A DIRECT-type Approach for Derivative-Free Constrained Global Optimization

G. Di Pillo\textsuperscript{1}, G. Liuzzi\textsuperscript{2}, S. Lucidi\textsuperscript{1}, V. Piccialli\textsuperscript{3} and F. Rinaldi\textsuperscript{4}

\textsuperscript{1}Sapienza University of Rome
\textsuperscript{2}National Council of Research, Rome
\textsuperscript{3}Tor Vergata University of Rome
\textsuperscript{4}University of Padua

SIAM Conference on Optimization
San Diego, California, USA
May 19-22, 2014
Outline

Introduction

A DIRECT-type Algorithm

Using Global Information

Local Search

Numerical Results

Conclusions

G. Di Pillo, G. Liuzzi, S. Lucidi, V. Piccialli and F. Rinaldi

A DIRECT-type Approach for Derivative-Free Constrained Global Optimization
Problem Definition

We assume that

- $f$ is a *general nonlinear* function
- $g$ and $h$ are sets of *general nonlinear* inequality and equality constraints
- no global information (convexity, Lipschitz constants,....) available.

Feasible set of Problem (P) denoted by $\mathcal{F}$:

$$\mathcal{F} = \{ x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0, l \leq x \leq u \}.$$ 

Problem stated as finding a point $x^*$ such that

$$f(x^*) \leq f(x) \quad \forall x \in \mathcal{F}.$$ 

We denote by $\mathcal{X}^*$ the set of global solutions of Problem (P)

$$\mathcal{X}^* = \{ x^* : f(x^*) \leq f(x) \quad \forall x \in \mathcal{F} \}.$$
Why a derivative-free approach?

In recent years, we have been involved with industrial contracts concerning the optimal design of devices whose performances (objective functions) are provided by simulation programs:

- the design of electro-mechanic devices,
- the design of electromagnetic diagnostic equipments,
- the shape optimization in ship design,
- the design of nanoelectronic circuits.

In these cases, even if you may consider that the objective function depends continuously on the decision variables, you cannot evaluate its derivative.

On the other hand, in these problems usually the number $n$ of decision variables is not large, thus making viable a derivative-free approach.
Why a derivative-free approach?

A derivative-free approach is required in all cases where a so-called black box optimization is only possible, that is cases where only the values of objective functions and constraints functions for given values of the decision variables are available.

This motivates the effort we have devoted to derivative-free optimization and the development of a derivative-free library:

www.dis.uniroma1.it/~lucidi/DFL/
Why Global Optimization?

- Many real-world problems in Engineering, Economics and Applied Sciences modeled as nonlinear global minimization problems
- Growing attention in the search for global rather than local solutions of nonlinear optimization problems
- Unconstrained or simply-constrained problems: many algorithmic approaches (deterministic or probabilistic) available
- General constraints: various approaches described (see e.g. [Floudas, 1999], [Tawarmalani, Sahinidis, 2002], [Neumaier et al., 2005], [Birgin et al., 2009])
Our aim is to combine

- an efficient derivative-free global optimization algorithm for problems with simple bounds,
- an exact penalty approach for feasibility of general constraints
- a bilevel approach for optimality of the overall problem

in order to develop a derivative-free algorithm for the global solution of optimization problems with general constraints.

In particular we will make use of DIRECT as algorithm for global optimization with simple bounds.
Constraints Violation

Let us define the following penalty function

\[ w(x) = \left\| \begin{pmatrix} \max\{0, g(x)\} \\ h(x) \end{pmatrix} \right\|. \]
A Bilevel Approach

Begin

Find set $W$ of global minimizers for Problem 1

Choose the point in $W$ with minimum $f(x)$

End
Constraints Violation

Let us define the following penalty function

\[ w(x) = \left\| \left( \max\{0, g(x)\} \right) \frac{1}{h(x)} \right\|. \]

The original problem can be thought as the sum of two conflicting problems:

**Problem 1**

\[ \text{globmin} \quad w(x) \quad l \leq x \leq u \]

**Problem 2**

\[ \text{globmin} \quad f(x) \quad l \leq x \leq u \]

We denote by \( \mathcal{D} \) the hyperinterval

\[ \mathcal{D} = \{ x \in \mathbb{R}^n : l \leq x \leq u \} \]

Of course \( X^* \subseteq \mathcal{D} \).
A Bilevel Approach

1. Begin
2. Find set \( W \) of global minimizers for Problem 1
3. Choose the point in \( W \) with minimum \( f(x) \)
4. End
Iteration $k$ of the DIRECT Algorithm

1. Begin iteration $k$, Partition $\{D^i, i \in I_k\}$
2. Choose $I_k^* \subseteq I_k$ related to Potentially Optimal Hyperintervals
3. Do $A(\bar{x}^i) \forall i \in I_k^*$ and Possibly Improve Best Solution
4. Do Partition Procedure on $D^i, i \in I_k^*$

End iteration $k$
Potentially Optimal Hyperintervals for $w$

We introduce the following definition of \textit{potentially optimal hyperintervals} for the penalty function $w$.

\textbf{Definition 1}

An hyperinterval $D^n$, $h \in I_k$ is said to be \textit{potentially optimal} for $w$ if a constant $\overline{l}^n$ exists such that:

$$w(x^n) - \frac{\overline{l}^n}{2} \|u^n - l^n\| \leq w(x') - \frac{\overline{l}^n}{2} \|u' - l'\|, \quad \forall i \in I_k, \quad (1)$$

$$w(x^n) - \frac{\overline{l}^n}{2} \|u^n - l^n\| \leq w_{\min} - \epsilon |w_{\min}| \quad (2)$$

We further define the set of indices

$$I_w^k = \{i \in I_k : D' \text{ is potentially optimal with respect to } w\}$$
Partitioning Strategy

- evaluate the function \( w \) on the \( 2n \) points \( x' \pm \frac{2}{3} \delta \)

G. Di Pillo, G. Liuzzi, S. Lucidi, V. Piccialli and F. Rinaldi
A DiRECT-type Approach for Derivative-Free Constrained Global Optimization
Convergence Results

Proposition 2

Given $\epsilon > 0$, let $\hat{x} \in \mathcal{D}$ and $B(\hat{x}, \epsilon) = \{x \in \mathbb{R}^n : \|x - \hat{x}\| \leq \epsilon\}$. Then, there exists an iteration $k$ and an index $h \in I_k$ such that $x^h \in B(\hat{x}, \epsilon)$.

Corollary 1

For every feasible point $\hat{x} \in \mathcal{F}$ and for every $\epsilon > 0$, there exists an iteration $k$ and an index $h \in I_k$ such that point $x^h \in B(\hat{x}, \epsilon)$.

Corollary 2

For every global minimum point $x^*$ of $f(x)$ on $\mathcal{F}$ and for every $\epsilon > 0$, there exists an iteration $k$ and an index $h \in I_k$ such that point $x^h \in B(x^*, \epsilon)$.
Using Global Information

- Estimate of Lipschitz constants available for some classes of constraints (e.g. linear, quadratic)
- Use these estimates to define a new version of the DIRECT Algorithm
- New way to choose the hyperintervals to be partitioned
Strongly Potentially Optimal Hyperintervals for $w$

We introduce the following definition of strongly potentially optimal hyperintervals for the penalty function $w$.

**Definition 3**

Let $\bar{L} > 0$ be an estimate of the Lipschitz constant of function $w$. An hyperinterval $D^h$, $h \in I_k$ is said to be strongly potentially optimal for $w$ if one of the following conditions is satisfied:

- a constant $\bar{L}^h \in (0, \bar{L})$ exists such that:

$$w(x^h) - \frac{\bar{L}^h}{2} \| u^h - r^h \| \leq w(x^i) - \frac{\bar{L}^h}{2} \| u^i - r^i \|, \quad \forall i \in I_k,$$

$$w(x^h) - \frac{\bar{L}^h}{2} \| u^h - r^h \| \leq w_{\min} - \varepsilon \max\{w_{\min}, \eta\}, \quad \eta \text{ small}$$

- $$w(x^h) - \frac{L}{2} \| u^h - r^h \| \leq w(x^i) - \frac{L}{2} \| u^i - r^i \|, \quad \forall i \in I_k.$$
Convergence to Feasible Points

If Strongly Potentially Optimal Hyperintervals Used
Strongly Potentially Optimal Hyperintervals for Problem (P)

We introduce the following definition of strongly potentially optimal hyperintervals for Problem (P).

**Definiton 4**

Let $\bar{L} > 0$ be an estimate of the Lipschitz constant of function $f$. An hyperinterval $D^h$, $h \in l_k^w$ is said to be strongly potentially optimal with respect to $f$ if one of the following conditions is satisfied:

- a constant $\bar{L}^h \in (0, \bar{L})$ exists such that:
  \[ f(x^h) - \frac{\bar{L}^h}{2} \| u^h - r^h \| \leq f(x^i) - \frac{\bar{L}^h}{2} \| u^i - r^i \|, \quad \forall i \in l_k^w, \tag{8} \]

- \[ f(x^h) - \frac{\bar{L}^h}{2} \| u^h - r^h \| \leq f_{\min} - \epsilon \max\{ |f_{\min}|, \eta \}, \quad \eta \text{ small} \tag{9} \]

- \[ f(x^h) - \frac{\bar{L}^h}{2} \| u^h - r^h \| \leq f(x^i) - \frac{\bar{L}^h}{2} \| u^i - r^i \|, \quad \forall i \in l_k^w, \tag{10} \]
Convergence to Global Solutions

If Strongly Potentially Optimal Hyperintervals for Problem (P) used
Local Search

- **IDEA**: Replace random generation in multistart approach with deterministic partitioning of DIRECT
- Use the most promising points generated by DIRECT as starting points for the local algorithm
- Perform a DF local minimization from each centroid of the potentially optimal intervals
Local Search

In this way we can

- exploit the capability of the DIRECT approach of producing partitions which cluster around promising regions in order to locate a point in an *attraction region* of a global solution

- avoid the need that DIRECT produces a very fine partition to obtain a *good* approximation of a global solution

- exploit the efficiency of local algorithms in solving ill-conditioned and/or large scale problems
Convergence Result (with Local Searches)

The map $\mathcal{A}: \mathbb{R}^n \to \mathbb{R}^n$ represents a local minimization

**Assumption 1**

For any starting point $x \in \mathcal{D}$:
- $\bar{x} = \mathcal{A}(x)$ is a stationary point of the original constrained problem;
- For every global minimum $x^*$ an open set $\mathcal{L}$ exists such that if $x \in \mathcal{L}$ then $x^* = \mathcal{A}(x)$.

**Proposition 3**

An iteration $k$ and an index $h \in I^*_k$ exist such that $\mathcal{A}(x^h) = x^*$ where $x^*$ is a global minimum of the problem.
Numerical Results

- Numerical experimentation carried out on an Intel Core 2 Duo 3.16 GHz processor with 3.25 GB RAM
- Performance evaluated on a set of 84 problems from GLOBALLib [Shcherbina, Neumaier, Sam-Haroud, Vu, Nguyen, 2003] collection with dimension $n \leq 10$
- Estimates of the Lipschitz constant not available
- Local minimization carried out with DFN$_{con}$ [Fasano, Liuzzi, Lucidi, Rinaldi, 2013: A DF algorithm for constrained optimization problems, to appear in SIOPT]
- Comparison with the results of [Di Pillo, Lucidi, Rinaldi, 2013: A DF algorithm for constrained global optimization based on exact penalty functions, to appear in JOTA]
Numerical Results

- Maximum number of hyperintervals $500 \times n \times (me + mi)$, with $n$ number of variables, $me$ number of equality constraints and $mi$ number of inequality constraints.

- Constraints violation given by

$$cv(x) = \left\| \left( \max\{0, g(x)\} \right) \right\|_\infty$$

- Relative gap given by

$$\frac{|f(x^*) - f^*|}{\max\{1, |f^*|\}}$$

where $f^*$ is the best known solution.

- For gap evaluation considered solutions $x^*$ s.t.

$$cv(x^*) \leq 10^{-4}$$
Performance of the Algorithm (Constraints Violation)

Constraints violation: blue line SIOPT2014, red line JOTA paper

G. Di Pillo, G. Liuzzi, S. Lucidi, V. Piccialli and F. Rinaldi
A DIRECT-type Approach for Derivative-Free Constrained Global Optimization
Performance of the Algorithm (Gap)

Optimality gap: blue line SI OPT 2014, red line JOTA paper

G. Di Pillo, G. Liuzzi, S. Lucidi, V. Piccialli and F. Rinaldi
A DIRECT-type Approach for Derivative-Free Constrained Global Optimization
Numerical Results

Algorithm gives good solutions in reasonable time

Average CPU time 10.37 seconds
Conclusions and Future Work

Conclusions
- Development of a DIRECT-type approach for constrained optimization
- Preliminary numerical results very promising

Future Work
- Development of a more efficient version of the DIRECT algorithm to increase the dimension of the problems that can be solved
- Numerical experimentation when estimates of the Lipschitz constants are available
- Comparison with other codes
Many thanks for your attention !!!

G. Di Pillo, G. Liuzzi, S. Lucidi, V. Piccialli and F. Rinaldi

A DIRECT-type Approach for Derivative-Free Constrained Global Optimization