Global Optimization of Mixed-Integer Nonlinear Problems Using Interval Analysis

Presentation 11f, AIChE Annual Meeting 2005

Gang (Gary) Xu — Invensys/SimSci-Esscor
Mark A. Stadtherr — U. of Notre Dame

Motivation:

Explore the potential use of interval-Newton/generalized- bisection (INGB) approach (with or without branch-and-bound) for deterministic global optimization in mixed-integer nonlinear programming (MINLP).

Outline:

- Problem definition and background;
- Introduction and discussion of interval-Newton/generalized bisection (INGB) and its application to MINLP;
- Combination of INGB and branch-and-bound
- Examples

MINLP Problem:

$$\begin{aligned} \min \quad & f(x,y) \\ \text{s.t.} \quad & h(x,y) = 0, \\ & g(x,y) \leq 0, \end{aligned}$$

with x indicating real variables and y indicating integer variables

- Want to determine global minimum, or
- May want to determine all KKT points

Considerable work has been done on this problem – for example:

- ECP combined with a general branch and bound;
- DICOPT++ coupled with zero-one branch-and-bound;
- Outer Approximation (OA)
- SMIN- and GMIN αBB (Floudas and coworkers)

INGB Applied to MINLP:

- Basic Idea
  - Formulate MINLP problem as continuous (NLP) problem
  - Solve for KKT points using interval-Newton approach -- Guaranteed to find all KKT points
  - Enforce integers using a special bisection rule during the interval-Newton algorithm
  - Can combine with branch-and-bound -- Find only the KKT point that is the global minimum

Interval-Newton/Generalized Bisection Method (INGB)

- Given a system of equations, an initial interval (bounds on all variables), and a solution tolerance:
  - IN/GB can find (enclose), with mathematical and computational certainty, all solutions to the equation system, or it can determine that no solutions exist
  - The equation system must have a finite number of real roots in the initial interval
  - No strong assumptions or simplifications to the equation system are needed

INGB Method

Problem: Solve $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ for all roots in the interval $\mathbf{X}^{(0)}$

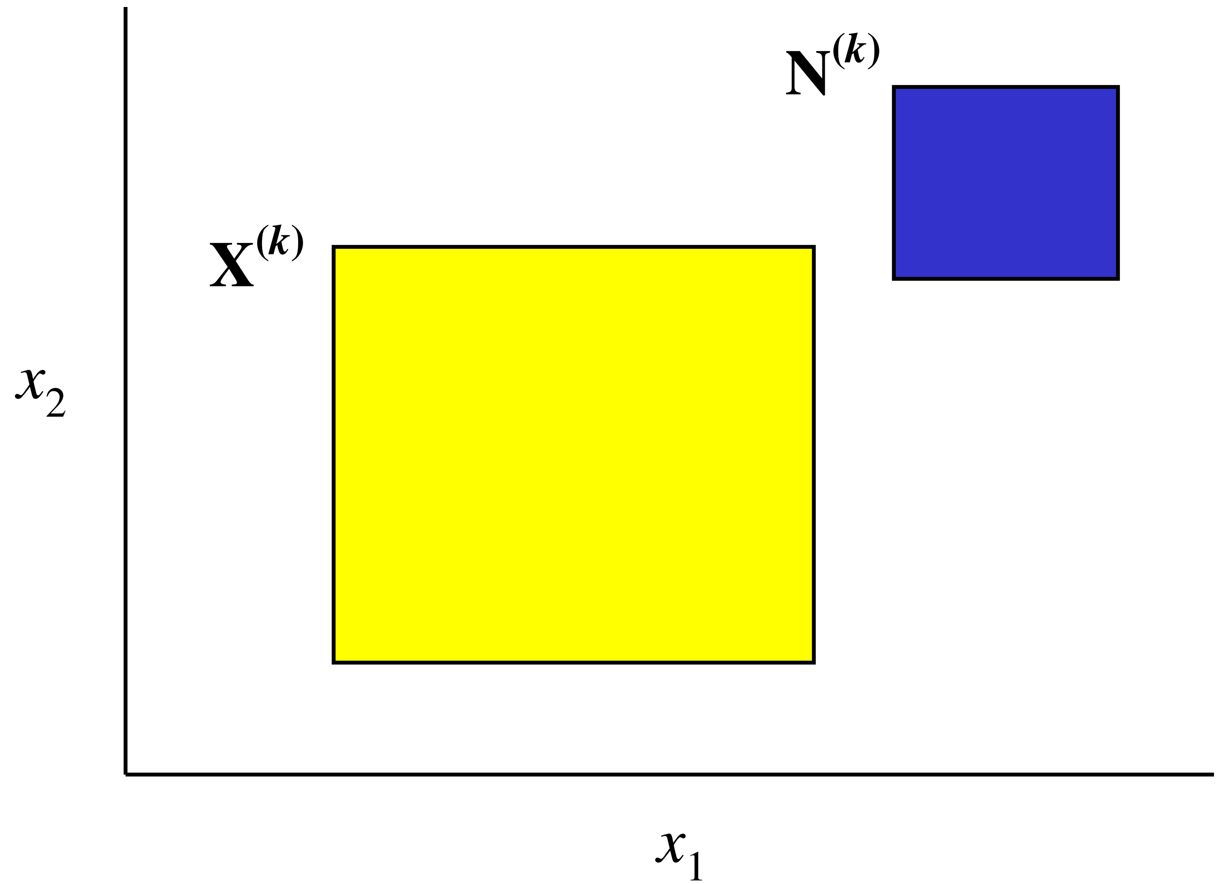
Basic iteration scheme: For a particular subinterval (box), $\mathbf{X}^{(\mathbf{k})}$, perform root inclusion test:

- Range test: Compute an interval extension (bounds on range) for each function in the system: $\mathbf{F}(\mathbf{X}^{(k)})$
  - If $\mathbf{0}$ is not in $\mathbf{F}(\mathbf{X}^{(k)})$, delete the box
- Interval Newton test: Compute the image, $\mathbf{N}^{(k)}$, of the box by solving the linear interval equation system

$$F'(\mathbf{X}^{(k)})(\mathbf{N}^{(k)} - \boldsymbol{x}^{(k)}) = -\boldsymbol{f}(\boldsymbol{x}^{(k)})$$

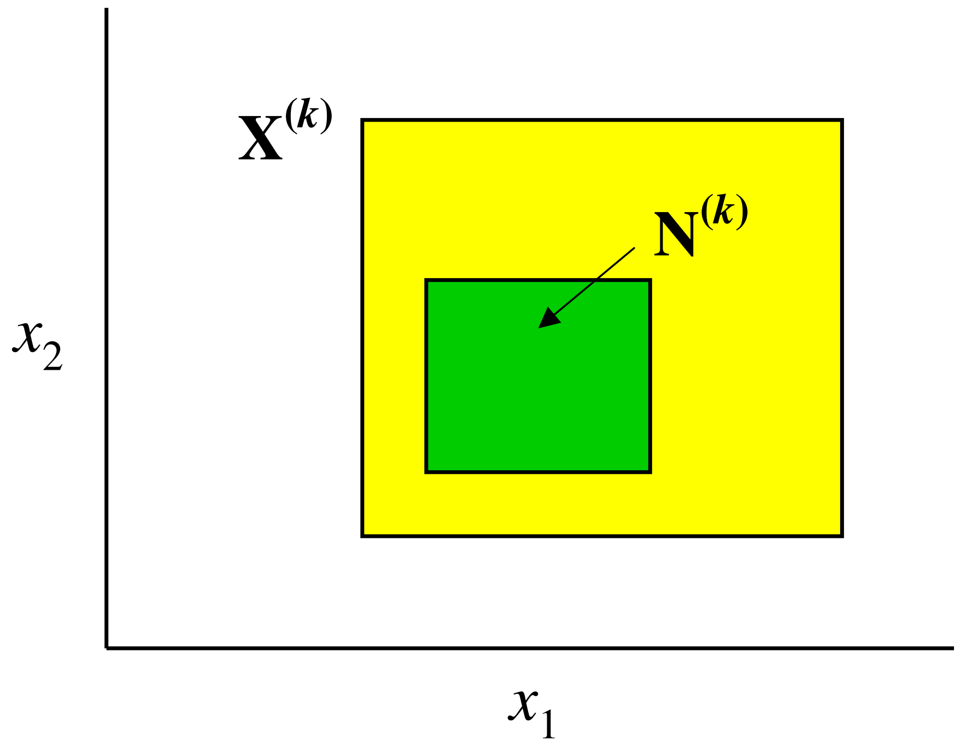
  - $\boldsymbol{x}^{(k)}$ is a point in $\mathbf{X}^{(k)}$
  - $F'(\mathbf{X}^{(k)})$ is the interval extension of the Jacobian matrix of $\boldsymbol{f}(\boldsymbol{x})$ over the interval $\mathbf{X}^{(k)}$

INGB Method: Interval Newton Test



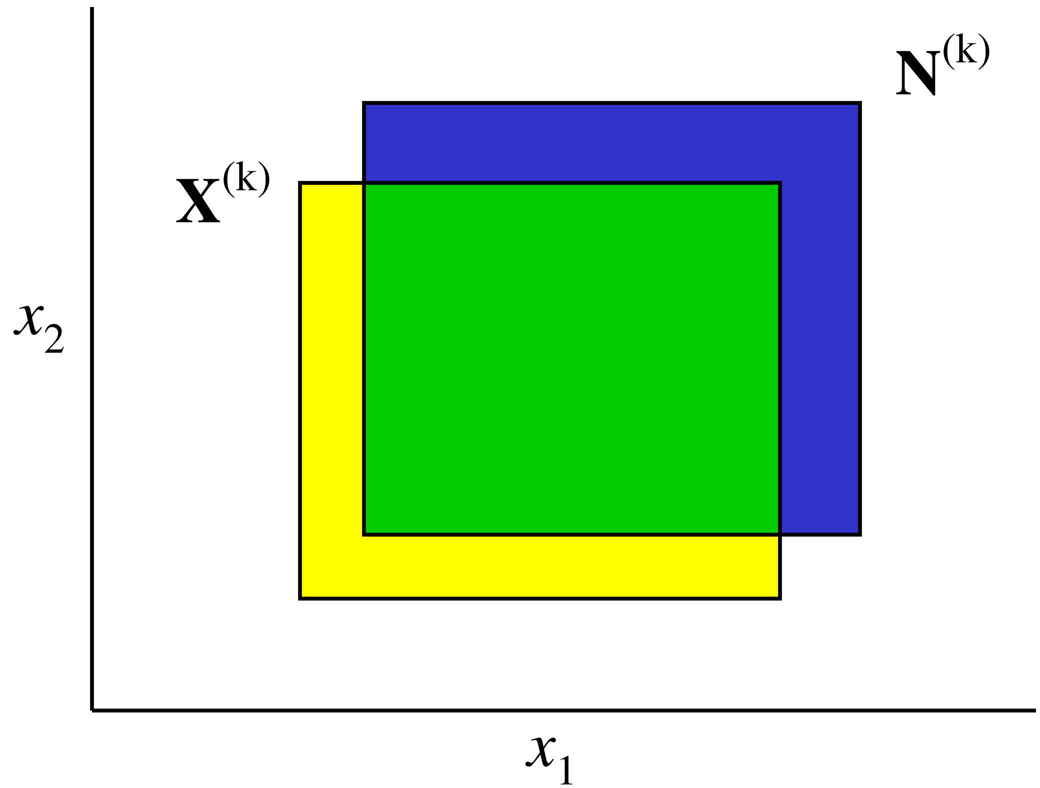
- There is no solution in $\mathbf{X}^{(k)}$

INGB Method: Interval Newton Test



- There is a unique solution in $\mathbf{X}^{(k)}$ that is also in $\mathbf{N}^{(k)}$
- Additional interval-Newton steps will tightly enclose the solution with quadratic convergence

INGB Method: Interval Newton Test



- Any solutions in $\mathbf{X}^{(k)}$ are in $\mathbf{X}^{(k)} \cap \mathbf{N}^{(k)}$
- If the intersection is sufficiently small, repeat the root inclusion test
- Otherwise, bisect the intersection and apply the root inclusion test to each resulting subinterval

New bisection rule on integer variables

- In a bisection for an integer variable, “cut-off” unnecessary real regions
- For example, if the interval for y is currently [0, 5], and it is bisected, the result will be subintervals [0, 2] and [3, 5] (not [0, 2.5] and [2.5, 5])
- Similarly, if the interval for y is currently [0, 1], and it is bisected, the result will be subintervals [0, 0] and [1, 1]
- In final results, integer variables will be represented by degenerate (tight) intervals, e.g. [1, 1]

INGB with interval branch-and-bound (INGB-BB)

- For a given candidate interval, compute interval bounds on objective function
- If lower bound is greater than known upper bound on global minimum, then delete the candidate interval
- Update of upper bound on the global minimum can be done in several ways, e.g.,
  - Use local NLP solver at intervals for which the integer variables are tight
- Adds computational overhead (computation of bounds on objective and update of upper bound), thus INGB-BB may or may not be more efficient than INGB alone

Some pros and cons

- Easy to apply – just solve KKT conditions
- No need for problem reformulations
- Potential to find all KKT points
- However, the need to solve for Lagrange multipliers means increase in problem dimensionality (but bisection on such variables can have lower priority than real and integer variables)

Examples:

- Six small examples from *Adjiman, et. al., 2000, AIChE J.* were selected for test (DELL 2GHz)

Example	# of roots	Global minimizer	CPU (seconds) No B&B	CPU (seconds) B&B
1	7	x=(1.12,1.31), y=(0,1,1)	0.109	0.016
2	2	x=(0.94194,-2.1), y=1	5.531	1.234
3	7	x=(0.2,0.8,1.908), y=(1,1,0,1)	852.844	29.925
4	81	x=(0.97,0.9925,0.98) y=(0,1,1,1,0,1,1,0)	5.266	1.376
5	8	y=(3,1)	0.734	0.234
6	5	x=4, y=1	0.547	0.359

Correct results were obtained in all cases.

Application: VLE Phase Stability Calculation

- Asymmetric modeling for VLE calculation
  - EOS model for vapor phase
  - Activity coefficient model for liquid phase
- Objective function in phase stability analysis (tangent plane distance) has slope discontinuity
- By introducing a binary variable the discontinuity is eliminated, but the VLE phase stability calculation becomes MINLP

VLE Phase Stability Calculation with Asymmetric Model

- Tangent plane distance analysis: minimize $D(x)$
  - $D(x) = g(x) - g_0 - \nabla g(x_0) \bullet (x - x_0)$
- For asymmetric model:
  - $D(x) = \min\{D^V(x), D^L(x)\}$;
- Make $D(x)$ continuous by introducing a binary variable θ,
  - $D(x) = \theta D^V(x) + (1 - \theta) D^L(x)$;
  - θ is then determined as part of the optimization problem (MINLP)
- See G. Xu, W. D. Haynes, M. A. Stadtherr, *Fluid Phase Equilbria*, 235, 152-165 (2005)

Concluding Remarks:

- We have explored the use of INGB for solving MINLP problem;
- Use of branch-and-bound (INGB-BB) improves performance on example problems;
- There are many ways to improve the efficiency of the current algorithm, e.g., constraint propagation using interval Taylor models (Lin & Stadtherr, paper #89b, Monday 12:48pm);
- Future work may be targeted to medium and larger scale mixed-integer nonlinear problems, such as reactor networks, pump networks, heat-exchange networks, etc.

Acknowledgement:

- Invensys/SimSci-Esscor
- Indiana 21st Century Research and Technology Fund

Example 1: Kocis and Grossmann (1988)

$$
\begin{aligned}
\min_{x,y} \quad & 2x_1 + 3x_2 + 1.5y_1 + 2y_2 - 0.5y_3 \\
\text{s.t.} \quad & x_1^2 + y_1 = 1.25 \\
& x_2^{1.5} + 1.5y_2 = 3 \\
& x_1 + y_1 \leq 1.6 \\
& 1.333x_2 + y_2 \leq 3 \\
& -y_1 - y_2 + y_3 \leq 0 \\
& x_1, x_2, x_3 \geq 0 \\
& (y_1, y_2, y_3) \in \{0,1\}^3
\end{aligned}
$$

return to table

Example 2: Floudas (1995)

$$\min_{x,y} \quad -0.7y + 5(x_1 - 0.5)^2 + 0.8$$

$$\begin{aligned} \text{s.t.} \quad & -e^{(x_1-0.2)} - x_2 \le 0 \\ & x_2 + 1.1y \le -1 \\ & x_1 - 1.2y \le 0.2 \\ & 0.2 \le x_1 \le 1 \\ & -2.22554 \le x_2 \le 1 \\ & y \in \{0,1\} \end{aligned}$$

return to table

Example 3: Yuan et. al. (1988)

$$\min_{x,y} \; (y_1-1)^2+(y_2-2)^2+(y_3-1)^2-\ln(y_4+1)$$
$$+(x_1-1)^2+(x_2-2)^2+(x_3-3)^2$$

s.t.

$$y_1+y_2+y_3+x_1+x_2+x_3 \le 5$$
$$y_3^2+x_1^2+x_2^2+x_3^2 \le 5.5$$
$$y_1+x_1 \le 1.2 \qquad y_2^2+x_2^2 \le 1.64$$
$$y_2+x_2 \le 1.8 \qquad y_3^2+x_3^2 \le 4.25$$
$$y_3+x_3 \le 2.5 \qquad y_2^2+x_3^2 \le 4.64$$
$$y_4+x_1 \le 1.2 \qquad x_1,x_2,x_3 \ge 0$$
$$y \in \{0,1\}^4$$

return to table

Example 4: Berman and Ashrafi (1993)

$$\begin{aligned}
\min_{x,y} \quad & -x_1 x_2 x_3 \\
\text{s.t.} \quad & x_1 + 0.1^{y_1} 0.2^{y_2} 0.15^{y_3} = 1 \\
& x_2 + 0.05^{y_4} 0.2^{y_5} 0.15^{y_6} = 1 \\
& x_3 + 0.02^{y_7} 0.06^{y_8} = 1 \\
& -y_1 - y_2 - y_3 \leq -1 \\
& -y_4 - y_5 - y_6 \leq -1 \\
& -y_7 - y_8 \leq -1 \\
& 3y_1 + y_2 + 2y_3 + 3y_4 + 2y_5 + y_6 + 3y_7 + 2y_8 \leq 10 \\
& 0 \leq x_1, x_2, x_3 \leq 1 \\
& y \in \{0,1\}^8
\end{aligned}$$

return to table

Example 5: Pörn et al. (1997)

$$
\begin{aligned}
\min_{x,\,y} \quad & 7y_1 + 10y_2 \\
\text{s.t.} \quad & y_1^{1.2} y_2^{1.7} - 7y_1 - 9y_2 \le -24 \\
& -y_1 - 2y_2 \le -5 \\
& -3y_1 + y_2 \le 1 \\
& 4y_1 - 3y_2 \le 11 \\
& y_1, y_2 \in [1,5] \cap \mathrm{N}
\end{aligned}
$$

return to table

Example 6: Pörn et al. (1997)

$$
\begin{aligned}
\min_{x,y} \quad & 3y - 5x \\
\text{s.t.} \quad & 2y^2 - 2y^{0.5} - 2x^{0.5}y^2 + 11y + 8x \leq 39 \\
& -y + x \leq 3 \\
& 2y + 3x \leq 24 \\
& 1 \leq x \leq 10 \\
& y \in [1,6] \cap \mathrm{N}
\end{aligned}
$$

return to table