A Brief Introduction to Mixed-Integer Nonlinear Programming (MINLP)

Eduardo Camponogara

Department of Automation and Systems Engineering
Federal University of Santa Catarina

October 2016

Continuous Optimization

Mixed-Integer Optimization

Mixed-Integer Nonlinear Optimization

Global Optimization Strategy

Summary

Continuous Optimization

Mixed-Integer Optimization

Mixed-Integer Nonlinear Optimization

Global Optimization Strategy

Linear Programming

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} : \; & Ax \leq b \\ & x \geq 0 \end{aligned}$$

Remarks:

- Solvable in polynomial-time.
- Efficient algorithms: SIMPLEX and primal-dual interior-point method.

Linear Programming

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.}: \quad & Ax \leq b \\
& x \geq 0
\end{aligned}
$$

Remarks:

- Solvable in polynomial-time.
- Efficient algorithms: SIMPLEX and primal-dual interior-point method.

Quadratic Programming

$$\begin{aligned} \min \quad & x^T Q x + c^T x \\ \text{s.t.}: \ & Ax \leq b \end{aligned}$$

in which $Q \succ 0$.

Remarks:

- Solved efficiently with active-set and interior-point methods.

Quadratic Programming

$$\begin{aligned} \min \quad & x^T Q x + c^T x \\ \text{s.t.}: \ & Ax \leq b \end{aligned}$$

in which $Q \succ 0$.

Remarks:

- Solved efficiently with active-set and interior-point methods.

Convex Optimization

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.}: \; & Ax = b \\ & f_j(x) \leq 0,\; j = 1, \ldots, m \end{aligned}$$

in which $f_j : \mathbb{R}^n \to \mathbb{R}, j = 0, \ldots, m,$ are convex functions.

Remarks:

- Solved efficiently with interior-point method and Newton's algorithm.

Convex Optimization

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.}: \ & Ax = b \\ & f_j(x) \leq 0, \ j = 1, \dots, m \end{aligned}$$

in which $f_j : \mathbb{R}^n \to \mathbb{R}, j = 0, \dots, m$, are convex functions.

Remarks:

- Solved efficiently with interior-point method and Newton's algorithm.

Properties of Convex Programs

- Any local solution x induces a global minimum.
- Any solution that satisfies first-order KKT optimality conditions is a local minimum, and therefore a global minimum.
- Several classes of problems can be shown to be convex.

Summary

Continuous Optimization

Mixed-Integer Optimization

Mixed-Integer Nonlinear Optimization

Global Optimization Strategy

Mixed-Integer Linear Programming

$$
\begin{aligned}
\min \quad & c_{\mathrm{x}}^T x + c_{\mathrm{y}}^T y \\
\text{s.t.}: \; & A_{\mathrm{x}} x + A_{\mathrm{y}} y \leq b \\
& x \geq 0,\; y \text{ is integer}
\end{aligned}
$$

Remarks:

- NP-Hard problem.
- Can be solved in practice with branch-and-bound, branch-and-cut, and other strategies.

Mixed-Integer Linear Programming

$$\begin{aligned} \min \quad & c_{\mathrm{x}}^T x + c_{\mathrm{y}}^T y \\ \text{s.t.} : \; & A_{\mathrm{x}} x + A_{\mathrm{y}} y \leq b \\ & x \geq 0, \; y \text{ is integer} \end{aligned}$$

Remarks:

- NP-Hard problem.
- Can be solved in practice with branch-and-bound, branch-and-cut, and other strategies.

Mixed-Integer Convex Programming (MICP)

$$
\begin{aligned}
\min \quad & f_0(x) \\
\mathrm{s.t.}: \quad & Ax = b \\
& f_j(x) \le 0,\ j = 1, \ldots, m
\end{aligned}
$$

in which $x = (x_{\mathrm{C}}, x_{\mathrm{I}})$ and x_{I} is integer.

Mixed-Integer Convex Optimization

Remarks:

- MILP and MICP are conceptually the same, in the sense that lower bounds can be obtained in polynomial-time.
- If we relax the integrality constraint, MILP becomes an LP and MICP becomes a convex program, both of which can be efficiently computed.

Implications:

- Branch-and-bound can be applied to MICP, much like the way it is applied to MILP.
- Of course, MICP will be much harder and the effectiveness will depend on how fast bounds are computed.

Mixed-Integer Convex Optimization

Remarks:

- MILP and MICP are conceptually the same, in the sense that lower bounds can be obtained in polynomial-time.
- If we relax the integrality constraint, MILP becomes an LP and MICP becomes a convex program, both of which can be efficiently computed.

Implications:

- Branch-and-bound can be applied to MICP, much like the way it is applied to MILP.
- Of course, MICP will be much harder and the effectiveness will depend on how fast bounds are computed.

Mixed-Integer Convex Optimization

Difficulties:

- For MILP, the dual Simplex algorithm provides hot start optimization from infeasible solutions produced by branch-and-bound.
  - Introducing a bound on a variable renders the current dictionary infeasible in the primal, but suboptimal in the dual.
  - a constraint in the primal becomes a variable in the dual.
- Such hot start mechanisms are not generally available for convex problems.

Mixed-Integer Convex Optimization

Difficulties:

- For MILP, the dual Simplex algorithm provides hot start optimization from infeasible solutions produced by branch-and-bound.
  - Introducing a bound on a variable renders the current dictionary infeasible in the primal, but suboptimal in the dual.
  - a constraint in the primal becomes a variable in the dual.
- Such hot start mechanisms are not generally available for convex problems.

Mixed-Integer Convex Optimization

Much of the literature explores the problem structure:

- Bonmin is a general solver for MICP.
- CPLEX can handle Mixed-Integer Quadratically Constrained Programs, in which the quadratic functions are convex.

Summary

Continuous Optimization

Mixed-Integer Optimization

Mixed-Integer Nonlinear Optimization

Global Optimization Strategy

Mixed-Integer Nonlinear Optimization

The general MINLP is a problem of the form:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.}: \; & g_j(x) \leq 0, \; j = 1, \ldots, p \\ & h_j(x) = 0, \; j = 1, \ldots, q \end{aligned}$$

in which $x = (x_{\mathrm{C}}, x_{\mathrm{I}})$ and x_{I} is integer.

Remarks:

- Very hard problem, for which there does not exist general methods.
- Can be solved for particular problems, and particular sizes.

Mixed-Integer Nonlinear Optimization

The general MINLP is a problem of the form:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.}: \ & g_j(x) \leq 0, \ j = 1, \ldots, p \\ & h_j(x) = 0, \ j = 1, \ldots, q \end{aligned}$$

in which $x = (x_{\mathrm{C}}, x_{\mathrm{I}})$ and x_{I} is integer.

Remarks:

- Very hard problem, for which there does not exist general methods.
- Can be solved for particular problems, and particular sizes.

Mixed-Integer Nonlinear Optimization

What will we do here?

- We address the general class in which the nonlinear terms are bilinear.
- For instance, $x \cdot y$, x^2, *etc.*
- Several applications are found in theory and practice.

Mixed-Integer Nonlinear Optimization

What will we do here?

- We address the general class in which the nonlinear terms are bilinear.
- For instance, $x \cdot y$, x^2, *etc.*
- Several applications are found in theory and practice.

Bilinear Optimization: Polynomials

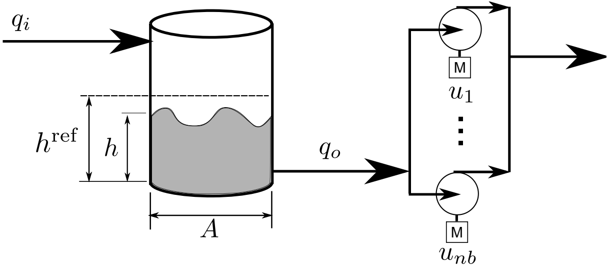
- Polynomials can be converted into bilinear models.
- For instance,

$$f(x_1, x_2, x_3) = x_1^2 x_2 x_3$$

can be converted to

$$\begin{aligned} f &= w_{1,1} \cdot w_{2,3} \\ w_{1,1} &= x_1^2 \\ w_{2,3} &= x_2 x_3 \end{aligned}$$

Bilinear Optimization: Blending



Bilinear Optimization: Blending

For stable operations, the ratio of the "qualities" are held constant:

$$\frac{q_i}{q} = \frac{\Delta q_i}{\Delta q} \iff q_i \cdot \Delta q = \Delta q_i \cdot q$$

in which:

- q is the volume of the mixture in the tank.
- q_i is the volume of a product i in the mixture.
- Δq is the volume being pumped out.
- Δq_i is the volume of quality i being pumped out.

Bilinear Optimization: Blending

For stable operations, the ratio of the "qualities" are held constant:

$$\frac{q_i}{q} = \frac{\Delta q_i}{\Delta q} \iff q_i \cdot \Delta q = \Delta q_i \cdot q$$

in which:

- q is the volume of the mixture in the tank.
- q_i is the volume of a product i in the mixture.
- Δq is the volume being pumped out.
- Δq_i is the volume of quality i being pumped out.

Bilinear Optimization: Modeling Trigonometric Functions

Polar coordinates,

$$x = r\cos(\theta)$$
$$y = r\sin(\theta)$$

can be represented in Cartesian form

$$x^2 + y^2 = r^2$$

without the need of trigonometric functions.

McCormick Envelopes

Consider the bilinear function

$$w = xy$$

and suppose that bounds are known:

$$x^L \leq x \leq x^U$$
$$y^L \leq y \leq y^U$$

McCormick Envelopes

First we manipulate the bounds

$$x - x^L \geq 0 \qquad y - y^L \geq 0$$
$$x^U - x \geq 0 \qquad y^U - y \geq 0$$

Mulitplying the first of each bound, we obtain

$$(x - x^L)(y - y^L) \geq 0 \iff xy - xy^L - yx^L + x^L y^L \geq 0$$
$$\iff w \geq xy^L + yx^L - x^L y^L$$

which is a linear inequality.

McCormick Envelopes

First we manipulate the bounds

$$x - x^L \geq 0 \qquad y - y^L \geq 0$$
$$x^U - x \geq 0 \qquad y^U - y \geq 0$$

Mulitplying the first of each bound, we obtain

$$\begin{aligned}(x - x^L)(y - y^L) \geq 0 \iff & xy - xy^L - yx^L + x^L y^L \geq 0 \\ \iff & w \geq xy^L + yx^L - x^L y^L\end{aligned}$$

which is a linear inequality.

McCormick Envelopes

- The underestimators of the function are represented by:

$$w \geq x^L y + x y^L - x^L y^L$$
$$w \geq x^U y + x y^U - x^U y^U$$

- The overestimators of the function are represented by:

$$w \leq x^U y + x y^L - x^U y^L$$
$$w \leq x y^U + x^L y - x^L y^U$$

McCormick Envelopes

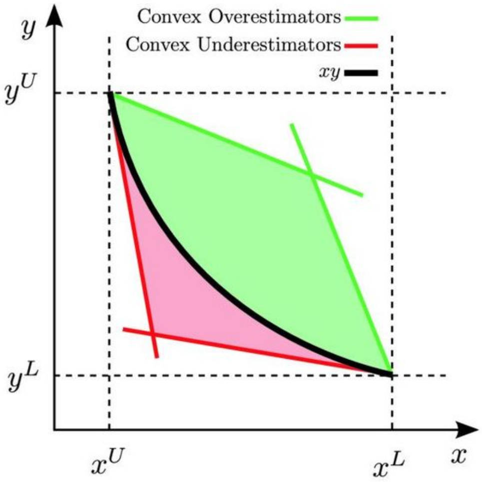
- The underestimators of the function are represented by:

$$w \geq x^L y + x y^L - x^L y^L$$
$$w \geq x^U y + x y^U - x^U y^U$$

- The overestimators of the function are represented by:

$$w \leq x^U y + x y^L - x^U y^L$$
$$w \leq x y^U + x^L y - x^L y^U$$

McCormick Envelopes



McCormick Envelopes

What do we gain from McCormick Envelopes?

Consider an NLP with only bilinear terms.

- Replacing the bilinear terms with McCormick Envelopes converts the NLP into an LP.
- The LP solution gives us a lower bound.
- Any feasible solution gives us an upper bound.

McCormick Envelopes

For an MINLP with only bilinear terms.

- Replacing the bilinear terms with McCormick Envelopes renders MINLP an MILP.
- The MILP solution gives us a lower bound.
- Any feasible solution gives us an upper bound.

McCormick Envelopes

What if we improve the bounds iteratively?

- We can produce a series of increasing lower bounds: $w^{L,k}$, $k = 0, 1, \ldots$
- And a series of decreasing upper bounds: $w^{U,k}$, $k = 0, 1, \ldots$

so that:

$$w^{L,0} \leq w^{L,1} \leq \cdots \leq w^{L,k}$$
$$w^{U,0} \geq w^{U,1} \geq \cdots \geq w^{U,k}$$

and we stop when $w^{U,k} - w^{L,k} \leq \epsilon$.

McCormick Envelopes

What if we improve the bounds iteratively?

- We can produce a series of increasing lower bounds: $w^{L,k}$, $k = 0, 1, \ldots$
- And a series of decreasing upper bounds: $w^{U,k}$, $k = 0, 1, \ldots$

so that:

$$w^{L,0} \leq w^{L,1} \leq \cdots \leq w^{L,k}$$
$$w^{U,0} \geq w^{U,1} \geq \cdots \geq w^{U,k}$$

and we stop when $w^{U,k} - w^{L,k} \leq \epsilon$.

McCormick Envelopes

What if we improve the bounds iteratively?

- We can produce a series of increasing lower bounds: $w^{L,k}$, $k = 0, 1, \dots$
- And a series of decreasing upper bounds: $w^{U,k}$, $k = 0, 1, \dots$

so that:

$$w^{L,0} \leq w^{L,1} \leq \cdots \leq w^{L,k}$$
$$w^{U,0} \geq w^{U,1} \geq \cdots \geq w^{U,k}$$

and we stop when $w^{U,k} - w^{L,k} \leq \epsilon$.

McCormick Envelopes

How can we improve the bounds using McCormick Envelopes?

- We can splice the domain of the variables and impose McCormick Envelopes in each subdomain of the domain.

McCormick Envelopes

How can we improve the bounds using McCormick Envelopes?

- We can splice the domain of the variables and impose McCormick Envelopes in each subdomain of the domain.

Summary

Continuous Optimization

Mixed-Integer Optimization

Mixed-Integer Nonlinear Optimization

Global Optimization Strategy

Problem Definition

$$
\begin{aligned}
\min \quad & f_0(x) \\
\text{s.t. :} \quad & f_q(x) \leq 0,\ q \in \mathcal{Q} \setminus \{0\} \\
& f_q(x) = \sum_{(i,j) \in \mathcal{BL}} a_{i,j,q} x_i x_j + h_q(x),\ q \in \mathcal{Q} \\
& x^L \leq x \leq x^U \\
& x \in \mathbb{R}^m
\end{aligned}
$$

McCormick Envelopes

- In case of an NLP, an LP relaxation is derived by replacing each bilinear term with a new variable:

$$w_{ij} = x_i \cdot x_j$$

and adding four sets of constraints.

- In case of an MINLP, an MILP relaxation is derived.

This strategy is known as McCormick relaxation.

McCormick Envelopes

- In case of an NLP, an LP relaxation is derived by replacing each bilinear term with a new variable:

$$w_{ij} = x_i \cdot x_j$$

  and adding four sets of constraints.
- In case of an MINLP, an MILP relaxation is derived.

This strategy is known as McCormick relaxation.

Piecewise McCormick Relaxation

In either case (NLP or MINLP), a tighter MILP relaxation can be obtained as follows:

- Partition the domain of one of the variables, let us say x_j of the bilinear term into n disjoint regions.
- Add new binary variables to the formulation to select the optimal partition x_j.

This strategy is known as piecewise McCormick relaxation.

Piecewise McCormick Relaxation

In either case (NLP or MINLP), a tighter MILP relaxation can be obtained as follows:

- Partition the domain of one of the variables, let us say x_j of the bilinear term into n disjoint regions.
- Add new binary variables to the formulation to select the optimal partition x_j.

This strategy is known as piecewise McCormick relaxation.

Piecewise McCormick Relaxation

- Let x_{jn}^{L} and x_{jn}^{U} be lower and upper bounds for x_j in partition n.
- If the value of x_j does belong to such a partition, then the binary variable $y_{jn} = 1$ and the McCormick envelope hold.
- The piecewise McCormick relaxation can be formulated as a Generalized Disjunctive Program (Raman and Grossmann, 1994).
- PR-GDP is tighter because the use of partition-dependent x_{jn}^{L} and x_{jn}^{U} inside the disjunction, instead of the global bounds x_j^{L} and x_j^{U}.

Piecewise McCormick Relaxation

- Let x_{jn}^L and x_{jn}^U be lower and upper bounds for x_j in partition n.
- If the value of x_j does belong to such a partition, then the binary variable $y_{jn} = 1$ and the McCormick envelope hold.
- The piecewise McCormick relaxation can be formulated as a Generalized Disjunctive Program (Raman and Grossmann, 1994).
- PR-GDP is tighter because the use of partition-dependent x_{jn}^L and x_{jn}^U inside the disjunction, instead of the global bounds x_j^L and x_j^U.

Formulation

$$\min \quad z^R = f_0(x) = \sum_{(i,j)\in\mathcal{BL}} a_{ij0} w_{ij} + h_0(x)$$

$$\text{s.t.}: f_q(x) = \sum_{(i,j)\in\mathcal{BL}} a_{ijq} w_{ij} + h_q(x) \le 0,\ q \in \mathcal{Q} \setminus \{0\}$$

$$\forall j, (i,j) \in \mathcal{BL}:$$

$$\bigvee_{n=1}^{N} \left[\begin{array}{c} y_{jn} \\ \left.\begin{array}{l} w_{ij} \ge x_{jn}^L x_i + x_j x_i^L - x_{jn}^L x_i^L \\ w_{ij} \ge x_{jn}^U x_i + x_j x_i^U - x_{jn}^U x_i^U \\ w_{ij} \le x_{jn}^U x_i + x_j x_i^L - x_{jn}^U x_i^L \\ w_{ij} \le x_{jn} x_i^U + x_{jn}^L x_i - x_{jn}^L x_i^U \end{array}\right\} \forall i, (i,j) \in \mathcal{BL} \\ x_{jn}^L \le x_j \le x_{jn}^U \end{array} \right]$$

A Brief Introduction to MINLP Optimization

- End!
- Thank you for your attention.