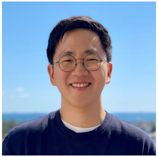
On the solution of large-scale mathematical programs with many complementarity constraints

François Pacaud

Centre Automatique et Systèmes (CAS), Mines Paris - PSL

Conférence SMAI-MODE
March 18, 2026

Who are we?

- Alexis Montoison @ Argonne National Laboratory
- Armin Nurkanović @ Freiburg University
- François Pacaud @ Mines Paris-PSL
- Anton Pozharskiy @ Freiburg University
- Sungho Shin @ MIT

Context

This research stems from an inspirational talk given in 2020

The Era of "Non"-Optimization Problems

Jong-Shi Pang

Research question

How mature are "non"-optimization solvers in the large-scale regime?

Cui, Pang. Modern Nonconvex Nondifferentiable Optimization. 2021.

Structurally nonconvex problems

For $a, b \in \mathbb{R}$, we define the complementarity constraint

$$0 \leq a \perp b \geq 0 \quad \iff \quad (a, b) \geq 0 \,, \; a \times b = 0$$

Disjunctive behavior: either $a = 0$ or $b = 0$

Mathematical program with complementarity constraints (MPCC)

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \; f(x) \\ &\text{subject to } c(x) = 0 \,, \\ &\qquad\qquad 0 \leq x_1 \perp x_2 \geq 0 \end{aligned}$$

N.B. We partition the variables as $x = (x_0, x_1, x_2) \in \mathbb{R}^{n-2p} \times \mathbb{R}^p \times \mathbb{R}^p$ to isolate the complementarity contributions

Complementarities make explicit (part of) the nonconvex structure

Scholtes. Nonconvex structures in nonlinear programming. 2004.

Practical examples

Complementarity constraints are very handy to reformulate *nonsmooth* functions!

- $y = sgn(x)$

$$
\begin{aligned}
& x = x_p - x_n \; , \\
& 0 \leq x_p \perp 1 - y \geq 0 \; , \\
& 0 \leq x_n \perp 1 + y \geq 0
\end{aligned}
$$

- $y = |x|$

$$
\begin{aligned}
& y = x_p + x_n \; , \\
& x = x_p - x_n \; , \\
& 0 \leq x_p \perp x_n \geq 0
\end{aligned}
$$

- $y = \max\{0, x\}$

$$
\begin{aligned}
& y = x_p \; , \\
& x = x_p - x_n \; , \\
& 0 \leq x_p \perp x_n \geq 0
\end{aligned}
$$

- $y = \min_x c^\top x \text{ s.t. } Ax = b \; , \; x \geq 0$

$$
\begin{aligned}
& c + A^\top \lambda - \nu = 0 \; , \\
& Ax - b = 0 \; , \\
& 0 \leq x \perp \nu \geq 0
\end{aligned}
$$

Baumrucker, Biegler. MPEC problem formulations in chemical engineering applications. 2009.

Making explicit the combinatorial structure

Fact: We do not know beforehand which complementarity constraints are active

For $i = 1, \cdots, p$, either $x_{1i} = 0$ or $x_{2i} = 0$ $\rightarrow 2^p$ different possibilities!

MINLP formulation

For $M > 0$ large-enough, the MPCC is equivalent to the MINLP

$$\begin{aligned} \min_{x,y} \ & f(x) \\ \text{subject to } & c(x) = 0 \ , \\ & 0 \leq x_1 \leq M y_i \ , && \forall i = 1, \cdots, p \\ & 0 \leq x_2 \leq M(1 - y_i) \ , && \forall i = 1, \cdots, p \\ & y \in \{0, 1\}^p \end{aligned}$$

Problem: MINLP reformulation gets intractable as p increases

Important notations

- **Feasible set**: $\Omega = \{x \in \mathbb{R}^n \mid c(x) = 0\ ,\ 0 \leq x_1 \perp x_2 \geq 0\}$.

- **Index sets**: for $x \in \Omega$, build a partition of $\{1, \cdots, p\}$ as

$$\mathcal{I}_{+0}(x) = \{i \in \{1, \cdots, p\} \mid x_{1,i} > 0\, ,\ x_{2,i} = 0\}\ ,$$
$$\mathcal{I}_{0+}(x) = \{i \in \{1, \cdots, p\} \mid x_{1,i} = 0\, ,\ x_{2,i} > 0\}\ ,$$
$$\mathcal{I}_{00}(x) = \{i \in \{1, \cdots, p\} \mid x_{1,i} = 0\, ,\ x_{2,i} = 0\}\ .$$

- **Strict complementarity:** $\mathcal{I}_{00}(x) = \emptyset$ (easy case)

- **MPCC Lagrangian:** $\mathcal{L}^{MPCC}(x, \lambda, \mu) = f(x) + \lambda^\top c(x) - \mu_1^\top x_1 - \mu_2^\top x_2$

- **MPCC linearized feasible cone:** For $x \in \Omega$,

$$\mathcal{F}_\Omega(x) = \Big\{ d \in \mathbb{R}^n \mid \nabla c_i(x)^\top d = 0\ ,\ \forall i = 1, \cdots, m,$$
$$d_{1,i} = 0\ ,\ \forall i \in \mathcal{I}_{0+}(x),$$
$$d_{2,i} = 0\ ,\ \forall i \in \mathcal{I}_{+0}(x),$$
$$0 \leq d_{1,i} \perp d_{2,i} \geq 0\ ,\ \forall i \in \mathcal{I}_{00}(x) \Big\}\ .$$

Stationarity conditions

What is the equivalent of the KKT conditions for MPCCs?

Bouligand stationarity (B)

The point $x \in \Omega$ is B-stationary if $d = 0$ is a solution to the LPCC

$$\min_{d \in \mathbb{R}^n} \nabla f(x)^\top d \quad \text{subject to} \quad d \in \mathcal{F}_\Omega(x) \, .$$

Strong stationarity (S)

The point $x \in \Omega$ is S-stationary if there exists (λ, μ) such that

$$\begin{aligned}
&\nabla_x \mathcal{L}^{\mathrm{MPCC}}(x, \lambda, \mu) = 0 \, , \\
&c(x) = 0 \, , \\
&x_{1,i} \geq 0 \, , \; \mu_{1,i} = 0 \, , \; x_{2,i} = 0 \, , \; \mu_{2,i} \in \mathbb{R} \, , \; \forall i \in \mathcal{I}_{+0}(x) \, , \\
&x_{1,i} = 0 \, , \; \mu_{1,i} \in \mathbb{R} \, , \; x_{2,i} \geq 0 \, , \; \mu_{2,i} = 0 \, , \; \forall i \in \mathcal{I}_{0+}(x) \, , \\
&x_{1,i} = 0 \, , \; \mu_{1,i} \geq 0 \, , \; x_{2,i} = 0 \, , \; \mu_{2,i} \geq 0 \, , \; \forall i \in \mathcal{I}_{00}(x) \, .
\end{aligned}$$

$\rightarrow$ encodes the KKT conditions of a regular nonlinear program

Proposition

If MPCC-LICQ holds, then Bouligand-stationarity is equivalent to strong-stationarity.

Let's do a break

What do we want to do?

Find a S-stationary point

To that goal, we are currently developing a suite of solvers based on three methods:

- Scholtes relaxation method ;
- ℓ_1-penalty reformulation ;
- Augmented Lagrangian method.

Scholtes relaxation method

Observe that the MPCC is equivalent to the (degenerate) nonlinear program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & f(x) \\ \text{subject to } & c(x) = 0 \, , \\ & (x_1, x_2) \geq 0 \\ & x_{1i} x_{2i} \leq 0 \quad \forall i = 1, \cdots, p \end{aligned}$$

- Smooth, but it violates MFCQ at all feasible points!
- Multipliers are unbounded: failure in optimization solvers

Scholtes relaxation (homotopy method)

For $\tau > 0$, solve

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & f(x) \\ \text{subject to } & c(x) = 0 \, , \\ & (x_1, x_2) \geq 0 \\ & x_{1i} x_{2i} \leq \tau \quad \forall i = 1, \cdots, p \end{aligned}$$

- If MPCC-MFCQ holds, we recover a solution for MPCC as $\tau \to 0$
- If IPM is used, τ should be tuned proportionally to the barrier parameter

Scholtes. Convergence Properties of a Regularization Scheme for Mathematical Programs with Complementarity Constraints. 2001.

The ℓ_1-penalty method

Multipliers are unbounded? Use ℓ_1 penalty to cap them!

Scholtes relaxation (homotopy method)

For a penalty $\rho > 0$, solve

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & f(x) + \rho x_1^\top x_2 \\ \text{subject to } & c(x) = 0 \, , \\ & (x_1, x_2) \geq 0 \end{aligned}$$

- Converge under MPCC-LICQ
- Work very well together with IPM
- But falls behind the Scholtes relaxation numerically
- And can converge to spurious solutions...

A more recent alternative: augmented Lagrangian method

Mixing the Scholtes regularization with a penalty-based approach

We use a recent variant of Augmented Lagrangian called *Nonlinearly constrained Lagrangian method* (NCL)

NCL method

For penalty $\rho_k > 0$ and multiplier estimates (λ_k, ν_k), solve

$$\begin{aligned} \min_{x,r,t} \; & f(x) + \lambda_k^\top r + \nu_k^\top t + \frac{1}{2}\rho_k(\|r\|^2 + \|t\|^2) \\ \text{subject to } & c(x) + r = 0 \,, \\ & x_{1i} x_{2i} \leq t_i \\ & (x_1, x_2) \geq 0 \end{aligned}$$

The new variables (r, t) regularize the problem

- Converge also under MPCC-LICQ
- Inner solves are akin to a stabilized IPM
- The regularization t plays the role of a dynamic Scholtes relaxation
- Efficient with warmstart!

Izmailov, Solodov, Uskov. Global Convergence of Augmented Lagrangian Methods Applied to Optimization Problems with Degenerate Constraints, Including Problems with Complementarity Constraints. 2012.

Why is it difficult for IPM?

IPM is mostly Newton method applied to the problem's stationarity conditions

At each iteration, IPM solves a *KKT system* with the structure

$$\begin{bmatrix} W + \Sigma & J^\top \\ J & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

with

- W: Hessian of the Lagrangian
- Σ: diagonal terms corresponding to the bound constraints
- J: Jacobian of the constraints

Observations

- The complementarity terms increase the indefinitess in W
- The terms in Σ can blow up quickly if we decrease τ (resp. increase ρ) too fast
- If $\mathcal{I}_{00}(x) \neq \emptyset$ at the solution, the Jacobian J is not full row-rank

Trade-off

1. Do not shrink the feasible set too fast
2. Apply a proper primal-dual regularization to the KKT system

Example: SCOPF problem with $N - 1$ security criterion

How to compute a solution that is robust w.r.t. the loss of one critical element for K distinct contingencies?

Complementarity constraints

- **Droop control:** adjust the power generation in contingency k from the base case generation, $\forall k = 1, \cdots, K$:

$$\pi_{g+}^k - \pi_{g-}^k = p_g^k - (p_g^0 + \alpha_g \Delta) \, ,$$
$$0 \leq \pi_{g-}^k \perp \overline{p}_g - p_g^k \geq 0 \, ,$$
$$0 \leq \pi_{g+}^k \perp p_g^k - \underline{p}_g \geq 0 \, .$$

- **PV/PQ switches:** keep the voltage magnitudes at the PV buses at their nominal values $v_b^k = v_b^0$ by injecting or absorbing reactive power, $\forall k = 1, \cdots, K$:

$$\nu_{b+}^k - \nu_{b-}^k = v_b^k - v_b^0 \, ,$$
$$0 \leq \nu_{b-}^k \perp \overline{q}_g - q_g^k \geq 0 \, ,$$
$$0 \leq \nu_{b+}^k \perp q_g^k - \underline{q}_g \geq 0 \, .$$

$\rightarrow$ We get a total of $\mathcal{O}(K)$ contingencies

Numerical results

Here, using the Scholtes relaxation implemented in CCopt.jl

K	n	p	Objective	#Iter	Time (s)
10	43k	2k	7.4737	117	7
20	82k	4k	7.8295	200	26
50	201k	11k	7.8372	191	66
100	398k	22k	7.8428	169	151

Table: Time-to-solution on an instance with 500 buses (ACTIVSg500) as we increase the number of contingencies K

Take-away messages

Final answer

Yes, it is practical to solve large-scale MPCCs!

Applications

- Power systems and gas networks
- Optimal control and robotics with nonsmooth dynamics
- Bilevel optimization and equilibrium problems

New software are coming

- `ComplementOpt.jl`: improve the support of complementarity in JuMP (including automatic problem's reformulation)
- `CCopt.jl`: A custom solver for MPCCs (to be released soon).