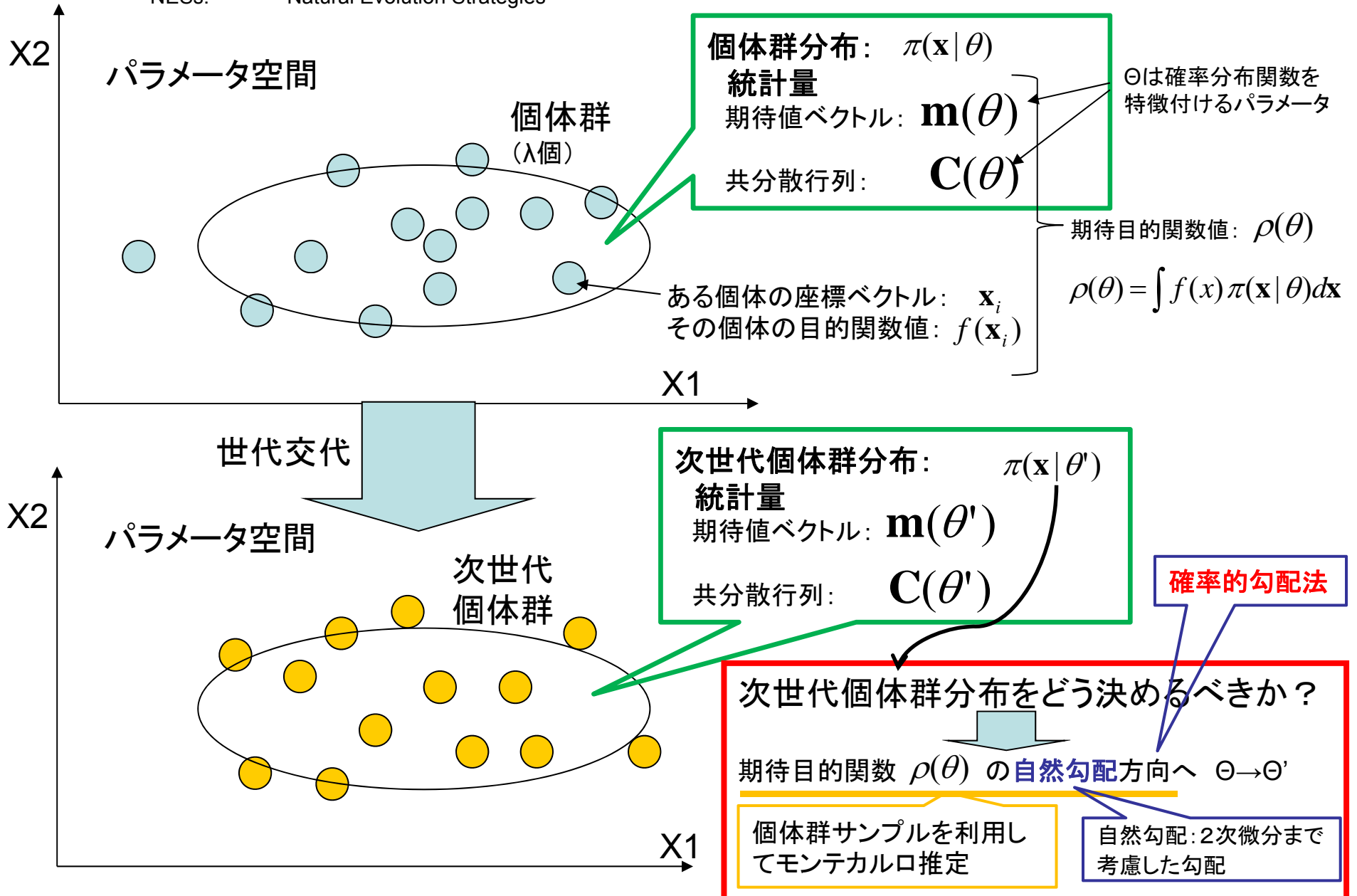
CMA-ES: Covariance Matrix Adaptation Evolution Strategies
NESSs: Natural Evolution Strategies



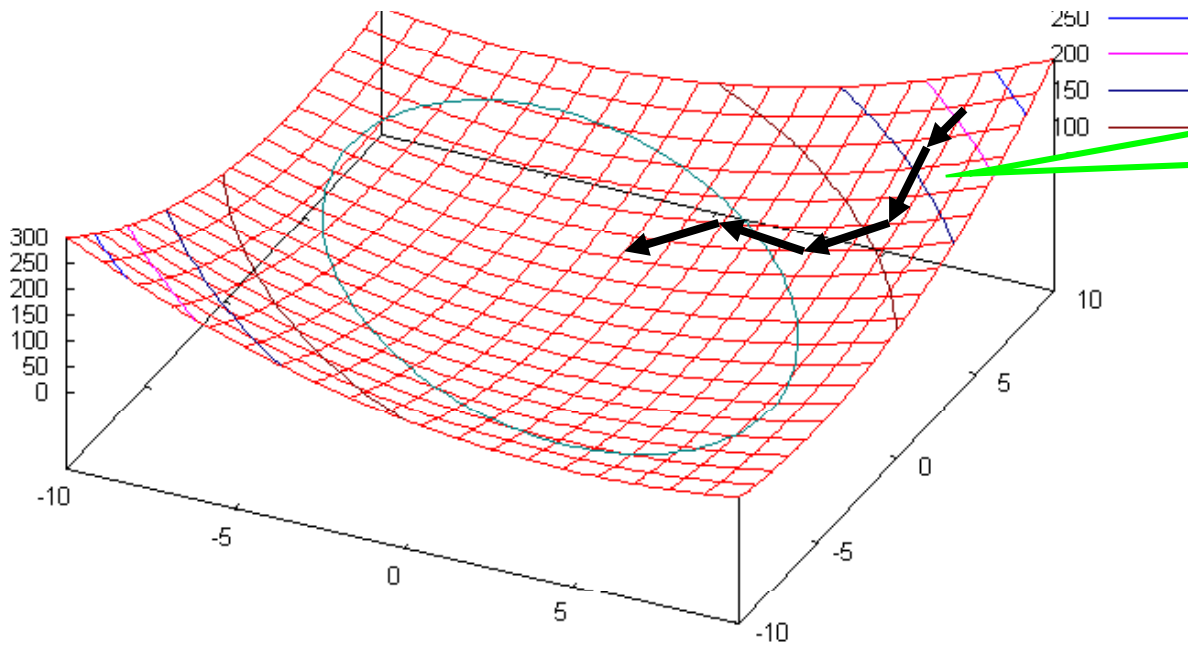
究極の実数値 GA: CMA-ES + NESSs

CMA-ES の学習率に関する理論的考察
 秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
 進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).

CMA-ES: Covariance Matrix Adaptation Evolution Strategies
 NESSs: Natural Evolution Strategies



通常の「勾配法」と CMA-ES + NESs との関係

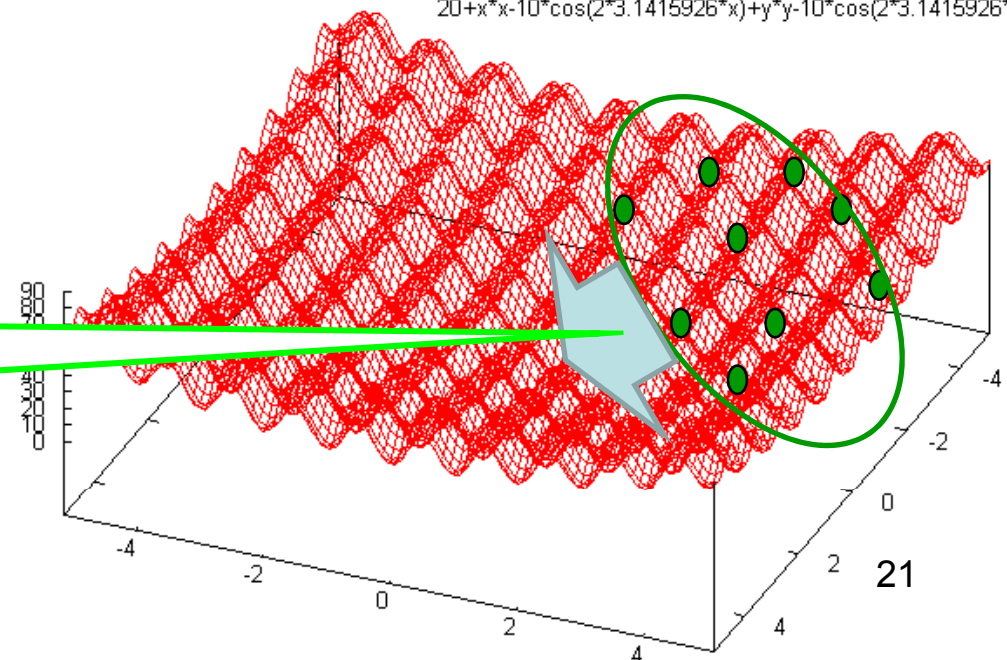


「勾配法」では1つの解候補がコスト関数の勾配を下る

解候補の移動方向の期待値
=コスト関数の(共役)勾配

$$20+x*x-10*\cos(2*3.1415926*x)+y*y-10*\cos(2*3.1415926*y)$$

「CMA-ES+NESs」では
解候補の集団全体がコスト関数の
期待値の自然勾配を下る



究極の実数値 GA: CMA-ES + NESs

期待目的関数値 $\rho(\theta)$ の自然勾配 = (Fisher情報行列の逆行列) $\cdot \nabla_{\theta} \rho(\theta)$

個体群分布 $\pi(\mathbf{x} | \theta)$ が(多次元)正規分布の場合、
モンテカルロ推定が簡単

準ニュートン法で出てくるA行列に類似

以降、本来なら情報幾何学に基づいた複雑な式展開が続くが、本講義の趣旨から逸脱するので省略

しかし、その結論の簡潔さは以下のとおり驚異に値する:



究極の実数値 GA: CMA-ES + NESSs

期待目的関数値 $\rho(\theta)$ の自然勾配 = (Fisher情報行列の逆行列) $\cdot \nabla_{\theta} \rho(\theta)$

個体群分布 $\pi(\mathbf{x} | \theta)$ が(多次元)正規分布の場合、モンテカルロ推定が簡単

準ニュートン法で出てくるA行列に類似

以降、本来なら情報幾何学に基づいた複雑な式展開が続くが、本講義の趣旨から逸脱するので省略

しかし、その結論の簡潔さは以下のとおり驚異に値する:

多次元(あるいは多変量)正規分布 $\pi(\mathbf{x} | \theta)$ ただし期待値座標 $\mathbf{m}(\theta)$ 共分散行列 $\mathbf{C}(\theta)$ に従って個体群 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\lambda}$ を生成し、各評価値 $f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_{\lambda})$ を計算

現世代

このとき多次元(あるいは多変量)正規分布の自然勾配方向へと次世代個体群分布 $\pi(\mathbf{x} | \theta')$ を生成:

$$\left\{ \begin{array}{l} \text{期待値座標} \\ \text{ベクトル:} \\ \text{共分散行列:} \end{array} \right. \quad \begin{array}{l} \mathbf{m}' = \mathbf{m} + \eta \sum_{i=1}^{\lambda} \frac{f(\mathbf{x}_i)}{\lambda} (\mathbf{x}_i - \mathbf{m}) \\ \mathbf{C}' = \mathbf{C} + \eta \sum_{i=1}^{\lambda} \frac{f(\mathbf{x}_i)}{\lambda} ((\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T - \mathbf{C}) \end{array}$$

ただし η は学習率パラメータで $0 < \eta \leq 1$ このアルゴリズムは $f(\mathbf{x})$ が大きくなる方向へ更新

$\nabla_{\theta} \rho(\theta)$ のモンテカルロ推定

$$N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

正規分布

準備

$$\begin{cases} \frac{\partial}{\partial \mu} \ln N(\mu, \sigma^2) = \frac{x-\mu}{\sigma^2} \\ \frac{\partial}{\partial \sigma} \ln N(\mu, \sigma^2) = \frac{(x-\mu)^2 - \sigma^2}{\sigma^3} \end{cases}$$

なぜ正規分布にlnを施したものを微分するのか？理由は下の式

$$\frac{\partial}{\partial \mu} \rho(\theta) = \frac{\partial}{\partial \mu} \int f(x) N(\mu, \sigma^2) dx = \int f(x) \frac{\partial}{\partial \mu} N(\mu, \sigma^2) dx$$

$$= \int f(x) N(\mu, \sigma^2) \left(\frac{\partial}{\partial \mu} \ln N(\mu, \sigma^2) \right) dx = \int f(x) N(\mu, \sigma^2) \left(\frac{x-\mu}{\sigma^2} \right) dx$$

$$= E \left\{ f(x) \frac{x-\mu}{\sigma^2} \right\}$$

つまり $f(x) \frac{x-\mu}{\sigma^2}$ をたくさんサンプルして平均をとれば、それが $\frac{\partial}{\partial \mu} \rho(\theta)$

【復習】 多変量正規分布

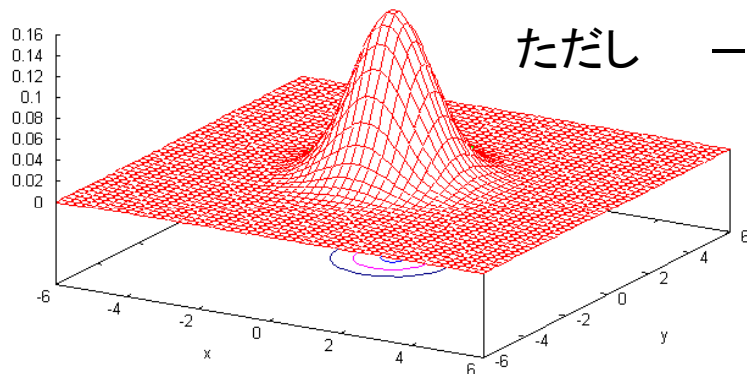
確率密度 $\pi(\mathbf{x} | \theta) = \pi(\mathbf{x} | \mathbf{m}, \mathbf{C}) = N(\mathbf{m}, \mathbf{C})$

期待値ベクトル
↓
分散共分散行列

$$= \left(\frac{1}{\sqrt{2\pi}} \right)^n \left(\frac{1}{\sqrt{\det(\mathbf{C})}} \right) \exp \left(-\frac{1}{2} (\mathbf{x} - \mathbf{m})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{m}) \right)$$

2変数正規分布

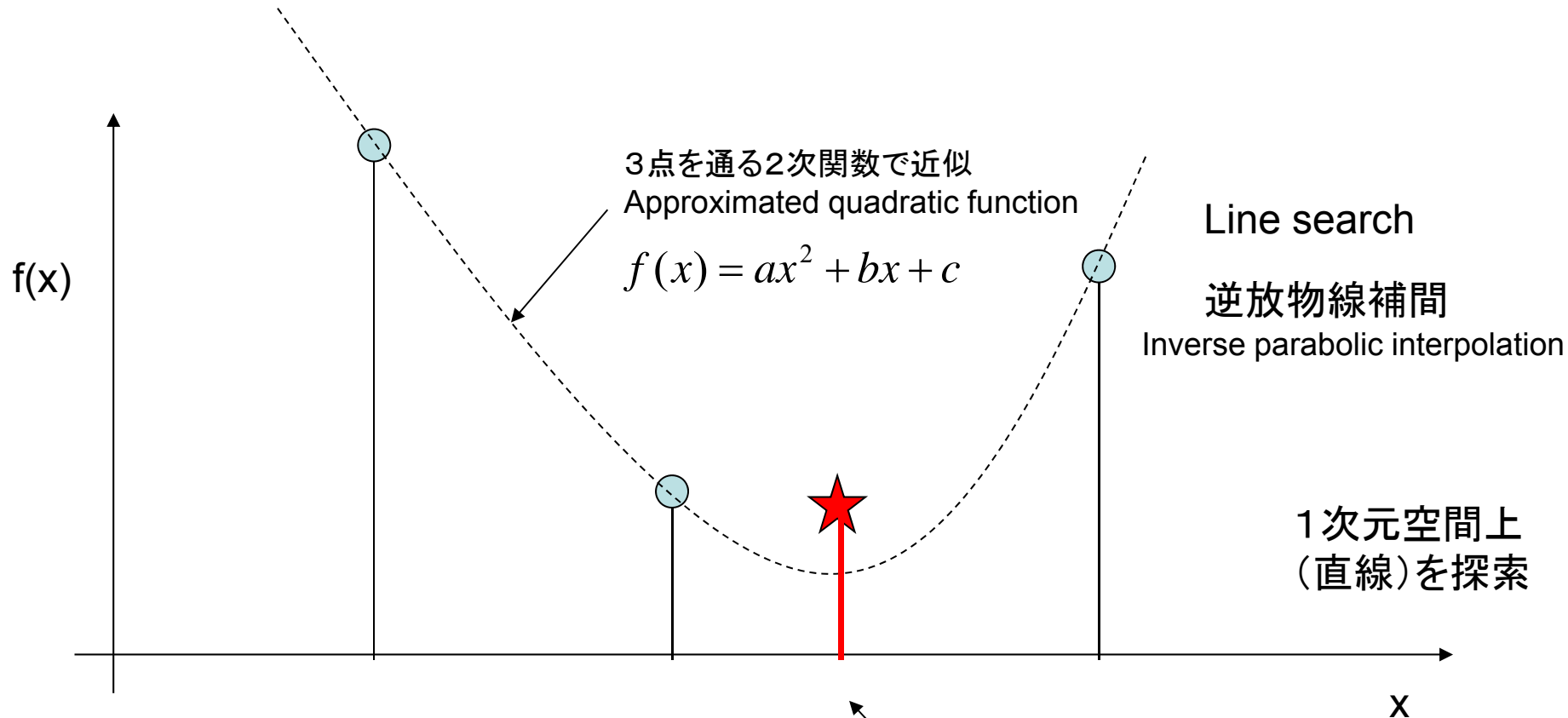
$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left(-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\} \right)$$



ただし $-1 < \rho < 1$

- | | | |
|---|----------------------|----------------|
| } | μ_1, μ_2 | x, y の期待値 |
| | σ_1, σ_2 | x, y の分散 |
| | ρ | x と y との相関 |

【復習】 ニュートン法 Newton method



$$\frac{df}{dx} = 0 \quad \text{より、} 2ax + b = 0$$

$$x = (2a)^{-1}(-b)$$

近似した2次関数の極小をサーチ
Search this point where gradient of the approximated function is zero.

【復習】 準ニュートン法

Quasi-Newton method

関数 $f(\mathbf{x})$ をテイラー級数で近似
(Taylor series)

$$f(\mathbf{x}) = f(\mathbf{P}) + \sum_i \frac{\partial f}{\partial x_i} x_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} x_i x_j + \dots$$

この行列 \mathbf{A} は $f(\mathbf{x})$ の Hesse 行列
(Hessian matrix)

$$= f(\mathbf{P}) + \nabla f|_{\mathbf{P}} \cdot \mathbf{x} + \frac{1}{2} \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} + \dots$$

各要素は \mathbf{P} における 2 階偏導関数

$$[\mathbf{A}]_{ij} \equiv \frac{\partial^2}{\partial x_i \partial x_j} \Big|_{\mathbf{P}}$$

2次形式
(quadratic)

極小点では勾配ゼロだから上の式を微分すると

$$\nabla f(\mathbf{x}) = \nabla f|_{\mathbf{P}} + \mathbf{A} \cdot \mathbf{x} = \mathbf{0}$$

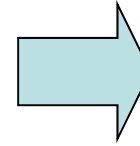
これを解くと

$$\mathbf{x} = \mathbf{A}^{-1} (-\nabla f|_{\mathbf{P}}) \quad \text{次はこの点を探査すべし}$$

しかし、行列 \mathbf{A} は未知 なので、探索の過程から逐次的に近似する
Since matrix \mathbf{A} is unknown, it is estimated from the process of the optimization incrementally.

数値最適化を行う遺伝的アルゴリズム:まとめ

交叉方法: **正規分布交叉(UNDX)**
世代交代: **MGG**



レンズ系設計の最適化
プロペラ形状の最適化
船形最適化など多数の実績

しかし今後はアルゴリズムがずっと簡単かつ強力な

交叉方法: **多親交叉REX (一様分布使用)**
世代交代: **多親世代交代モデルJGG**

OR

最も洗練され、かつ最強の **CMA-ES + NESs**

を使用すべし

遺伝的アルゴリズム(GA)の研究課題

●いかにして致死遺伝子が出にくいようにパラメータを
遺伝子へコード化するか？

→ 問題毎に、ある程度の専門知識を要する

修正オペレータ: 致死遺伝子を実行可能解へ修正

●どのように子個体を生成するか？

連続関数の最適化問題 → ユークリッド空間中？

組合せ最適化問題 → ？

●どのように世代交代すればよいか？

進化遅い: 無駄な計算量大 ← → 多様な解候補

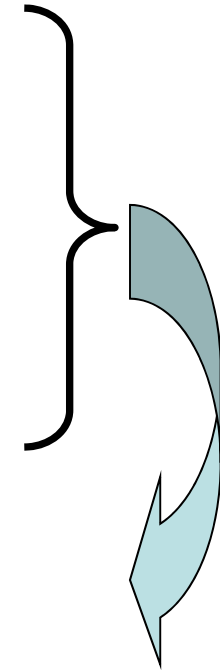
進化早い: 解候補のバリエーション減少、探索の停滞

遺伝的アルゴリズム(GA)の適用における注意点

- 数値最適化問題において適用を試みる前に、勾配法やシンプレックス法で解けないか？
- 組合せ最適化問題に適用を試みる前に、動的計画法(DP)で解ける問題ではないか？有効なヒューリスティクスが存在しないか？

単純な問題においては遺伝的手法は**非効率**

- 数値最適化と組合せ最適化の**複合問題**
→ 遺伝的アルゴリズムは有望な選択肢
- **多目的最適化問題**
→ 遺伝的アルゴリズムは有効



参考文献

- 1) 長尾 智晴: 最適化アルゴリズム, 昭晃堂(2000).
- 2) W.H.Press, B.P.Flannery, S.A.Teukolsky and W.T.Vetterling:
ニューメリカルレシピ・イン・シー C言語による数値計算のレシピ, 技術評論社(1993).
- 3) 小野 功, 佐藤 浩, 小林 重信:
単峰性正規分布交叉UNDXを用いた実数値GAによる関数最適化,
人工知能学会誌 Vol.14, No.6, pp.1146—1155 (1999).
- 4) 樋口隆英, 筒井茂義, 山村雅幸: 実数値GAにおけるシンプレックス交叉の提案,
人工知能学会誌, Vol.16, No.1, pp146-155, (2001).
- 5) 小林重信:
実数値GAのフロンティア,
人工知能学会誌Vol.24, No.1, pp.147--162 (2009)
- 6) 秋本洋平, 永田裕一, 小野功, 小林重信(東京工業大学)
CMA-ES の学習率に関する理論的考察,
進化計算シンポジウム2010 講演論文集 pp.202--207 (2010).