Mathematical Programs with Complementarity Constraints and Related Problems

Prof. Dr. Alexandra Schwartz
schwartz@gsc.tu-darmstadt.de

TECHNISCHE UNIVERSITÄT DARMSTADT

Ressources

TECHNISCHE UNIVERSITÄT DARMSTADT

You can find

- my slides
- and some more extensive lecture notes

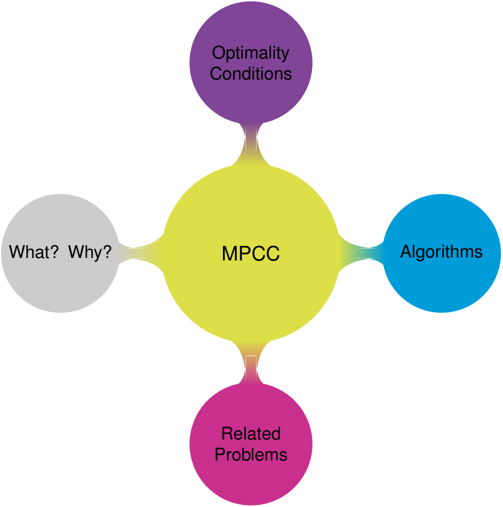
at

github.com/alexandrabschwartz/Winterschool2018

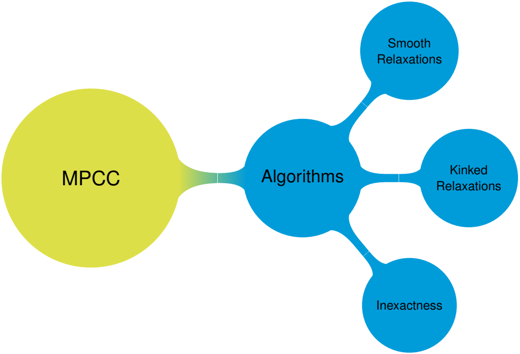
If you have any questions, please come to me during the week or contact me at

schwartz@gsc.tu-darmstadt.de

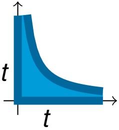
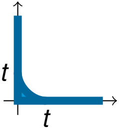
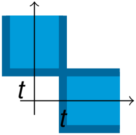
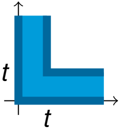
Contents of the Course



What did we do yesterday?

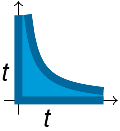
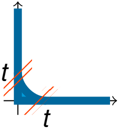
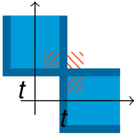
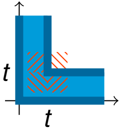


Relaxation Methods: Exact KKT Points

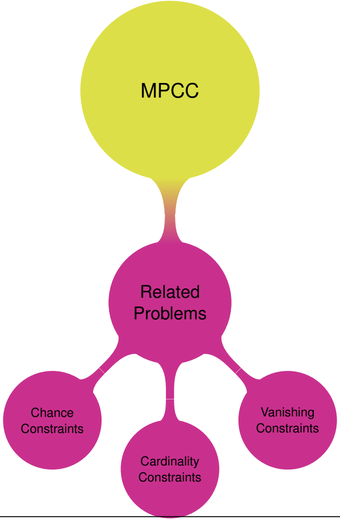
relaxation	global	local	nonsmooth	kinked
authors	Scholtes	Steff./Ulbr.	Kadrani et al.	Kanzow, S.
geometry				
stationarity under CQ	C-stat. MPEC-MFCQ	C-stat. MPEC-CPLD	M-stat. MPEC-CPLD	M-stat. MPEC-CPLD

Relaxation Methods: Inexact KKT Points

Let $t_k \downarrow 0$ and $\varepsilon_k = o(t_k)$. Let x^k be ε_k-KKT points of NLP(t_k) and assume $x^k \to x^*$. Then the limit x^* satisfies:

relaxation	global	local	nonsmooth	kinked
authors	Scholtes	Steff./Ulbr.	Kadrani et al.	Kanzow, S.
geometry				
stationarity	C-stat.	W-stat.	W-stat.	W-stat.
under CQ	MPEC-MFCQ	MPEC-MFCQ	MPEC-MFCQ	MPEC-MFCQ

What is the plan for today?



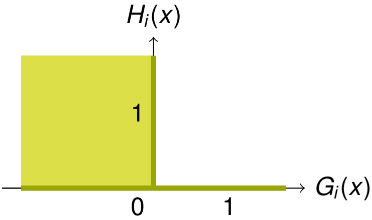
Vanishing Constraints

Mathematical Program with Vanishing Constraints (MPVC)

TECHNISCHE UNIVERSITÄT DARMSTADT

A mathematical program with vanishing constraints (MPVC) is of the form

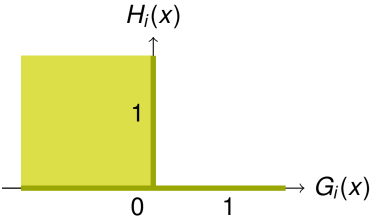
$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0,$$
$$H(x) \geq 0, \quad G(x) \circ H(x) \geq 0.$$



Mathematical Program with Vanishing Constraints (MPVC)

A mathematical program with vanishing constraints (MPVC) is of the form

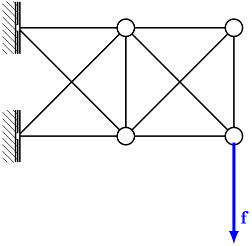
$$\min_x f(x) \quad \text{s.t.} \quad g(x) \le 0, \quad h(x) = 0,$$
$$H(x) \ge 0, \quad G(x) \circ H(x) \ge 0.$$



Two cases are possible:

- $H_i(x^*) > 0$ and thus $G_i(x^*) \le 0$
- $H_i(x^*) = 0$ and then $G_i(x^*) \in \mathbb{R}$ arbitrarily.

Example: Truss Design



Given some anchor points and a load point, how should you construct a load bearing truss?

Example: Truss Design (continued)

TECHNISCHE UNIVERSITÄT DARMSTADT

$\min_{a,u} \quad \sum_i l_i a_i$	minimize volume of the truss
s.t. $\quad 0 \leq a \leq a_{\max}$	bounds on cross sections
$K(a)u = f$	force equilibrium
$f^T u \leq c$	bound on the deformation
$(\sigma_i(u)^2 - \sigma_{\max}^2) a_i \leq 0$	bound on the stress on bar i

Variables

- a_i cross sectional area of bar i
- u_i nodal displacements at the hinges

Constants

- f applied force
- l_i length of bar i
- $K(a)$ stiffness matrix of the truss
- $\sigma_i(u)$ stress on bar i

Relation between Vanishing and Complementarity Constraints

TECHNISCHE UNIVERSITÄT DARMSTADT

One can reformulate an MPVC as an MPCC using slack variables as follows:

$$\begin{aligned} \min_{x,y} f(x) \quad \text{s.t.} \quad & g(x) \leq 0, h(x) = 0, \\ & G(x) - y \leq 0, \\ & 0 \leq H(x) \perp y \geq 0. \end{aligned}$$

Relation between Vanishing and Complementarity Constraints

One can reformulate an MPVC as an MPCC using slack variables as follows:

$$\begin{aligned} \min_{x,y} f(x) \quad \text{s.t.} \quad & g(x) \leq 0, h(x) = 0, \\ & G(x) - y \leq 0, \\ & 0 \leq H(x) \perp y \geq 0. \end{aligned}$$

Exercise

How is the relation between the MPCC reformulation and the original MPVC?

Relation between Vanishing and Complementarity Constraints

TECHNISCHE UNIVERSITÄT DARMSTADT

One can reformulate an MPVC as an MPCC using slack variables as follows:

$$\min_{x,y} f(x) \quad \text{s.t.} \quad \begin{aligned} &g(x) \le 0, h(x) = 0, \\ &G(x) - y \le 0, \\ &0 \le H(x) \perp y \ge 0. \end{aligned}$$

Then x^* is a solution of the MPVC if and only if (x^*, y^*) with

$$y_i^* \begin{cases} = 0 & \text{if } H_i(x^*) > 0, \\ \ge \max\{0, G_i(x^*)\} & \text{if } H_i(x^*) = 0 \end{cases}$$

if a solution of the corresponding MPCC. But y^* is not uniquely determined.

MPVCs versus Constraint Qualifications

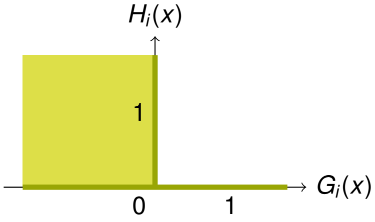
TECHNISCHE UNIVERSITÄT DARMSTADT

Vanishing Constraints:

$$H(x)_i \geq 0, \quad G(x)_i H(x)_i \geq 0 \qquad \forall i = 1, \dots, q$$

Let x^* be feasible for MPVC.

- If $H_i(x^*) = 0$ for at least one index $i \in \{1, \dots, q\}$, then LICQ is violated in x^*.
- If there exists an index i such that $H_i(x^*) = 0$ and $G_i(x^*) \geq 0$, then MFCQ is violated in x^*.
- The vanishing constraints i with $H_i(x^*) > 0$ are “well-behaved”.



How to handle MPVCs?

- One can try to apply NLP theory.

How to handle MPVCs?

- One can try to apply NLP theory.
- One can use the MPCC reformulation.

How to handle MPVCs?

- One can try to apply NLP theory.
- One can use the MPCC reformulation.
- One can define $V := \{(a, b) \in \mathbb{R}^2 \mid b \geq 0, ab \leq 0\}$ and use the condition

$$(\gamma_i, \nu_i) \in N_V^F(G_i(x^*), H_i(x^*)) \subseteq N_V^M(G_i(x^*), H_i(x^*))$$

  to obtain possible optimality conditions.

How to handle MPVCs?

- One can try to apply NLP theory.
- One can use the MPCC reformulation.
- One can define $V := \{(a,b) \in \mathbb{R}^2 \mid b \geq 0, ab \leq 0\}$ and use the condition

$$(\gamma_i, \nu_i) \in N_V^F(G_i(x^*), H_i(x^*)) \subseteq N_V^M(G_i(x^*), H_i(x^*))$$

  to obtain possible optimality conditions.
- One can locally decompose the MPVC into TNLP(x^*, I) for all $I \subseteq \{i \mid G_i(x^*) = H_i(x^*) = 0\}$

$$\begin{aligned} \min_x f(x) \quad \text{s.t.} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & G_i(x) \leq 0, \quad H_i(x) \geq 0 \quad \forall i \in \{i \mid G_i(x^*) < 0\} \cup I, \\ & H_i(x) = 0 \quad \forall i \in \{i \mid G_i(x^*) > 0\} \cup I^c \end{aligned}$$

How to handle MPVCs?

- One can try to apply NLP theory.
- One can use the MPCC reformulation.
- One can define $V := \{(a, b) \in \mathbb{R}^2 \mid b \geq 0, ab \leq 0\}$ and use the condition

$$(\gamma_i, \nu_i) \in N_V^F(G_i(x^*), H_i(x^*)) \subseteq N_V^M(G_i(x^*), H_i(x^*))$$

  to obtain possible optimality conditions.
- One can locally decompose the MPVC into TNLP(x^*, I) for all $I \subseteq \{i \mid G_i(x^*) = H_i(x^*) = 0\}$

$$\begin{aligned} \min_x f(x) \quad \text{s.t.} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & G_i(x) \leq 0, \quad H_i(x) \geq 0 \quad \forall i \in \{i \mid G_i(x^*) < 0\} \cup I, \\ & H_i(x) = 0 \quad \forall i \in \{i \mid G_i(x^*) > 0\} \cup I^c \end{aligned}$$

- One can use a nonsmooth reformulation

$$\varphi(a, b) = \max\{-b, \min\{a, b\}\} \leq 0.$$

Cardinality Constraints

Cardinality Constrained Optimization Problems

- A cardinality constrained problem is of the form

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0, \quad \|x\|_0 \leq \kappa$$

with the "sparsity"-term

$$\|x\|_0 := \text{number of nonzero components of } x.$$

- These problems appear e.g. in
  - portfolio optimization,
  - compressed sensing,
  - feature selection,
  - communications engineering,
  - truss design.

Example: Portfolio Optimization

Consider a basic portfolio selection problem

$$\min_x x^T Q x \quad \text{s.t.} \quad \begin{aligned} &\mu^T x \geq \rho, \\ &e^T x = 1, \\ &0 \leq x_i \leq u_i \quad \forall i = 1, \dots, n, \\ &\|x\|_0 \leq \kappa, \end{aligned}$$

where

- Q is the covariance matrix of the n possible assets,
- μ is the expected revenue,
- $e = (1, \dots, 1)^T$.

Relation to Sparse Optimization Problems

TECHNISCHE UNIVERSITÄT DARMSTADT

For some weight $\rho > 0$ consider

$$\min_x f(x) + \rho \|x\|_0 \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0.$$

Relation to Sparse Optimization Problems

TECHNISCHE UNIVERSITÄT DARMSTADT

For some weight $\rho > 0$ consider

$$\min_x f(x) + \rho \|x\|_0 \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0.$$

- Sparse $\rightarrow$ Cardinality: A solution x^* of the sparse problem is always a solution of the cardinality constrained problem with $\kappa := \|x^*\|_0$.

Relation to Sparse Optimization Problems

TECHNISCHE UNIVERSITÄT DARMSTADT

For some weight $\rho > 0$ consider

$$\min_x f(x) + \rho\|x\|_0 \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0.$$

- Sparse $\rightarrow$ Cardinality: A solution x^* of the sparse problem is always a solution of the cardinality constrained problem with $\kappa := \|x^*\|_0$.
- Cardinality $\not\rightarrow$ Sparse: Consider the following example by O. Burdakov

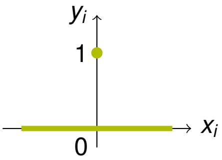
$$f(x) = \|Ax - b\|_2^2, \quad \text{where} \quad A = \begin{pmatrix} 0 & 3 & -3 \\ 3 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}.$$

There is no $\rho \geq 0$ for which one obtains the same solution as for $\kappa = 2$.

Continuous Reformulation using Complementarity Constraints

- **Idea**: Introduce binary variables y counting the zeros of x:

$$x \circ y = 0, \quad \|x\|_0 \leq n - e^T y \leq \kappa.$$

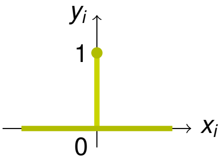


Continuous Reformulation using Complementarity Constraints

- **Idea**: Introduce binary variables y counting the zeros of x:

$$x \circ y = 0, \quad \|x\|_0 \leq n - e^T y \leq \kappa.$$

Relax the binary variables to $y \in [0, e]$.



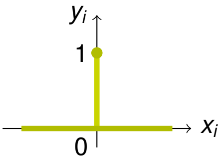
Continuous Reformulation using Complementarity Constraints

- **Idea**: Introduce binary variables y counting the zeros of x:

$$x \circ y = 0, \quad \|x\|_0 \leq n - e^T y \leq \kappa.$$

Relax the binary variables to $y \in [0, e]$.

- **Advantages**: No bounds on x needed; relaxation possible; solves the "right" problem.
- **Disadvantages**: Introduces a complementarity-type condition.



Relation between the Original Problem and the Continuous Reformulation

TECHNISCHE UNIVERSITÄT DARMSTADT

Exercise

What is the relation between continuous reformulation

$$\min_{x,y} f(x) \quad s.t. \quad g(x) \leq 0, \quad h(x) = 0,$$
$$0 \leq y \leq e, \quad e^T y \geq n - \kappa,$$
$$x \circ y = 0.$$

and the original problem

$$\min_{x} f(x) \quad s.t. \quad g(x) \leq 0, \quad h(x) = 0, \quad \|x\|_0 \leq \kappa?$$

Relation between the Original Problem and the Continuous Reformulation (continued)

Global Solutions:

- x^* is a global solution of the original problem if and only if there exists y^* such that (x^*, y^*) is a global solution of the continuous reformulation.

Relation between the Original Problem and the Continuous Reformulation (continued)

TECHNISCHE UNIVERSITÄT DARMSTADT

Global Solutions:

- x^* is a global solution of the original problem if and only if there exists y^* such that (x^*, y^*) is a global solution of the continuous reformulation.

Local Solutions:

- If x^* is a local solution of the original problem, then for every y such that (x^*, y) is feasible, it is a local solution of the continuous reformulation.

TECHNISCHE UNIVERSITÄT DARMSTADT

Relation between the Original Problem and the Continuous Reformulation (continued)

Global Solutions:

- x^* is a global solution of the original problem if and only if there exists y^* such that (x^*, y^*) is a global solution of the continuous reformulation.

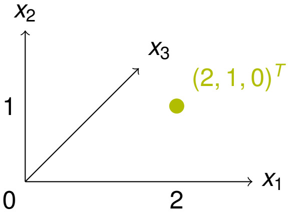
Local Solutions:

- If x^* is a local solution of the original problem, then for every y such that (x^*, y) is feasible, it is a local solution of the continuous reformulation.
- If (x^*, y^*) is a local solution of the continuous reformulation and $\|x^*\|_0 = \kappa$, then x^* is a local solution of the original problem.

Example for Spurious Local Solutions

Consider the 3-dimensional cardinality constrained problem

$$\min_x \|x - (2,1,0)^T\|_2^2 \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$



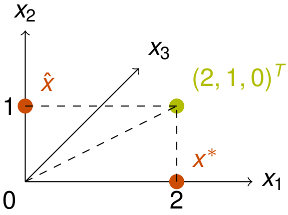
Example for Spurious Local Solutions

TECHNISCHE UNIVERSITÄT DARMSTADT

Consider the 3-dimensional cardinality constrained problem

$$\min_x \|x - (2,1,0)^T\|_2^2 \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$

The global minimizer is $x^* = (2,0,0)^T$ and the local minimizer $\hat{x} = (0,1,0)^T$.

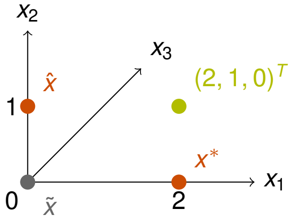


Example for Spurious Local Solutions

Consider the 3-dimensional cardinality constrained problem

$$\min_x \|x - (2,1,0)^T\|_2^2 \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$

The global minimizer is $x^* = (2,0,0)^T$ and the local minimizer $\hat{x} = (0,1,0)^T$.

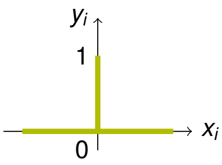


However, the continuous relaxation has additional local minimizer such as $(\tilde{x}, \tilde{y})$ with $\tilde{x} = (0,0,0)^T$ and $\tilde{y} = (1,1,0)$.

Half and Full Complementarity Formulation

- The continuous reformulation results in

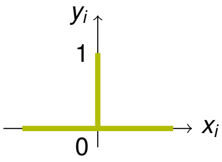
$$\min_{x,y} f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0,$$
$$0 \leq y \leq e, \quad e^T y \geq n - \kappa,$$
$$x \circ y = 0.$$



Half and Full Complementarity Formulation

- The continuous reformulation results in

$$\min_{x,y} f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0,$$
$$0 \leq y \leq e, \quad e^T y \geq n - \kappa,$$
$$x \circ y = 0.$$



- The full complementarity formulation is given by

$$\min_{x^+ - x^-, y} f(x^+ - x^-) \quad \text{s.t.} \quad g(x^+ - x^-) \leq 0, \quad h(x^+ - x^-) = 0,$$
$$y \leq e, \quad e^T y \geq n - \kappa,$$
$$x^+, x^- \geq 0, \quad y \geq 0, \quad (x^+ + x^-) \circ y = 0.$$

Why should you not treat the continuous reformulation as an MPCC?

- In all points of interest MPCC-LICQ and MPCC-MFCQ are violated.

Why should you not treat the continuous reformulation as an MPCC?

- In all points of interest MPCC-LICQ and MPCC-MFCQ are violated.
- MPCC theory generates conditions on the gradients with respect to y which are either violated or meaningless.

Why should you not treat the continuous reformulation as an MPCC?

TECHNISCHE UNIVERSITÄT DARMSTADT

- In all points of interest MPCC-LICQ and MPCC-MFCQ are violated.
- MPCC theory generates conditions on the gradients with respect to y which are either violated or meaningless.
- The analogues of M-, C- and W-stationarity coincide.

Why should you not treat the continuous reformulation as an MPCC?

- In all points of interest MPCC-LICQ and MPCC-MFCQ are violated.
- MPCC theory generates conditions on the gradients with respect to y which are either violated or meaningless.
- The analogues of M-, C- and W-stationarity coincide.
- All local minima of the continuous reformulation are S-stationary (= KKT points) under a Guignard-type CQ.

Why should you not treat the continuous reformulation as an MPCC?

- In all points of interest MPCC-LICQ and MPCC-MFCQ are violated.
- MPCC theory generates conditions on the gradients with respect to y which are either violated or meaningless.
- The analogues of M-, C- and W-stationarity coincide.
- All local minima of the continuous reformulation are S-stationary (= KKT points) under a Guignard-type CQ.
- The Scholtes relaxation converges to S-stationary points under an MFCQ-type CQ.

Chance Constraints

Optimization Problems with Chance Constraints

TECHNISCHE UNIVERSITÄT DARMSTADT

An optimization problems with chance constraints is of the form

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0,$$
$$P(G(x,\xi) \leq 0) \geq \eta,$$

where

- $\eta \in [0,1]$ is the level of confidence, e.g. $\eta = 0.9,\ 0.95,\ 0.99,$
- ξ is a random variable with finitely many possible realization $\xi_1 \dots, \xi_N$ and known probabilities $p_1, \dots, p_N$.

Example: Portfolio Optimization

TECHNISCHE UNIVERSITÄT DARMSTADT

Assume that

- we can invest in n assets,
- N possible scenarios i can occur with probabilities p_i and prognosed returns $\xi_i \in \mathbb{R}^n$,
- we want to invest a budget of 1 to maximize the expected return
- but limit the probability of having losses to $1 - \eta \in (0, 1)$.

We have to solve

$$\max_x \sum_{i=1}^{N} p_i \xi_i^T x \quad \text{s.t.} \quad x \geq 0, \quad \sum_{i=1}^{n} x_i = 1,$$
$$P(\xi^T x \geq 0) \geq \eta$$

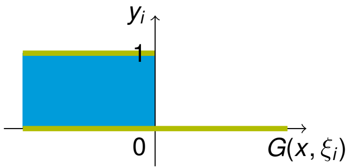
Reformulation of Chance Constraints

The chance constrained problem

$$\min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad h(x) = 0, \quad P(G(x,\xi) \leq 0) \geq \eta,$$

can equivalently be rewritten as

$$\begin{aligned} \min_{x,y} f(x) \quad \text{s.t.} \quad & g(x) \leq 0, \quad h(x) = 0, \\ & y \in \{0,1\}^N, \\ & p^T y \geq \eta, \\ & y_i G(x,\xi_i) \leq 0. \end{aligned}$$



Similarly to cardinality constraints, we can relax the binary variables to $y \in [0,1]^N$.

What should you remember?

TECHNISCHE UNIVERSITÄT DARMSTADT

- There are many classes of optimization problems with similar structures/problems as MPCCs.
- Reformulating them as an MPCC is not always the best approach.
- But one can use ideas from MPCCs to handle these classes such as
  - decomposition of the feasible set into easier pieces,
  - nonsmooth reformulation,
  - using normal cones to obtain conditions on multipliers.

What did we do this week?