Workshop on Integer Programming and Continuous Optimization
Chemnitz, November 2004

Combinatorial Structures in Nonlinear Programming

Stefan Scholtes

Judge Institute of Management
University of Cambridge

Background

NLP community is primarily interested in **local** optimization

- Generic optimization framework for large problems
- No convexity guarantee necessary - but works well if problems are convex

Non-convexity in optimization problems arises from

- mixed curvature in functions, e.g. polynomials, trigonometric, etc.
- combinatorial selections, e.g. min/max, median, etc.

Little work to date on **local optimization** of NLPs with combinatorial constraints

- Local optimization doesn't make sense for discrete problems (e.g. Integer Programming)
- It does make sense for connected constraint sets that have an exploitable combinatorial description

Outline of the Talk

Motivation

$\longrightarrow$ Examples of combinatorial constraints

Traditional Approaches

$\longrightarrow$ Reformulation as MIP

$\longrightarrow$ Reformulation as NLP

Theoretical Framework

$\longrightarrow$ Primal-dual framework for NLPs with combinatorial constraints

Algorithmic Framework

$\longrightarrow$ SQP for NLPs with combinatorial constraints

Conclusion and Outlook

What is a combinatorial constraint?

Traditional NLP constraint: $g(x) \geq 0$, $h(x) = 0$

Combinatorial constraint involves a selection process from a finite (possibly very large) number of constraint candidates

Examples:

Complementarity constraints: $0 \leq g(x) \perp h(x) \geq 0$

Exclusion constraints: $x \notin X_i = \{y | g_j(y) > 0,\ j \in I_i\}$

$$\Leftrightarrow \min_{j \in I_i} g_j(x) \leq 0$$

$\rightarrow$ More general: Finite max-min constraints

Conditional constraints: If $g(x) > 0$ then $h(x) \geq 0$

$\rightarrow$ More general: Boolean formulas, involving nonlinear constraints as boolean arguments

Combinations thereof...

Example: Robust Optimization

Two-stage stochastic program - Minimize expected cost

$$\min_{x \in X} \sum_i p_i \min_{y \in Y_i} f(x, y_i)$$

Robust counterpart - Minimize maximal cost under all scenarios

$$\min_{x \in X} \max_i \min_{y \in Y_i} f(x, y_i)$$

Finite recourse: $\min_{x \in X} \max_i \min_{j \in I_i} f_j(x)$

NLP with combinatorial constraints:

$$\begin{aligned} \min_{(x,\alpha)} \quad & \alpha \\ \text{s.t.} \quad & \min_{j \in I_i} f_j(x) \le \alpha, \forall i \end{aligned}$$

Traditional Approaches: MIP reformulation

Example: $\min\{g(x), h(x)\} = 0$

Big-M reformulation:

$$\begin{aligned} 0 \le g(x) &\le M(1-y) \\ 0 \le h(x) &\le My \\ y &\in \{0,1\} \end{aligned}$$

Problem: May have artificially disconnected a connected set within which a descent algorithm might be able to make good progress

$\longrightarrow$ May have introduced unnecessary combinatorial curse

$\longrightarrow$ Inefficient for local optimization of large complementarity constrained problems

Traditional Approaches: NLP reformulation

Whitney's theorem: Every closed set is the set of roots of a C^∞-function

→ E.g.: If $g(x) > 0$ then $h(x) \geq 0 \;\Leftrightarrow\; h(x)\max\{g(x)^3, 0\} \geq 0$ (C^2)

Problem: Regularity conditions violated

→ No strictly feasible point close to a point with $g(x) < 0$ (no MFCQ)

→ Potential cause of numerical problems

→ Primal-dual optimality framework doesn't apply

Interesting aside: A combination of the "right" NLP-reformulation and the "right" NLP method can work well (Leyffer et al. for MPECs)

→ "Strike of luck" for a special case?

Traditional Approaches: Nonsmooth optimization

Combinatorial selections introduce natural non-smoothness

→ max-min selections

Nonsmooth optimization seems appropriate

→ Bundle methods

But nonsmooth optimization is too general

→ Focus on analytically quite complicated derivative surrogates

→ Not obvious how to exploit combinatorial structure

→ Based on often fairly weak and "ill-understood" stationarity conditions

Towards Local Optimization of NLPs with Combinatorial Constraints

Success of local optimization in NLP is based on

- Well-understood primal-dual stationarity conditions
- Link with Newton's method
- Strategies to globalize Newton's method

Aim: Provide a framework for local optimization of NLPs with combinatorial constraints which maintains these features

NLPs with combinatorial structure

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Motivation:

Decompose analytical source of non-convexity ($g(.)$) from combinatorial source (Z)

Hope to get a better idea

$\longrightarrow$ where and how to apply NLP tools

$\longrightarrow$ where and how to apply combinatorial tools

NLPs with combinatorial structure

$$(\text{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

NLP structure set $Z = \{(s,t) \mid s \geq 0, t = 0\} = R^n_+ \times 0^m$

Convex structure sets Z have a trivial combinatorial structure

$\longrightarrow$ Not of interest in this talk

NLPs with combinatorial structure

$$(\mathrm{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

MPEC structure set

$$Z = \{z = (s, t, u, v) \mid s \geq 0, t = 0, \min\{u, v\} = 0\}$$

NLPs with combinatorial structure

$$(\mathrm{P}) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

MPEC structure set

$$\begin{aligned} Z &= \{z = (s, t, u, v) \mid s \geq 0, t = 0, \min\{u, v\} = 0\} \\ &= R^n_+ \times 0^m \times L^p \end{aligned}$$

NLPs with combinatorial structure

$$(\mathrm{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

MPEC structure set

$$\begin{aligned} Z &= \{z = (s,t,u,v) \mid s \geq 0, t = 0, \min\{u,v\} = 0\} \\ &= R_+^n \times 0^m \times L^p \end{aligned}$$

New structural element is $L = \{(u_i, v_i) \mid \min\{u_i, v_i\} = 0\}$

MPEC structure set is non-convex but easy to analyse

NLPs with combinatorial structure

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Further example of a combinatorial structure set

$x_i \geq 0$ and $g_i(x) \geq 0$ if $x_i > 0$

NLPs with combinatorial structure

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Further example of a combinatorial structure set

$$x_i \geq 0 \text{ and } g_i(x) \geq 0 \text{ if } x_i > 0$$

NLP formulation: $x_i \geq 0, \;\; x_i g_i(x) \geq 0$ violates MFCQ if $x_i = 0$ and $g_i(x) < 0$

NLPs with combinatorial structure

$$(\text{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Further example of a combinatorial structure set

$$x_i \geq 0 \text{ and } g_i(x) \geq 0 \text{ if } x_i > 0$$

NLP formulation: $x_i \geq 0, \ x_i g_i(x) \geq 0$ violates MFCQ if $x_i = 0$ and $g_i(x) \leq 0$

Full structure set $Z = R^n_+ \times 0^m \times Y^p$

New structural element

$$(x_i, g_i) \in Y = \{(u_i, v_i) \in R^2 \mid u_i \geq 0 \text{ and } (v_i \geq 0 \text{ or } u_i = 0)\}$$

Structure set is non-convex but easy to analyse

NLPs with combinatorial structure

$$(P) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Main message of the talk:

NLPs with combinatorial structures can be dealt with in much the same way as standard NLPs

$\longrightarrow$ Optimality conditions based on primal-dual structure

$\longrightarrow$ Methods related to Newton's method

Critical points

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Lagrangian $L(x, \lambda) = f(x) - \lambda g(x)$

Critical points

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Lagrangian $L(x, \lambda) = f(x) - \lambda g(x)$

x is **critical point** if $\nabla_x L(x, \lambda) := \nabla f(x) - \nabla g(x)^\top \lambda = 0$ and

$\lambda_i = 0$ if g_i is **inactive**

Critical points

(P) $$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Lagrangian $L(x, \lambda) = f(x) - \lambda g(x)$

x is **critical point** if $\nabla_x L(x, \lambda) := \nabla f(x) - \nabla g(x)^\top \lambda = 0$ and

$$\lambda_i = 0 \text{ if } g_i \text{ is inactive}$$

g_i is **inactive** at $\bar{z} = g(\bar{x})$ if the validity of $z \in Z$ is independent of the value of z_i for z close to $\bar{z}$

Critical points

$$(P) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Lagrangian $L(x, \lambda) = f(x) - \lambda g(x)$

x is *critical point* if $\nabla_x L(x, \lambda) := \nabla f(x) - \nabla g(x)^\top \lambda = 0$ and

$$\lambda_i = 0 \text{ if } g_i \text{ is inactive}$$

g_i is *inactive* at $\bar{z} = g(\bar{x})$ if the validity of $z \in Z$ is independent of the value of z_i for z close to $\bar{z}$

Optimality condition: A local optimizer x^* is a critical point if the gradients of active g_i's are linearly independent (LICQ)

Primal dual structure

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \min & \nabla f(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Primal dual structure

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \text{min} & \nabla f(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Necessary optimality condition if (LICQ) holds at $g(x)$ and Z is locally $*-$shaped at $z = g(x)$

$\exists$ n.h. U of z s.t. $\forall z' \in Z \cap U$ and $\forall \alpha \in [0,1] : z + \alpha(z' - z) \in Z$

Primal dual structure

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \text{min} & \nabla f(x)d \\ \text{s.t.} & g(x) + \nabla g(x)d \in Z \end{array}$$

Necessary optimality condition if (LICQ) holds at $g(x)$ and Z is locally $*-$shaped at $z = g(x)$

$\exists$ n.h. U of z s.t. $\forall z' \in Z \cap U$ and $\forall \alpha \in [0,1] : z + \alpha(z' - z) \in Z$

Examples: convex sets, roots and lower-level sets of piecewise linear functions, finite unions of locally $*-$shaped sets

Primal dual structure

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \text{min} & \nabla f(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

If x is a local minimiser and LICQ holds then x is a critical point

$$\nabla f(x) = \nabla g(x)^\top \lambda$$

Primal dual structure

(P)
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \min & \nabla f(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Stationarity condition: $d = 0$ is local optimum of

$$\begin{array}{ll} \min & \lambda^\top \nabla g(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Primal dual structure

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is loc. optimum of

$$\begin{array}{ll} \min & \lambda^\top \nabla g(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Primal dual structure

$$(\text{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is loc. optimum of

$$\begin{array}{ll} \text{min} & \lambda^\top \nabla g(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Replace $z = g(x), v = \nabla g(x) d$ (LICQ allows us to recover d)

Primal dual structure

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $d = 0$ is loc. optimum of

$$\begin{array}{ll} \text{min} & \lambda^\top \nabla g(x) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Replace $z = g(x), v = \nabla g(x) d$ (LICQ allows us to recover d)

Stationarity condition: $v = 0$ is loc. optimum of

$$\begin{array}{ll} \text{min} & \lambda^\top v \\ \text{s.t.} & z + v \in Z \end{array}$$

Primal dual structure

(P)
$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $v = 0$ is loc. optimum of

$$\begin{array}{ll} \text{min} & \lambda^\top v \\ \text{s.t.} & z + v \in Z \end{array}$$

Primal dual structure

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

Stationarity condition: $v = 0$ is loc. optimum of

$$\begin{array}{ll} \min & \lambda^\top v \\ \text{s.t.} & z + v \in Z \end{array} \tag{1}$$

Primal dual structure:

$$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0 \text{ is a local optimum of } (1)\}$$

Note: $(z, \lambda) \in PD_Z \Rightarrow \lambda_i = 0$ if z_i is inactive at z

→ o/w decrease objective by small change of v_i w/o violating $z + v \in Z$

The primal dual problem

(P)

$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

(PD)

$$\begin{array}{rcl} \nabla f(x) + \nabla g(x)^\top \lambda & = & 0 \\ (g(x), \lambda) & \in & PD_Z \end{array}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

The primal dual problem

(P)

$$\begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

(PD)

$$\begin{array}{rcl} \nabla f(x) + \nabla g(x)^\top \lambda & = & 0 \\ (g(x), \lambda) & \in & PD_Z \end{array}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

Great Help: $PD_{Z_1 \times \cdots \times Z_n} = PD_{Z_1} \times \cdots \times PD_{Z_n}$

Examples of primal dual problems

(PD)

$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

NLP structure set: $Z = R^n_+ \times 0^m$

$PD_{R_+} = \{(z, \lambda) \in R^2 \mid z \geq 0, \lambda \geq 0, z\lambda = 0\}$

$PD_0 = \{(z, \lambda) \in R^2 \mid z = 0\}$

Examples of primal dual problems

(PD)
$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

MPEC structure set: $Z = R^n_+ \times 0^m \times L^p$, $L = \{z \in R^2 \mid \min\{z_1, z_2\} = 0\}$

$PD_{R_+} = \{(z, \lambda) \in R^2 \mid z \geq 0, \lambda \geq 0, z\lambda = 0\}$

$PD_0 = \{(z, \lambda) \in R^2 \mid z = 0\}$

Examples of primal dual problems

(PD)
$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

MPEC structure set: $Z = R^n_+ \times 0^m \times L^p$, $L = \{z \in R^2 \mid \min\{z_1, z_2\} = 0\}$

$PD_{R_+} = \{(z, \lambda) \in R^2 \mid z \geq 0, \lambda \geq 0, z\lambda = 0\}$

$PD_0 = \{(z, \lambda) \in R^2 \mid z = 0\}$

$PD_L = \{(z, \lambda) \in R^{2+2} \mid \min\{z_1, z_2\} = 0, z_i \lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0\}$

Examples of primal dual problems

(PD)
$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

MPEC structure set: $Z = R^n_+ \times 0^m \times L^p$, $L = \{z \in R^2 \mid \min\{z_1, z_2\} = 0\}$

$PD_{R_+} = \{(z, \lambda) \in R^2 \mid z \geq 0, \lambda \geq 0, z\lambda = 0\}$

$PD_0 = \{(z, \lambda) \in R^2 \mid z = 0\}$

$PD_L = \{(z, \lambda) \in R^{2+2} \mid \min\{z_1, z_2\} = 0, z_i \lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0\}$

Notice: PD_L is not closed!

Examples of primal dual problems

(PD)

$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

$PD_Z = \{(z, \lambda) \mid z \in Z,\ v = 0$ is a local optimum of $\min\{\lambda v \mid z + v \in Z\}\}$

MPEC structure set: $Z = R^n_+ \times 0^m \times L^p$, $L = \{z \in R^2 \mid \min\{z_1, z_2\} = 0\}$

$PD_{R_+} = \{(z, \lambda) \in R^2 \mid z \geq 0, \lambda \geq 0, z\lambda = 0\}$

$PD_0 = \{(z, \lambda) \in R^2 \mid z = 0\}$

$PD_L = \{(z, \lambda) \in R^{2+2} \mid \min\{z_1, z_2\} = 0, z_i\lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0\}$

$\overline{PD}_L = \{(z, \lambda) \mid \min\{z_1, z_2\} = 0, z_i\lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0$ or $\lambda_1\lambda_2 = 0\}$

Strong stationarity vs. M-stationarity

(PD)
$$\begin{aligned} \nabla f(x) + \nabla g(x)^\top \lambda &= 0 \\ (g(x), \lambda) &\in PD_Z \end{aligned}$$

MPEC structure set $Z = R^n_+ \times 0^m \times L^p$

Strong stationarity: $PD_Z = (PD_{R_+})^n \times (PD_0)^m \times (PD_L)^p$ with

$PD_L = \{(z, \lambda) \in R^{2+2} | \min\{z_1, z_2\} = 0, z_i \lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0\}$

M-stationarity: Replace PD_L by

$\overline{PD}_L = \{(z, \lambda) | \min\{z_1, z_2\} = 0, z_i \lambda_i = 0, z = 0 \Rightarrow \lambda \geq 0 \text{ or } \lambda_1 \lambda_2 = 0\}$

Intermediate summary

Lagrangian framework extends straight-forwardly to NLPs with combinatorial constraints

Leads to primal dual problem

$\rightarrow$ Primal-dual structure depends **only** on Z, not on $g(.)$

No need for complicated concepts from non-smooth analysis

$\rightarrow$ Tighter optimality conditions than non-smooth analysis approaches

Now to methods: Exploit NLP features and combinatorial structure of Z in algorithms

Extends "composite optimization" approach of Fletcher et al. to non-convex structure sets Z

Combinatorial SQP

$$(\mathrm{P}) \qquad \begin{array}{ll} \text{min} & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

$$L(x, \lambda) = f(x) - g(x)\lambda$$

$$(\mathrm{QP}_Z(x, \lambda)) \qquad \begin{array}{ll} \text{min} & \nabla f(x) d + \frac{1}{2} d^\top \nabla^2_{xx} L(x, \lambda) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Combinatorial SQP

$$(\mathrm{P}) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \in Z \end{array}$$

$$L(x,\lambda) = f(x) - g(x)\lambda$$

$$(\mathrm{QP}_Z(x,\lambda)) \qquad \begin{array}{ll} \min & \nabla f(x) d + \frac{1}{2} d^\top \nabla^2_{xx} L(x,\lambda) d \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array}$$

Input from Combinatorial Optimization:

Find **local** solution of "combinatorial" QP in an efficient way

Combinatorial SQP

Local quadratic convergence of combinatorial SQP

Assumptions:

- LICQ
- strict complementarity
- second order condition
- Z locally $*-$shaped at x^*

Conclusions: If (x, λ) is close to (x^*, λ^*) then

- $\text{QP}_Z(x, \lambda)$ is feasible
- $\text{QP}_Z(x, \lambda)$ has a unique stationary point $d^*(x, \lambda)$ in a neighbourhood of 0
- If "combinatorial" QP-solver detects these stationary points then primal dual iterates converge quadratically to (x^*, λ^*)

Solving QPs with combinatorial constraints

Must exploit structure of Z

Active set approach: $Z = \bigcup_I Z_I$

$$\begin{array}{rl} \text{min} & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array} \quad = \quad \min_I \begin{array}{rl} \text{min} & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z_I \end{array}$$

Solving QPs with combinatorial constraints

Must exploit structure of Z

Active set approach: $Z = \bigcup_I Z_I$

$$\begin{array}{ll} \min & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array} \quad = \quad \min_I \; \begin{array}{ll} \min & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z_I \end{array}$$

MPEC example

$$\begin{aligned} Z &= R_+^n \times 0^m \times L^p \\ Z_I &= R_+^n \times 0^m \times (0, R_+)^I \times (R_+, 0)^{I^c} \end{aligned}$$

If LICQ holds at the solution of QP$_{Z_I}$ then QP$_{Z_I}$-LMs tell us whether we are stationary for the QP$_Z$ and, if not, how to change I to improve

Solving QPs with combinatorial constraints

Must exploit structure of Z

Active set approach: $Z = \bigcup_I Z_I$

$$\begin{array}{ll} \text{min} & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z \end{array} \quad = \quad \min_I \begin{array}{ll} \text{min} & q(d) \\ \text{s.t.} & g(x) + \nabla g(x) d \in Z_I \end{array}$$

MPEC example

$$\begin{aligned} Z &= R_+^n \times 0^m \times L^p \\ Z_I &= R_+^n \times 0^m \times (0, R_+)^I \times (R_+, 0)^{I^c} \end{aligned}$$

If LICQ holds at the solution of QP$_{Z_I}$ then QP$_{Z_I}$-LMs tell us whether we are stationary for the QP$_Z$ and, if not, how to change I to improve

Can change active complementarity constraints within subproblem solve

Active set approach extends to other non-convex structure sets Z

Globalization

Aim: Guarantee convergence to a stationary point / local minimizer from any starting point

General penalty-trust-region results extend to NLPs with combinatorial constraints (S & Stöhr '99)

Assumption: Z is lower level set of a piecewise affine function

Problem: Convergence assumes that ratio of objective value of detected local solution of subproblem to global solution of subproblem is bounded away from zero

$\longrightarrow$ Has been overcome by Stöhr '00 for MPECs

$\longrightarrow$ No results so far for more general combinatorial NLP

Other approaches: Avoiding the devil

Can use the primal-dual framework to analyse alternative approaches

Smoothing: Replace combinatorial structure set Z by nearby smooth manifold and apply NLP

Example: $\min\{x, y\} = 0 \;\approx\; x, y \geq 0, \quad x_i y_i = \mu$

Regularization: Blow Z up to a nearby standard NLP-set with non-empty interior and apply NLP

Example: $\min\{x, y\} = 0 \;\approx\; x, y \geq 0, \quad x_i y_i \leq \mu$

Penalization: Formulate Z as NLP constraint set, move degenerate constraints into objective function and apply NLP

Example: $x, y \geq 0, xy = 0 \Rightarrow$ augment objective by $\frac{1}{\mu} xy$

Other approaches: Avoiding the devil

MPEC experience:

Convergence analyses as $\mu \to 0+$

$\to$ Local path of NLP-solutions converging to strongly stationary solution of MPEC

Good practical performance reported for MPECs for many NLP solvers

Regularization and penalization are the basis for promising interior point approaches

Conclusion

NLP primal-dual framework extends straight-forwardly to NLPs with combinatorial constraints, providing natural primal-dual optimality conditions

→ Nonsmooth analysis is unnecessary "barrier to entry"

SQP extends naturally within this framework

→ Efficient local solution of combinatorial subproblems is key

→ Can we get global convergence without solving subproblems "nearly globally" as we converge?

"Combinatorial" primal-dual framework useful to analyse effectiveness of alternative approaches

→ Smoothing, Regularization, Penalization, Interior Point Methods, etc.

Where from here?

Theory: Further extend NLP framework to combinatorial constraints

→ Sensitivity Analysis

→ Stochastic Problems

Methods / Codes: Need to combine NLP and combinatorial expertise

→ Want to make best use of existing NLP methods / codes

→ Generic solver module, controlling global and fast local convergence PLUS specialized combinatorial subproblem modules?

Applications: Key driver for further work

→ Need high-impact show-cases to motivate further research

Main reference for this talk:

S., Nonconvex Structures in Nonlinear Programming, *Ops. Res.* (2004)