Mathematically Rigorous Feasible Points

Some Observations on Exclusion Regions in Interval Branch and Bound Algorithms

Ralph Baker Kearfott

Department of Mathematics
University of Louisiana at Lafayette

SCAN 2014, September, 2014

Outline

Overall Context

- In branch and bound (B&B) algorithms for constrained global optimization, an acceleration technique is to construct regions $\boldsymbol{x}^*$ around local optimizing points $\check{x}$, then delete these *exclusion regions* from further search.

Mathematically Rigorous Feasible Points

Overall Context

- In branch and bound (B&B) algorithms for constrained global optimization, an acceleration technique is to construct regions $\boldsymbol{x}^*$ around local optimizing points $\check{x}$, then delete these *exclusion regions* from further search.
- $\boldsymbol{x}^*$ too small $\Longrightarrow$
  adjacent regions not easily rejected. In this case, the B&B algorithm may produce a large *cluster* of small boxes sharing a boundary with $\boldsymbol{x}^*$.

Overall Context

- In branch and bound (B&B) algorithms for constrained global optimization, an acceleration technique is to construct regions $\boldsymbol{x}^*$ around local optimizing points $\check{x}$, then delete these *exclusion regions* from further search.
- $\boldsymbol{x}^*$ too small $\Longrightarrow$ adjacent regions not easily rejected. In this case, the B&B algorithm may produce a large *cluster* of small boxes sharing a boundary with $\boldsymbol{x}^*$.
- $\boldsymbol{x}^*$ too large $\Longrightarrow$ solution bounds are not accurate.

History

(selected work)

- Rump (1994 and probably earlier) proposes using interval slopes to construct a small box in which a unique solution exists and, with *epsilon-inflation*, a large box in which no other solutions can exist.

History

(selected work)

- Rump (1994 and probably earlier) proposes using interval slopes to construct a small box in which a unique solution exists and, with *epsilon-inflation*, a large box in which no other solutions can exist.
- Kearfott (Math. Comp., 1987) analyzes an abstract B&B algorithm for nonlinear systems, proving under certain conditions that if the diameter of $\boldsymbol{x}^*$ is proportional to $\sqrt{\epsilon}$, where ϵ is the smallest box diameter produced in branch and bound, then clustering of boxes will not occur.

History

(selected work)

- Rump (1994 and probably earlier) proposes using interval slopes to construct a small box in which a unique solution exists and, with *epsilon-inflation*, a large box in which no other solutions can exist.
- Kearfott (Math. Comp., 1987) analyzes an abstract B&B algorithm for nonlinear systems, proving under certain conditions that if the diameter of $\boldsymbol{x}^*$ is proportional to $\sqrt{\epsilon}$, where ϵ is the smallest box diameter produced in branch and bound, then clustering of boxes will not occur.
- Mayer (1995) surveys epsilon-inflation techniques.

History

(selected work)

- Rump (1994 and probably earlier) proposes using interval slopes to construct a small box in which a unique solution exists and, with *epsilon-inflation*, a large box in which no other solutions can exist.
- Kearfott (Math. Comp., 1987) analyzes an abstract B&B algorithm for nonlinear systems, proving under certain conditions that if the diameter of $\boldsymbol{x}^*$ is proportional to $\sqrt{\epsilon}$, where ϵ is the smallest box diameter produced in branch and bound, then clustering of boxes will not occur.
- Mayer (1995) surveys epsilon-inflation techniques.
- van Iwaarden (1996, Ph.D. dissertation) independently discovers a technique he calls *backboxing* to construct exclusion regions

History

(selected work)

- Rump (1994 and probably earlier) proposes using interval slopes to construct a small box in which a unique solution exists and, with *epsilon-inflation*, a large box in which no other solutions can exist.
- Kearfott (Math. Comp., 1987) analyzes an abstract B&B algorithm for nonlinear systems, proving under certain conditions that if the diameter of $\boldsymbol{x}^*$ is proportional to $\sqrt{\epsilon}$, where ϵ is the smallest box diameter produced in branch and bound, then clustering of boxes will not occur.
- Mayer (1995) surveys epsilon-inflation techniques.
- van Iwaarden (1996, Ph.D. dissertation) independently discovers a technique he calls *backboxing* to construct exclusion regions
- Schichl and Neumaier (2004, SINUM) study using higher order information for constructing exclusion

History
(continued)

- Schichl and Neumaier (2013, accepted), study exclusion regions for optimization problems.

History
(continued)

- Schichl and Neumaier (2013, accepted), study exclusion regions for optimization problems.
- Domes and Neumaier (March, 2014) propose a new technique for verification of feasibility, for construction of exclusion regions.

Scope of This Work

Contrast with Previous Work

- Previous work depended on uniqueness verification using Kuhn–Tucker conditions and interval Newton methods, or more sophisticated higher-order methods (in the case of pure nonlinear systems, especially.
- Motivation for this work is to imagine what conditions are necessary in a branch and bound algorithm for boxes adjacent to a small solution-containing box to be verified feasible, then construct the solution-containing box in such a way that adjacent boxes are sure to be verified infeasible. The exclusion region is then constructed to fit exactly this imagined set of adjacent boxes, eliminating the need for them to be created and processed by branch-and-bound.

The Present Work

Similarities with Other Ongoing Work

- This work involves taking a linear combination of the constraints that is likely to be easy to prove infeasible.

The Present Work

Similarities with Other Ongoing Work

- This work involves taking a linear combination of the constraints that is likely to be easy to prove infeasible.
- Work in progress by Domes and Neumaier also takes linear combinations of constraints, but a different linear combination related to approximate Lagrange multipliers.

The Present Work

Similarities with Other Ongoing Work

- This work involves taking a linear combination of the constraints that is likely to be easy to prove infeasible.
- Work in progress by Domes and Neumaier also takes linear combinations of constraints, but a different linear combination related to approximate Lagrange multipliers.
- The Domes / Neumaier work provides a rejection criterion, while this work provides an exclusion region construction criterion.

The Present Work

Similarities with Other Ongoing Work

- This work involves taking a linear combination of the constraints that is likely to be easy to prove infeasible.
- Work in progress by Domes and Neumaier also takes linear combinations of constraints, but a different linear combination related to approximate Lagrange multipliers.
- The Domes / Neumaier work provides a rejection criterion, while this work provides an exclusion region construction criterion.
- We are also working on use of a third linear combination, for use in a rejection criterion.

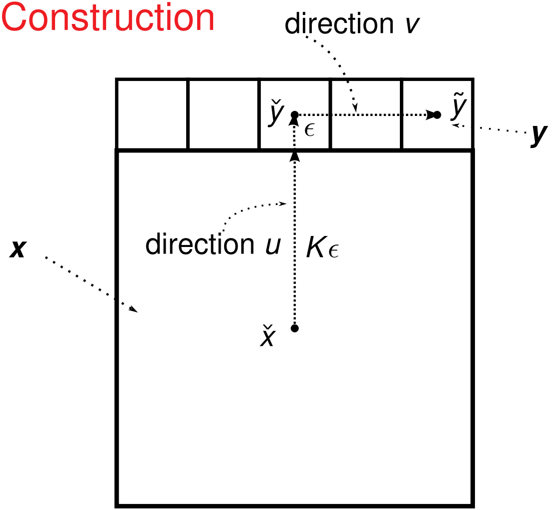
The Present Work

Similarities with Other Ongoing Work

- This work involves taking a linear combination of the constraints that is likely to be easy to prove infeasible.
- Work in progress by Domes and Neumaier also takes linear combinations of constraints, but a different linear combination related to approximate Lagrange multipliers.
- The Domes / Neumaier work provides a rejection criterion, while this work provides an exclusion region construction criterion.
- We are also working on use of a third linear combination, for use in a rejection criterion.
- None of these is a catch-all to work on all problems.

Mathematically Rigorous Feasible Points

The Construction



For fixed ϵ, K is chosen so boxes $\boldsymbol{y}$ will certainly be rejected.

Overall Analysis

- The problem is assumed to have been converted to one with equality constraints only*.

* *Inequality constraints may also be included later, at some additional complication.*

Overall Analysis

- The problem is assumed to have been converted to one with equality constraints only*.
- We expand an equality constraint in a Taylor series, and look at increments in a constraint proceeding from the center $\check{x}$ to the boundary of the central box in a direction u orthogonal to the box face.

* *Inequality constraints may also be included later, at some additional complication.*

Mathematically Rigorous Feasible Points

Overall Analysis

- The problem is assumed to have been converted to one with equality constraints only*.
- We expand an equality constraint in a Taylor series, and look at increments in a constraint proceeding from the center $\check{x}$ to the boundary of the central box in a direction u orthogonal to the box face.
- We subtract from that possible decrements in directions parallel to the box face.

* *Inequality constraints may also be included later, at some additional complication.*

Overall Analysis

- The problem is assumed to have been converted to one with equality constraints only*.
- We expand an equality constraint in a Taylor series, and look at increments in a constraint proceeding from the center $\check{x}$ to the boundary of the central box in a direction u orthogonal to the box face.
- We subtract from that possible decrements in directions parallel to the box face.
- We bound the increments and decrements with interval evaluations, to prove the constraint is infeasible on a shell abutting the box.

* *Inequality constraints may also be included later, at some additional complication.*

Problem Notation

Problem Notation

minimize $\varphi(x)$

subject to:

$$C(x) = (c_1(x), \ldots c_{m_1}(x)) = 0,$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}$ and

$c_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots m_1$.

Problem Notation

minimize $\varphi(x)$
subject to:
$$C(x) = (c_1(x), \ldots c_{m_1}(x)) = 0,$$
where $\varphi : \mathbb{R}^n \to \mathbb{R}$ and
$c_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots m_1$.

The technique will work best if φ is simple, such as if φ consists of a single independent variable.

Selecting the Constraint

- A separate constraint combination is used for each coordinate direction u.

Selecting the Constraint

- A separate constraint combination is used for each coordinate direction u.
- The idea is to maximize the rate of change in the direction u and minimize the rate of change in orthogonal directions.

Selecting the Constraint

- A separate constraint combination is used for each coordinate direction u.
- The idea is to maximize the rate of change in the direction u and minimize the rate of change in orthogonal directions.
- The heuristic we use here to achieve this is to choose that constraint combination

$$\tilde{c}_i = \sum_{j=1}^{m_1} \alpha_j c_j(x) \quad \text{which minimizes } \|(\nabla C(\check{x}))\alpha - e_i\|_2$$

over the parameters $\alpha = (\alpha_1, \ldots, \alpha_{m_1})$, where $\nabla C(\check{x})$ is the n by m_1 matrix whose i-th column is the gradient of c_i at $\check{x}$ and where $e_i \in \mathbb{R}^n$ is the i-th coordinate vector.

Computing the Optimal Constraint Combination

Observations

- Minimizing $\|(\nabla C(\check{x}))\alpha - e_i\|_2$ is a floating point linear least squares problem, that can be done approximately.

Computing the Optimal Constraint Combination

Observations

- Minimizing $\|(\nabla C(\check{x}))\alpha - e_i\|_2$ is a floating point linear least squares problem, that can be done approximately.
- High-quality library software for efficiently doing this is available.

Computing the Optimal Constraint Combination

Observations

- Minimizing $\|(\nabla C(\check{x}))\alpha - e_i\|_2$ is a floating point linear least squares problem, that can be done approximately.
- High-quality library software for efficiently doing this is available.
- Once the optimal $\tilde{c}_i$ is obtain, we use it to compute an expansion factor K_i (depending on ϵ) for the i-th coordinate direction.

Computing an Overall Expansion Factor $\overline{K}$

- It is in general not possible to compute box width factors K_i for each coordinate direction, even with optimal $\tilde{c}_i$.

Computing an Overall Expansion Factor $\overline{K}$

- It is in general not possible to compute box width factors K_i for each coordinate direction, even with optimal $\tilde{c}_i$.
- The boundary-box rejection criterion depends on each coordinate width, not just the i-th.

Computing an Overall Expansion Factor $\overline{K}$

- It is in general not possible to compute box width factors K_i for each coordinate direction, even with optimal $\tilde{c}_i$.
- The boundary-box rejection criterion depends on each coordinate width, not just the i-th.
- We combine the expansion factors K_i of reasonable width to an overall expansion factor, used for every coordinate:

$$\overline{K} = \begin{cases} \displaystyle\max_{\substack{1\le i\le n, p\in\{-1,1\}\\ K_{i,p}<K_{\max}}} K_{i,p} & \text{if } \exists K_{i,p} < K_{\max}, \\ K_{\max} & \text{otherwise.} \end{cases}$$

$\overline{K}\epsilon$ will be the radius of the central box in each coordinate direction.

Experimental Setup

- We implemented the techniques using libraries from GlobSol.

Experimental Setup

- We implemented the techniques using libraries from GlobSol.
- We tried the techniques on a subset of the COCONUT Lib-1 test set.

Experimental Setup

- We implemented the techniques using libraries from GlobSol.
- We tried the techniques on a subset of the COCONUT Lib-1 test set.
- We automatically converted the inequality constraints to equality constraints with bound constraints, for these problems.

Experimental Setup

- We implemented the techniques using libraries from GlobSol.
- We tried the techniques on a subset of the COCONUT Lib-1 test set.
- We automatically converted the inequality constraints to equality constraints with bound constraints, for these problems.
- For each problem, we measured the number of faces of the central box upon which adjacent small boxes can be eliminated.

Experimental Results

- Of 26 problems tried, 15 had faces upon which adjacent small boxes could be eliminated.

Experimental Results

- Of 26 problems tried, 15 had faces upon which adjacent small boxes could be eliminated.
- Of the problems with box faces upon which adjacent small boxes could be eliminated, the number of faces ranged from $1/34$ to $32/44$.

Additional Observations

- The use of interval-valued partial derivatives in a Taylor expansion is similar to what is done in the interval Gauss–Seidel method.

Additional Observations

- The use of interval-valued partial derivatives in a Taylor expansion is similar to what is done in the interval Gauss–Seidel method.
- The linear combination of constraints is a kind of preconditioner for the system of constraints.

Additional Observations

- The use of interval-valued partial derivatives in a Taylor expansion is similar to what is done in the interval Gauss–Seidel method.
- The linear combination of constraints is a kind of preconditioner for the system of constraints.
- We are presently investigating an alternate linear combination, motivated by the same ideas, for filtering boxes during the branch and bound algorithm.

Additional Observations

- The use of interval-valued partial derivatives in a Taylor expansion is similar to what is done in the interval Gauss–Seidel method.
- The linear combination of constraints is a kind of preconditioner for the system of constraints.
- We are presently investigating an alternate linear combination, motivated by the same ideas, for filtering boxes during the branch and bound algorithm.
- We are comparing the techniques, theoretically and in practice, to our optimally preconditioned interval Gauss–Seidel method.