Topologically relevant stationarity concepts

Oliver Stein

Institute of Operations Research
Karlsruhe Institute of Technology

ICCP 2014
Humboldt-Universität zu Berlin
August 4–8, 2014

Survey

1. The unconstrained smooth case
2. The constrained smooth case
3. Mathematical programs with complementarity constraints
4. Mathematical programs with vanishing constraints

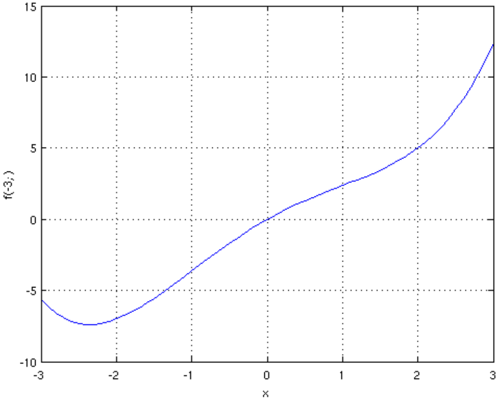
Four reasons to look at stationary points

- Candidates for local minimizers
- Design of homotopy methods
- Understanding the problem structure (Morse theory)
- Convergence results for KKT points

Four reasons to look at stationary points

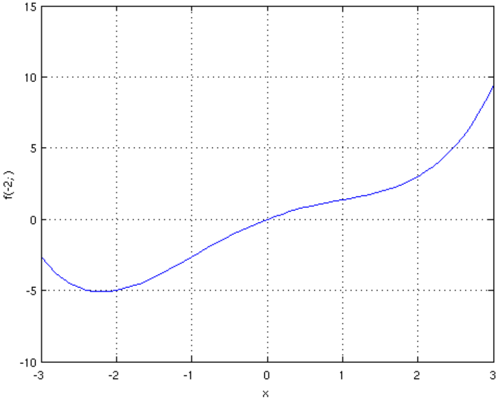
- Candidates for local minimizers
- Design of homotopy methods
- Understanding the problem structure (Morse theory)
- Convergence results for KKT points

Example: parametric unconstrained smooth optimization



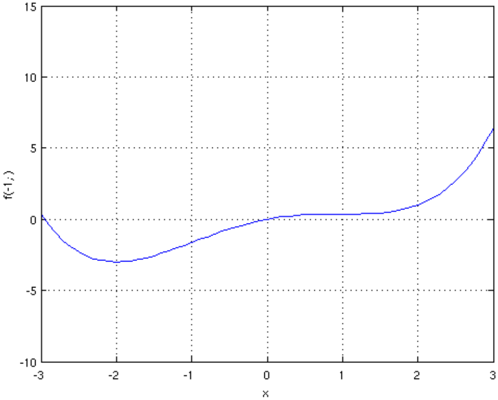
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = -3$$

Example: parametric unconstrained smooth optimization



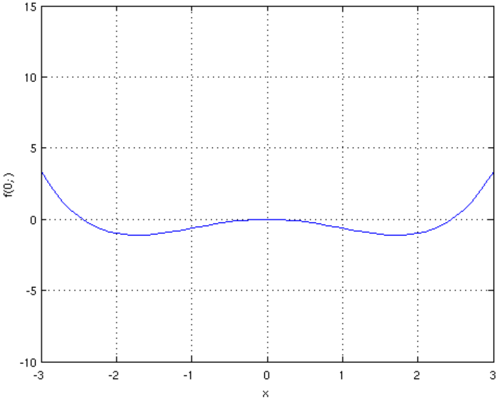
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = -2$$

Example: parametric unconstrained smooth optimization



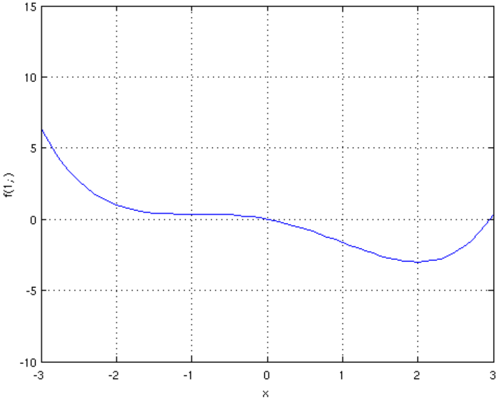
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = -1$$

Example: parametric unconstrained smooth optimization



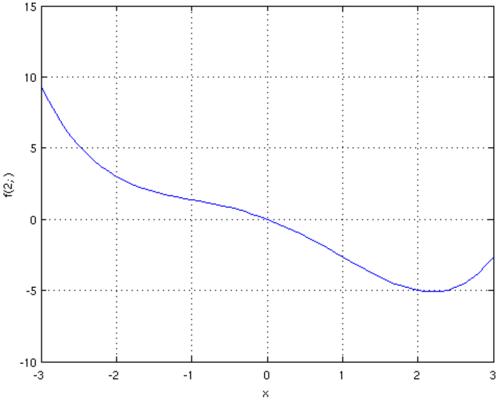
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = 0$$

Example: parametric unconstrained smooth optimization



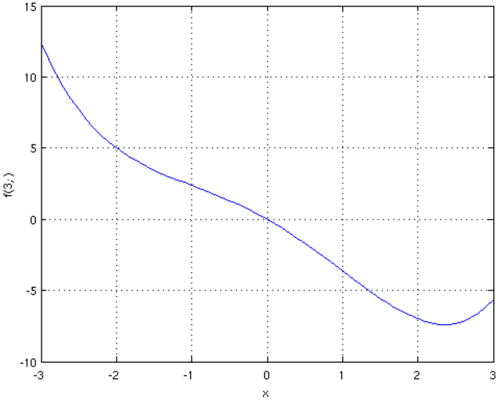
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = 1$$

Example: parametric unconstrained smooth optimization



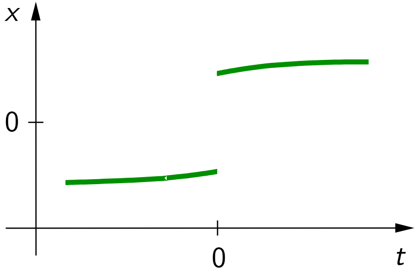
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = 2$$

Example: parametric unconstrained smooth optimization



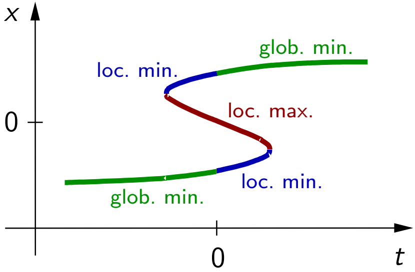
$$f(t,x) = \frac{x^4}{8} - \frac{3}{4}x^2 - tx \quad \text{for} \quad t = 3$$

Example: parametric unconstrained smooth optimization



Unfolded set of global minimizers

Example: parametric unconstrained smooth optimization



Unfolded set of critical points

Nondegenerate critical points

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of $f \;\Rightarrow\; \nabla f(\bar{x}) = 0$.

Definitions

$\bar{x} \in \mathbb{R}^n$ is called nondegenerate critical point of $f \in C^2(\mathbb{R}^n, \mathbb{R})$, if $\nabla f(\bar{x}) = 0$, and $D^2 f(\bar{x})$ is nonsingular.

The number of negative eigenvalues of $D^2 f(\bar{x})$ is called the Morse index or quadratic index of $\bar{x}$, briefly $QI(\bar{x})$.

Theorem (Jongen/Jonker/Twilt, 1983)

Generically, all critical points of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ are nondegenerate.

Nondegenerate critical points

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of $f \;\Rightarrow\; \nabla f(\bar{x}) = 0$.

Definitions

$\bar{x} \in \mathbb{R}^n$ is called nondegenerate critical point of $f \in C^2(\mathbb{R}^n, \mathbb{R})$, if $\nabla f(\bar{x}) = 0$, and $D^2 f(\bar{x})$ is nonsingular.

The number of negative eigenvalues of $D^2 f(\bar{x})$ is called the Morse index or quadratic index of $\bar{x}$, briefly $QI(\bar{x})$.

Theorem (Jongen/Jonker/Twilt, 1983)

Generically, all critical points of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ are nondegenerate.

Nondegenerate critical points

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of $f \;\Rightarrow\; \nabla f(\bar{x}) = 0$.

Definitions

$\bar{x} \in \mathbb{R}^n$ is called nondegenerate critical point of $f \in C^2(\mathbb{R}^n, \mathbb{R})$, if $\nabla f(\bar{x}) = 0$, and $D^2 f(\bar{x})$ is nonsingular.

The number of negative eigenvalues of $D^2 f(\bar{x})$ is called the Morse index or quadratic index of $\bar{x}$, briefly $QI(\bar{x})$.

Theorem (Jongen/Jonker/Twilt, 1983)

Generically, all critical points of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ are nondegenerate.

Nondegenerate critical points

Characterization of local minimality

For any nondegenerate critical point $\bar{x}$ of f we have

$$\bar{x} \text{ is a local minimizer of } f \;\Leftrightarrow\; QI(\bar{x}) = 0.$$

Theorem (Morse Lemma - local structure)

Let $\bar{x}$ be a nondegenerate critical point of $f \in C^2(\mathbb{R}^n, \mathbb{R})$. Then, modulo a local C^1 diffeomorphism, locally around $\bar{x}$ we have

$$f(x) \;=\; -x_1^2 - x_2^2 - \ldots - x_{QI(\bar{x})}^2 + x_{QI(\bar{x})+1}^2 + \ldots + x_n^2 \,.$$

Nondegenerate critical points

Characterization of local minimality

For any nondegenerate critical point $\bar{x}$ of f we have

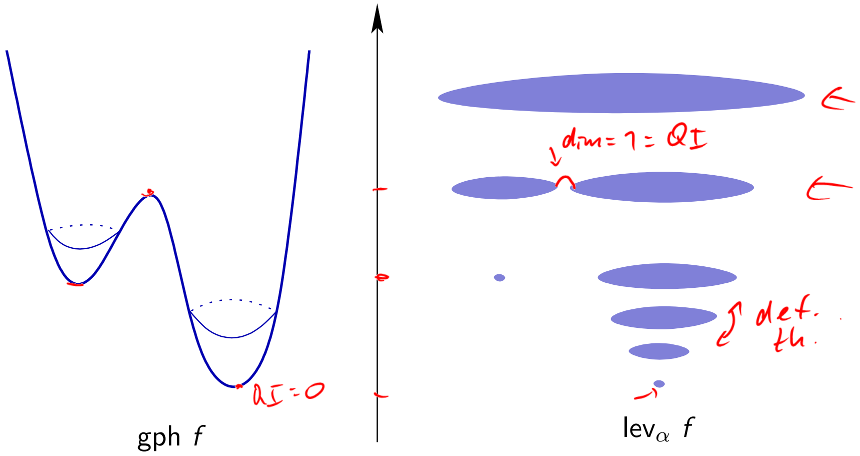
$$\bar{x} \ \text{ is a local minimizer of } f \ \Leftrightarrow \ QI(\bar{x}) = 0.$$

Theorem (Morse Lemma - local structure)

Let $\bar{x}$ be a nondegenerate critical point of $f \in C^2(\mathbb{R}^n, \mathbb{R})$.
Then, modulo a local C^1 diffeomorphism, locally around $\bar{x}$ we have

$$f(x) \ = \ -x_1^2 - x_2^2 - \ldots - x_{QI(\bar{x})}^2 + x_{QI(\bar{x})+1}^2 + \ldots + x_n^2 \,.$$

Deformation and cell attachment - global structure



Constrained smooth optimization

Consider the restriction of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ to

$$M = \{g(x) \geq 0,\ h(x) = 0\}$$

with $g \in C^2(\mathbb{R}^n, \mathbb{R}^p)$, $h \in C^2(\mathbb{R}^n, \mathbb{R}^q)$, and let

$$L(x, \lambda, \mu) = f(x) - \lambda^\intercal g(x) - \mu^\intercal h(x)$$

be the Lagrangian of f on M.

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of f on M with some CQ
$\Rightarrow\ \bar{x}$ KKT point of f on M.

Constrained smooth optimization

Consider the restriction of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ to

$$M \;=\; \{g(x) \geq 0,\; h(x) = 0\}$$

with $g \in C^2(\mathbb{R}^n, \mathbb{R}^p)$, $h \in C^2(\mathbb{R}^n, \mathbb{R}^q)$, and let

$$L(x, \lambda, \mu) \;=\; f(x) - \lambda^\mathsf{T} g(x) - \mu^\mathsf{T} h(x)$$

be the Lagrangian of f on M.

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of f on M with some CQ
$\Rightarrow\;\; \bar{x}$ KKT point of f on M.

Nondegenerate critical points

Definitions

$\bar{x} \in M$ is called nondegenerate critical point of f on M with multipliers $\bar{\lambda}$ and $\bar{\mu}$ if

- $\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$,
- LICQ holds at $\bar{x}$ in M,
- $\bar{\lambda}_i \neq 0$ for all active g_i,
- $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})|_{T(\bar{x},M)}$ is nonsingular.

The number of negative $\bar{\lambda}_i$ is called linear index of $\bar{x}$ ($LI(\bar{x})$), and the number of negative eigenvalues of $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})|_{T(\bar{x},M)}$ is called quadratic index of $\bar{x}$ ($QI(\bar{x})$).

Nondegenerate critical points

Characterization of local minimality

For any nondegenerate critical point $\bar{x}$ of f on M we have

$$\bar{x} \text{ is a local minimizer of } f \iff LI(\bar{x}) + QI(\bar{x}) = 0.$$

Morse theory and homotopy

- The generalizations to the constrained case of genericity, Morse lemma, deformation theorem and cell attachment theorem have been shown by Jongen/Jonker/Twilt (1983).
- For deformation and cell attachment, only the nondegenerate KKT points are relevant, that is, the nondegenerate critical points with $LI(\bar{x}) = 0$.
- Homotopy methods have been studied by, e.g., Guddat/Guerra Vázquez/Jongen (1990) ($LI(\bar{x}) \geq 0$ is relevant).

Mathematical programs with complementarity constraints

Consider the restriction of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ to the set

$$M \;=\; \{G_i(x) \geq 0,\; H_i(x) \geq 0,\; G_i(x)H_i(x) = 0,\; i = 1, \ldots \ell\}$$

with $G \in C^2(\mathbb{R}^n, \mathbb{R}^\ell)$, $H \in C^2(\mathbb{R}^n, \mathbb{R}^\ell)$, and let

$$L(x, \gamma, \eta) \;=\; f(x) - \gamma^\intercal G(x) - \eta^\intercal H(x)$$

be the Lagrangian of f on M.

Applications of MPCCs

- Game theory
- Obstacle problems
- Truss topology design
- Network equilibria
- Bilevel optimization
- Semi-infinite optimization
- ...

C-stationarity

C-stationarity

$\bar{x} \in M$ is called C-stationary point of f on M with multipliers $\bar{\gamma}$ and $\bar{\eta}$ if

- $\nabla_x L(\bar{x}, \bar{\gamma}, \bar{\eta}) = 0$,
- $\bar{\gamma}_i = 0$ for all i with $G_i(\bar{x}) > 0$, $H_i(\bar{x}) = 0$,
- $\bar{\eta}_i = 0$ for all i with $G_i(\bar{x}) = 0$, $H_i(\bar{x}) > 0$,
- $\bar{\gamma}_i \bar{\eta}_i \geq 0$ for all i with $G_i(\bar{x}) = H_i(\bar{x}) = 0$.

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of f on M with some MPEC-CQ
$\Rightarrow$ $\bar{x}$ C-stationary for f on M.

C-stationarity

C-stationarity

$\bar{x} \in M$ is called C-stationary point of f on M with multipliers $\bar{\gamma}$ and $\bar{\eta}$ if

- $\nabla_x L(\bar{x}, \bar{\gamma}, \bar{\eta}) = 0$,
- $\bar{\gamma}_i = 0$ for all i with $G_i(\bar{x}) > 0$, $H_i(\bar{x}) = 0$,
- $\bar{\eta}_i = 0$ for all i with $G_i(\bar{x}) = 0$, $H_i(\bar{x}) > 0$,
- $\bar{\gamma}_i \bar{\eta}_i \geq 0$ for all i with $G_i(\bar{x}) = H_i(\bar{x}) = 0$.

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of f on M with some MPEC-CQ
$\Rightarrow$ $\bar{x}$ C-stationary for f on M.

Nondegenerate C-stationary points

Definitions (Ralph/St., 2006, Jongen/Rückmann/Shikhman, 2009)

A C-stationary point $\bar{x}$ of f on M with multipliers $\bar{\gamma}$ and $\bar{\eta}$ is called nondegenerate if

- MPEC-LICQ holds at $\bar{x}$,
- $D_x^2 L(\bar{x}, \bar{\gamma}, \bar{\eta})|_{T(\bar{x},M)}$ is nonsingular,
- $\bar{\gamma}_i \bar{\eta}_i > 0$ for all i with $G_i(\bar{x}) = H_i(\bar{x}) = 0$. $\quad (\star)$

The number of pairs $(\bar{\gamma}_i, \bar{\eta}_i)$ with negative entries in $(\star)$ is called biactive index of $\bar{x}$ ($BI(\bar{x})$), the number of negative eigenvalues of $D_x^2 L(\bar{x}, \bar{\gamma}, \bar{\eta})|_{T(\bar{x},M)}$ is called quadratic index of $\bar{x}$ ($QI(\bar{x})$), and their sum $BI(\bar{x}) + QI(\bar{x})$ is called C-index of $\bar{x}$.

Nondegenerate C-stationary points

Characterization of local minimality

For any nondegenerate C-stationary point $\bar{x}$ of f on M we have

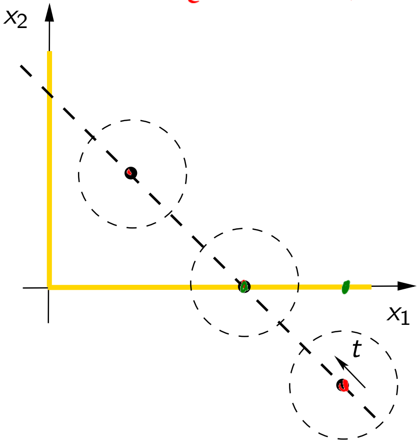
$$\bar{x} \text{ is a local minimizer of } f \;\Leftrightarrow\; BI(\bar{x}) + QI(\bar{x}) = 0.$$

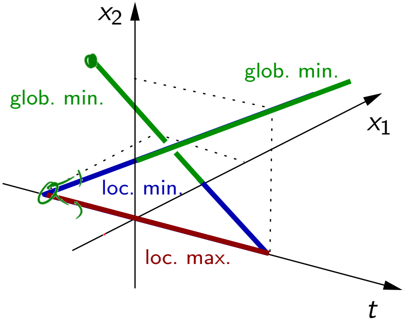
Morse theory and homotopy for MPCCs

- The (full) generalizations to MPCCs of genericity, Morse lemma, deformation theorem and cell attachment theorem have been shown by Jongen/Rückmann/Shikhman (2009).
- Homotopy methods for (special) generic MPCCs have been studied by Ralph/St. (2006).

An MPCC homotopy

$x_1 x_2 = 0, x_1 \geq 0, x_2 \geq 0$



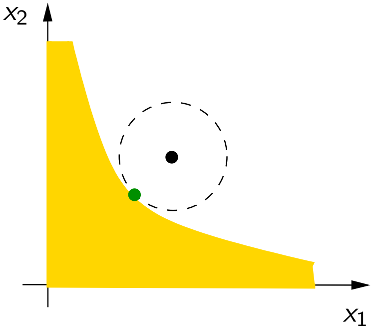


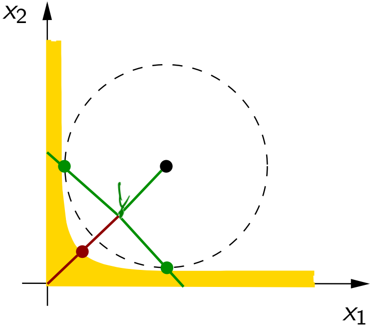
Limits of KKT points

Smoothed MPCCs (Scholtes 2001, Steffensen/Ulbrich 2010, Hoheisel/Kanzow/Schwartz 2011)

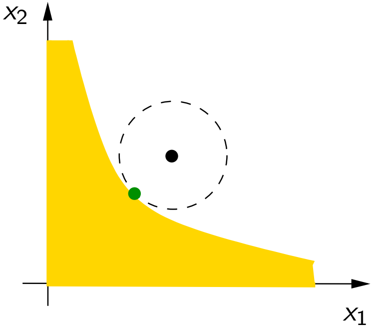
For a sequence of smoothing parameters $t_k \searrow 0$ and a sequence of KKT points x^k of some smoothing problem $NLP(t_k)$ with $x^k \to \bar{x}$, under some CQ the point $\bar{x}$ is C-stationary for MPCC.

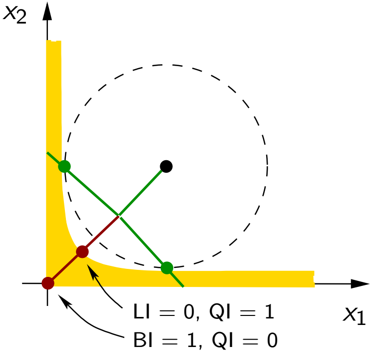
Example: Scholtes smoothing for an MPCC





Example: Scholtes smoothing for an MPCC





First order descent directions at C-stationary points

Nondegenerate C-stationary points with positive biactive index allow first order descent directions.

This is due to the nonsmoothness of MPCCs and cannot be avoided in a topologically relevant stationarity concept.

T-stationarity

Definition

For a given class of optimizations problems we call a set of conditions a stationarity concept, if these conditions hold (under some CQ) at each local minimizer, and we call the stationarity concept topologically relevant, if it admits

- a nondegeneracy concept
- the definition of a (Morse) index,
- a Morse lemma,
- a deformation theorem,
- and a cell attachment theorem.

The stationarity concept is then also called T-stationarity.

T-stationarity

Examples:

- Unconstrained smooth optimization:

  T-stationarity = stationarity

- Constrained smooth optimization:

  T-stationarity = KKT-stationarity

- MPCCs:

  T-stationarity = C-stationarity

- Disjunctive optimization:

  T-stationarity = stationarity (Jongen/Rückmann/St. 1997)

Limits of KKT points for MPCCs revisited

Smoothed MPCCs (Scholtes 2001, Steffensen/Ulbrich 2010, Hoheisel/Kanzow/Schwartz 2011)

For a sequence of smoothing parameters $t_k \searrow 0$ and a sequence of KKT-stationary points x^k of some smoothing problem $NLP(t_k)$ with $x^k \to \bar{x}$, under some CQ the point $\bar{x}$ is C-stationary for MPCC.

Limits of KKT points for MPCCs revisited

Smoothed MPCCs (Scholtes 2001, Steffensen/Ulbrich 2010, Hoheisel/Kanzow/Schwartz 2011)

For a sequence of smoothing parameters $t_k \searrow 0$ and a sequence of T-stationary points x^k of some smoothing problem $NLP(t_k)$ with $x^k \to \bar{x}$, under some CQ the point $\bar{x}$ is T-stationary for MPCC.

Mathematical programs with vanishing constraints

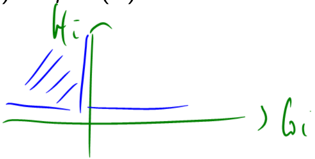
Consider the restriction of $f \in C^2(\mathbb{R}^n, \mathbb{R})$ to the set

$$M \;=\; \{H_i(x) \geq 0,\;\; G_i(x)H_i(x) \leq 0,\;\; i = 1, \ldots \ell\}$$

with $G \in C^2(\mathbb{R}^n, \mathbb{R}^\ell)$, $H \in C^2(\mathbb{R}^n, \mathbb{R}^\ell)$, and let

$$L(x, \gamma, \eta) \;=\; f(x) - \gamma^\mathsf{T} G(x) - \eta^\mathsf{T} H(x)$$

be the Lagrangian of f on M.



Application of MPVCs

- MPVC was introduced as a model for structural and topology optimization.
- It is motivated by the fact that the constraint G_i does not play any role whenever H_i is active.

T-stationarity for MPVCs

from scratch

T-stationarity (Dorsch/Shikhman/St., 2010)

$\bar{x} \in M$ is called T-stationary point of f on M with multipliers $\bar{\gamma}$ and $\bar{\eta}$ if

- $\nabla_x L(\bar{x}, \bar{\gamma}, \bar{\eta}) = 0$,
- $\bar{\gamma}_i = 0$ for all i with $G_i(\bar{x}) < 0$, $H_i(\bar{x}) \geq 0$,
- $\bar{\gamma}_i = 0$ for all i with $G_i(\bar{x}) > 0$, $H_i(\bar{x}) = 0$,
- $\bar{\gamma}_i \leq 0$ for all i with $G_i(\bar{x}) = 0$, $H_i(\bar{x}) \geq 0$,
- $\bar{\eta}_i = 0$ for all i with $H_i(\bar{x}) > 0$,
- $\bar{\eta}_i \geq 0$ for all i with $G_i(\bar{x}) < 0$, $H_i(\bar{x}) = 0$,
- $\bar{\gamma}_i \bar{\eta}_i \geq 0$ for all i with $G_i(\bar{x}) = H_i(\bar{x}) = 0$.

T-stationarity for MPVCs

Necessary condition

$\bar{x} \in \mathbb{R}^n$ local minimizer of f on M with some MPVC-CQ
$\Rightarrow \;\; \bar{x}$ T-stationary for f on M.

Nondegenerate T-stationary points

Definitions (Dorsch/Shikhman/St., 2010)

A T-stationary point $\bar{x}$ of f on M with multipliers $\bar{\gamma}$ and $\bar{\eta}$ is called nondegenerate if

- MPVC-LICQ holds at $\bar{x}$,
- $D_x^2 L(\bar{x}, \bar{\gamma}, \bar{\eta})|_{T(\bar{x},M)}$ is nonsingular,
- $\bar{\gamma}_i < 0$ for all i with $G_i(\bar{x}) = 0$, $H_i(\bar{x}) \geq 0$,
- $\bar{\eta}_i > 0$ for all i with $G_i(\bar{x}) < 0$, $H_i(\bar{x}) = 0$,
- $\bar{\gamma}_i \bar{\eta}_i > 0$ for all i with $G_i(\bar{x}) = H_i(\bar{x}) = 0$. $\quad (\star)$

The number of pairs $(\bar{\gamma}_i, \bar{\eta}_i)$ with negative entries in $(\star)$ is called biactive index of $\bar{x}$ ($BI(\bar{x})$), the number of negative eigenvalues of $D_x^2 L(\bar{x}, \bar{\gamma}, \bar{\eta})|_{T(\bar{x},M)}$ is called quadratic index of $\bar{x}$ ($QI(\bar{x})$), and their sum $BI(\bar{x}) + QI(\bar{x})$ is called T-index of $\bar{x}$.

Nondegenerate T-stationary points

Characterization of local minimality

For any nondegenerate T-stationary point $\bar{x}$ of f on M we have

$$\bar{x} \text{ is a local minimizer of } f \iff BI(\bar{x}) + QI(\bar{x}) = 0.$$

Morse theory for MPVCs

- The generalizations to MPVCs of genericity, Morse lemma, deformation theorem and cell attachment theorem have been shown by Dorsch/Shikhman/St. (2010).

Limits of KKT points

Smoothed MPVCs (Hoheisel/Kanzow/Schwartz 2011)

For a sequence of smoothing parameters $t_k \searrow 0$ and a sequence of KKT-stationary points x^k of some smoothing problem $NLP(t_k)$ with $x^k \to \bar{x}$, under some CQ the point $\bar{x}$ is T-stationary for MPVC.

Limits of KKT points

Smoothed MPVCs (Hoheisel/Kanzow/Schwartz 2011)

For a sequence of smoothing parameters $t_k \searrow 0$ and a sequence of T-stationary points x^k of some smoothing problem $NLP(t_k)$ with $x^k \to \bar{x}$, under some CQ the point $\bar{x}$ is T-stationary for MPVC.

Observation and a conjecture

In the known smoothing methods for MPCC, as well as for MPVC, any nondegenerate T-stationary point $\bar{x}$ of the nonsmooth problem

- locally 'unfolds' into a smooth curve $\{x(t) | \, t \in U(0)\}$ of KKT points of the smoothing problems (with smoothing parameter t, $x(0) = \bar{x}$, via the implicit function theorem),
- and the T-index of $\bar{x}$ coincides with the Morse index of $x(t)$, $t \in U(0)$ (via continuity arguments).

Conjecture 1

These effects occur for 'a large class of smoothing methods' for 'a large class of nonsmooth problems'.

Observation and a conjecture

In the known smoothing methods for MPCC, as well as for MPVC, any nondegenerate T-stationary point $\bar{x}$ of the nonsmooth problem

- locally 'unfolds' into a smooth curve $\{x(t) | \, t \in U(0)\}$ of KKT points of the smoothing problems (with smoothing parameter t, $x(0) = \bar{x}$, via the implicit function theorem),
- and the T-index of $\bar{x}$ coincides with the Morse index of $x(t)$, $t \in U(0)$ (via continuity arguments).

Conjecture 1

These effects occur for 'a large class of smoothing methods' for 'a large class of nonsmooth problems'.

Another conjecture

Conjecture 2

T-stationarity is uniquely defined for ‘a large class of nonsmooth problems’.

Conclusion

For any class of optimization problems, the T-stationarity concept is the natural one for

- topological considerations
- design of homotopy methods
- limits of KKT points
- (design of Newton methods).

Cited references

D. Dorsch, V. Shikhman, O. Stein, *Mathematical programs with vanishing constraints: critical point theory,* Journal of Global Optimization, Vol. 52 (2012), 591-605.

J. Guddat, F. Guerra Vázquez, H.Th. Jongen, *Parametric Optimization: Singularities, Pathfollowing and Jumps,* Wiley, Chichester, and Teubner, Stuttgart, 1990.

T. Hoheisel, C. Kanzow, A. Schwartz, *Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints,* Optimization Methods and Software, Vol. 27 (2012), 483-512.

H.Th. Jongen, P. Jonker, F. Twilt, *Nonlinear Optimization in* $\mathbb{R}^n$*. I. Morse Theory, Chebyshev Approximation,* Peter Lang Verlag, Frankfurt a.M., 1983.

H.Th. Jongen, J.-J. Rückmann, V. Shikhman, *MPCC: Critical point theory,* SIAM Journal on Optimization, Vol. 20 (2009), 473-484.

H.Th. Jongen, J.-J. Rückmann, O. Stein, *Disjunctive optimization: critical point theory,* Journal of Optimization Theory and Applications, Vol. 93 (1997), 321-336.

D. Ralph, O. Stein, *The C-index: a new stability concept for quadratic programs with complementarity constraints,* Mathematics of Operations Research, Vol. 36 (2011), 504-526.

S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints,* SIAM Journal on Optimization, Vol. 11 (2001), 918-936.

S. Steffensen, M. Ulbrich, *A New Relaxation Scheme for Mathematical Programs with Equilibrium Constraints,* SIAM Journal on Optimization, Vol. 20 (2010), 2504-2539.