Optimality conditions for mathematical programs with equilibrium constraints via variational analysis

Jane Ye

University of Victoria, Canada

Lecture 1 at the Forum on Developments and Origins of Operations Research
November 23, 2021
Organizers: The Mathematical Programming Branch of OR Society of China
& Southern University of Science and Technology

Outline

- Review of constraint qualifications and necessary optimality conditions for nonlinear programs
- Mathematical program with complementarity constraints
- Mathematical program with variational inequality constraints

Nonlinear program

Standard nonlinear program:

$$\begin{array}{lll} \min & f(x) & \\ s.t. & h_i(x) = 0 & i = 1, \ldots, p, \\ & g_i(x) \le 0 & i = 1, \ldots, q, \end{array}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $h_i : \mathbb{R}^n \to \mathbb{R}, g_i : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable.

Karush-Kuhn-Tucker condition

KKT condition (Kuhn and Tucker 1951, Karush 1934):

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^g \nabla g_i(\bar{x}) = 0,$$
$$\underbrace{\lambda^g \geq 0}_{\text{sign condition}} \qquad \underbrace{\lambda_i^g g_i(\bar{x}) = 0 \quad i = 1, \ldots, q}_{\text{complementary slackness condition}} \qquad .$$

Let $I_g(\bar{x}) := \{i : g_i(\bar{x}) = 0\}$ be the index set of all active constraints. KKT condition can also be written as:

$$0 = \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i \in I_g(\bar{x})} \lambda_i^g \nabla g_i(\bar{x})$$
$$\lambda^g \geq 0.$$

In general a constraint qualification is needed for the KKT condition to hold at a local minimum.

Fritz John condition

Fritz John condition (Fritz John 1948): For any local minimum $\bar{x}$, there exist $\lambda_0 \geq 0, \lambda^g \geq 0, \lambda^h$ NOT ALL equal to zero such that

$$0 = \lambda_0 \nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i \in I_g(\bar{x})} \lambda_i^g \nabla g_i(\bar{x}).$$

- Case I: $\lambda_0 = 0$ and $(\lambda^g, \lambda^h) \neq (0, 0)$. In this case we call (λ^g, λ^h) an abnormal multiplier.
- Case II: $\lambda_0 > 0$. In this case FJ condition becomes KKT condition.

NNAMCQ/PLICQ

- FJ condition says that either case I or case II will occur.
- The following condition is equivalent to saying that Case I of FJ condition does not hold:

$$\left\{ \begin{array}{l} \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i \in I_g(\bar{x})} \lambda_i^g \nabla g_i(\bar{x}) = 0 \\ \lambda^g \geq 0 \end{array} \right. \Longrightarrow (\lambda^h, \lambda^g) = 0.$$

  We call this condition Positive Linearly Independent Constraint Qualification (PLICQ) or No Nonzero Abnormal Multiplier Constraint Qualification (NNAMCQ).
- By FJ condition, PLICQ/NNAMCQ $\Rightarrow$ Case II of FJ condition $\Longleftrightarrow$ KKT condition holds at $\bar{x}$.

Magasarian Fromovitz constraint qualification (MFCQ)

- By using an alternative theorem (Motzkin's theorem), PLICQ is equivalent to the Mangasarian-Fromovitz constraint qualification (MFCQ):

$$\nabla h_1(\bar{x}), \ldots, \nabla h_p(\bar{x}) \text{ are linearly independent,}$$

$$\exists d \in R^n \text{ such that } \begin{cases} \nabla h_i(\bar{x}) \cdot d = 0 & \forall i = 1, \ldots, p, \\ \underbrace{\nabla g_i(\bar{x}) \cdot d < 0}_{\iff g_i'(\bar{x}; d) < 0} & \text{if } g_i(\bar{x}) = 0. \end{cases}$$

- MFCQ implies that relative interior of the feasible region is nonempty!
- MFCQ is equivalent to the boundedness of the multipliers.

Usual constraint qualifications

- Liner independence CQ: the gradient vectors $\nabla h_i(\bar{x}), \nabla g_i(\bar{x})(i \in I_g(\bar{x}))$ are linearly independent.
- Slater Condition: g_i is convex and there is a Slater point x_0 such that $h_i(x_0) = 0(i = 1, \ldots, p), g_i(x_0) < 0(i \in I_g(\bar{x}))$.

$$
\begin{array}{ccc}
\text{LICQ} & & \text{Slater Condition} \\
\Downarrow & & \Downarrow \\
& \text{Mangasarian Fromovitz CQ} & \\
& \Downarrow & \\
& \text{KKT condition holds at the local solution } \bar{x} &
\end{array}
$$

Mathematical program with complementarity constraints

$$\text{(MPCC)} \qquad \min \quad f(x)$$
$$\text{s.t.} \quad \underbrace{G(x) \le 0, H(x) \le 0, G(x)^T H(x) = 0}_{\text{complementarity constraints}}$$

where $G : \mathbb{R}^n \to \mathbb{R}^m$, $H : \mathbb{R}^n \to \mathbb{R}^m$. There may also be equality and/or inequality constraints. For simplicity, we will omit them from this talk.

This is a class of very difficult problems since

MFCQ never holds for MPCCs!

Nash Games

- In his 1949 thesis, John Nash introduced the concept of Nash equilibrium. The result, published in 1951, became part of the foundation of the modern game theory.
  Nash believes that people are acting on their own best interests. His insight was that the game would reach it's end after every player independently chose his best response to the other players' best strategies.
- Nash equilibrium for two players: $(\bar{x}, \bar{y})$ is a Nash equilibrium if two players optimize simultaneously:

$$\bar{x} \text{ solves } \min_x f(x, \bar{y}) \qquad \bar{y} \text{ solves } \min_y g(\bar{x}, y)$$

- If there are inequality constraints, by writing down the KKT condition we have a complementarity system due to the complementary slackness condition in the KKT condition.

Stackelberg games

Stackelberg equilibrium (von Stackelberg 1934, English translation 1952):
The leader moves first, and the follower follows. If the leader chooses x, then the follower will choose y to minimize $g(x, y)$.
The leader tries to find $\bar{x}$, anticipating the reaction of the follower so that his objective is optimized.
$(\bar{x}, \bar{y})$ is a Stackelberg equilibrium if it solves the following bilevel program:

$$\min_{x,y} f(x, y) \quad \text{s.t.} \quad y \in \arg\min_{y'} g(x, y').$$

Examples of Leader and Follower relationships: Employer and employee, government and company, wholesaler and retailer, etc.
If there are inequality constraints in the follower's program, then by replacing the follower's program by the KKT condition, one will obtain a MPCC.

Cournot-Nash equilibria

N profit-maximizing firms, q_i = production of the ith firm, $p(\cdot)$ is the price as a function of quantity.
A Cournot-Nash equilibrium is $q^* = (q_1^*, \ldots, q_N^*)$ such that

$$\max q_i p(\sum_{j \neq i} q_j^* + q_i) - c_i(q_i) \qquad \text{s.t. } q_i \geq 0$$

The KKT condition for the ith firm is:

$$0 = \underbrace{p(\sum q_i^*) + q_i^* p'(\sum q_i^*) - c_i'(q_i^*)}_{F_i(q^*)} - \eta_i$$

$$\eta_i \geq 0, q_i^* \eta_i = 0, q_i^* \geq 0$$

If each profit function is concave, then q^* is a Cournot-Nash equilibrium if and only if it solves the complementarity system:

$$F(q) \geq 0, q \geq 0, F(q)^T q = 0$$

Stackelberg-Cournot-Nash equilibria

Now consider the case in which the $N+1$ th firm (the Stackelberg firm) chooses its production level x to maximize its profit, taking into account the reactions of the other N firms and the effects their reactions have on price:

$$\begin{aligned} \max_{x,q} \quad & xp(\sum_{i=1}^{N} q_i + x) - c_{N+1}(x) \\ \text{s.t.} \quad & \underbrace{F(q+x) \geq 0, q \geq 0, F(q+x)^T q = 0}_{\text{complementarity constraints}}, \\ & x \geq 0. \end{aligned}$$

The solution is called a Stackelberg-Cournot-Nash equilibrium.

Difficulties of Solving MPCCs

(MPCC) $\quad \min f(x)$ s.t. $G(x) \leq 0, H(x) \leq 0, G(x)^T H(x) = 0.$

If MPCC is considered as a standard nonlinear program, then MFCQ fails to hold at each feasible point!

Proof: Let $\bar{x}$ be a feasible point of MPCC. Then $\bar{x}$ solves

$$\min G(x)^T H(x) \text{ s.t. } G(x) \leq 0, H(x) \leq 0.$$

The Fritz John condition $\Longrightarrow \exists(\lambda_0, \lambda^G, \lambda^H) \neq 0$ such that

$$0 = \lambda_0 \nabla(G^T H)(\bar{x}) + \sum_{i \in I_G(\bar{x})} \lambda_i^G \nabla G_i(\bar{x}) + \sum_{i \in I_H(\bar{x})} \lambda_i^H \nabla H_i(\bar{x})$$

$$\lambda^G \geq 0, \lambda^H \geq 0.$$

This means that the gradient vectors of the constraints of (MPCC):

$$\nabla(G^T H)(\bar{x}), \nabla G_i(\bar{x})(i \in I_G(\bar{x})), \nabla H_i(\bar{x})(i \in I_H(\bar{x}))$$

are Positively Linearly Dependent. Hence, PLICQ/MFCQ does not hold.

Reformulations of MPCCs

- The complementarity constraint

$$G(x) \leq 0, H(x) \leq 0, G(x)^\top H(x) = 0$$

can be reformulated as a geometric constraint

$$(G(x), H(x)) \in \Omega_C \textbf{ where } \Omega_C := \{(y, z) | 0 \geq y \perp z \leq 0\}$$

Reformulations of MPCCs

- The complementarity constraint

$$G(x) \leq 0, H(x) \leq 0, G(x)^\top H(x) = 0$$

can be reformulated as a geometric constraint

$$(G(x), H(x)) \in \Omega_C \text{ where } \Omega_C := \{(y, z) | 0 \geq y \perp z \leq 0\}$$

- or as a nonsmooth equation constraint

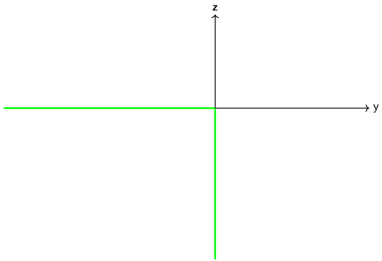
$$\max\{G(x), H(x)\} = 0$$

- But now the problem is not a standard NLP any more.
- So MPCCs is an intrinsic nonsmooth nonconvex optimization problem!

- The complementarity cone

$$\Omega_C := \{(y,z) | 0 \geq y \perp z \leq 0\}$$

is nonconvex.



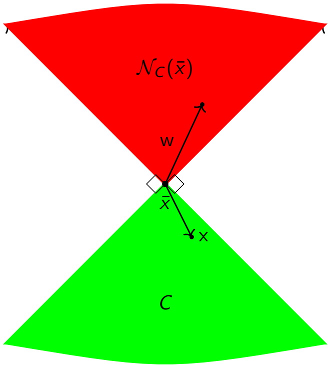
- And the function $\varphi(x) := \max\{G_i(x), H_i(x)\}$ is nonsmooth at $\bar{x}$ when $G_i(\bar{x}) = H_i(\bar{x})$.

We need the tool of Variational Analysis since

Variational Analysis provides tools to study optimization problems involving nonsmooth functions and nonconvex sets

Normal Cone for Convex Sets

If C is a closed and convex set and $\bar{x} \in C$ be given, then the normal cone is $\mathcal{N}_C(\bar{x}) := \{w : \langle w, x - \bar{x} \rangle \leq 0 \quad \forall x \in C\}$.



Optimality conditions

- Let $\bar{x}$ be a local optimal solution of

$$\min f(x) \text{ s.t. } \Phi(x) \in C,$$

where $f, \Phi : \mathbb{R}^n \to \mathbb{R}^m$ are C^1 and C is convex and closed.

- The KKT condition is

$$0 = \nabla f(\bar{x}) + \nabla \Phi(\bar{x})^T \lambda^\Phi, \text{ **for some** } \lambda^\Phi \in \mathcal{N}_C(\Phi(\bar{x}))$$

- If $C = \{0\}$, then the normal cone is the whole space, no sign condition is required $\Rightarrow$ the Lagrange multiplier rule holds.
- If $C = \mathbb{R}^m_-$, then

$$\mathcal{N}_C(\Phi(\bar{x})) = \mathcal{N}_{\mathbb{R}_-}(\Phi_1(\bar{x})) \times \cdots \times \mathcal{N}_{\mathbb{R}_-}(\Phi_m(\bar{x}))$$

and thus the multipliers for inequalities must be nonnegative $\Rightarrow$ and the complementary slackness condition holds $\Longrightarrow$ KKT condition holds.

Normal Cones for Nonconvex sets

- Let C be nonempty and closed, and $\bar{x} \in C$ be given. The regular normal cone is

$$\widehat{\mathcal{N}}_C(\bar{x}) := \{w : \langle w, x - \bar{x} \rangle \leq o\|x - \bar{x}\| \quad \forall x \in C\}.$$

- The limiting/Mordukhovich normal cone to C at $\bar{x}$ is

$$\mathcal{N}_C(\bar{x}) := \{\lim_{k\to\infty} w^k : \exists\{x^k\} : \lim_{k\to\infty} x^k = \bar{x}, w^k \in \widehat{\mathcal{N}}_C(x^k)\}$$

- and the Clarke normal cone is the closure of the convex hull of the limiting normal cone:

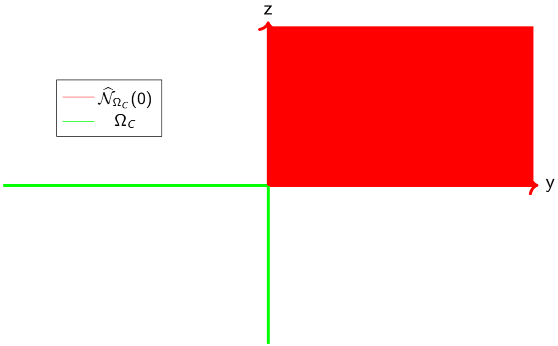
$$\mathcal{N}^c_C(\bar{x}) = clco\mathcal{N}_C(\bar{x})$$

- The following inclusions may be strict.

$$\widehat{\mathcal{N}}_C(\bar{x}) \subseteq \mathcal{N}_C(\bar{x}) \subseteq \mathcal{N}^c_C(\bar{x}).$$

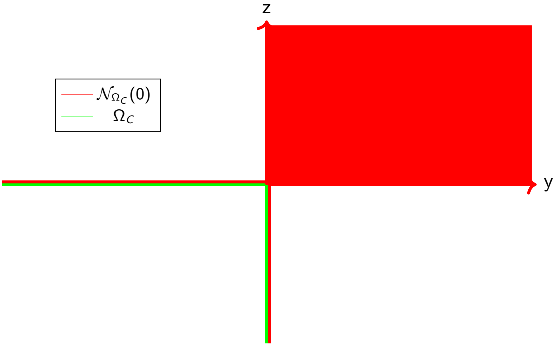
The regular normal cone to Ω_C at $(\bar{y}, \bar{z}) \in \Omega$:

$$\widehat{\mathcal{N}}_{\Omega_C}(\bar{y}, \bar{z}) = \left\{ (u, v) : \begin{array}{ll} u_i = 0 & \text{if } \ \bar{y}_i < 0, \bar{z}_i = 0 \\ v_i = 0 & \text{if } \ \bar{z}_i < 0, \bar{y}_i = 0 \\ u_i \geq 0, v_i \geq 0 & \text{if } \ \bar{y}_i = \bar{z}_i = 0 \end{array} \right\}.$$



The limiting normal cone of Ω_C at $(\bar{y}, \bar{z}) \in \Omega_C$ is

$$\mathcal{N}_{\Omega_C}(\bar{y}, \bar{z}) = \left\{ (u, v) : \begin{array}{lll} u_i = 0 & & \text{if} \quad \bar{y}_i < 0 \\ v_i = 0 & & \text{if} \quad \bar{z}_i < 0 \\ \left\{ \begin{array}{ll} \text{either} & u_i > 0, v_i > 0 \\ \text{or} & u_i v_i = 0 \end{array} \right. & & \text{if} \quad \bar{y}_i = \bar{z}_i = 0 \end{array} \right\}.$$



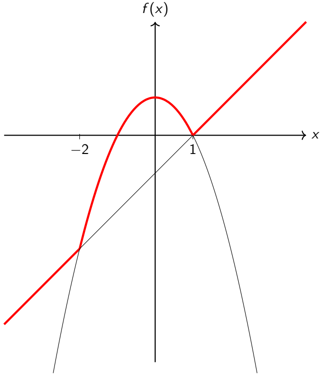
The Clarke normal cone of Ω_C is the whole space at $\bar{y}_i = \bar{z}_i = 0$.

Clarke generalized gradient

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz. By Rademacher's theorem, f is differentiable almost everywhere. $\partial^c f(x)$, the Clarke generalized gradient of f at x, is the convex hull of the set

$$\left\{ \lim_{k\to\infty} \nabla f(x_k) \;\middle|\; f \text{ is differentiable at } x_k \text{ and } x_k \to x \right\}.$$

- Example: $f(x) := \max\{1 - x^2, x - 1\}$ is nonsmooth if and only if $1 - x^2 = x - 1$ if and only if $x = 1$ or $x = -2$. When $x_k \neq 1, -2$, $\nabla f(x_k) = -2x_k$ or 1.
- Taking limits as $x_k \to 1, -2$ we have $\partial^c f(1) = co\{-2, 1\} = [-2, 1]$. $\partial^c f(-2) = co\{4, 1\} = [1, 4]$.



Example: $f(x) := \max\{1 - x^2, x - 1\}$.
$\partial^c f(1) = co\{-2, 1\} = [-2, 1]$. $\partial^c f(-2) = co\{4, 1\} = [1, 4]$.

Optimality conditions in terms of regular normal cone

Let $\bar{x}$ be a local optimal solution of

$$\min f(x) \textbf{ s.t. } \Phi(x) \in C$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable and C is closed.

If $\nabla\Phi(\bar{x})$ has maximal row rank m

then the KKT type necessary optimality condition holds:

$$0 = \nabla f(\bar{x}) + \nabla\Phi(\bar{x})^T \lambda^\Phi, \quad \lambda^\Phi \in \widehat{\mathcal{N}}_C(\Phi(\bar{x})).$$

We call such an optimality condition a strong (S-) stationary condition.

Proof. Since $\bar{x}$ is a local minimal solution of $\min f(x)$ s.t. $x \in \mathcal{F}$, where $\mathcal{F} := \{x : \Phi(x) \in C\}$ is the feasible region, we have

$$0 \in \nabla f(\bar{x}) + \widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x}).$$

By the change of coordinates formula in Exercise 6.7 of Rockafellar and Wets 1998, since the matrix $\nabla\Phi(\bar{x})$ has maximal rank, we have

$$\widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x}) = \nabla\Phi(\bar{x})^T \widehat{\mathcal{N}}_C(\Phi(\bar{x})).$$

Hence

$$0 \in \nabla f(\bar{x}) + \nabla\Phi(\bar{x})^T \widehat{\mathcal{N}}_C(\Phi(\bar{x})).$$

This means

$$0 = \nabla f(\bar{x}) + \nabla\Phi(\bar{x})^T \lambda^\Phi, \quad \lambda^\Phi \in \widehat{\mathcal{N}}_C(\Phi(\bar{x})).$$

Strong stationary condition for MPCC

Applying the optimality condition in terms of regular normal cone to the reformulated MPCC:

$$\begin{array}{ll} \min & f(x) \\ s.t. & (G(x), H(x)) \in \Omega_C \end{array}$$

we obtain the S-optimality condition for MPCC.

Theorem

Let $\bar{x}$ be a local solution of (MPCC). If MPCC LICQ holds at $\bar{x}$, i.e.

$$\nabla G_i(\bar{x})(i \in I_G(\bar{x})), \nabla H_i(\bar{x})(i \in I_H(\bar{x}))$$

are linearly independent, then there exist (λ^G, λ^H) such that

$$0 = \nabla f(\bar{x}) + \nabla G(\bar{x})^T \lambda^G + \nabla H(\bar{x})^T \lambda^H, \quad (\lambda^G, \lambda^H) \in \widehat{\mathcal{N}}_{\Omega_C}(G(\bar{x}), H(\bar{x}))$$

S-stationary condition for MPCC

Since

$$\widehat{\mathcal{N}}_{\Omega_C}(G(\bar{x}), H(\bar{x})) = \left\{ (u,v) : \begin{array}{lll} u_i = 0 & if & G_i(\bar{x}) < 0, H_i(\bar{x}) = 0 \\ v_i = 0 & if & G_i(\bar{x}) < 0, H_i(\bar{x}) = 0 \\ u_i, v_i \geq 0 & if & G_i(\bar{x}) = H_i(\bar{x})_i = 0 \end{array} \right\}.$$

S-stationary condition becomes

$$\begin{aligned} &0 = \nabla f(\bar{x}) + \nabla G(\bar{x})^\top \lambda^G + \nabla H(\bar{x})^\top \lambda^H \\ &\lambda_i^G = 0 \text{ for } i \text{ such that } G_i(\bar{x}) < 0 \\ &\lambda_i^H = 0 \text{ for } i \text{ such that } H_i(\bar{x}) < 0 \end{aligned}$$

$$\lambda_i^G \geq 0, \lambda_i^H \geq 0 \textbf{ if } H_i(\bar{x}) = G_i(\bar{x}) = 0.$$

Optimality conditions in terms of limiting normal cone

Let $\bar{x}$ be a local optimal solution of $\quad \min f(x) \quad \text{s.t. } \Phi(x) \in C$, where $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable and C is closed. Suppose that the map $M(x) := \Phi(x) - C$ is metric subregular at $(\bar{x}, 0) \in \mathrm{gph} M$ (or $M^{-1}(v)$ is calm at $(0, \bar{x})$) which means that the a local error bound holds at $\bar{x}$:

$\exists \mu \geq 0$ and $N(\bar{x})$, a neighborhood of $\bar{x}$ such that

$$dist_{\mathcal{F}}(x) \leq \mu dist_C(\Phi(x)) \quad \forall x \in N(\bar{x}),$$

where $\mathcal{F} := \{x : \Phi(x) \in C\}$.

Then the KKT type necessary optimality condition holds:

$$0 = \nabla f(\bar{x}) + \nabla \Phi(\bar{x})^T \lambda^\Phi, \quad \lambda^\Phi \in \mathcal{N}_C(\Phi(\bar{x})).$$

We call this condition an Mordukhovich (M-) stationary condition.

Clarke's Exact Penalty Principle

By Proposition 2.4.3 of Clarke 1983, the distance function is always an exact penalty function!

Let f be Lipschitz with constant L_f on set S. Suppose $\bar{x} \in C \subseteq S$ is a solution of $\min_{x \in C} f(x)$. Then $\bar{x}$ is also a solution of the exact penalty problem $\min_{x \in S} f(x) + L_f \mathrm{dist}_C(x)$.

Proof: For each $x \in S$, suppose $\tilde{x} = \Pi_C(x)$, the projection onto set C. Then

$$\begin{aligned} f(\bar{x}) &\le f(\tilde{x}) \quad \text{by optimality of } \bar{x} \text{ on } C \\ &= f(x) + (f(\tilde{x}) - f(x)) \\ &\le f(x) + L_f \|\tilde{x} - x\| \quad \text{by Lipschitz continuity of } f \\ &= f(x) + L_f \mathrm{dist}_C(x). \end{aligned}$$

Use an approximate projection if no projection exists.

Proof: Since $\bar{x}$ is a local optimal solution of $\min f(x)$ s.t. $x \in \mathcal{F} := \{x | \Phi(x) \in C\}$, by Clarke's exact penalty principle, it is a local solution of $\min f(x) + L_f \mathrm{dist}_{\mathcal{F}}(x)$. Since the local error bound holds at $\bar{x}$, i.e., $dist_{\mathcal{F}}(x) \leq \mu dist_C(\Phi(x))$ for all x sufficiently close to $\bar{x}$, we have for x sufficiently close to $\bar{x}$,

$$\begin{aligned} f(\bar{x}) &\leq f(x) + L_f d_{\mathcal{F}}(x) \\ &\leq f(x) + L_f \mu dist_C(\Phi(x)) \end{aligned}$$

This means that $\bar{x}$ is a local optimal solution to the problem

$$\min_x f(x) + \gamma dist_C(\Phi(x)),$$

where $\gamma \geq L_f \mu$. By optimality condition and the fact that $\partial d_C(\Phi(x)) = \mathcal{N}_C(\Phi(\bar{x}))$, we have

$$0 = \nabla f(\bar{x}) + \nabla \Phi(\bar{x})^T \lambda^\Phi, \quad \lambda^\Phi \in \mathcal{N}_C(\Phi(\bar{x})).$$

Constraint qualifications for M-stationary condition

We say NNAMCQ holds at $\bar{x}$ if

$$0 = \nabla \Phi(\bar{x})^T \lambda^\Phi, \quad \lambda^\Phi \in \mathcal{N}_C(\Phi(\bar{x})) \Longrightarrow \lambda^\Phi = 0.$$

$\nabla \Phi(\bar{x})$ has maximal row rank m

$\Downarrow$

NNAMCQ Φ is affine and C is finite union of convex polyhedral sets

$\Downarrow$ $\Downarrow$

Error Bound/Metric Subregularity/Calmness

$\Downarrow$

M-stationary condition holds at the local solution $\bar{x}$

M-stationary condition for MPCC

Theorem: Suppose that either G, H are affine or there is no λ^G, λ^H not all equal to zero such that

$$0 = \nabla G(\bar{x})^T \lambda^G + \nabla H(\bar{x})^T \lambda^H, \; (\lambda^G, \lambda^H) \in \mathcal{N}_{\Omega_C}(G(\bar{x}), H(\bar{x})).$$

Then

$$0 = \nabla f(\bar{x}) + \nabla G(\bar{x})^T \lambda^G + \nabla H(\bar{x})^T \lambda^H, \; (\lambda^G, \lambda^H) \in \mathcal{N}_{\Omega_C}(G(\bar{x}), H(\bar{x}))$$

Proof. Reformulate MPCC equivalently as

$$\begin{array}{ll} \min & f(x) \\ s.t. & (G(x), H(x)) \in \Omega_C \end{array}$$

and apply the necessary optimality condition in terms of limiting normal cones.

M-stationary condition

Since $\widehat{\mathcal{N}}_{\Omega_C}(0,0) = \{(u,v) | \text{either } u > 0, v > 0 \text{ or } uv = 0\}$. The M-stationary condition can be written as

$$0 = \nabla f(\bar{x}) + \nabla G(\bar{x})^\top \lambda^G + \nabla H(\bar{x})^\top \lambda^H$$
$$\lambda_i^G = 0 \text{ for } i \text{ such that } G_i(\bar{x}) < 0$$
$$\lambda_i^H = 0 \text{ for } i \text{ such that } H_i(\bar{x}) < 0$$

either $\lambda_i^G > 0, \lambda_i^H > 0$ **or** $\lambda_i^G \lambda_i^H = 0$ **if** $H_i(\bar{x}) = G_i(\bar{x}) = 0$**.**

Constraint qualifications for M-stationary condition of MPCC

We say NNAMCQ holds at $\bar{x}$ if

$$0 = \nabla G(\bar{x})^\top \lambda^G + \nabla H(\bar{x})^\top \lambda^H$$
$$\lambda_i^G = 0 \text{ for } i \text{ s.t . } G_i(\bar{x}) < 0, \lambda_i^H = 0 \text{ for } i \text{ s.t. } H_i(\bar{x}) < 0$$
$$\text{either } \lambda_i^G > 0, \lambda_i^H > 0 \text{ or } \lambda_i^G \lambda_i^H = 0 \text{ if } H_i(\bar{x}) = G_i(\bar{x}) = 0$$

$\Longrightarrow \lambda^G = 0, \lambda^H = 0$

MPCC LICQ

$\Downarrow$

NNAMCQ $\qquad$ $G(x), H(x)$ are affine

$\Downarrow$ $\qquad$ $\Downarrow$

Error Bound/Metric Subregularity/Calmness

$\Downarrow$

M-stationary condition holds at the local solution $\bar{x}$

The Clarke nonsmooth multiplier rule

Let $\bar{x}$ be a local optimal solution of

$$\min f(x) \text{ s.t. } \Phi(x) = 0,$$

where $f, \Phi : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous near $\bar{x}$. Suppose that the problem is calm at $\bar{x}$, i.e. $\bar{x}$ is also a local solution of the penalized problem

$$\min f(x) + \mu \|\Phi(x)\|$$

for some $\mu > 0$. Then, there exists a multiplier λ^Φ such that

$$0 \in \partial^c f(\bar{x}) + \sum_{i=1}^{m} \lambda_i^\Phi \partial^c \Phi_i(\bar{x}).$$

Clarke stationary condition

$\min f(x)$ s.t.$\Phi_i(x) := \max\{G_i(x), H_i(x)\} = 0 \quad i = 1, \ldots, m$. By the Clarke nonsmooth multiplier rule, if the problem is calm at the local optimal solution $\bar{x}$, then $\exists\ \lambda$ such that $0 \in \nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial^c \Phi_i(\bar{x})$. Since

$$\partial^c \Phi_i(\bar{x}) = \begin{cases} \nabla G_i(\bar{x}) & \text{if } H_i(\bar{x}) < 0 \\ \nabla H_i(\bar{x}) & \text{if } G_i(\bar{x}) < 0 \\ \{\alpha \nabla G_i(\bar{x}) + (1-\alpha)\nabla H_i(\bar{x}) : \alpha \in [0,1]\} & \text{if } G_i = H_i = 0, \end{cases}$$

we can find $\alpha_i \in [0,1]$ such that

$$0 = \nabla f(\bar{x}) + \sum_{i=1}^{m} \underbrace{\alpha_i \lambda_i}_{\lambda_i^G} \nabla G_i(\bar{x}) + \sum_{i=1}^{m} \underbrace{(1-\alpha_i)\lambda_i}_{\lambda_i^H} \nabla H_i(\bar{x}).$$

The sign condition for the multipliers is $\lambda_i^G \lambda_i^H = \alpha_i (1-\alpha_i) \lambda_i^2 \geq 0$.

W-,S-,M- and C- stationary condition

A feasible $\bar{x}$ is called a weak stationary point

$$0 = \nabla f(\bar{x}) + \nabla G(\bar{x})^\top \lambda^G + \nabla H(\bar{x})^\top \lambda^H$$
$$\lambda_i^G = 0 \text{ if } G_i(\bar{x}) < 0, \lambda_i^H = 0 \text{ if } H_i(\bar{x}) < 0.$$

$\bar{x}$ is a S-, M-, C- stationary point if it is W-stationary and

$$\underbrace{\lambda_i^G \geq 0, \lambda_i^H \geq 0}_{\text{S Stationary}} \text{ if } H_i(\bar{x}) = G_i(\bar{x}) = 0$$

$$\underbrace{\text{either } \lambda_i^G > 0, \lambda_i^H > 0 \text{ or } \lambda_i^G \lambda_i^H = 0}_{\text{M Stationary}} \text{ if } H_i(\bar{x}) = G_i(\bar{x}) = 0$$

$$\underbrace{\lambda_i^G \lambda_i^H \geq 0}_{\text{C Stationary}} \quad \text{if } H_i(\bar{x}) = G_i(\bar{x}) = 0$$

In general, we have S $\Longrightarrow$ M$\Longrightarrow$ C $\Longrightarrow$ W

- S-stationary condition is very strong, since it is the same as the classical KKT condition when the complementarity constraint is treated as an inequality constraint.
- M-stationary conditions hold under very reasonable conditions. Since the classical KKT conditions may not hold for MPECs, M- and C- stationary conditions are applicable stationary conditions, and many algorithms converge to either M- or C-stationary points.
- Successful applications of MPECs include: the electric power market, traffic, telecommunication networks, revenue management and many other areas of operations research.

More general classes of MPECs

Consider the mathematical program with variational inequality constraint:

$$(MPVIC) \quad \min_{x,y} \quad F(x,y)$$
$$s.t. \quad \langle \phi(x,y), y' - y \rangle \geq 0 \quad \forall y' \in \Gamma,$$

where $\Gamma := \{y | g(y) \leq 0\}$.
If Γ is convex, then (MPVIC) can be equivalent written as

$$(MPEC) \quad \min_{x,y} \quad F(x,y)$$
$$s.t. \quad 0 \in \phi(x,y) + \mathcal{N}_\Gamma(y).$$

MPCC reformulation

- If g is affine or the Slater condition holds, then

$$\mathcal{N}_\Gamma(y) = \nabla g(y)^T \mathcal{N}_{\mathbb{R}^p_-}(g(y)),$$

This suggests considering problem (MPCC)

$$\begin{array}{llll} (MPCC) & \min\limits_{x,y,\lambda} & F(x,y) \\ & s.t. & 0 = \phi(x,y) + \nabla g(y)^T \lambda \\ & & g(y) \le 0, \lambda \ge 0, g(y)^T \lambda = 0. \end{array}$$

Here λ is referred to as a multiplier.

- In the case where there is only a unique multiplier λ for each y, (MPEC) and (MPCC) are equivalent. But if the multiplier is not unique, the two problems are not necessarily equivalent if the local optimality is considered: Dempe and Dutta (2012).

Pseudo-Lipschitz continuity of Aubin and Upper-Lipschitz continuity of Robinson

Definition (Pseudo-Lipschitz continuity/Aubin continuity/, Aubin 1994)

A set-valued map $\mathcal{F}$ is said to be pseudo-Lipschitz continuous at $(\bar{v}, \bar{z}) \in gph\mathcal{F}$ if there exists a constant $\mu \geq 0$, U a neighborhood of $\bar{z}$ and V a neighborhood of $\bar{v}$ such that

$$\mathcal{F}(v) \cap U \subseteq \mathcal{F}(v') + \mu \|v - v'\| \mathbb{B} \quad \forall v, v' \in V,$$

where $\mathbb{B}$ is the closed unit ball.

Definition (Upper-Lipschitz continuity, Robinson 1975)

A set-valued map $\mathcal{F}$ is said to be upper-Lipschitz continuous at $\bar{v}$ if there $\exists\ \mu \geq 0$ and V a neighborhood of $\bar{v}$ such that

$$\mathcal{F}(v) \subseteq \mathcal{F}(\bar{v}) + \mu \|v - \bar{v}\| \mathbb{B} \quad \forall v \in V.$$

Calmness of a set-valued map

Definition (Upper pseudo-Lipschitz continuity (JY and Ye 1997)/calmness (Rockafellar and Wets 1998))

A set-valued map $\mathcal{F}$ *is said to be upper pseudo-Lipschitz continuous/calm at* $(\bar{v}, \bar{z}) \in gph\mathcal{F}$ *if there exists a constant* $\mu \geq 0$, *U a neighborhood of* $\bar{z}$ *and V a neighborhood of* $\bar{v}$ *such that*

$$\mathcal{F}(v) \cap U \subseteq \mathcal{F}(\bar{v}) + \mu \|v - \bar{v}\| \mathbb{B} \quad \forall v \in V.$$

The name "Upper pseudo-Lipschitz continuity" indicates that it is weaker than both upper-Lipschitz continuity and pseudo-Lipschitz continuity.

Theorem 3.1 of JY and Ye (1997)

Let $\bar{z}$ be a local optimal solution of $\quad \min f(z) \quad \text{s.t.} \ 0 \in \Phi(z)$,
where $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ is a set-valued map.
Suppose the perturbed feasible map $\mathcal{F}(v) := \{z | v \in \Phi(z)\}$ is calm/pseudo upper-Lipschitz continuous at $(0, \bar{z})$

there exists a constant $\mu \geq 0$ and U, a neighborhood of $\bar{z}$ and V a neighborhood of 0 such that

$$\mathcal{F}(v) \cap U \subseteq \mathcal{F}(0) + \mu \|v - 0\| \mathbb{B} \quad \forall v \in V.$$

Then M-stationary condition holds: there exists η such that

$$0 \in \nabla f(\bar{z}) + D^* \Phi(\bar{z}, 0)(\eta).$$

Here

$$\xi \in D^* \Phi(\bar{z}, 0)(\eta) \Longleftrightarrow (\xi, -\eta) \in \mathcal{N}_{gph \Phi}(\bar{z}, 0)$$

Proof of Theorem 3.1 of JY and Ye (1997)

$$\min f(z) \text{ s.t. } 0 \in \Phi(z) \Longleftrightarrow \min_{z,v} f(z) \text{ s.t. } v \in \Phi(z), v = 0.$$

Let $z \in \mathcal{F}(v) \cap U$, $v \in V$. Then by calmness there exists $z^* \in \mathcal{F}(0)$ such that $\|z - z^*\| \leq \mu\|v\|$. Hence

$$f(\bar{z}) \quad \leq \quad f(z^*) \leq f(z) + L_f\|z^* - z\| \leq f(z) + L_f\mu\|v\|.$$

Hence $(z, v) = (\bar{z}, 0)$ also locally solve (for $\gamma \geq L_f\mu$):

$$\min_{z,v} f(z) + \gamma\|v\| \text{ s.t. } v \in \Phi(z) \Longleftrightarrow \min_{z,v} f(z) + \gamma\|v\| \text{ s.t. } (z, v) \in gph\Phi.$$

Applying optimality conditions we have

$$0 \in \nabla f(\bar{z}) \times \gamma\mathbb{B}_m + \mathcal{N}_{gph\Phi}(\bar{z}, 0).$$

Therefore we can find $\eta \in \gamma\mathbb{B}_m$ such that

$$(-\nabla f(\bar{z}), -\eta) \in \mathcal{N}_{gph\Phi}(\bar{z}, 0) \Longleftrightarrow -\nabla f(\bar{z}) \in D^*\Phi(\bar{z}, 0)(\eta).$$

M-optimality conditions for (MPEC)

Recall

$$(MPEC) \quad \min_{x,y} \quad F(x,y) \quad \text{s.t. } 0 \in \phi(x,y) + \mathcal{N}_\Gamma(y).$$

Theorem (Theorem 3.2 of JY and Ye (1997))

Let $(\bar{x},\bar{y})$ be solution of (MPEC). If the perturbed feasible map $\Sigma(v) := \{(x,y)|v \in \phi(x,y) + \mathcal{N}_\Gamma(y)\}$ is calm at $(0,(\bar{x},\bar{y}))$, then M-stationary condition holds: there exists η such that

$$0 \in \nabla F(\bar{x},\bar{y}) + \nabla\phi(\bar{x},\bar{y})^T\eta + \{0\} \times D^*\mathcal{N}_\Gamma(\bar{y}, -\phi(\bar{x},\bar{y}))(\eta).$$

Proof Take $z := (x,y)$ and $\Phi(z) := \phi(x,y) + \mathcal{N}_\Gamma(y)$. The result follows from Theorem 3.1 by using the coderivative sum rule:

$$D^*\Phi(\bar{z},0)(\eta) = \nabla\phi(\bar{x},\bar{y})^T\eta + \{0\} \times D^*\mathcal{N}_\Gamma(\bar{y}, -\phi(\bar{x},\bar{y}))(\eta).$$

Comparing two formulations (MPEC) and (MPCC)

- Adam, Henrion and Outrata (2018) has compared (MPEC) and (MPCC) in terms of calmness. They pointed out that very often, the calmness condition for (MPEC) is satisfied at some $(\bar{x}, \bar{y})$ while the one for (MPCC) is not satisfied at $(\bar{x}, \bar{y}, \lambda)$ for a certain multiplier λ.
- In Gfrerer and JY (2017), an example is given for which MPCC-GCQ (Guignard CQ) is violated at $(\bar{x}, \bar{y}, \lambda)$ for any multiplier λ but the calmness holds for (MPEC) at $(\bar{x}, \bar{y})$.
- A verifiable sufficient conditions for calmness of (MPEC) was given in Gfrerer and JY (2017). In Gfrerer and JY (2020), by using directional variational analysis, a sharper necessary optimality condition for (MPEC) are derived.

Reference:

- JY, D.L. Zhu and Q.J. Zhu 1997, Exact penalization and necessary optimality conditions for generalized bilevel programming problems, SIAM Journal on Optimization, vol. 7, 481-507.
  The MFCQ was shown to fail at each feasible point of MPCCs in this paper.
- JY and X.Y. Ye 1997, Necessary optimality conditions for optimization problems with variational inequality constraints, Mathematics of Operations Research, vol. 20, 977-997.
  The calmness for a set-valued map was introduced under the name "pseudo upper-Lipschitz continuity". M-stationary condition was shown to be a necessary optimality condition under the calmness condition in this paper.

- L. Adam, R. Henrion and J. Outrata 2018, On M-stationarity conditions in MPECs and the associated qualification conditions, Math. Program. 168, 229-259.
- H. Gfrerer and JY 2017, New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis, SIAM Journal on Optimization, vol. 27, 842-865.
  A new sufficient condition for calmness that is easier to compute and yet weaker than NNAMCQ of (MPEC) was proposed in this paper.
- H. Gfrerer and JY 2020, New sharp necessary optimality conditions for mathematical programs with equilibrium constraints, Set-Valued and Variational Analsis, vol. 278, 395-426.
  A new necessary optimality condition which is sharper than M-stationary condition for (MPEC) was derived in this paper. This new optimality condition is applicable even when no constraint qualifications holds for the corresponding MPCC.

Thank you!