A New Relaxation Method for Mathematical Program with Complementarity Constraint

T.Migot[1]
tangi.migot@insa-rennes.fr
Joint work with J.-P. Dussault[2] and M.Haddou[1]

INFORMS 2016 Annual Meeting - Nashville

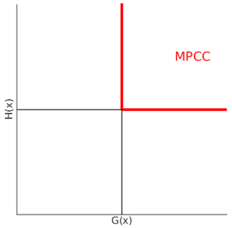
[1]IRMAR-INSA, Rennes, FRANCE
[2]BISOUS, Sherbrooke, QUÉBEC

Mathematical Program with Complementarity Constraint

f, h, g, G, H are continuously differentiable maps.

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ &\text{s.t. } h(x) = 0,\ g(x) \leq 0 \quad , \\ &\qquad 0 \leq G(x) \perp H(x) \geq 0 \end{aligned} \tag{MPCC}$$

where $G(x) \perp H(x)$ stands for $G_i(x)H_i(x) = 0$
$(G_i(x) > 0 \Longrightarrow H_i(x) = 0$ and $H_i(x) > 0 \Longrightarrow G_i(x) = 0)$.



Feasible set of $0 \leq G(x) \perp H(x) \geq 0$

Applications :

- bilevel programming
- optimal control
- ...

Major difficulty :

Classical CQs, fail to hold in general $\Longrightarrow$ no KKT.

Example

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \; & x_1^2 + x_2^2 - x_3 \\ s.t. \; & -4x_1 + x_3 \leq 0, \\ & -4x_2 + x_3 \leq 0, \\ & 0 \leq x_1 \perp x_2 \geq 0 \end{aligned}$$

Obviously the point $(0,0,0)^T$ *is the global minimum. There exists multipliers* $\lambda^{g_1}, \lambda^{g_2}, \lambda^G, \lambda^H, \lambda^{\perp} = (1, 0, -4, 0, 0)$ *but none with the correct signs regarding the KKT conditions.*

Major difficulty :

Classical CQs, fail to hold in general $\Longrightarrow$ no KKT.

Example

$$\begin{aligned} &\min_{x \in \mathbb{R}^2} x_1^2 + x_2^2 - x_3 \\ &s.t. \ -4x_1 + x_3 \leq 0, \\ &\quad\quad -4x_2 + x_3 \leq 0, \\ &\quad\quad 0 \leq x_1 \perp x_2 \geq 0 \end{aligned}$$

Obviously the point $(0,0,0)^T$ *is the global minimum. There exists multipliers* $\lambda^{g_1}, \lambda^{g_2}, \lambda^G, \lambda^H, \lambda^{\perp} = (1, 0, -4, 0, 0)$ *but none with the correct signs regarding the KKT conditions.*

What is a stationary point in the MPCC sense ?

MPCC-Lagrangian function of (MPCC) as

$$\mathcal{L}_{MPCC}(x,\lambda) = f(x) + g(x)^T\lambda^g + h(x)^T\lambda^h - G(x)^T\lambda^G - H(x)^T\lambda^H$$

$$\mathcal{I}^{00} := \{i \mid G_i(x) = 0, H_i(x) = 0\}$$
$$\mathcal{I}^{+0} := \{i \mid G_i(x) > 0, H_i(x) = 0\}$$
$$\mathcal{I}^{0+} := \{i \mid G_i(x) = 0, H_i(x) > 0\}$$

Definition

x^* feasible for (MPCC) is said

- *Weak-stationary if there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{p+q+2m}$ such that*

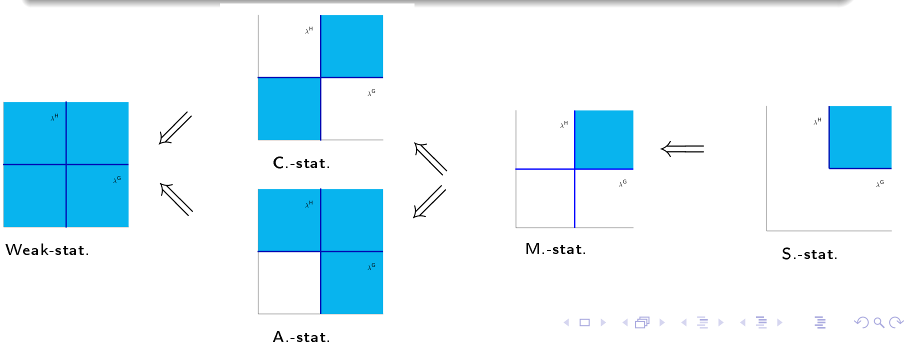
$$\nabla_x \mathcal{L}_{MPCC}(x^*, \lambda^g, \lambda^h, \lambda^G, \lambda^H) = 0,$$
$$\lambda^g_{\mathcal{I}_g} \geq 0,\ \lambda^G_{\mathcal{I}^{+0}} = 0,\ \lambda^H_{\mathcal{I}^{0+}} = 0$$

Definition

x^* feasible for (MPCC) is said

- *C.-stationary :* $\lambda_i^G \lambda_i^H \geq 0$.
- *A.-stationary :* $\lambda_i^G \geq 0$ *or* $\lambda_i^H \geq 0$.
- *M.-stationary : either* $\lambda_i^G > 0$, $\lambda_i^H > 0$ *or* $\lambda_i^G \lambda_i^H = 0$.
- *S.-stationary :* $\lambda_i^G \geq 0$, $\lambda_i^H \geq 0$.

for all $i \in \mathcal{I}^{00} := \{i \mid G_i(x^*) = H_i(x^*) = 0\}$.



Necessary optimality conditions for (MPCC) :

Theorem (Flegel-Kanzow, 06')

A local minimum of (MPCC) that satisfies MPCC-GCQ or any stronger MPCC-CQ is an M-stationary point.

- A classical KKT-point is an S-starionary point.
- We will not get into the details of MPCC-CQs here.

Goal/Motivation :

- Numerical methods should converge to M-stationary points

Relax the constraint : $0 \leq G(x) \perp H(x) \geq 0$

- Pro : Improved regularity (= satisfy a CQ)
- Con : Convergence properties ?

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ &\text{s.t } h(x) = 0,\ g(x) \leq 0, \\ &\qquad G(x) \geq 0,\ H(x) \geq 0, \\ &\qquad \Phi_{t_k}(G(x), H(x)) \leq 0, \end{aligned} \qquad (Relax_t)$$

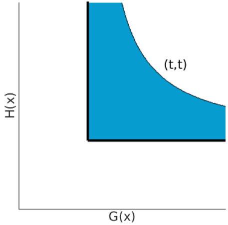
It is also possible to relax the positivity constraints

$$G_i(x) \geq -t_k,\ H_i(x) \geq -t_k\ .$$

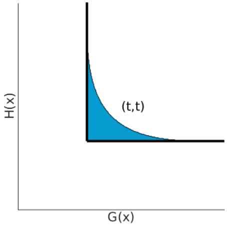
Relax the constraint : $0 \leq G(x) \perp H(x) \geq 0$

- Pro : Improved regularity (= satisfy a CQ)
- Con : Convergence properties ?

Relaxation methods that converge to C-stationary points ($t \downarrow 0$) :



Scheel-Scholtes,2000

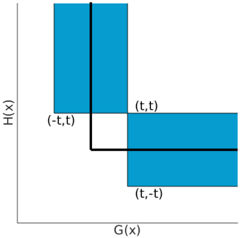


Steffensen-Ulbrich,2010

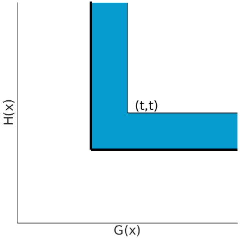
Relax the constraint : $0 \leq G(x) \perp H(x) \geq 0$

- Pro : Improved regularity (= satisfy a CQ)
- Con : Convergence properties ?

Relaxation methods that converge to M-stationary points ($t \downarrow 0$) :



Kadrani-Dussault-Benchakroun,2009



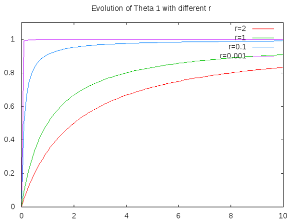
Kanzow-Schwarz,2011

Butterfly relaxations schemes

$$\min_{x\in\mathbb{R}^n} f(x) \text{ s.t } h(x)=0,\ g(x)\leq 0,$$

$$G(x)\geq -\alpha(r,t),\ H(x)\geq -\alpha(r,t), \quad (Butterfly_{t,r})$$

$$(H(x)-t\theta_r(G(x)))(G(x)-t\theta_r(H(x)))\leq 0$$



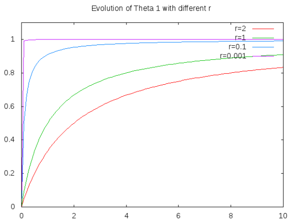
$\theta_r(x):\mathbb{R}\to]-\infty,1]$ are : C^2, increasing, concave, $\theta_r(0)=0$,
$\theta_r(<0)<0$ and
$\lim_{r\downarrow 0}\theta_r(>0)=1$

Butterfly relaxations schemes

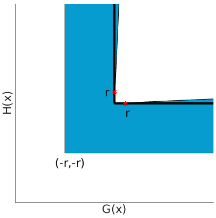
$$\min_{x\in\mathbb{R}^n} f(x) \text{ s.t } h(x)=0,\ g(x)\leq 0,$$

$$G(x)\geq -\alpha(r,t),\ H(x)\geq -\alpha(r,t), \quad (Butterfly_{t,r})$$

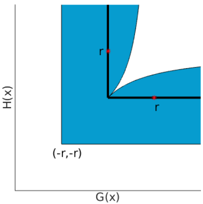
$$(H(x)-t\theta_r(G(x)))(G(x)-t\theta_r(H(x)))\leq 0$$



$\theta_r(x):\mathbb{R}\to]-\infty,1]$ are : C^2, increasing, concave, $\theta_r(0)=0$, $\theta_r(<0)<0$ and $\lim_{r\downarrow 0}\theta_r(>0)=1$



Butterfly $t=r^{3/2}$



Butterfly $t=r$

Convergence Theorem

1. $\{t_k, r_k\} \downarrow_{k\to\infty} 0$ such that for all $k \in \mathbb{N}$: $r_k > 0$ and $t_k\theta'(0) \leq r_k$.
2. $\{x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k\}$ a sequence of stationary (KKT-) points of $(Butterfly_{t,r})$ for all $k \in \mathbb{N}$ with $x^k \to x^*$.
3. MPCC-LICQ holds at x^*

Theorem

x^* is an A-stationary point.
Furthermore if $t_k = o(r_k)$ for k sufficiently large. Then, x^* is an M-stationary point.

Convergence Theorem

1. $\{t_k, r_k\} \downarrow_{k\to\infty} 0$ such that for all $k \in \mathbb{N}$: $r_k > 0$ and $t_k\theta'(0) \leq r_k$.
2. $\{x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k\}$ a sequence of stationary (KKT-) points of $(Butterfly_{t,r})$ for all $k \in \mathbb{N}$ with $x^k \to x^*$.
3. MPCC-LICQ holds at x^*

Theorem

x^* is an A-stationary point.
Furthermore if $t_k = o(r_k)$ for k sufficiently large. Then, x^* is an M-stationary point.

Existence of stationary point : If MPCC-LICQ holds at x^*, some Constraint Qualification holds for the relaxation in a neighborood for some sufficiently small t.

Generic algorithm : Relaxation methods for (MPCC).

Data: starting vector x^0; initial relaxation parameter t_0; update parameter $(\sigma_t, \sigma_r) \in (0,1)^2$ and p_{min} the minimum parameter value, ϵ the precision tolerance ;

Begin ;

Set $k := 0$;

while $\max(t_k, r_k) > p_{\min}$ ***and*** *min_local_crit* $> \epsilon$ **do**

  x^{k+1} solution of R_{t_k, r_k} with x^k initial point;

  $(t_{k+1}, r_{k+1}) := (t_k \sigma_t, r_k \sigma_r)$;

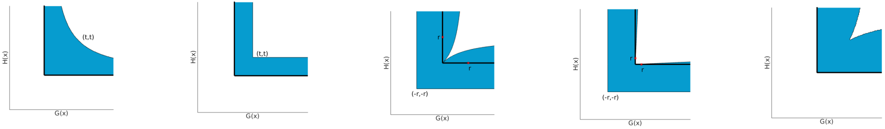
return: f_{opt} the optimal value at the solution x_{opt} or a decision of infeasibility or unboundedness.

Numerical comparison between methods on MacMPEC test problems.

[Sven Leyffer. Macmpec: Ampl collection of mpecs. Argonne National Laboratory. Available at www.mcs.sanl.gov/leyfier/MacMPEC, 2000.]

Relaxation methods used for the comparison :

SS KS $B_{(t=r)}$ $B_{(t=r^{3/2})}$ $B_{(s=t,2t=r)}$



A sensitivity analysis on several values of the parameters $T \in \{100, 25, 10, 5, 1, 0.5, 0.1\}$ and $S \in \{0.1, 0.05, 0.01\}$, which corresponds to t_0 and σ_t as described in Table 1.

Relaxation	NL	SS	KS	Butterfly
t_0	none	T^2	T	T
σ_t	none	S^2	S	S

Parameter links among the methods so that the starting "area" is similar.

We take into account three criteria :

a) Feasibility of the last relaxed non-linear program: $\max(-g(x), |h(x)|, -\Phi(x)) \leq 10^{-7}$

b) Feasibility of the complementarity constraint: $\min(G(x), H(x)) \leq \sqrt{10^{-7}}$

(c) The complementarity between the Lagrange multipliers and the constraints of the last relaxed non-linear program.

Results using SNOPT to solve the non-linear programs.

101 pb						
snopt	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t,2t=r)}$	$B_{(t=r^{3/2})}$
best	92.1	94.1	94.1	96.0	93.1	95.0
average	92.1	90.4	90.3	91.7	89.4	91.6
worst	92.1	83.2	86.1	87.1	86.1	87.1

min : % worst set of parameter. **average** : average % of success. **max** : % best set of parameter

Definition of ϵ-stationary points

$$\left\| \nabla \mathcal{L}_R(x, \lambda^g, \lambda^h, \lambda^G, \lambda^H, \lambda^\Phi) \right\|_\infty \leq \epsilon_k$$

with

$$|h_i(x)| \leq \epsilon_k, \ \forall i \in \{1, \dots, m\}$$

$$g_i(x) \leq \epsilon_k, \ \lambda_i^g \geq 0, \ \left|\lambda_i^g g_i(x)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, p\}$$

$$G_i(x) + \alpha(r_k, t_k) \geq -\epsilon_k, \ \lambda_i^G \geq 0, \ \left|\lambda_i^G (G_i(x) + r_k)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, q\}$$

$$H_i(x) + \alpha(r_k, t_k) \geq -\epsilon_k, \ \lambda_i^H \geq 0, \ \left|\lambda_i^H (H_i(x) + r_k)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, q\}$$

$$\Phi^B_{\hat{t}_k,i}(x) \leq \epsilon_k, \ \lambda_i^\Phi \geq 0, \ \left|\lambda_i^\Phi \Phi^B_{\hat{t}_k,i}(x)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, q\}$$

Christian Kanzow and Alexandra Schwartz.
The Price of Inexactness: Convergence Properties of Relaxation Methods for Mathematical Programs with Complementarity Constraints Revisited.
Mathematics of Operations Research, 40(2):253–275, may 2015.

Definition of ϵ-stationary points

$$\left\|\nabla \mathcal{L}_R(x, \lambda^g, \lambda^h, \lambda^G, \lambda^H, \lambda^\Phi)\right\|_\infty \leq \epsilon_k$$

with

$$|h_i(x)| \leq \epsilon_k, \ \forall i \in \{1, \dots, m\}$$

$$g_i(x) \leq \epsilon_k, \ \lambda_i^g \geq 0, \ \left|\lambda_i^g g_i(x)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, p\}$$

$$G_i(x) + \alpha(r_k, t_k) \geq -\epsilon_k, \ \lambda_i^G \geq 0, \ \left|\lambda_i^G (G_i(x) + r_k)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, q\}$$

$$H_i(x) + \alpha(r_k, t_k) \geq -\epsilon_k, \ \lambda_i^H \geq 0, \ \left|\lambda_i^H (H_i(x) + r_k)\right| \leq \epsilon_k \ \forall i \in \{1, \dots, q\}$$

$$\Phi^B_{\hat{t}_k,i}(x) \leq 0, \ \lambda_i^\Phi \geq 0, \ \left|\lambda_i^\Phi \Phi^B_{\hat{t}_k,i}(x)\right| \leq 0 \ \forall i \in \{1, \dots, q\}$$

Christian Kanzow and Alexandra Schwartz.
The Price of Inexactness: Convergence Properties of Relaxation Methods for Mathematical Programs with Complementarity Constraints Revisited.
Mathematics of Operations Research, 40(2):253–275, may 2015.

ϵ-Convergence Theorem

1. $\{t_k, r_k\} \downarrow_{k\to\infty} 0$ such that for all $k \in \mathbb{N}:\ r_k > 0$ and $t_k\theta'(0) \leq r_k$.
2. $\{x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k\}$ a sequence of ϵ-stationary (KKT-) points of $(Butterfly_{t,r})$ for all $k \in \mathbb{N}$ with $x^k \to x^*$.
3. MPCC-LICQ holds at x^*

Theorem

Assume that $t_k = o(r_k)$ for $k \in \mathbb{N}$ sufficiently large and $\epsilon = o(r_k)$. Then, x^* is an M-stationary point.

Conclusions :

- A new family of relaxation schemes that extend existing methods
- Best known theoretical results for these methods
- Encouraging perspectives from the numerical results.

Perspectives :

- Convergence of more realistic ϵ-stationary sequences.
- Implementation of the non-linear solver to compute such sequences.
- Intelligent updating strategy.

Michael L Flegel and Christian Kanzow.
A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints.
Springer, 2006.

Abdeslam Kadrani, Jean-Pierre Dussault, and Abdelhamid Benchakroun.
A new regularization scheme for mathematical programs with complementarity constraints.
SIAM Journal on Optimization, 20(1):78–103, 2009.

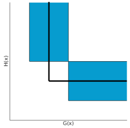
Christian Kanzow and Alexandra Schwartz.
A New Regularization Method for Mathematical Programs with Complementarity Constraints with Strong Convergence Properties.
SIAM Journal on Optimization, 23(2):770–798, apr 2013.

Relaxation Methods for MPCC

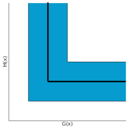
Mathematical Program with Complementarity Constraint

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ &\text{s.t. } h(x) = 0,\ g(x) \leq 0 \\ &\qquad 0 \leq G(x) \perp H(x) \geq 0 \end{aligned} \qquad \text{(MPCC)}$$

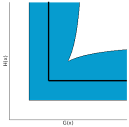
Relaxation methods that converge to M-stationary points :



Kadrani-Dussault-Benchakroun,2009



Kanzow-Schwarz,2011



Dussault-Haddou-Migot,2016