Global Exploration Strategy for Derivative Free Optimization

N. Soualmi

joint work : Y. Diouane, S. Gratton, L. N. Vicente

APO Journee des Doctorants 2014

Outline

1. Introduction
2. A framework for global exploration
3. Application to a class of evolution strategy
4. Summary

Outline

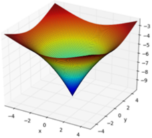
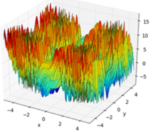
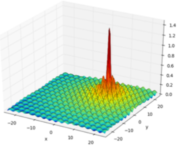
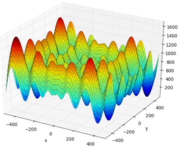
1. Introduction
2. A framework for global exploration
3. Application to a class of evolution strategy
4. Summary

Introduction

Global Optimization and Derivative-Free Optimization (DFO)

Objective function is supposed to be :

- Differentiable, but derivative can't be used
- Noisy
- Multimodal
- Evaluations costly

Global Exploration Strategies

Some existing strategies

Global Exploration Strategies

Some existing strategies

- Branch and bound inspired methods (MCS [Neumaier, 1999], DIRECT [Jones, et al. 1993],..)

Global Exploration Strategies

Some existing strategies

- Branch and bound inspired methods (MCS [Neumaier, 1999], DIRECT [Jones, et al. 1993],..)
- Multi-start strategies (MLSL [Rinnooy, et al. 1987], GLODS [Custòdio, Madeira, 2013],..)

Global Exploration Strategies

Some existing strategies

- Branch and bound inspired methods (MCS [Neumaier, 1999], DIRECT [Jones, et al. 1993],..)
- Multi-start strategies (MLSL [Rinnooy, et al. 1987], GLODS [Custòdio, Madeira, 2013],..)
- Stochastic strategies (CMA-ES [Hansen, Ostermeier, 2001], ASA [Ingber, 1993],..)
- Others (PSWARM [Vaz, Vicente, 2009], HMLSL [Pal, 2013],..)

Main objectives

- Propose a possible framework for global exploration.
- Application to various algorithms.
- Novel application of the proposed framework in stochastic strategies.

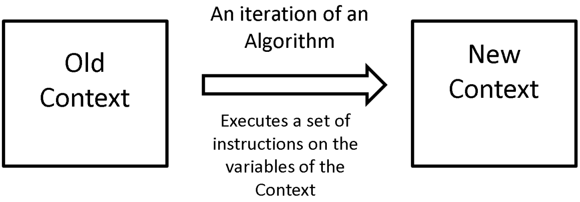
Outline

1. Introduction
2. A framework for global exploration
3. Application to a class of evolution strategy
4. Summary

Context of execution of an algorithm

Context

It gathers all variables and control parameters used by the algorithm (e.g. step size, historic, current point..)



The proposed framework

The proposed framework

The List_context

This object contains a list of execution contexts for *algorithm*

The proposed framework

The List_context

This object contains a list of execution contexts for *algorithm*

The building procedure

- From an existing List_context, create new contexts :
  - Split a context into many new contexts : enable to follow different tracks (e.g. splitting procedure in MCS)
  - Sample new contexts for exploring search space (e.g. sampling step in MLSL)
- Freezing contexts : either the context has converged, or two (or more) are converging toward same local minimum (e.g. merging in GLODS)

The proposed Framework

The proposed Framework

The Activation procedure

Select contexts from list_context (e.g. clustering technique in MLSL, activation of boxes of different levels in MCS).

The proposed Framework

The Activation procedure

Select contexts from list_context (e.g. clustering technique in MLSL, activation of boxes of different levels in MCS).

The Running procedure

Given an initial context, execute a number of instructions of *algorithm* (e.g. n iterations, n function evaluations..)

The proposed framework

Framework

Initialize List_context(1)

For i = 1,..

    Select some contexts to be active

    For K in list_context(i) which is active

        Set context K and run *algorithm*

    End

    Build List_context(i+1) from List_context(i)

End

Outline

1. Introduction
2. A framework for global exploration
3. Application to a class of evolution strategy
4. Summary

Evolution Strategies (ES)

The algorithm

1) Offspring generation : compute new sample of points

$X_{k+1} = \{x^1_{k+1}, x^2_{k+1}, .., x^\lambda_{k+1}\}$, such that

$$x^i_{k+1} \sim x^{mean}_k + \sigma \cdot N(0, C)$$

2) Evaluate $f(x^i_{k+1})$, $i = 1, .., \lambda$

3) Reorder $X_{k+1} = \{\tilde{x}^1_{k+1}, .., \tilde{x}^\lambda_{k+1}\}$ such that $f(\tilde{x}^1_{k+1} \leq .. \leq f(\tilde{x}^i_{k+1})$

$$x^{mean}_{k+1} = \sum_{i=1}^{\mu} w_i \cdot \tilde{x}^i_{k+1}$$

4) Update σ, C, go to step 1)

Application to ES

Creating new contexts :

- Use of the notion of dispersion metrics introduced in [Lunacek, Whithley, 2006]

> *"The dispersion metric introduced by Lunacek and Whithley attempts such a classification by quantifying the notion of underlying unimodality. Using this metric it was observed that CMA-ES was originally developed as a local search strategy, whereas the concept of multi-funnel functions is based on global information."* [Muller, et al. 2009]

Application to ES

Creating new contexts :

- Use of the notion of dispersion metrics introduced in [Lunacek, Whithley, 2006]

  "The dispersion metric introduced by Lunacek and Whithley attempts such a classification by quantifying the notion of underlying unimodality. Using this metric it was observed that CMA-ES was originally developed as a local search strategy, whereas the concept of multi-funnel functions is based on global information." [Muller, et al. 2009]

- Our approach : We compute an approximation of the dispersion metric of the function, using clustering techniques on the sample of points.

Clustering technique

- Decompose the sample S into two groups :

  C_b : μ_1 "best" points $\quad C_w$: μ_2 "worst" points

- Uses clustering techniques : Apply kmeans [MacQueen, 1967] for partitioning C_b into p clusters :

  $C_b = \bigcup_{j=1}^{p} C_j$

- Approximate dispersion metric :

  Compute average pairwise distance (apd) for C_1,.., C_p, C_b, C_w, and S.

If $\frac{\sum_{j=1}^{p} apd(C_j)}{p \cdot apb(S)} \geq \alpha$ **or** $\frac{apd(C_w)}{apb(S)} \leq \beta$ **then**

the dispersion is declared high : generate new contexts

else

the dispersion is low, do not generate new contexts

end

Application to ES

The Activation procedure

Select N contexts among :

- Contexts with best minimum values found (1)
- Best progression (2)
- Newly created contexts (3)

Activated contexts $= \theta_1 \cdot N$ from $(1) + \theta_2 \cdot N$ from (2) $+ \theta_3 \cdot N$ from (3)
where $\theta_1 + \theta_2 + \theta_3 = 1$

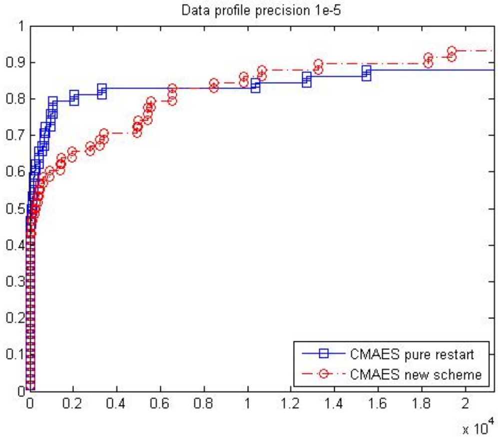
Merging

Consider distance between means of active contexts, say $mean_1$ and $mean_2$
If $distance(mean_1, mean_2) < threshold$ then freeze one of the two context.

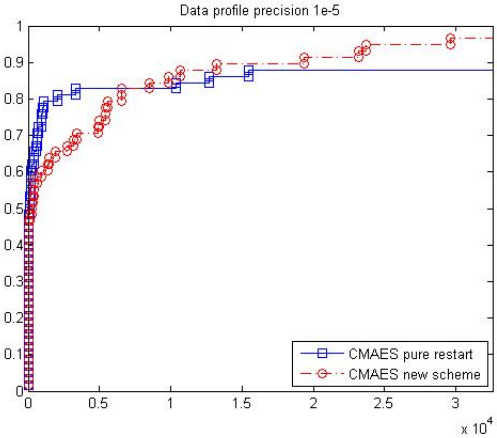
Preliminary numerical results

- Test set of global optimization problems used by GLODS [Custodio, Madeira, 2013].
- Budget of 20000 evaluation functions.
- Competitor : pure multi-start strategy CMAES [Auger, Hansen, 2005].
- Test strategy : data profile.

Numerical results



Numerical results



Outline

1. Introduction
2. A framework for global exploration
3. Application to a class of evolution strategy
4. Summary

Summary

- We propose a framework for global optimization, based on the update and activation of contexts.
- Possible application to other algorithms
- A novel application of this framework on a class of evolution strategy.
- On going work focuses on theory in parallel asynchronous algorithms, and implementation on parallel machines