Self-Adjusting Population Sizes for Non-Elitist Evolutionary Algorithms Why Success Rates Matter

黄道煜

2023年3月11日

目录

Algorithm1

Algorithm 1 Self-adjusting $(1, \{F^{1/s}\lambda, \lambda/F\})$ EA.

Initialization: Choose $x \in \{0,1\}^n$ uniformly at random (u.a.r.) and $\lambda := 1$
Optimization: **for** $t \in \{1, 2, \dots\}$ **do**
  Mutation: **for** $i \in \{1, \dots, \lambda\}$ **do**
    Create $y'_i \in \{0,1\}^n$ by copying x and flipping each bit independently with probability $1/n$.
  Selection: Choose $y \in \{y'_1, \dots, y'_\lambda\}$ with $f(y) = \max\{f(y'_1), \dots, f(y'_\lambda)\}$ u.a.r.
  Update:
  if $f(y) > f(x)$ **then** $x \leftarrow y$; $\lambda \leftarrow \max\{1, \lambda/F\}$;
  else $x \leftarrow y$; $\lambda \leftarrow F^{1/s}\lambda$;

F 为进化强度，s 为成功率

当迭代成功时，λ 缩小，变为 λ/F
当迭代失败时，λ 增加，变为 $F^{1/s}\lambda$

Analysis1

定义 $f(x)$ 为 fitness function

定义 $g(x)$ 为 potential function 且满足 $g(X_t) = f(x_t) + h(\lambda_t)$.

定义以下符号：

$$p_{i,\lambda}^{-} = \Pr\left(f(x_{t+1}) < i \mid f(x_t) = i\right)$$

$$p_{i,\lambda}^{0} = \Pr\left(f(x_{t+1}) = i \mid f(x_t) = i\right)$$

$$p_{i,\lambda}^{+} = \Pr\left(f(x_{t+1}) > i \mid f(x_t) = i\right)$$

$$\Delta_{i,\lambda}^{-} = \mathrm{E}\left(i - f(x_{t+1}) \mid f(x_t) = i \text{ and } f(x_{t+1}) < i\right)$$

$$\Delta_{i,\lambda}^{+} = \mathrm{E}\left(f(x_{t+1}) - i \mid f(x_t) = i \text{ and } f(x_{t+1}) > i\right)$$

Analysis1

给出这四个变量的取值范围

$$1-\frac{en}{en+\lambda(n-i)} \leq 1-\left(1-\frac{n-i}{en}\right)^{\lambda} \leq p_{i,\lambda}^{+}$$

$$p_{i,\lambda}^{+} \leq 1-\left(1-1.14\left(\frac{n-i}{n}\right)\left(1-\frac{1}{n}\right)^{n-1}\right)^{\lambda} \leq 1-\left(1-\frac{n-i}{n}\right)^{\lambda}$$

$$1 \leq \Delta_{i,\lambda}^{+} \leq \sum_{j=1}^{\infty}\left(1-\left(1-\frac{1}{j!}\right)^{\lambda}\right)$$

$$\left(\frac{i}{n}-\frac{1}{e}\right)^{\lambda} \leq p_{i,\lambda}^{-} \leq \left(1-\frac{n-i}{en}-\left(1-\frac{1}{n}\right)^{n}\right)^{\lambda} \leq \left(\frac{e-1}{e}\right)^{\lambda}$$

$$1 \leq \Delta_{i,\lambda}^{-} \leq \frac{e}{e-1}$$

Analysis1

Theorem 3.1 *Consider the elitist* $(1+\{F^{1/s}\lambda, \lambda/F\})$ *EA on* OneMax *with* $f(x_0) \geq a$ *and an initial offspring population size of* λ_0*. For every integer* $b \leq n$*, the expected number of evaluations for it to reach a fitness of at least* b *is at most*

$$\lambda_0 \cdot \frac{F}{F-1} + \left(\frac{1}{e} + \frac{1-F^{-1/s}}{\ln(F^{1/s})}\right) \cdot \frac{F^{\frac{s+1}{s}}-1}{F-1} \sum_{i=a}^{b-1} \frac{en}{n-i}.$$

给出了 fitness value 从 a 到 b 的期望计算次数

Analysis1

Lemma 3.3 *Consider the elitist* $(1+\{F^{1/s}\lambda, \lambda/F\})$ *EA on* OneMax *with* $f(x_0) \geq a$ *and an initial offspring population size of* λ_0*. Fix an integer* $b \leq n$ *and, for all* $0 \leq i \leq n-1$*, let* U_i *denote the number of function evaluations made during all unsuccessful generations on fitness level* i*. Then the number of evaluations to reach a fitness of at least* b *is at most*

$$\lambda_0 \cdot \frac{F}{F-1} + \frac{F^{\frac{s+1}{s}} - 1}{F-1} \sum_{i=a}^{b-1} U_i.$$

建立总计算次数与迭代失败时计算次数的联系

Analysis1

定义函数 $\phi(\lambda_t) := \frac{F}{F-1}\lambda_t$ 表示“计算次数银行账户”。当 λ 增加时向账户中存钱，λ 减小时向账户中取钱，因此可以将迭代成功时的计算次数通过该函数在迭代失败时进行计算。

根据 $\phi(\lambda_t) - \phi(\lambda_t/F) = \lambda_t$ 通过待定系数法得到函数表达式。

每次失败时需要的计算次数为：

$$\lambda_t + \phi(\lambda_t F^{1/s}) - \phi(\lambda_t) = \lambda_t \left(1 + \frac{F^{\frac{s+1}{s}}}{F-1} - \frac{F}{F-1}\right) = \lambda_t \cdot \frac{F^{\frac{s+1}{s}} - 1}{F-1}$$

Analysis1

Lemma 3.4 *Consider the* $(1+\{F^{1/s}\lambda, \lambda/F\})$ *EA starting on fitness level* i *with an offspring population size of* λ*. For every initial* λ*, the expected number of evaluations in unsuccessful generations for the* $(1+\{F^{1/s}\lambda, \lambda/F\})$ *EA on fitness level* i *is at most*

$$E(U_i) \leq \frac{1}{p_{i,1}^{+}} \cdot \left(\frac{1}{e} + \frac{1 - F^{-1/s}}{\ln(F^{1/s})} \right).$$

给出 fitness value 为 i 时，迭代失败时期望的计算次数

Analysis1

根据期望定义：

$$\sum_{j=0}^{\infty} \lambda F^{j/s} \cdot \Pr\left(\text{no success in } \lambda \cdot \frac{F^{\frac{j+1}{s}}-1}{F^{1/s}-1} \text{ evaluations}\right)$$

$$\leq \sum_{j=0}^{\infty} \lambda F^{j/s} \cdot \exp\left(-p_{i,1}^{+} \lambda \cdot \frac{F^{\frac{j+1}{s}}-1}{F^{1/s}-1}\right)$$

对于任意最大值在 α 上的单峰函数$f(x)$，有：

$$\sum_{i=0}^{\infty} f(i) \leq f(\alpha) + \int_{0}^{\infty} f(i)\, \mathrm{d}i.$$

根据结论放缩得到：

$$\frac{1}{ep_{i,1}^{+}} + \exp\left(\frac{p_{i,1}^{+}\lambda}{F^{1/s}-1}\right) \int_{j=0}^{\infty} \lambda F^{j/s} \cdot \exp\left(-p_{i,1}^{+} \lambda \cdot \frac{F^{\frac{j+1}{s}}}{F^{1/s}-1}\right) \mathrm{d}j$$

使用换元积分得到：

$$\frac{1}{ep_{i,1}^{+}} + \exp\left(-p_{i,1}^{+}\lambda\right) \cdot \frac{F^{1/s}-1}{p_{i,1}^{+} \cdot F^{1/s} \ln\left(F^{1/s}\right)}.$$

Analysis1

Theorem 3.1 *Consider the elitist* $(1+\{F^{1/s}\lambda, \lambda/F\})$ *EA on* OneMax *with* $f(x_0) \geq a$ *and an initial offspring population size of* λ_0*. For every integer* $b \leq n$*, the expected number of evaluations for it to reach a fitness of at least* b *is at most*

$$\lambda_0 \cdot \frac{F}{F-1} + \left(\frac{1}{e} + \frac{1-F^{-1/s}}{\ln(F^{1/s})}\right) \cdot \frac{F^{\frac{s+1}{s}}-1}{F-1} \sum_{i=a}^{b-1} \frac{en}{n-i}.$$

分析 $(1+\{F^{1/s}\lambda, \lambda/F\})$ EA 计算次数总体思路：

1. 使用摊还分析，建立总计算次数与迭代失败时计算次数的联系
2. 运用放缩技巧，直接根据期望定义推出迭代失败时计算次数的期望

Analysis2

Theorem 4.1 *Let the update strength $F > 1$ and the success rate $0 < s < 1$ be constants. Then for any initial search point and any initial λ the expected number of generations of the self-adjusting $(1, \lambda)$ EA on* OneMax *is $O(n)$.*

$0 < s < 1$ 时，期望迭代次数为 $O(n)$

Analysis2

根据期望的定义，计算每次迭代 potential function 期望的增加量

$$
\begin{aligned}
&\mathrm{E}\left(g(X_{t+1})-g(X_t)\mid X_t\right)\\
&=\left(\Delta_{i,\lambda}^{+}+h(\lambda/F)-h(\lambda)\right)p_{i,\lambda}^{+}+\left(h(\lambda F^{1/s})-h(\lambda)\right)(p_{i,\lambda}^{0}+p_{i,\lambda}^{-})-\Delta_{i,\lambda}^{-}p_{i,\lambda}^{-}\\
&=\left(\Delta_{i,\lambda}^{+}+h(\lambda/F)-h(\lambda)\right)p_{i,\lambda}^{+}+\left(h(\lambda F^{1/s})-h(\lambda)\right)(1-p_{i,\lambda}^{+})-\Delta_{i,\lambda}^{-}p_{i,\lambda}^{-}\\
&=\left(\Delta_{i,\lambda}^{+}+h(\lambda/F)-h(\lambda F^{1/s})\right)p_{i,\lambda}^{+}+h(\lambda F^{1/s})-h(\lambda)-\Delta_{i,\lambda}^{-}p_{i,\lambda}^{-}
\end{aligned}
$$

给出 h(x), $g(x)$ 具体形式：

$$
h(\lambda_t) \quad = \quad -\tfrac{2s}{s+1}\log_F\left(\max\left(\tfrac{enF^{1/s}}{\lambda_t},1\right)\right)
$$

$$
g_1(X_t)=f(x_t)-\frac{2s}{s+1}\log_F\left(\max\left(\frac{enF^{1/s}}{\lambda_t},1\right)\right).
$$

Analysis2

分情况讨论，考虑 $\lambda_t \le en$ 时：

$$
\begin{aligned}
&\mathrm{E}\left(g(X_{t+1}) - g(X_t) \mid X_t\right) \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda)\right) p_{i,\lambda}^{+} + \left(h(\lambda F^{1/s}) - h(\lambda)\right) (p_{i,\lambda}^{0} + p_{i,\lambda}^{-}) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-} \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda)\right) p_{i,\lambda}^{+} + \left(h(\lambda F^{1/s}) - h(\lambda)\right) (1 - p_{i,\lambda}^{+}) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-} \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda F^{1/s})\right) p_{i,\lambda}^{+} + h(\lambda F^{1/s}) - h(\lambda) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-} \\
&= \frac{2}{s+1} + \left(\Delta_{i,\lambda}^{+} - 2\right) p_{i,\lambda}^{+} - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}. \\
&\ge 1 + \frac{1-s}{s+1} - p_{i,\lambda}^{+} - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}
\end{aligned}
$$

$$h(\lambda_t) \quad = \quad -\tfrac{2s}{s+1} \log_F \left(\max\left(\tfrac{enF^{1/s}}{\lambda_t}, 1\right)\right)$$

$$g_1(X_t) = f(x_t) - \frac{2s}{s+1} \log_F \left(\max\left(\frac{enF^{1/s}}{\lambda_t}, 1\right)\right).$$

Analysis2

分情况讨论，考虑 $\lambda_t \leq en$ 时：

$$\begin{aligned}
&\mathrm{E}\left(g(X_{t+1})-g(X_t)\mid X_t\right)\\
&\geq 1+\frac{1-s}{s+1}-p_{i,\lambda}^{+}-\Delta_{i,\lambda}^{-}p_{i,\lambda}^{-}\\
&\geq \frac{1-s}{s+1}+\left(1-\frac{1.14}{e\left(1-\frac{1}{n}\right)}\left(\frac{n-i}{n}\right)\right)^{\lfloor\lambda\rceil}-\left(\frac{e}{e-1}\right)\left(1-\frac{n-i}{en}-\frac{1}{e}\left(1-\frac{1}{n}\right)\right)^{\lfloor\lambda\rceil}\\
&=\frac{1-s}{s+1}+\left(1-\frac{1.14}{e}\left(\frac{n-i}{n-1}\right)\right)^{\lfloor\lambda\rceil}-\left(\frac{e}{e-1}\right)\left(\frac{e-1}{e}-\frac{n-i-1}{en}\right)^{\lfloor\lambda\rceil}.
\end{aligned}$$

$$h(\lambda_t) = -\tfrac{2s}{s+1}\log_F\left(\max\left(\tfrac{enF^{1/s}}{\lambda_t},1\right)\right)$$

$$g_1(X_t)=f(x_t)-\frac{2s}{s+1}\log_F\left(\max\left(\frac{enF^{1/s}}{\lambda_t},1\right)\right).$$

继续分类讨论 $\lfloor\lambda\rceil \geq 2$ 和 $\lfloor\lambda\rceil = 1$ 的情况，得到

$$\mathrm{E}\left(g(X_{t+1})-g(X_t)\mid X_t\right)\geq \tfrac{1-s}{2e}$$

Analysis2

分情况讨论，考虑 $\lambda_t \le en$ 时：

$$\mathrm{E}\left(g(X_{t+1}) - g(X_t) \mid X_t\right)$$

$$= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda)\right) p_{i,\lambda}^{+} - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}$$

$$= \left(\Delta_{i,\lambda}^{+} + \frac{2s}{s+1}\log_F\left(\frac{\lambda}{F}\right) - \frac{2s}{s+1}\log_F(\lambda)\right) p_{i,\lambda}^{+} - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}$$

$$= \left(\Delta_{i,\lambda}^{+} - \frac{2s}{s+1}\right) p_{i,\lambda}^{+} - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}$$

$$\ge \left(1-\frac{1}{e}\right)\left(1-\frac{2s}{s+1}\right) - \left(\frac{e-1}{e}\right)^{en-1}$$

$$= \frac{1}{e}\left(1-\frac{2s}{s+1}\right) + \left(1-\frac{2}{e}\right)\left(1-\frac{2s}{s+1}\right) - \left(\frac{e-1}{e}\right)^{en-1}$$

$$h(\lambda_t) \quad = \quad -\tfrac{2s}{s+1}\log_F\left(\max\left(\tfrac{enF^{1/s}}{\lambda_t}, 1\right)\right)$$

$$g_1(X_t) = f(x_t) - \frac{2s}{s+1}\log_F\left(\max\left(\frac{enF^{1/s}}{\lambda_t}, 1\right)\right).$$

比较后两项，得到

$$\mathrm{E}\left(g(X_{t+1}) - g(X_t) \mid X_t\right) \ge \frac{1-s}{2e}$$

Analysis2

Theorem 4.1 *Let the update strength $F > 1$ and the success rate $0 < s < 1$ be constants. Then for any initial search point and any initial λ the expected number of generations of the self-adjusting $(1, \lambda)$ EA on* OneMax *is $O(n)$.*

运用漂移分析得到期望迭代次数为

$$\frac{n + \frac{2s}{s+1} \log_F \left(enF^{1/s}\right)}{\frac{1-s}{2e}} = O(n)$$

Analysis2

Theorem 4.1 *Let the update strength $F > 1$ and the success rate $0 < s < 1$ be constants. Then for any initial search point and any initial λ the expected number of generations of the self-adjusting $(1, \lambda)$ EA on* ONEMAX *is $O(n)$.*

分析 $(1, \{F^{1/s}\lambda, \lambda/F\})$ EA 迭代次数总体思路:
直接使用 Additive Drift Analysis

Analysis3

Theorem 4.5 *Let the update strength $F > 1$ and the success rate $0 < s < 1$ be constants. The expected number of function evaluations of the self-adjusting $(1, \lambda)$ EA on* OneMax *is $O(n \log n)$.*

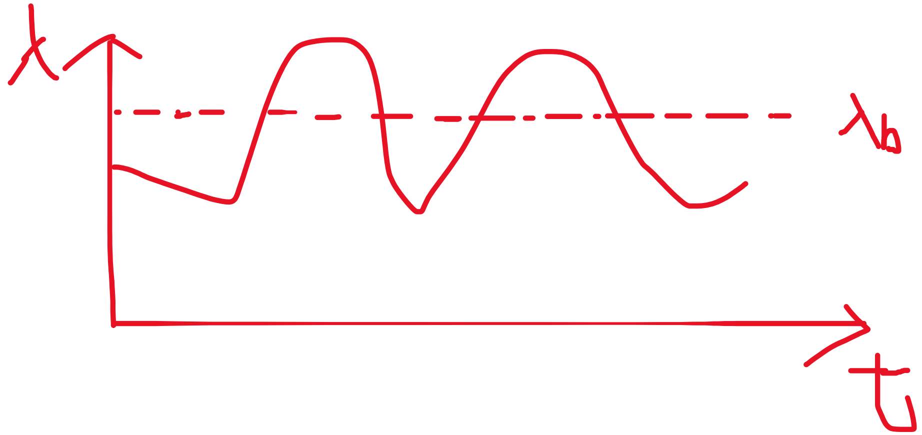
$0 < s < 1$ 时，期望计算次数为 $O(n)$

Analysis3

Lemma 4.8 *Consider the self-adjusting* $(1,\lambda)$ *EA as in Theorem 4.5. Fix a fitness value* b *and denote the current offspring population size by* λ_0. *Define* $\overline{\lambda_b} := CF^{1/s}/p^{+}_{b-1,1}$ *for a constant* $C > 0$ *that may depend on* F *and* s *that satisfies*

$$\left(\frac{s+1}{s} \cdot e^{1-C}\right)^{\frac{s}{s+1}} \leq \frac{F^{-1/s}}{2}. \tag{13}$$

Define a phase *as a sequence of generations that ends in the first generation where* λ *attains a value of at most* $\overline{\lambda_b}$ *or a fitness of at least* b *is reached. Then the expected number of evaluations made in that phase is* $O(\lambda_0)$.



Analysis3

先计算一个 phase 时间为 z 时的期望计算次数

$$\mathrm{E}(T \mid Z = z) \le \lambda_0 \cdot \sum_{i=1}^{z} F^{i/s} = \lambda_0 \cdot \frac{F^{\frac{z+1}{s}} - F^{1/s}}{F^{1/s} - 1} \le \lambda_0 \cdot \frac{F^{\frac{z+1}{s}}}{F^{1/s} - 1}$$

Analysis3

再计算一个 phase 时间为 z 时的概率
定义变量 Y'_i 表示第 i 次迭代是否失败
对迭代成功的概率进行放缩：

$$1-\left(1-p_{i,1}^{+}\right)^{\lambda_t} \geq 1-\left(1-p_{b-1,1}^{+}\right)^{\overline{\lambda_b}F^{-1/s}} \geq 1-e^{-\overline{\lambda_b}F^{-1/s}\cdot p_{b-1,1}^{+}} = 1-e^{-C}.$$

定义变量 Y_i 满足 $\Pr(Y_i = 1) = e^{-c}$，$\Pr(Y_i = 0) = 1 - e^{-c}$

定义 $\mathrm{Y} = \sum_{\mathrm{i}=1}^{z} Y_i$, $\mathrm{Y}' = \sum_{\mathrm{i}=1}^{z} Y'_i$ ，有 $Y' \preccurlyeq \mathrm{Y}$
$\Pr(\mathrm{Z}=\mathrm{z}) < \Pr(\mathrm{Z}>\mathrm{z}) \leq \Pr\left(Y' \geq z\frac{s}{s+1}\right) \leq \Pr(Y \geq z\frac{s}{s+1})$
根据 E(Y =ze^{-c})，应用 Chernoff bound 得到上界

Analysis3

最终得到每个 phase 期望计算次数

$$
\begin{aligned}
\mathrm{E}(T) &= \sum_{z=1}^{\infty} \Pr(Z=z) \cdot \mathrm{E}(T \mid Z=z) \\
&\leq \sum_{z=1}^{\infty} F^{-z/s} \cdot 2^{-z} \cdot \lambda_0 \cdot \frac{F^{\frac{z+1}{s}}}{F^{1/s}-1} \\
&= \lambda_0 \cdot \frac{F^{1/s}}{F^{1/s}-1} \cdot \sum_{z=1}^{\infty} 2^{-z} = \lambda_0 \cdot \frac{F^{1/s}}{F^{1/s}-1}.
\end{aligned}
$$

fitness value 从 a 到 b 到期望计算次数

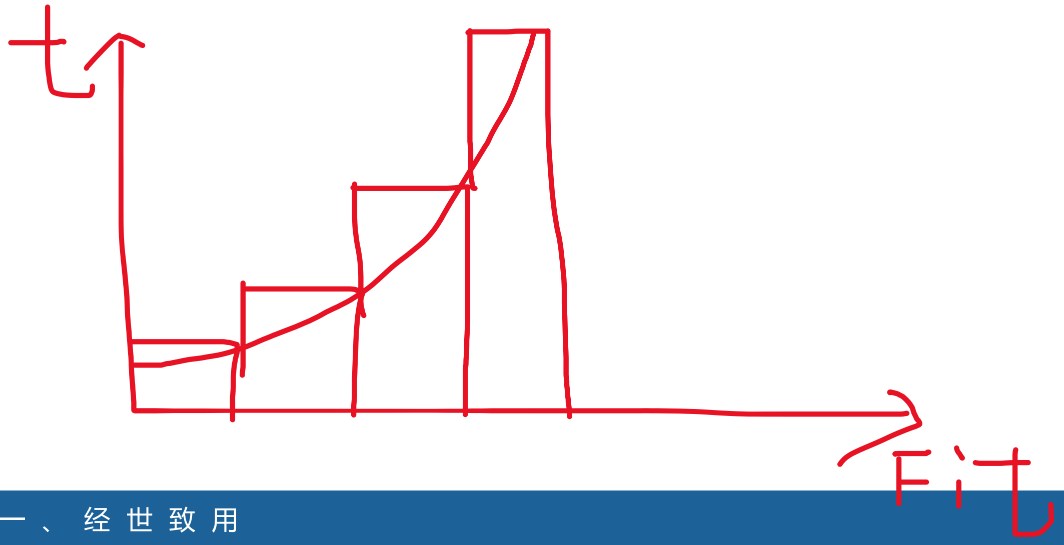
$$
E(\lambda_0 + \cdots + \lambda_t \mid \lambda_0) \leq O(\lambda_0) + O(b-a+\log n) \cdot \frac{1}{p^{+}_{b-1,1}}.
$$

Analysis3

Lemma 4.9 *Consider the self-adjusting* $(1, \lambda)$ *EA as in Theorem* 4.5. *Starting with a fitness of* a *and an offspring population size of* λ_0, *the expected number of function evaluations until a fitness of at least* b *is reached for the first time is at most*

$$E(\lambda_0 + \cdots + \lambda_t \mid \lambda_0) \leq O(\lambda_0) + O(b - a + \log n) \cdot \frac{1}{p^+_{b-1,1}}.$$

fitness value 从 a 到 b 到期望计算次数

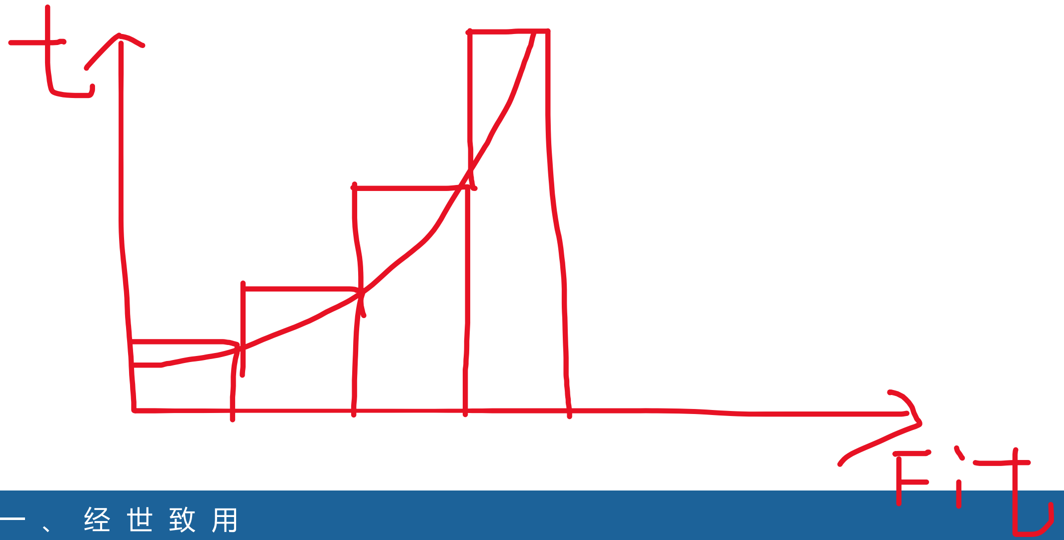


Analysis3

Lemma 4.9 with $a := n - (i+1)\log n$, $b := n - i\log n$ and $\lambda_0 := \lambda^{(i)}$,

$$\mathrm{E}(T_i) \leq O(\lambda^{(i)}) + O(\log n) \cdot \frac{1}{p^{+}_{n-i\log(n)-1,1}}.$$

根据当前种群所处 fitness value 分成大小为 log n 的段，计算分别每段的期望计算次数



Analysis3

Lemma 4.9 with $a := n - (i+1)\log n$, $b := n - i\log n$ and $\lambda_0 := \lambda^{(i)}$,

$$\mathrm{E}(T_i) \leq O(\lambda^{(i)}) + O(\log n) \cdot \frac{1}{p^{+}_{n-i\log(n)-1,1}}.$$

根据当前种群所处 fitness value 分成大小为 log n 的段，计算分别每段的期望计算次数

$$\begin{aligned}\sum_{i=0}^{\lceil n/\log n\rceil - 1} \mathrm{E}(T_i) &\leq O(n\log n) \cdot \sum_{i=0}^{\lceil n/\log n\rceil - 1} \frac{1}{1 + i\log n} \\ &\leq O(n\log n) \cdot \left(1 + \sum_{i=1}^{\lceil n/\log n\rceil - 1} \frac{1}{i\log n}\right) \\ &= O(n\log n) \cdot \left(1 + \frac{H_{\lceil n/\log n\rceil - 1}}{\log n}\right) = O(n\log n)\end{aligned}$$

Analysis3

Lemma 4.10 *Call a generation t* excessive *if, for a current search point with fitness i, at the end of the generation* λ *is increased beyond* $5F^{1/s}\ln(n)/p_{i,1}^{+}$. *Let* T *denote the expected number of function evaluations before a global optimum is found. Let* $\overline{T}$ *denote the number of evaluations made before a global optimum is found or until the end of the first excessive generation. Then*

$$E(T) \leq E\left(\overline{T}\right) + O(1).$$

说明 $\bar{T}, T$ 期望时间相同，可以认为算法结束前 λ 不过超过限制

Analysis3

根据λ 大小，将算法运行时间分为三部分

第一部分：$\lambda < 5F^{1/s}\ln(n)/p_{i,1}^{+}$

第二部分：$5F^{1/s}\ln(n)/p_{i,1}^{+} \leq \lambda < F^{1/s}n^{3}$

第三部分：$F^{1/s}n^{3} \leq \lambda$

每一部分对应的事件为 B，对应计算次数为 T，对应迭代次数为 G

$$\mathrm{E}(T) \leq \mathrm{E}\left(T^{(1)}\right) + \Pr\left(\overline{B^{(1)}}\right)\left(\mathrm{E}\left(T^{(2)}\right) + \Pr\left(\overline{B^{(2)}}\right) \cdot \mathrm{E}\left(T^{(3)}\right)\right)$$

Analysis3

$$\mathrm{E}(T) \leq \mathrm{E}\left(T^{(1)}\right) + \Pr\left(\overline{B^{(1)}}\right)\left(\mathrm{E}\left(T^{(2)}\right) + \Pr\left(\overline{B^{(2)}}\right) \cdot \mathrm{E}\left(T^{(3)}\right)\right)$$

$\mathrm{B}^{(1)}$ 时，每次迭代 λ 超出限制的概率

$$(1 - p_{i,1}^{+})^{\lambda_t} \leq (1 - p_{i,1}^{+})^{5\ln(n)/p_{i,1}^{+}} \leq e^{-5\ln(n)} = n^{-5}.$$

总体 λ 超出限制的概率

$$\Pr\left(\overline{B^{(1)}}\right) \leq \sum_{t=1}^{\infty} t \cdot n^{-5} \cdot \Pr\left(G^{(1)} = t\right) = n^{-5} \cdot \mathrm{E}\left(G^{(1)}\right) = O(n^{-4})$$

期望计算次数

$$\mathrm{E}\left(T^{(2)}\right) \leq \mathrm{E}\left(G^{(2)}\right) \cdot \lambda^{(2)} = O(n^4)$$

Analysis3

$$\mathrm{E}(T) \le \mathrm{E}\left(T^{(1)}\right) + \Pr\left(\overline{B^{(1)}}\right)\left(\mathrm{E}\left(T^{(2)}\right) + \Pr\left(\overline{B^{(2)}}\right) \cdot \mathrm{E}\left(T^{(3)}\right)\right)$$

$\mathrm{B}^{(2)}$ 时，每次迭代 λ 超出限制的概率

$$(1 - p_{i,1}^{+})^{\lambda_t} \le (1 - p_{n-1,1}^{+})^{\lambda^{(2)} F^{-1/s}} \le e^{-\Omega(n^2)}$$

总体 λ 超出限制的概率

$$\Pr\left(\overline{B^{(2)}}\right) \le \sum_{t=1}^{\infty} t \cdot e^{-\Omega(n^2)} \cdot \Pr\left(G^{(2)} = t\right) = e^{-\Omega(n^2)} \cdot \mathrm{E}\left(G^{(2)}\right) = e^{-\Omega(n^2)}$$

期望计算次数

$$\mathrm{E}\left(T^{(3)}\right) \le n^n$$

Analysis3

$$\mathrm{E}(T) \leq \mathrm{E}\left(T^{(1)}\right) + \Pr\left(\overline{B^{(1)}}\right)\left(\mathrm{E}\left(T^{(2)}\right) + \Pr\left(\overline{B^{(2)}}\right) \cdot \mathrm{E}\left(T^{(3)}\right)\right)$$

$$\mathrm{E}(T) \leq \mathrm{E}\left(T^{(1)}\right) + O(n^{-4})\left(O(n^4) + e^{-\Omega(n^2)} \cdot n^n\right) = \mathrm{E}\left(T^{(1)}\right) + O(1).$$

Analysis3

Theorem 4.5 *Let the update strength $F > 1$ and the success rate $0 < s < 1$ be constants. The expected number of function evaluations of the self-adjusting $(1, \lambda)$ EA on* OneMax *is $O(n \log n)$.*

分析 $(1, \{F^{1/s}\lambda, \lambda/F\})$ EA 迭代次数总体思路:

1. 对于每个 fitness value 划分 phase，分析每个 phase 的期望计算次数，得出期望计算次数与期望迭代次数的联系。
2. 通过计算不同时间上界的关系，规避掉种群过大时产生的干扰

Analysis4

Theorem 5.1 *Let the update strength $F \leq 1.5$ and the success rate $s \geq 18$ be constants. With probability $1 - e^{-\Omega(n/\log^4 n)}$ the self-adjusting $(1, \lambda)$ EA needs at least $e^{\Omega(n/\log^4 n)}$ evaluations to optimise* OneMax.

分析 $(1, \{F^{1/s}\lambda, \lambda/F\})$ EA 计算次数总体思路:
直接使用 Negative Drift Analysis

Analysis4

Theorem 2.6 (Negative drift theorem with scaling [36]) *Let X_t, $t > 0$ be real-valued random variables describing a stochastic process over some state space. Suppose there exists an interval $[a, b] \subseteq \mathbb{R}$ and, possibly depending on $\ell := b - a$, a drift bound $\varepsilon := \varepsilon(\ell) > 0$ as well as a scaling factor $r := r(\ell)$ such that for all $t \geq 0$ the following three conditions hold:*

1. $E(X_{t+1} - X_t \mid X_0, \ldots, X_t;\ a < X_t < b) \geq \varepsilon$.
2. $\Pr\left(|X_{t+1} - X_t| \geq jr \mid X_0, \ldots, X_t;\ a < X_t\right) \leq e^{-j}$ *for* $j \in \mathbb{N}_0$.
3. $1 \leq r^2 \leq \varepsilon\ell/(132\log(r/\varepsilon))$.

Then for the first hitting time $T^* := \min\{t \geq 0\colon X_t < a \mid X_0, \ldots, X_t;\ X_0 \geq b\}$ *it holds that* $\Pr\left(T^* \leq e^{\varepsilon\ell/(132r^2)}\right) = O(e^{-\varepsilon\ell/(132r^2)})$.

说明了在 potential value 期望减少的情况下， potential value 从 a 到 b 的期望时间为指数级别

第一个条件说明 potential value 期望减少

第二三个条件说明 potential value 大幅变化概率极小

Analysis4

根据期望的定义，计算每次迭代 potential function 期望的增加量

$$
\begin{aligned}
& \mathrm{E}\left(g(X_{t+1}) - g(X_t) \mid X_t\right) \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda)\right) p_{i,\lambda}^{+} + \left(h(\lambda F^{1/s}) - h(\lambda)\right)\left(p_{i,\lambda}^{0} + p_{i,\lambda}^{-}\right) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-} \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda)\right) p_{i,\lambda}^{+} + \left(h(\lambda F^{1/s}) - h(\lambda)\right)\left(1 - p_{i,\lambda}^{+}\right) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-} \\
&= \left(\Delta_{i,\lambda}^{+} + h(\lambda/F) - h(\lambda F^{1/s})\right) p_{i,\lambda}^{+} + h(\lambda F^{1/s}) - h(\lambda) - \Delta_{i,\lambda}^{-} p_{i,\lambda}^{-}
\end{aligned}
$$

给出 h(x), $g(x)$ 具体形式：

$$h(\lambda_t) := 2.2 \log_F^2(\lambda_t)$$

$$g_2(X_t) := f(x_t) + 2.2 \log_F^2 \lambda_t .$$

Analysis4

根据期望的定义，计算每次迭代 potential function 期望的增加量

$$
\begin{aligned}
&\mathrm{E}\left(g(X_{t+1}) - g(X_t) \mid X_t\right) \\
&= \left(\Delta^+_{i,\lambda} + h(\lambda/F) - h(\lambda)\right) p^+_{i,\lambda} + \left(h(\lambda F^{1/s}) - h(\lambda)\right)(p^0_{i,\lambda} + p^-_{i,\lambda}) - \Delta^-_{i,\lambda} p^-_{i,\lambda} \\
&= \left(\Delta^+_{i,\lambda} + h(\lambda/F) - h(\lambda)\right) p^+_{i,\lambda} + \left(h(\lambda F^{1/s}) - h(\lambda)\right)(1 - p^+_{i,\lambda}) - \Delta^-_{i,\lambda} p^-_{i,\lambda} \\
&= \left(\Delta^+_{i,\lambda} + h(\lambda/F) - h(\lambda F^{1/s})\right) p^+_{i,\lambda} + h(\lambda F^{1/s}) - h(\lambda) - \Delta^-_{i,\lambda} p^-_{i,\lambda}
\end{aligned}
$$

给出 h(x), $g(x)$ 具体形式：

$$h(\lambda_t) := 2.2 \log_F^2(\lambda_t)$$

$$g_2(X_t) := f(x_t) + 2.2 \log_F^2 \lambda_t .$$

Analysis 4

Lemma 5.3 *Consider the self-adjusting* $(1, \lambda)$ *EA as in Theorem* 5.1. *Then there is a constant* $\delta > 0$ *such that for every* $0.84n + 2.2\log^2(4.5) < g_2(X_t) < 0.85n$,

$$E(g_2(X_{t+1}) - g_2(X_t) \mid X_t) \leq -\delta.$$

选取 fitness value 为 0.84n~0.85n 的区间，每次 potential value 都期望减小

Analysis 4

$$\Delta_{g_2} \leq \left(\Delta^+_{i,\lambda} - \left(1 + \frac{1}{s}\right) \cdot 4.4 \log_F(\lambda) + 2.2 - \frac{2.2}{s^2}\right) p^+_{i,\lambda} + \frac{4.4 \log_F \lambda}{s} + \frac{2.2}{s^2} - \Delta^-_{i,\lambda} p^-_{i,\lambda}.$$

$$\Delta_{g_2} \leq \left(\Delta^+_{i,\lambda} - \frac{19}{18} \cdot 4.4 \log_F \lambda + 2.2 - \frac{2.2}{324}\right) p^+_{i,\lambda} + \frac{4.4 \log_F \lambda}{18} + \frac{2.2}{324} - \Delta^-_{i,\lambda} p^-_{i,\lambda}.$$