Additive noise in discrete optimization

Y. Akimoto, S. Astete, O. Teytaud

Our acknowledgements to: A. Pruguel-Bennett, J.Rowe, J. Shapiro in Dagstuhl seminar Theory of Evolutionary Algorithms 2013, who participated in the discussions which initiated this work

Introduction

The problem

- Evolutionary Algorithm designed to optimize discrete functions
- What happens if we add *noise*?

Case 1

Gaussian noise

Case 2

Any noise with finite variance

Initial setting

Setting

Let:

- *EA* an Evolutionary Algorithm
- G_n an *instance*, i.e. a set of fitness functions from $\{0,1\}^n$ to $\mathbb{N}$.
- Every $g \in G_n$ has an optimum $x^* := \arg\max g$

Definition (Complexity $f(n, \delta)$ of an algorithm EA)

The **complexity** $f(n, \delta)$ of an algorithm EA is the number of fitness evaluations needed to find the optimum of any function $g \in G_n$, with probability at least $1 - \delta$.

The Solution

Add noise to G_n

EA

Noise free instance

G_n

$\longrightarrow\longrightarrow\longrightarrow$

Modify EA:
Revaluation individuals

Averaging their fitness

Modified-EA

Noisy instance

$G_n + noise$

Revaluation-EA

Definition (Algorithm Revaluation-EA)

Let α and β two individuals. Then Revaluation-EA:

- To compute the fitness value of α:
  Average between k fitness evaluations of α, where

$$k = max\left(1, \left\lceil \sigma^2 \log\left(\frac{1 - \exp(\log(1-\delta)/f(n,\delta))}{A}\right) / \log(B) \right\rceil\right)$$

- To compare the individual α and the individual β:
  - α better β if average of α greater by at least $\frac{1}{2}$
  - β better α if average of β greater by at least $\frac{1}{2}$
  - they have the same fitness value otherwise.

Theorem 1

Theorem

Assume that EA solves G_n with complexity $f(n, \delta)$. Then Revaluation-EA solves $G_n + \sigma\mathcal{N}$ with probability $(1-\delta)^2$ and number of fitness evaluations

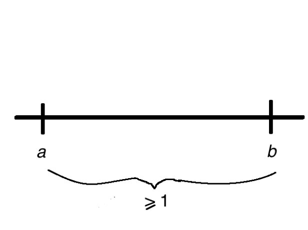
$$O\left(f(n,\delta)\left\lceil \sigma^2 \log\left(\frac{f(n,\delta)}{-\log(1-\delta)}\right)\right\rceil\right) \quad (1)$$

Proof Theorem 1

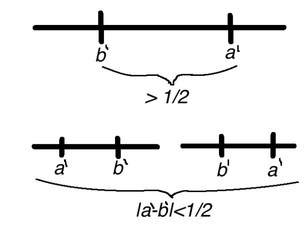
Step 1

Compute probability of "misranking" a *given* individual, using deviation of a Gaussian.

EA: $b > a$



Revaluation-EA: $b' \leq a'$



Proof Theorem 1

Step 1

Compute probability of "misranking" a *given* individual, using deviation of a Gaussian.

Step 2

Compute probability of misranking *any* individual, using union bound and Step 1.

Step 3

Compute probability of EA and Revaluation-EA having the same ranking $\rightsquigarrow$ We have the result

ManyRevaluation-EA

Definition (Algorithm ManyRevaluation-EA)

Analogous to Revaluation-EA, but with:

$$k = max(1, \lceil 4\sigma^2 f(n,\delta)/\delta \rceil) \tag{2}$$

Theorem 2

Theorem

Assume that EA solves G_n with complexity $f(n,\delta)$. Then ManyRevaluation-EA solves $G_n + \sigma N$ (with N some arbitrary noise with variance 1) with probability $(1-\delta)^2$ and number of fitness evaluations

$$O\left(f(n,\delta) max(1, 4\sigma^2 f(n,\delta)/\delta)\right). \tag{3}$$

Proof Theorem 2

Step 1

Compute probability of "misranking" a *given* individual, using **Chebyshev's inequality**

Step 2

Compute probability of misranking *any* individual, using union bound and Step 1.

Step 3

Compute probability of EA and Revaluation-EA having the same ranking
$\rightsquigarrow$ We have the result

Application: OneMax

OneMax_n

For every $z \in \{0,1\}^n$ we define:

$$\begin{aligned} \mathrm{OM_z} : \{0,1\}^n &\rightarrow \{0,\ldots,n\} \\ x &\mapsto |\{j \in \{0,\ldots,n\} | x_j = z_j\}| \end{aligned}$$

So, OneMax_n is the set of all $\mathrm{OM_z}, z \in \{0,1\}^n$

Algorithm and complexity

- $(1+1)-\mathrm{EA} : \Theta(n \log n)$ [Droste, Jansen, Wegener, 2002]

Conclusions

Contribution

From noise free case to noisy case modifying the original algorithm

- Find optimum with high probability
- Complexity of the algorithm for noisy instances

Influence of Noise

Gaussian

$$O\left(f(n,\delta)\left\lceil \sigma^2 \log\left(\frac{f(n,\delta)}{-\log(1-\delta)}\right)\right\rceil\right)$$

Any noise finite variance

$$O\left(f(n,\delta) max(1, \frac{4\sigma^2 f(n,\delta)}{\delta})\right)$$

Perspectives

Perspectives

1. Finding the complexity? Not free! [Oliveto, Witt, 2010]
2. Can we discard the need of the complexity *a priori*?
3. Revaluate vs not revaluate [Shapiro, Pruguel − Bennet, Rowe]
4. Complexity analysis in noisy settings