Optimality-based Domain Reductions for Global Optimization

A. Caprara, M. Locatelli, M. Monaci

Aussois, January 2013

Outline of the talk

- Settings:
  - global optimization problems
  - domain reduction strategies
- Optimality-based domain reduction strategies:
  - some relevant strategies
  - properties on a subclass of GO problems
  - computational results

Global Optimization problems

We consider the following problem

$$(GO) \qquad \max_{\mathbf{x} \in F \cap B} f(\mathbf{x})$$

where:

- f is a nonconvex function;
- $F \subseteq \mathbb{R}^n$ is a closed convex set;
- $B = [\mathbf{l}, \mathbf{u}]$ is a box.

- Difficult (although non integer) problems;
- often these problems have enough structure to derive an optimal (or approximate) solution in a reasonable amount of time;
- among solution methods, enumerative algorithms play a relevant role.

Domain Reduction strategies

- The performances of solution algorithms (e.g., branch-and-bound) can be strongly enhanced by procedures that try to reduce as much as possible the domain for the variables.
- These procedures are usually referred to as *Domain Reduction* (DR) strategies.
- Reducing the domain for variables
  - reduces the solution space (hence, smaller decision trees);
  - yields to tighter upper bounds at decision nodes.
- DR strategies are crucial in most optimization solvers for global optimization (e.g., BARON).

Domain reduction strategies

There are two main classes of DR strategies:

- feasibility-based DR: shrink the initial box B to a smaller box $\widetilde{B}$ that includes all feasible solutions (i.e., $B \cap F = \widetilde{B} \cap F$)
- optimality-based DR: assume that a lower bound LB is known, and define a box B_{DR} that includes all solutions (if any) whose value is larger than LB.

We will restrict our attention to

- optimality-based DR strategies,
- applied to variables that appear in the objective function only; $f = f(x_1, \ldots, x_t)$

Domain reduction strategies

By definition, for any optimality-based DR, the following relation holds:

$$B_{LB} \subseteq B_{DR} \subseteq B,$$

i.e., B_{LB} is a lower limit for the reduction that can be attained by any optimality-based DR.

Questions:

- Is there a significative subclass of non trivial GO problems for which we are able to find a DR which always guarantees that $B_{DR} = B_{LB}$?
- From a computational viewpoint, which is the right compromise between the accuracy in the reduction (i.e., tightness of the upper bound) and the computational cost?

Concave overestimator

An *overestimator* of f over B is a function g such that

$$g(\mathbf{x}) \geq f(\mathbf{x}) \quad \forall\, \mathbf{x} \in B.$$

- Typically, function g depends on the box B.
- Once an overestimator g is known, an upper bound on the optimal solution value can be computed as $\max_{\mathbf{x} \in F \cap B} g(\mathbf{x})$.
- Usually, one is interested in a *concave overestimator*.
- The *best* (i.e., smallest) concave overestimator of f over the box B is called the *concave envelope* of f over B.
- The closest g to the concave envelope, the tighter the resulting upper bound. However, improving the overestimator is difficult.
- DR strategies: try to improve g by reducing box B.

Standard DR (SDR)

- Choose a variable x_k ($k \in [1, t]$), and solve the following convex optimization problem:
- $$\begin{aligned} l'_k = \min \quad & x_k \\ \text{s.t.} \quad & g(x_1, \ldots, x_t) \geq LB, \\ & (x_1, \ldots, x_n) \in F \\ & l_j \leq x_j \leq u_j, \qquad j = 1, \ldots, n. \end{aligned}$$
- then determine $u'_k = \max x_k$ ...

- Each time some the lower and/or upper bound is improved, we possibly also improve the overestimating function g.
- In this case, it makes sense to iterate the process.
- We call Iterated SDR (ISDR) the strategy obtained by iteratively applying SDR until no further range reductions are possible for variable x_k.

Nonlinear Removal DR (NRDR)

- Fix variable x_k to some value $\alpha \in [\ell_k, u_k]$;
- compute an upper bound when $x_k = \alpha$ as follows

$$\begin{aligned} h_k(\alpha) = \max \quad & g_\alpha^k(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_t) \\ \text{s.t.} \quad & (x_1, \ldots, x_{k-1}, \alpha, x_{k+1}, \ldots, x_n) \in F \\ & l_j \leq x_j \leq u_j, \qquad \forall j \neq k. \end{aligned}$$

- Define the new lower bound for x_k as

$$l'_k = \inf\{\alpha \ : \ h_k(\alpha) \geq LB\}$$

and the new upper bound for x_k as

$$u'_k = \sup\{\alpha \ : \ h_k(\alpha) \geq LB\}.$$

Comparison between the DR strategies

Caprara and Locatelli proved that:

- if the same concave overestimator g is used for f in both ISDR and NRDR, then $B_{ISDR} = B_{NRDR}$
  i.e., the domain reductions for variable x_k produced by ISDR and NRDR coincide;

- if the DR strategy is applied to all variables, in turn, the resulting ranges for the variables converge to some limit which does not depend on the order in which variables are processed.

A subclass of GO problems

Consider GO problems that satisfy the following conditions:

- C1 the objective function f depends on two variables only, i.e., $f(x) = f(x_1, x_2)$;
- C2 for each $\alpha \in [l_1, u_1]$, $f_\alpha(x_2) = f(\alpha, x_2)$ is a convex, increasing (or decreasing) one-dimensional function;
- C3 for each $\beta \in [l_2, u_2]$, $f_\beta(x_1) = f(x_1, \beta)$ is a convex, increasing (or decreasing) one-dimensional function;
- C4 f is Lipschitzian with Lipschitz constant L.

A subclass of GO problems

The class of problems that satisfy C1-C4 includes some hard problems:

- the problem

$$\max\{x_1^2 - x_2 : Ax \leq b; \;\; l \leq x \leq u\}$$

  is NP-hard (Pardalos and Vavasis)
- minimizing the product of two affine functions over a polyhedron (see, e.g., Konno et al., and Sahinidis) can be transformed into an equivalent problem that satisfies C1-C4

$$\begin{array}{ll} \min & y_1 y_2 \\ \text{s.t.} & \sum_{j=1}^{n} c_j^1 x_j + c_{01} = y_1 \\ & \sum_{j=1}^{n} c_j^2 x_j + c_{02} = y_2 \\ & x \in P \end{array}$$

Properties of the subclass of GO problems

Theorem

For each problem that satisfies C1-C4, as soon as LB is not updated any more, we have

$$(B_{ISDR} =)\, B_{NRDR} = B_{LB}$$

Proof.

(sketch)

- we define a weaker DR strategy: apply the feasibility-based DR over the upper bounds of x_1 and x_2 and of NRDR over the lower bounds of the same variables (assuming both f_α and f_β increasing);
- we show that the iterative application of this strategy
  - either leads to an improved lower bound LB,
  - or identifies a rectangle $[\bar{l}_1, \overline{u}_1] \times [\bar{l}_2, \overline{u}_2] = B_{LB}$.

□

Special case

In particular, if

- only one optimal pair (x_1, x_2) exists (i.e., all globally optimal solutions have the same x_1 and x_2 values), and
- LB is equal to the global optimal value

then

- the rectangle B will be shrunk to a single point
- i.e., the procedure will converge to the globally optimal solution of the problem.

Enlarging the subclass?

Remove C1: the objective function involves more than two variables

$$\begin{array}{ll} \max & x_1^2 + x_2^2 + x_3^2 \\ \text{s.t.} & (6+8\varepsilon)x_1 + (6+4\varepsilon)x_2 + 6x_3 = 9 + 8\varepsilon \\ & 0 \le x_1, x_2 \le 1 \\ & 0 \le x_3 \le 1 + \varepsilon, \end{array}$$

where $\varepsilon > 0$.

- The feasible region is the polytope whose vertices are:
  $(1, 0, 1/2)$ $(1/2, 1, 0)$ $(0, 1/2, 1+\varepsilon)$
  $(0, 1, (3+4\varepsilon)/6)$ $(1, 3/(6+4\varepsilon), 0)$ $((3+2\varepsilon)/(6+8\varepsilon), 0, 1+\varepsilon)$.

- if LB is equal to the global optimal value $LB = 1/4 + (1+\varepsilon)^2$, then $h_1(0) > h_1(1) > LB$ $\rightarrow$ no reduction is possible for x_1.

The same applies in case f is linearly dependent on x_3.

Enlarging the subclass?

Remove C2: the objective function obtained when fixing x_1 is not monotone in x_2.

$$\begin{array}{ll} \max & (x_1 - 1/2)^2 + (x_2 - 1/2)^2 \\ \text{s.t.} & 2x_1 + 2x_2 \geq 1 \\ & -2x_1 + 2x_2 \leq 1 \\ & 2x_1 + 6x_2 \leq 7 \\ & 6x_1 + 2x_2 \leq 7 \\ & 2x_1 - 2x_2 \leq 1 \\ & 0 \leq x_1, x_2 \leq 1. \end{array}$$

- The feasible region is the polytope whose vertices are:

$$(1/2, 1) \quad (0, 1/2) \quad (1/2, 0) \quad (1, 1/2) \quad (7/8, 7/8).$$

- if LB is equal to the global optimal value $LB = 9/32$, then $h_1(0) = h_1(1) = 1/2 > LB \quad \rightarrow$ no reduction is possible for x_1.

Computational impact of DR strategies

- DR strategies allow to reduce the search box as soon as a lower bound LB is available
- what is the computational impact of these DR strategies in an enumerative scheme?
- which is the best compromise between the quality of the reduction and the required computing time?

Instances

- random instances on Linear Multiplicative Programming problems

$$\min \quad \prod_{i=1}^{p} (c_i x + c_{0i})$$
$$x \in P$$

  where $p \geq 2$, P is a polytope and
  $c_i x + c_{0i} > 0 \quad \forall\, x \in P,\ i = 1, \ldots, p.$

- Reformulate the problem as a concave separable problem as follows

$$\begin{array}{ll} \min & \sum_{i=1}^{p} \log(y_i) \\ & y_i = c_i x + c_{0i} \quad i = 1, \ldots, p \\ & x \in P \\ & \ell_i \leq y_i \leq u_i \qquad i = 1, \ldots, p, \end{array} \tag{1}$$

  where

$$\ell_i / u_i \;=\; \min / \max \quad c_i x + c_{0i}$$
$$x \in P.$$

Test settings

- DR strategies embedded within a branch-and-bound algorithm.
- At each node:
  - an upper bound is computed by replacing the objective function with its concave envelope;
  - the optimal solution of the relaxation also provides a feasible solution (i.e., a lower bound);
  - branching is performed by splitting an interval $[\bar{l}_i, \overline{u}_i]$ into two sub-intervals $[\bar{l}_i, y_i^*]$ and $[y_i^*, \overline{u}_i]$
- possible strengthening of the model using DR strategies: an upper bound v denotes the maximum number of iterations of the external loop in any iterated strategy.

Computational experiments: linear relaxation

Instance class			$v = 0$		$v = 1$		$v = \infty$	
p	m	n	Time	gap	Time	gap	Time	gap
2	500	500	0.01	2.94	0.67	0.22	1.72	0.00
5	500	500	0.01	8.14	1.41	6.15	82.27	3.77
10	200	200	0.00	11.44	0.24	10.58	4.07	6.06
20	100	100	0.00	12.75	0.14	11.68	1.89	5.97

- 10 instances for each class;
- times expressed in seconds on an Intel 6600 @ 2.40GHz.

Computational experiments: branch-and-bound

Instance class			$v = 0$		$v = 1$		$v = \infty$	
p	m	n	#opt	#nodes	#opt	#nodes	#opt	#nodes
2	500	500	10	33	10	4	10	1
5	500	500	10	1,716	10	114	10	29
10	200	200	10	124,265	10	1,433	9	204
20	100	100	0	$\simeq$ 1.5M	5	8,125	5	390

- 10 instances for each class;
- time limit for each instance: 3,600 seconds on an Intel 6600 @ 2.40GHz.

Conclusions

- Domain Reduction strategies play a crucial role in enumerative methods for Global Optimization problems.
- We have considered optimality-based domain reduction strategies,
- defining conditions under which we are guaranteed to produce the best possible reduction (for a given lower bound).
- Computational analysis to evaluate the impact of these procedures in a standard branch-and-bound algorithm.
- DR strategies, while expensive, offer a significative reduction of the size of the enumerative tree.