大偏差原理を用いた遺伝的アルゴリズムの Markov 連鎖的解析

鈴木讓

大阪大学

2009 年 10 月 2 日

あらまし

1. 遺伝的アルゴリズム (GA)
2. GA の定常確率
3. 大偏差原理からの結果
4. 主結果

遺伝的アルゴリズム (GA)

記号	GA
$L\ (\geq 1)$	個体長
$I := \{0,1\}^L$	個体の集合
$f : I \to (0,\infty)$	適合度関数
$M (\geq 2)$	集団サイズ
$S := I^M$	集団の集合

遺伝的アルゴリズム (GA)

$f(i)$ を最大にする $i \in I$ を求める

突然変異

$\rho(i,j),\ i,j \in I$

2 個体 i,j で (L ビット中) 何ビット違うか

$d(x,y),\ x,y \in S$

2 集団 x,y で (LM ビット中) 何ビット違うか

パラメータ $0 < \mu < 1$ の突然変異 M_μ

LM ビットを確率 μ で反転

$x,y \in S$ として、$x \xrightarrow{M_\mu} y$ の推移確率:

$$\mathcal{M}_\mu[x,y] := \mu^{d(x,y)}(1-\mu)^{LM-d(x,y)} > 0$$

交叉

交叉 C

$(i_1, \cdots, i_L), (i'_1, \cdots, i'_L) \in I$, $\pi \subseteq \{1, \cdots, L\}$: ランダム
i_k, i'_k, $k \in \pi$ を交換
これを有限回繰り返す
$x, y \in S$ として、$x \xrightarrow{C} y$ の推移確率: $\mathcal{C}[x, y]$

(例) $\pi = \{1, \cdots, k\}(1 \leq k \leq L) \Longrightarrow$ 一点交叉

交叉の性質 (仮定)

集団内の $k \in \{1, \cdots, L\}$ の位置の "1" の数の合計は変わらない

選択

$x = (x_1, \cdots, x_M), y = (y_1, \cdots, y_M) \in S$
$x_k, y_k \in I,\ k = 1, \cdots, M$

$x \succ y$

$\{x_1, \cdots, x_M\} \supseteq \{y_1, \cdots, y_M\}$

パラメータ $\beta > 0$ の選択 S_β

y_k を確率 $g_\beta(y_k) := f(y_k)^\beta / \sum_{k=1}^{M} f(x_k)^\beta$ で独立に選択 (β: 選択圧力)

$x, y \in S$ として、$x \xrightarrow{S_\beta} y$ の推移確率:

$$\mathcal{S}_\beta[x, y] := \begin{cases} \prod_{k=1}^{M} g_\beta(y_k), & x \succ y \\ 0, & x \not\succ y \end{cases}$$

パラメータ μ, β の GA

パラメータ μ, β の GA の世代交代

$x, u, v, y \in S$ として、$x \xrightarrow{M_\mu} u \xrightarrow{C} v \xrightarrow{S_\beta} y$ の推移確率:

$$Q_{\mu,\beta}[x, y] := \sum_{u,v \in S} \mathcal{M}_\mu[x, u] \mathcal{C}[u, v] \mathcal{S}_\beta[v, y]$$

世代交代を繰り返して、$f(i)$ を最大にする $i \in I$ を得る

GA の定常確率

$Q_{\mu,\beta} = (Q_{\mu,\beta}[x,y])_{x,y\in S}$: 有限 Markov 連鎖 (X_n) の推移行列

- S の要素が有限 (有限 Markov 連鎖)
- $0 < \mu < 1 \Longrightarrow Q_{\mu,\beta}[x,y] > 0 \Longrightarrow \exists k$ s.t. $Q^k_{\mu,\beta}[x,y] > 0$ (既約)
- 各 $x \in S$ で、周期 $gcd\{n | P(X_n = x | X_0 = x) > 0\}$ が 1 (非周期的)

エルゴードな有限 Markov 連鎖

既約で非周期的な有限 Markov 連鎖 $\Longrightarrow$ 定常確率が一意

GA の Markov 連鎖 $(0 < \mu < 1) \Longrightarrow$ 定常確率 $(q_{\mu,\beta}(x))_{x\in S}$ が一意

$$\begin{bmatrix} q_{\mu,\beta}(1) & \cdots & q_{\mu,\beta}(\sigma) \end{bmatrix}$$

$$= \begin{bmatrix} q_{\mu,\beta}(1) & , \cdots , & q_{\mu,\beta}(\sigma) \end{bmatrix} \begin{bmatrix} Q_{\mu,\beta}[1,1] & \cdots & Q_{\mu,\beta}[1,\sigma] \\ \vdots & \vdots & \vdots \\ Q_{\mu,\beta}[\sigma,1] & \cdots & Q_{\mu,\beta}[\sigma,\sigma] \end{bmatrix}$$

$(S = \{1, \cdots, \sigma\})$

問題

定常確率 $(q_{\mu,\beta}(x))_{x\in S}$ は、どのような性質をもつか。

- 選択圧力 β 大
- 突然変異確率 μ 小

重要性

GA が優れた性質をもつことを、数学的に証明できるのか。

仮定

μ, β はアルゴリズムの途中で変えない

従来の結果

$$U := \{x = (x_1, \cdots, x_M) \in S | x_1 = \cdots = x_M\}$$
$$U^* := \{x = (x_1, \cdots, x_M) \in U | f(x_k) = \max_{i \in I} f(i), k = 1, \cdots, M\}$$

Davis-Princepe, 1991

交叉を行わない場合、任意の $\beta > 0$ について、

$$\lim_{\mu \to 0} q_{\mu,\beta}(x) > 0 \Longrightarrow x \in U$$

Suzuki, 1998

交叉を行わない場合、

$$\lim_{\beta \to \infty} \lim_{\mu \to 0} q_{\mu,\beta}(x) > 0 \Longrightarrow x \in U^*$$

(交叉を含めると数学的な解析が難しくなる)

エルゴードな有限 Markov 連鎖の解析

$(Q_\alpha[x,y])_{x,y\in S}$: パラメータ $\alpha > 0$ の推移確率
$(q_\alpha(x))_{x\in S}$: 推移行列 Q_α の定常確率
$q_\infty(x) := \lim_{\alpha\to\infty} q_\alpha(x)$, $x \in S$

$x \in S$ から $y \in S$ へのコミュニケーションコスト

$$V(x \to y) := \lim_{\alpha\to\infty} -\frac{1}{\alpha}\log Q_\alpha[x,y] \in [0,\infty]$$

$S = \{1,\cdots,\sigma\}$ を頂点とする (有向非巡回グラフ)
$G(y)$: $y \in S$ を根にもつ極大木 (有向辺の集合で表現) の集合

$$V(\gamma) := \sum_{(x\to y)\in\gamma} V(x \to y)\ ,\ \gamma \in G(x)$$

$$W(x) := \min_{\gamma\in G(x)} V(\gamma)$$

$$W_{min} := \min_{x\in S} W(x)$$

Freidlin-Wentzell, 1984

$$Q_\alpha(x) := \sum_{\gamma \in G(x)} \prod_{(u,z)\in\gamma} Q_\alpha[u,z]$$

Freidlin-Wentzell, 1984

1. 推移行列 $(Q_\alpha[x,y])_{x,y\in S}$ の定常分布が

$$q_\alpha(x) = \frac{Q_\alpha(x)}{\sum_{y\in S} Q_\alpha(y)}, \quad x \in S$$

2. $\displaystyle \lim_{\alpha\to\infty} -\frac{1}{\alpha} \log q_\alpha(x) = W(x) - W_{min}$

各 $x \in S$ について、

$$q_\infty(x) > 0 \Longrightarrow W(x) = W_{min}$$

Cerf, 1998

Cert, 1998

- $S^* \subseteq U$ を証明
- $\alpha = \alpha(\mu, \beta)$ を $\alpha \to \infty \iff \mu \to 0, \beta \to \infty$ として、
  $S^* \subseteq U^*$ となる L, M, f, μ, β の条件を求める

十分条件の導出

$W(x), x \in S \setminus U^*$ の下界 $>$ $W(x), x \in U^*$ の上界 $\Longrightarrow S^* \subseteq U^*$

Albuquerque-Mazza, 2001

$$\Delta := \max_{i,j \in I} |\log f(i) - \log f(j)|$$

仮定

$\mu(\beta) = \epsilon \exp(-\lambda\beta)$ なる $0 < \epsilon < 1, \lambda > 0$ が存在
$f : I \to (0, \infty)$ が 1 対 1
交叉がない

Albuquerque-Mazza, 2001

1. $S^* \subseteq U$
2. $\lambda > M\Delta \Longrightarrow U^* \subseteq S^*$

本論文で証明したこと

$$\Delta := \max_{i,j \in I} |\log f(i) - \log f(j)|$$

仮定

$\mu(\beta) = \epsilon \exp(-\lambda\beta)$ なる $0 < \epsilon < 1, \lambda > 0$ が存在
$f : I \to (0, \infty)$ が 1 対 1

定理 1

$$\lambda > \frac{M}{M-1}\Delta \Longrightarrow U^* \subseteq S^*$$

- $\lambda > 2\Delta$ とおけば、任意の $M \geq 2$ で成立
- 交叉を行う通常の GA

交叉: もう 1 度

交叉の性質 (仮定)

集団内の $k \in \{1, \cdots, L\}$ の位置の "1" の数の合計は変わらない
(一様集団の場合、"1" の数の合計は、M 個か 0 個になる)

補題 3

1. 一様集団は、交叉によって非一様集団には移りえない
2. 非一様集団は、交叉によって一様集団には移りえない

$$x \sim y \overset{def}{\Longleftrightarrow} C[x, y] > 0$$

$V(x \to y)$

$x = (x_1, \cdots, x_M) \in S$
$V(x) := \sum_{k=1}^{M} - \log f(x_k)$

$$
\begin{aligned}
V(x \to y) &= \lim_{\beta \to \infty} -\frac{1}{\beta} \log Q_\beta[x, y] \\
&= \min_{u \sim v, v \succ y} \{\lambda d(x, u) + V(y) - \min_{v \succ r} V(r)\}
\end{aligned}
$$

$V(x \to y)$ (続)

証明:

$$\mathcal{M}_\beta[x,y] := \mu(\beta)^{d(x,y)}(1-\mu(\beta))^{LM-d(x,y)} > 0$$

$$g_\beta(y_k) := \frac{f(y_k)^\beta}{\sum_{k=1}^{M} f(x_k)^\beta}$$

$$\mathcal{C}[x,y] > 0$$

$$\mathcal{S}_\beta[x,y] := \begin{cases} \prod_{k=1}^{M} g_\beta(y_k), & x \succ y \\ 0, & x \not\succ y \end{cases}$$

$V(x \to y)$ (続)

$\lim_{\beta\to\infty} \frac{o(\beta)}{\beta} = 0$ とおくと、

$$Q_\beta[x,y] := \sum_{u,v\in S} \mathcal{M}_\beta[x,u]\mathcal{C}[u,v]\mathcal{S}_\beta[v,y]$$

$$\begin{aligned}
-\frac{1}{\beta}\log Q_\beta[x,y] &:= -\frac{1}{\beta}\log\{\sum_{u,v\in S} \mathcal{M}_\beta[x,u]\mathcal{C}[u,v]\mathcal{S}_\beta[v,y]\} \\
&= -\frac{1}{\beta}\log \sum_{u\sim v, v\succ y} \exp\{\delta(u,v)\beta + o(\beta)\} \\
&\to \max_{u\sim v, v\succ y} \delta(u,v) \\
&= \min_{u\sim v, v\succ y} \{\lambda d(x,u) + V(y) - \min_{v\succ r} V(r)\}
\end{aligned}$$

$S^* \subseteq U$ の証明

$S_- \subseteq S$
$S_+ := S \backslash S_-$

補題 2

1. 各 $x \in S_+$ について、$V(x \to y) = 0$ なる $y \in S_-$ が存在
2. 各 $x \in S_+, y \in S_-$ について、$V(y \to x) > 0$

$\Longrightarrow q_\infty(x) > 0 \Longrightarrow x \in S_-$

$S_- := U$
$S_+ := S \backslash U$

$$V(x \to y) = \min_{u \sim v, v \succ y} \{\lambda d(x, u) + V(y) - \min_{v \succ r} V(r)\}$$

を補題 2 に適用して、$q_\infty(x) > 0 \Longrightarrow x \in U$

$S^* \subseteq U^*$ の証明

$\varphi : S \to [0, \infty)$ がポテンシャル

$\overset{def}{\iff} V(x \to y) \in [0, \infty)$ となる各 $(x, y) \in S^2$ について、

$$V(x \to y) - V(y \to x) = \varphi(y) - \varphi(x)$$

補題 1

ポテンシャル $\varphi : S \to (-\infty, \infty)$ が存在するとき、各 $x \in S$ について、

$$q_\infty(x) = 0 \Longrightarrow \varphi(x) > \min_{y \in S} \varphi(y)$$

また、このとき、W もポテンシャルになる。

証明すべきこと

$V(x) = -M \log f(i)$, $x = (i) \in U$, $i \in I$ が $S := U$ 上のポテンシャル

$S^* \subseteq U^*$ の証明 (続)

$$V((i) \to (j)) = \min_{u \sim v, v \succ (j)} \{\lambda d((i), u) + V(y) - \min_{v \succ r} V(r)\}$$

$f(i) > f(j)$ のとき、

1. 突然変異と交叉で個体 j を生成し、選択で劣勢な j を M 個選択
2. 突然変異で個体 j を M 個生成し、j を M 個選択
3. 突然変異と交叉で個体 j を 1 個と、適合度が $f(i), f(j)$ 以下の個体を $M-1$ 個を生成し、選択で優勢な j を M 個選択

$$\begin{aligned} & \min\{\lambda\rho(i,j) + V((j)) - V((i)), M\lambda\rho(i,j), \lambda[\rho(i,j) + M - 1]\} \\ \leq \quad & V((i) \to (j)) \\ \leq \quad & \min\{\lambda\rho(i,j) + V((j)) - V((i)), M\lambda\rho(i,j)\} \end{aligned}$$

$\lambda > \dfrac{M}{M-1}\Delta$ より、$V((i) \to (j)) = \lambda\rho(i,j) + V((j)) - V((i))$

$S^* \subseteq U^*$ の証明 (続)

$$V((i) \to (j)) - V((j) \to (i)) = M(\log f(i) - \log f(j)) = V(j) - V(i)$$

$f(i) \leq f(j)$ のとき

$$V((i) \to (j)) \quad = \quad \lambda\rho(i,j)$$

$f(i) > f(j)$,$f(i) \leq f(j)$ の両方で、
$V((i) \to (j)) - V((j) \to (i)) = V(j) - V(i)$ (ポテンシャル)

例: $M = 2$ のとき

β 大では、f が 1 対 1 であることより、選択の後一様集団になる

$$\boxed{\begin{array}{c} i_1 \cdots i_L \\ \hline i_1 \cdots i_L \end{array}} \xrightarrow{\mathcal{M}_\beta} \boxed{\begin{array}{c} i'_1 \cdots i'_L \\ \hline i''_1 \cdots i''_L \end{array}} \xrightarrow{\mathcal{C}} \boxed{\begin{array}{c} j_1 \cdots j_L \\ \hline j'_1 \cdots j'_L \end{array}} \xrightarrow{\mathcal{S}_\beta} \boxed{\begin{array}{c} j_1 \cdots j_L \\ \hline j_1 \cdots j_L \end{array}}$$

$\{i'_h, i''_h\} = \{j_h, j'_h\}, h = 1, \cdots, L$

まとめ

GA の Markov 連鎖を大偏差原理を用いて解析

1. 突然変異確率 $\mu \to 0$, 選択圧力 $\beta \to \infty \iff$ 絶対温度 $T \to 0$
2. μ 小、β 大なら、もっと解明される可能性

証明できたこと

1. Albuquerque-Mazza, 2001 が、交叉ありでも成立
2. 「突然変異 → 交叉 → 選択」でも
「交叉 → 突然変異 → 選択」でも成立

今後の課題

1. 「$f : I \to (0, \infty)$ が 1 対 1」という仮定の除去
2. Cerf, 1998 との比較