Integer Nonlinear Optimization

Sven Leyffer & Jeff Linderoth

Mathematics and Computer Division
Argonne National Laboratory

Lawrence Livermore National Laboratory, May 23-25, 2007

Overview

Integer Nonlinear Optimization

Mixed Integer Nonlinear Program (MINLP)

minimize $f(x, y)$ subject to $c(x, y) \leq 0$, and y_i integer

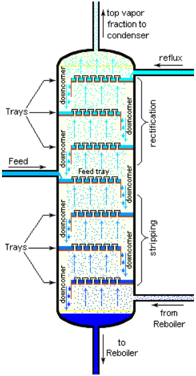
Small process design example:

- synthesis of distillation column

Integer Nonlinear Optimization

Mixed Integer Nonlinear Program (MINLP)

$$\text{minimize } f(x,y) \quad \text{subject to } c(x,y) \le 0, \text{ and } y_i \text{ integer}$$



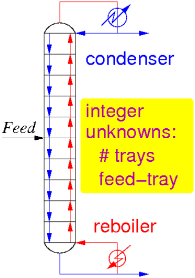
Small process design example:

- synthesis of distillation column
- nonlinear physics: phase equilibrium, component material balance

Integer Nonlinear Optimization

Mixed Integer Nonlinear Program (MINLP)

$$\text{minimize } f(x,y) \quad \text{subject to } c(x,y) \le 0, \text{ and } y_i \text{ integer}$$



Small process design example:

- synthesis of distillation column
- nonlinear physics: phase equilibrium, component material balance
- integers model number of trays in columns
- $y_i \in \{0,1\}$ models position of feeds

Applications of Integer Nonlinear Optimization

Mixed Integer Nonlinear Program (MINLP)

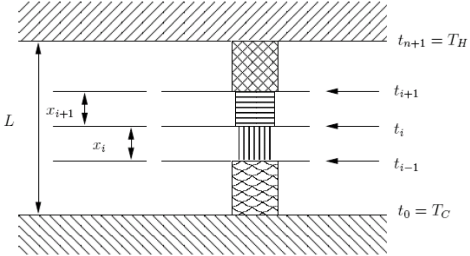
$$\boxed{\underset{x,y}{\text{minimize}} \; f(x,y) \quad \text{subject to } c(x,y) \le 0, \text{ and } y_i \text{ integer}}$$

- process design for FutureGen (zero CO_2-emissions fossil plant)
- radiation therapy treatment planning
- emergency evacuation planning; routing and dispatch
- blackout prevention of national power grid
- nuclear reactor core-reload operation
- design of thermal insulation layer for superconductors

Categorical Variables

Consider discrete optimization problems with categorical variables

- discrete choice, e.g. type of insulator material: m_i from $\mathcal{M} = \{$ nylon, teflon, epoxy, ... $\}$
- non-numerical, discrete variables $\Rightarrow$ no relaxation
- limited to heuristic search techniques



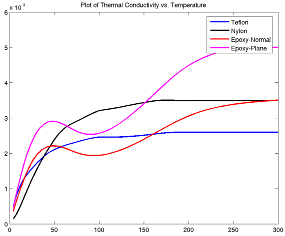
Design of thermal insulation system [Abramson:04]

Modeling Physics of System

Given t_{i-1}, t_i, heat transfer is $q_i = \dfrac{a_i}{x_i} \displaystyle\int_{t_{i-1}}^{t_i} k(t, m_i) dt$

... from Fourier's law where

- $k(t, m_i)$ thermal conductivity of insulator m_i at temperature t
- $k(t, m_i)$ given as tabulated data
- interpolate with cubic splines
- integrate with Simpson's rule
  $\Rightarrow$ consistent with cubic splines



Modeling Categorical Variables m_i

$z_{ij} \in \{0,1\}$ where $z_{ij} = 1 \;\Leftrightarrow\;$ layer i has j^{th} material

$$\sum_{j=1}^{|\mathcal{M}|} z_{ij} = y_i, \quad i = 1, \ldots, N+1.$$

... only existing layers ($y_i = 1$) can choose material

Heat transfer equation with categorical $m_i \in \mathcal{M}$

$$q_i = \frac{a_i}{x_i} \int_{t_{i-1}}^{t_i} k(t, m_i) dt \;\Leftrightarrow\; x_i q_i = a_i \int_{t_{i-1}}^{t_i} \sum_{j=1}^{\mathcal{M}} z_{ij} k(t, m_j) dt$$

Form $\hat{k}(t) := \sum_{j=1}^{\mathcal{M}} z_{ij} k(t, m_j)$ from data look-up

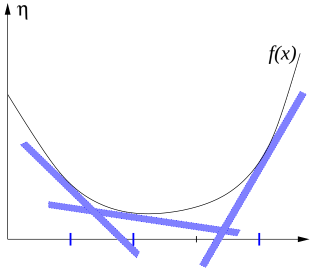
Integrate $\hat{k}(t)$ ... convex combination of materials

Overview

Outer Approximation (Duran & Grossmann, 86)

NLP subproblem y_j fixed:

$$\text{NLP}(y_j) \begin{cases} \min\limits_x & f(x, y_j) \\ \text{s.t.} & c(x, y_j) \le 0 \\ & x \in X \end{cases}$$



linearize f, c about $(x_j, y_j) =: z_j$

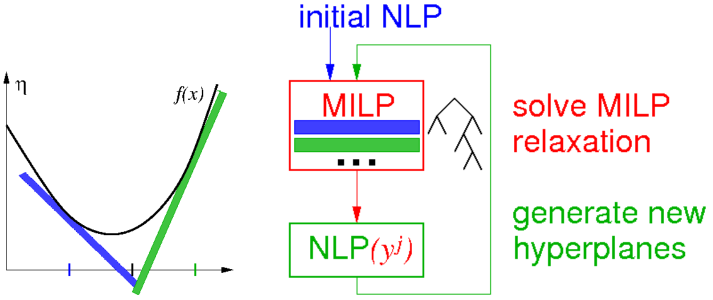
$\Rightarrow$ MINLP $(P) \equiv$ MILP (M)

$$(M) \begin{cases} \underset{z=(x,y),\eta}{\text{minimize}} & \eta \\ \text{subject to} & \eta \ge f_j + \nabla f_j^T (z - z_j) \quad \forall y_j \in Y \\ & 0 \ge c_j + \nabla c_j^T (z - z_j) \quad \forall y_j \in Y \\ & x \in X,\ y \in Y \text{ integer} \end{cases}$$

but need linearizations $\forall y_j \Rightarrow$ solve relaxations of (M)

Outer Approximation (Duran & Grossmann, 86)

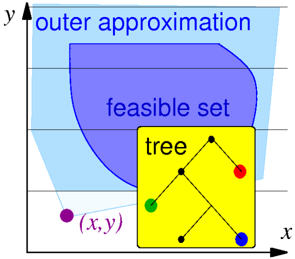
Alternate between solve NLP(y_j) and MILP relaxation



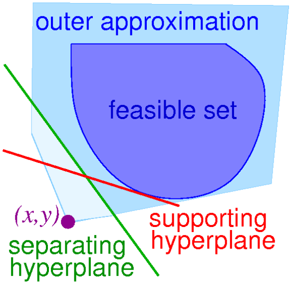
MILP ⇒ lower bound; NLP ⇒ upper bound

... MILP solution is bottleneck ... no hot-starts for MILP

Branch-and-Cut for MINLP (Quesada & Grossmann, 92)



branching +



cutting planes

- interrupt MILP branch-and-cut & add linearizations
  e.g. solve NLP(y_j) $\Rightarrow$ separates y_j ... infeasible
- New Solver: FilterSQP + MINTO = FilMINT [with Linderoth]

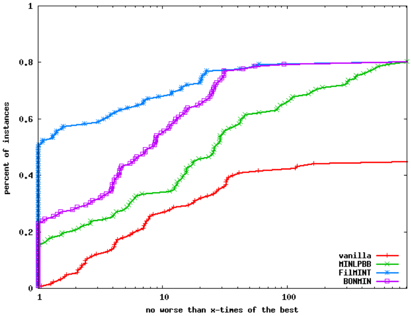
Important MIP Tricks

Important MIP tricks based on numerical experiments:

- diving-based primal heuristic (get good incumbent)
- pseudo-cost branching & adaptive node selection good
- add only violated linearizations to (M)aster
- use MINTO's row management
  ... remove cuts that are inactive for 15 LPs
- range of cuts: extended cutting plane to full NLP
  surprise: ECP is best ... Kelly's cutting plane method!!!

... from extensive runs on 120+ MINLPs [Kumar Abhishek, 2007]

Compare to MINLP-BB & BONMIN [IBM/CMU]



Performance profile

- fraction of problems solved within factor x of best solver
- time-limit: 4 desktop-hours

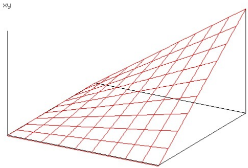
Overview

Nonconvex MINLPs

Nonconvex functions $f(x, y)$ or $c(x, y)$ add layer of complexity
$\Rightarrow$ linearizations $\neq$ outer approximations

Baron: convex envelopes

- bilinear terms $x \cdot y$
- convex & concave envelope
- McCormick; Sahinidis & Tawarmalani



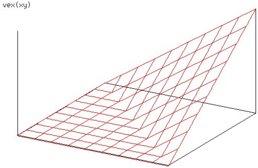
Alternative: piecewise linear approximation (Martin, 2004)
... special-ordered sets, use automatic-differentiation?

Nonconvex MINLPs

Nonconvex functions $f(x, y)$ or $c(x, y)$ add layer of complexity
$\Rightarrow$ linearizations $\neq$ outer approximations

Baron: convex envelopes

- bilinear terms $x \cdot y$
- convex & concave envelope
- McCormick; Sahinidis & Tawarmalani



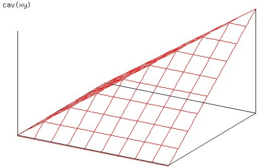
Alternative: piecewise linear approximation (Martin, 2004)
... special-ordered sets, use automatic-differentiation?

Nonconvex MINLPs

Nonconvex functions $f(x, y)$ or $c(x, y)$ add layer of complexity
$\Rightarrow$ linearizations $\neq$ outer approximations

Baron: convex envelopes

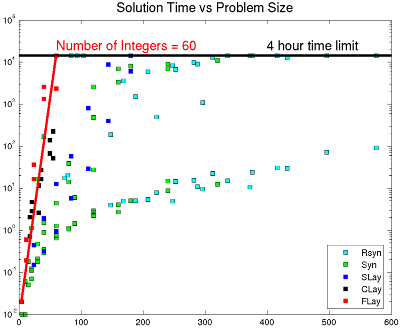
- bilinear terms $x \cdot y$
- convex & concave envelope
- McCormick; Sahinidis & Tawarmalani



Alternative: piecewise linear approximation (Martin, 2004)
... special-ordered sets, use automatic-differentiation?

The Curse of Exponentiality

Integer optimization has exponential complexity



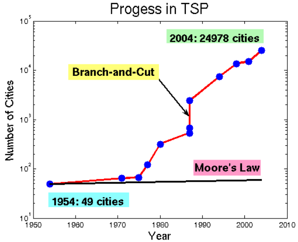
Time vs number of integers

Parallel MINLP

- 100s of processors get 80% efficiency
- 100,000 processors . . . research issues
- perfect speed-up only doubles problem size

Parallel computing alone not enough: need new methods!

The Curse of Exponentiality



Traveling Salesman Problem (TSP):

- shortest route through n cities; complexity $(n-1)!/2$
- applications: transportation, genome-sequencing, ...
- benchmark problem

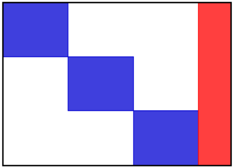
Historical progress on TSP:

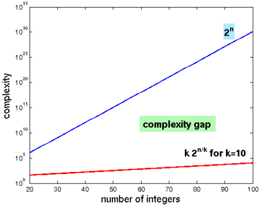
- 1954: Dantzig solves 54 cities problem
- 2004: Applegate et al. solve 25k cities in 84 CPU years

... projected increase from Moore's law: only 6 cities
Impact of 100,000 node BG with perfect speedup:
25k cities in 30 hours; 27k cities in a week ... pitiful

Decomposition for MINLP

- network interdiction under uncertainty
- scenario-approach leads to large problem
- decompose into small subproblems
- coordinate linking variables in master





Decomposition for integer optimization:

- 2^n versus $k \cdot 2^{n/k}$ ... huge gap
- more complex master problem
- Chen vs IBM's feasibility pump:

	TR-7		TR-12	
	UBD	time	UBD	time
Chen	26.7	60	138.8	324
IBM	27.5	390	-	7200

Simulation-Based MINLP

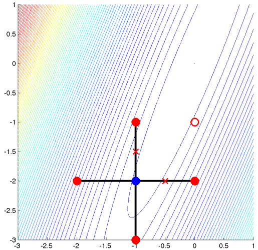
Many DOE applications are simulation-based

- min $f(x, y)$ where function evaluated from simulation
- no explicit functional form for $f(x, y) \Rightarrow$ no gradients

... pattern-search methods can be used

Pattern-search fails:

- search "discrete neighbors"
- fails for convex quadratic



Simulation-Based MINLP

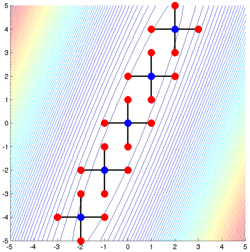
Many DOE applications are simulation-based

- min $f(x, y)$ where function evaluated from simulation
- no explicit functional form for $f(x, y) \Rightarrow$ no gradients

... pattern-search methods can be used

Pattern-search fails:

- search "discrete neighbors"
- fails for convex quadratic



Simulation-Based MINLP

Alternative: model-based optimization

- minimize model $m_k(x, y) \approx f(x, y)$ in trust-region
- models based on quadratic interpolation (UOBYQA et al.)
- models based on radial-basis functions (ORBIT, Shoemaker)
- more problem information: model defined everywhere

Open questions:

- lower bounds available
  $\Rightarrow$ convergence for convex/quadratic MINLP?
- how to include constraints $c(x, y) \leq 0$?
- efficient implementation for MINLP?

Conclusions & Outlook

Optimization is becoming more important as we
move from simulation to design of complex system.

Discrete design choices create new challenges:

- theory & solvers for integer nonlinear optimization
  - rigorous decomposition methods for MINLP
  - global solution of nonconvex MINLP
  - derivative-free (simulation-based) MINLP
- support broad range scientific & engineering applications
  - modeling challenges of discrete choices
  - modeling challenges of nonlinearities
- apply to complex systems (CO_2, fossil plants, ...)
- opportunities to leverage parallel resources