Perturbation analysis of matrix optimization

Chao Ding
Institute of Applied Mathematics
Academy of Mathematics and Systems Science, CAS

ICCOPT2019, **Berlin**
2019.08.06

Acknowledgements

Based on the joint work with **Ying Cui** at *USC*:

- Nonsmooth composite matrix optimizations: strong regularity, constraint nondegeneracy and beyond, *arXiv:1907.13253* (July, 2019).

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi \circ \lambda(g(\mathbf{x})) \\ \text{subject to} & h(\mathbf{x}) = 0, \end{array}$$

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi \circ \lambda(g(\mathbf{x})) \\ \text{subject to} & h(\mathbf{x}) = 0, \end{array}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} \quad & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi \circ \lambda(g(\mathbf{x})) \\ \text{subject to} \quad & h(\mathbf{x}) = 0, \end{aligned}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f : \mathbb{X} \to \mathbb{R}$, $g : \mathbb{X} \to \mathbb{S}^n$ and $h : \mathbb{X} \to \mathbb{Y}$: twice continuously differentiable functions

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} \quad & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi\circ\lambda(g(\mathbf{x})) \\ \text{subject to} \quad & h(\mathbf{x}) = 0, \end{aligned}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f : \mathbb{X} \to \mathbb{R}$, $g : \mathbb{X} \to \mathbb{S}^n$ and $h : \mathbb{X} \to \mathbb{Y}$: twice continuously differentiable functions
- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$: a **symmetric function**, i.e., for any $u \in \mathbb{R}^n$ and any $n \times n$ permutation matrix P,

$$\phi(Pu) = \phi(u)$$

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} \quad & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi\circ\lambda(g(\mathbf{x})) \\ \text{subject to} \quad & h(\mathbf{x}) = 0, \end{aligned}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f : \mathbb{X} \to \mathbb{R}$, $g : \mathbb{X} \to \mathbb{S}^n$ and $h : \mathbb{X} \to \mathbb{Y}$: twice continuously differentiable functions
- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$: a **symmetric function**, i.e., for any $u \in \mathbb{R}^n$ and any $n \times n$ permutation matrix P,

$$\phi(Pu) = \phi(u)$$

- λ: the vector of eigenvalues for any symmetric matrix

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi\circ\lambda(g(\mathbf{x})) \\ \text{subject to} & h(\mathbf{x}) = 0, \end{array}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f : \mathbb{X} \to \mathbb{R}$, $g : \mathbb{X} \to \mathbb{S}^n$ and $h : \mathbb{X} \to \mathbb{Y}$: twice continuously differentiable functions
- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$: a **symmetric function**, i.e., for any $u \in \mathbb{R}^n$ and any $n \times n$ permutation matrix P,

$$\phi(Pu) = \phi(u)$$

- λ: the vector of eigenvalues for any symmetric matrix

★ We focus on the symmetric case just for **simplicity**;

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} \quad \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi \circ \lambda(g(\mathbf{x}))$$
$$\text{subject to} \quad h(\mathbf{x}) = 0,$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f : \mathbb{X} \to \mathbb{R}$, $g : \mathbb{X} \to \mathbb{S}^n$ and $h : \mathbb{X} \to \mathbb{Y}$: twice continuously differentiable functions
- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$: a **symmetric function**, i.e., for any $u \in \mathbb{R}^n$ and any $n \times n$ permutation matrix P,

$$\phi(Pu) = \phi(u)$$

- λ: the vector of eigenvalues for any symmetric matrix

★ We focus on the symmetric case just for **simplicity**;

★ The obtained results can be extended to **non-symmetric** cases;

Nonsmooth Composite Matrix Optimization Problem

CMatOP:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & \Phi(\mathbf{x}) \triangleq f(\mathbf{x}) + \phi\circ\lambda(g(\mathbf{x})) \\ \text{subject to} & h(\mathbf{x}) = 0, \end{array}$$

- $\mathbb{X}$ and $\mathbb{Y}$: two given finite dimensional Euclidean spaces
- $f:\mathbb{X}\to\mathbb{R}$, $g:\mathbb{X}\to\mathbb{S}^n$ and $h:\mathbb{X}\to\mathbb{Y}$: twice continuously differentiable functions
- $\phi:\mathbb{R}^n\to(-\infty,+\infty]$: a **symmetric function**, i.e., for any $u\in\mathbb{R}^n$ and any $n\times n$ permutation matrix P,

$$\phi(Pu) = \phi(u)$$

- λ: the vector of eigenvalues for any symmetric matrix

★ We focus on the symmetric case just for **simplicity**;

★ The obtained results can be extended to **non-symmetric** cases;

★ This is a general model which includes many "**non-polyhedral**" OPs: **SDP**, **Eigenvalue optimization**, etc

More applications

- Fastest mixing Markov chain problem (fast load balancing of paralleled systems)
- Fastest distributed linear averaging problem
- The reduced rank approximations of transition matrices
- The low rank approximations of doubly stochastic matrices
- Low-rank approximation of matrices with linear structures
- Unsupervised learning
- ......

Spectral functions

$\phi \circ \lambda$: **spectral function** (**Friedland, 1981**)

Spectral functions

$\phi \circ \lambda$: **spectral function** (**Friedland, 1981**)

- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$ is a **symmetric** **convex piecewise linear function**

Spectral functions

$\phi \circ \lambda$: **spectral function** (**Friedland, 1981**)

- $\phi : \mathbb{R}^n \to (-\infty, +\infty]$ is a **symmetric** **convex piecewise linear function**
- a **convex piecewise linear function**: a **polyhedral convex function** (**Rockafellar, 1970**)

Convex piecewise linear functions

Theorem (Rockafellar & Wets, 1998)

ϕ can be expressed in the form of

$$\phi(\mathbf{x}) = \phi_1(\mathbf{x}) + \phi_2(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n,$$

with $\phi_1 : \mathbb{R}^n \to \mathbb{R}$ and $\phi_2 : \mathbb{R}^n \to (-\infty, +\infty]$ are defined by

$$\phi_1(\mathbf{x}) := \max_{1 \leq i \leq p} \left\{ \langle \mathbf{a}^i, \mathbf{x} \rangle - c^i \right\} \quad \text{and} \quad \phi_2(\mathbf{x}) := \delta_{\operatorname{dom} \phi}(\mathbf{x}),$$

- $\mathbf{a}^1, \ldots, \mathbf{a}^p \in \mathbb{R}^n$, $c^1, \ldots, c^p \in \mathbb{R}$ *with some positive integer* $p \geq 1$*;*
- $\operatorname{dom} \phi$ *is a polyhedral set:*

$$\operatorname{dom} \phi := \left\{ x \in \mathbb{R}^n \mid \max_{1 \leq i \leq q} \{ \langle \mathbf{b}^i, \mathbf{x} \rangle - d^i \} \leq 0 \right\}$$

- $\mathbf{b}^1, \ldots, \mathbf{b}^q \in \mathbb{R}^n$ *and* $d^1, \ldots, d^q \in \mathbb{R}$ *for some positive integer* $q \geq 1$.

Examples

SDP:

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1 \leq i \leq n} \{\langle \mathbf{e}^i, \lambda(X) \rangle\} \leq 0\}$$

Examples

SDP:

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1 \le i \le n} \{\langle \mathbf{e}^i, \lambda(X) \rangle\} \le 0\}$$

- $\mathbf{e}^i \in \mathbb{R}^n$: the canonical basis of $\mathbb{R}^n$

Examples

SDP:

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$$

- $\mathbf{e}^i \in \mathbb{R}^n$: the canonical basis of $\mathbb{R}^n$

$$g(\mathbf{x}) \in \mathbb{S}^n_- \quad \Longleftrightarrow \quad \phi_2 \circ \lambda(g(\mathbf{x})) = \delta_{\mathrm{dom}\,\phi}(g(\mathbf{x}))$$

Examples

SDP:

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1\leq i\leq n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \leq 0\}$$

- $\mathbf{e}^i \in \mathbb{R}^n$: the canonical basis of $\mathbb{R}^n$

$$g(\mathbf{x}) \in \mathbb{S}^n_- \quad \Longleftrightarrow \quad \phi_2 \circ \lambda(g(\mathbf{x})) = \delta_{\operatorname{dom}\phi}(g(\mathbf{x}))$$

Eigenvalue optimizations:

$$s_k(X) = \sum_{i=1}^{k} \lambda_i(X) = \max_{1\leq i\leq p} \left\{\langle \mathbf{a}^i, \lambda(X)\rangle\right\}$$

Examples

SDP:

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1 \leq i \leq n} \{\langle \mathbf{e}^i, \lambda(X) \rangle\} \leq 0\}$$

- $\mathbf{e}^i \in \mathbb{R}^n$: the canonical basis of $\mathbb{R}^n$

$$g(\mathbf{x}) \in \mathbb{S}^n_- \quad \Longleftrightarrow \quad \phi_2 \circ \lambda(g(\mathbf{x})) = \delta_{\operatorname{dom}\phi}(g(\mathbf{x}))$$

Eigenvalue optimizations:

$$s_k(X) = \sum_{i=1}^{k} \lambda_i(X) = \max_{1 \leq i \leq p} \left\{ \langle \mathbf{a}^i, \lambda(X) \rangle \right\}$$

- $\mathbf{a}^i \in \mathbb{R}^n$: the vector contains k ones and $n-k$ zeros

Perturbation analysis of CMatOPs

Canonically perturbed CMatOPs with parameters $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathbb{X} \times \mathbb{Y} \times \mathbb{S}^n$:

$$\begin{aligned} \underset{\mathbf{x} \in \mathbb{X}}{\text{minimize}} \quad & f(\mathbf{x}) - \langle \mathbf{a}, \mathbf{x} \rangle + \phi \circ \lambda(g(\mathbf{x}) + \mathbf{c}) \\ \text{subject to} \quad & h(\mathbf{x}) + \mathbf{b} = 0 \end{aligned}$$

Perturbation analysis of CMatOPs

Canonically perturbed CMatOPs with parameters $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathbb{X} \times \mathbb{Y} \times \mathbb{S}^n$:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & f(\mathbf{x}) - \langle \mathbf{a}, \mathbf{x} \rangle + \phi \circ \lambda(g(\mathbf{x}) + \mathbf{c}) \\ \text{subject to} & h(\mathbf{x}) + \mathbf{b} = 0 \end{array}$$

The Karush-Kuhn-Tucker (KKT) optimality condition for perturbed problem takes the following form:

$$\begin{cases} \mathbf{a} = \nabla f(\mathbf{x}) + h'(\mathbf{x})^*\mathbf{y} + g'(\mathbf{x})^*Y + g'(\mathbf{x})^*Z \\ \mathbf{b} = -h(\mathbf{x}) \\ \mathbf{c} \in -g(\mathbf{x}) + \partial\theta_1^*(Y) \\ \mathbf{c} \in -g(\mathbf{x}) + \partial\theta_2^*(Z) \end{cases}$$

with $\theta_1 = \phi_1 \circ \lambda$ and $\theta_2 = \phi_2 \circ \lambda$ are two spectral functions

Perturbation analysis of CMatOPs

Canonically perturbed CMatOPs with parameters $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathbb{X} \times \mathbb{Y} \times \mathbb{S}^n$:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & f(\mathbf{x}) - \langle \mathbf{a}, \mathbf{x} \rangle + \phi \circ \lambda(g(\mathbf{x}) + \mathbf{c}) \\ \text{subject to} & h(\mathbf{x}) + \mathbf{b} = 0 \end{array}$$

The Karush-Kuhn-Tucker (KKT) optimality condition for perturbed problem takes the following form:

$$\begin{cases} \mathbf{a} = \nabla f(\mathbf{x}) + h'(\mathbf{x})^*\mathbf{y} + g'(\mathbf{x})^*Y + g'(\mathbf{x})^*Z \\ \mathbf{b} = -h(\mathbf{x}) \\ \mathbf{c} \in -g(\mathbf{x}) + \partial\theta_1^*(Y) \\ \mathbf{c} \in -g(\mathbf{x}) + \partial\theta_2^*(Z) \end{cases}$$

with $\theta_1 = \phi_1 \circ \lambda$ and $\theta_2 = \phi_2 \circ \lambda$ are two spectral functions

Strong regularity:

When the solution mapping $S_{\mathrm{KKT}}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is Lipschitz continuous?

Why it matters?

Why it matters?

- **Perturbation theory**

Why it matters?

- **Perturbation theory**
- **Algorithm**

How?

Variational analysis

How?

Variational analysis but in slightly different way:

How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

- Tangent sets

How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

- Tangent sets
- Critical cones

How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

- Tangent sets
- Critical cones
- Second-order tangent sets

How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

- Tangent sets
- Critical cones
- Second-order tangent sets
- The "σ**-term**":

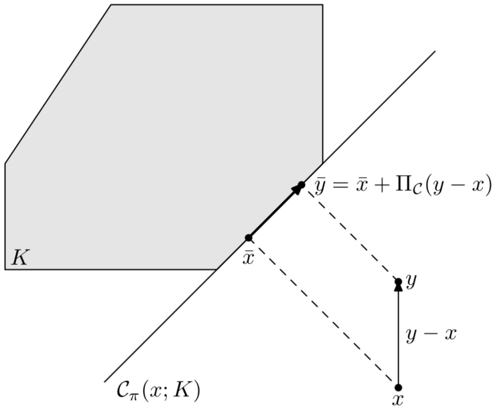
How?

Variational analysis but in slightly different way:

Variational analysis of spectral functions

- Tangent sets
- Critical cones
- Second-order tangent sets
- The "σ**-term**": the **key** difference between NLPs (**polyhedral**) and CMatOPs (**non-polyhedral**)

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral



The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

Metric projection operator $\Pi_{\mathcal{K}}$:

$$\overline{A} := \Pi_{\mathcal{K}}(C) := \operatorname{argmin}\left\{\frac{1}{2}\|Y - C\|^2 \mid Y \in \mathcal{K}\right\}$$

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

Metric projection operator $\Pi_{\mathcal{K}}$:

$$\overline{A} := \Pi_{\mathcal{K}}(C) := \operatorname{argmin}\left\{\frac{1}{2}\|Y - C\|^2 \mid Y \in \mathcal{K}\right\}$$

If $\mathcal{K}$ is a polyhedral closed convex set,

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

Metric projection operator $\Pi_{\mathcal{K}}$:

$$\overline{A} := \Pi_{\mathcal{K}}(C) := \operatorname{argmin} \left\{ \frac{1}{2} \|Y - C\|^2 \mid Y \in \mathcal{K} \right\}$$

If $\mathcal{K}$ is a polyhedral closed convex set,

- $\Pi_{\mathcal{K}}$ is directional differentiable **(Facchinei & Pang, 2003)**[1]

$$\Pi_{\mathcal{K}}(C + H) - \Pi_{\mathcal{K}}(C) = \Pi_{\mathcal{C}_{\mathcal{K}}(C)}(H) =: \Pi'_{\mathcal{K}}(C; H) \quad \forall\, H$$

[1] F. Facchinei and J. S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems: Volume I*, Springer-Verlag, New York, 2003.

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

Metric projection operator $\Pi_{\mathcal{K}}$:

$$\overline{A} := \Pi_{\mathcal{K}}(C) := \operatorname{argmin}\left\{\frac{1}{2}\|Y - C\|^2 \mid Y \in \mathcal{K}\right\}$$

If $\mathcal{K}$ is a polyhedral closed convex set,

- $\Pi_{\mathcal{K}}$ is directional differentiable **(Facchinei & Pang, 2003)**[1]

$$\Pi_{\mathcal{K}}(C + H) - \Pi_{\mathcal{K}}(C) = \Pi_{\mathcal{C}_{\mathcal{K}(C)}}(H) =: \Pi'_{\mathcal{K}}(C; H) \quad \forall\, H$$

- $\mathcal{C}_{\mathcal{K}}(C)$ is the critical cone of $\mathcal{K}$ at C

[1] F. Facchinei and J. S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems: Volume I*, Springer-Verlag, New York, 2003.

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

If $\mathcal{K}$ is a non-polyhedral closed convex set

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

If $\mathcal{K}$ is a non-polyhedral closed convex set but C^2-cone reducible,

- $\Pi_{\mathcal{K}}$ is directional differentiable and $\Pi'_{\mathcal{K}}(C;H)$ is the unique optimal solution to (**Bonnans et al., 1998**)[2]:

$$\min\left\{\|D-H\|^2-\sigma(\overline{B},\mathcal{T}^2_{\mathcal{K}}(\overline{A},D))\mid D\in\mathcal{C}_{\mathcal{K}}(C)\right\}$$

[2] J.F. Bonnans, R. Cominetti and A. Shapiro. Sensitivity analysis of optimization problems under second order regular constraints. *Mathematics of Operations Research* 23 (1998) 806–831.

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

If $\mathcal{K}$ is a non-polyhedral closed convex set but C^2-cone reducible,

- $\Pi_{\mathcal{K}}$ is directional differentiable and $\Pi'_{\mathcal{K}}(C; H)$ is the unique optimal solution to (**Bonnans et al., 1998**)[2]:

$$\min\left\{\|D - H\|^2 - \sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D)) \mid D \in \mathcal{C}_{\mathcal{K}}(C)\right\}$$

- $\overline{B} := C - \overline{A}$ and $\sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D))$ is the "σ-term" of $\mathcal{K}$

[2] J.F. Bonnans, R. Cominetti and A. Shapiro. Sensitivity analysis of optimization problems under second order regular constraints. *Mathematics of Operations Research* 23 (1998) 806–831.

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

If $\mathcal{K}$ is a non-polyhedral closed convex set but C^2-cone reducible,

- $\Pi_{\mathcal{K}}$ is directional differentiable and $\Pi'_{\mathcal{K}}(C; H)$ is the unique optimal solution to (**Bonnans et al., 1998**)[2]:

$$\min \left\{ \|D - H\|^2 - \sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D)) \mid D \in \mathcal{C}_{\mathcal{K}}(C) \right\}$$

- $\overline{B} := C - \overline{A}$ and $\sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D))$ is the "σ-term" of $\mathcal{K}$

polyhedral:

$$\begin{array}{ll} \min & \|D - H\|^2 \\ \text{s.t.} & D \in \mathcal{C}_{\mathcal{K}}(C) \end{array}$$

[2] J.F. Bonnans, R. Cominetti and A. Shapiro. Sensitivity analysis of optimization problems under second order regular constraints. *Mathematics of Operations Research* 23 (1998) 806–831.

The "σ-term": polyhedral $\Longrightarrow$ non-polyhedral (cont'd)

If $\mathcal{K}$ is a non-polyhedral closed convex set but C^2-cone reducible,

- $\Pi_{\mathcal{K}}$ is directional differentiable and $\Pi'_{\mathcal{K}}(C; H)$ is the unique optimal solution to (**Bonnans et al., 1998**)[2]:

$$\min \left\{ \|D - H\|^2 - \sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D)) \mid D \in \mathcal{C}_{\mathcal{K}}(C) \right\}$$

- $\overline{B} := C - \overline{A}$ and $\sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D))$ is the "σ-term" of $\mathcal{K}$

polyhedral:

$$\begin{array}{ll} \min & \|D - H\|^2 \\ \text{s.t.} & D \in \mathcal{C}_{\mathcal{K}}(C) \end{array}$$

non-polyhedral:

$$\begin{array}{ll} \min & \|D - H\|^2 - \sigma(\overline{B}, \mathcal{T}^2_{\mathcal{K}}(\overline{A}, D)) \\ \text{s.t.} & D \in \mathcal{C}_{\mathcal{K}}(C) \end{array}$$

---

[2] J.F. Bonnans, R. Cominetti and A. Shapiro. Sensitivity analysis of optimization problems under second order regular constraints. *Mathematics of Operations Research* 23 (1998) 806–831.

Convex piecewise linear + Symmetric

Convex piecewise linear + Symmetric

(**Rockafellar & Wets, 1998**): $\phi = \phi_1 + \phi_2$ with $\phi_2 = \delta_{\operatorname{dom}\phi}$

$$\phi_1(\mathbf{x}) = \max_{1\le i\le p}\left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\},\ \operatorname{dom}\phi = \left\{\mathbf{x}\in\mathbb{R}^n \mid \max_{1\le i\le q}\left\{\langle \mathbf{b}^i, \mathbf{x}\rangle - d^i\right\}\le 0\right\}$$

Convex piecewise linear + Symmetric

(Rockafellar & Wets, 1998): $\phi = \phi_1 + \phi_2$ with $\phi_2 = \delta_{\mathrm{dom}\,\phi}$

$$\phi_1(\mathbf{x}) = \max_{1\le i\le p}\left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\},\ \mathrm{dom}\,\phi = \left\{\mathbf{x}\in\mathbb{R}^n \mid \max_{1\le i\le q}\left\{\langle \mathbf{b}^i, \mathbf{x}\rangle - d^i\right\} \le 0\right\}$$

Proposition

Let $\phi = \phi_1 + \phi_2 : \mathbb{R}^n \to (-\infty, \infty]$ be a given proper convex piecewise linear function. ϕ is **symmetric** *over $\mathbb{R}^n$ if and only if the functions $\phi_1 : \mathbb{R}^n \to \mathbb{R}$ and $\phi_2 : \mathbb{R}^n \to (-\infty, \infty]$ satisfy the following conditions: for any $\mathbf{x} \in \mathbb{R}^n$,*

$$\phi_1(\mathbf{x}) = \max_{1\le i\le p}\left\{\max_{Q\in\mathbb{P}^n}\left\{\langle Q\mathbf{a}^i, \mathbf{x}\rangle - c^i\right\}\right\} \quad \text{and} \quad \phi_2(\mathbf{x}) = \delta_{\mathrm{dom}\,\phi}(\mathbf{x}),$$

where $\mathrm{dom}\,\phi = \left\{\mathbf{x}\in\mathbb{R}^n \mid \max_{1\le i\le q}\left\{\max_{Q\in\mathbb{P}^n}\left\{\langle Q\mathbf{b}^i, \mathbf{x}\rangle - d^i\right\}\right\} \le 0\right\}.$

Convex piecewise linear + Symmetric (cont'd)

- For $i = 1, \ldots, p$, define

$$\mathcal{D}_i := \left\{ \mathbf{x} \in \operatorname{dom} \phi \mid \langle \mathbf{a}^j, \mathbf{x} \rangle - c^j \leq \langle \mathbf{a}^i, \mathbf{x} \rangle - c^i \ \forall\, j = 1, \ldots, p \right\},$$

then $\operatorname{dom} \phi = \bigcup_{i=1,\ldots,p} \mathcal{D}_i$

- any $\overline{\mathbf{x}} \in \operatorname{dom} \phi$, we have two active index sets:

$$\iota_1(\overline{\mathbf{x}}) := \{1 \leq i \leq p \mid \overline{\mathbf{x}} \in \mathcal{D}_i\}, \quad \iota_2(\overline{\mathbf{x}}) := \{1 \leq i \leq q \mid \langle \mathbf{b}^i, \overline{\mathbf{x}} \rangle - d^i = 0\}.$$

Convex piecewise linear + Symmetric (cont'd)

- For $i = 1, \ldots, p$, define

$$\mathcal{D}_i := \left\{ \mathbf{x} \in \operatorname{dom} \phi \mid \langle \mathbf{a}^j, \mathbf{x} \rangle - c^j \le \langle \mathbf{a}^i, \mathbf{x} \rangle - c^i \ \forall\, j = 1, \ldots, p \right\},$$

  then $\operatorname{dom} \phi = \bigcup_{i=1,\ldots,p} \mathcal{D}_i$

- any $\overline{\mathbf{x}} \in \operatorname{dom} \phi$, we have two active index sets:

$$\iota_1(\overline{\mathbf{x}}) := \{1 \le i \le p \mid \overline{\mathbf{x}} \in \mathcal{D}_i\}, \quad \iota_2(\overline{\mathbf{x}}) := \{1 \le i \le q \mid \langle \mathbf{b}^i, \overline{\mathbf{x}} \rangle - d^i = 0\}.$$

Proposition

For any $i \in \iota_1(\overline{\mathbf{x}})$, $j \in \iota_2(\overline{\mathbf{x}})$ and $Q \in \mathbb{P}^n_{\overline{\mathbf{x}}}$ (i.e., $Q\overline{\mathbf{x}} = \overline{\mathbf{x}}$), there exist $i' \in \iota_1(\overline{\mathbf{x}})$ and $j' \in \iota_2(\overline{\mathbf{x}})$ such that $\mathbf{a}^{i'} = Q\mathbf{a}^i$ and $\mathbf{b}^{j'} = Q\mathbf{b}^j$, respectively.

Convex piecewise linear + Symmetric (cont'd)

Rockafellar & Wets, 1998, **Mordukhovich & Sarabi, 2016**:

- the **subgradients**:

  $$\partial\phi_1(\overline{\mathbf{x}}) = \operatorname{conv}\{\mathbf{a}^i,\ i \in \iota_1(\overline{\mathbf{x}})\}, \partial\phi_2(\overline{\mathbf{x}}) = \mathcal{N}_{\operatorname{dom}\phi}(\overline{\mathbf{x}}) = \operatorname{cone}\{\mathbf{b}^i,\ i \in \iota_2(\overline{\mathbf{x}})\}$$

Convex piecewise linear + Symmetric (cont'd)

Rockafellar & Wets, 1998, **Mordukhovich & Sarabi, 2016**:

- the **subgradients**:

$$\partial\phi_1(\overline{\mathbf{x}}) = \operatorname{conv}\{\mathbf{a}^i,\ i \in \iota_1(\overline{\mathbf{x}})\}, \partial\phi_2(\overline{\mathbf{x}}) = \mathcal{N}_{\operatorname{dom}\phi}(\overline{\mathbf{x}}) = \operatorname{cone}\{\mathbf{b}^i,\ i \in \iota_2(\overline{\mathbf{x}})\}$$

$\phi_1(\mathbf{x}) = \max\limits_{1\le i\le p} \left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\}$ is finite everywhere,

Convex piecewise linear + Symmetric (cont'd)

Rockafellar & Wets, 1998, **Mordukhovich & Sarabi, 2016**:

- the **subgradients**:

$$\partial\phi_1(\overline{\mathbf{x}}) = \text{conv}\{\mathbf{a}^i,\ i \in \iota_1(\overline{\mathbf{x}})\}, \partial\phi_2(\overline{\mathbf{x}}) = \mathcal{N}_{\text{dom}\,\phi}(\overline{\mathbf{x}}) = \text{cone}\{\mathbf{b}^i,\ i \in \iota_2(\overline{\mathbf{x}})\}$$

$\phi_1(\mathbf{x}) = \max\limits_{1\le i\le p} \left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\}$ is finite everywhere,

- ϕ_1 is directionally differentiable
- the **directional derivate**:

$$\phi_1'(\overline{\mathbf{x}};\mathbf{h}) = \max_{i\in\iota_1(\overline{\mathbf{x}})} \langle \mathbf{a}^i, \mathbf{h}\rangle, \quad \mathbf{h} \in \mathbb{R}^n.$$

Convex piecewise linear + Symmetric (cont'd)

Rockafellar & Wets, 1998, **Mordukhovich & Sarabi, 2016**:

- the **subgradients**:

$$\partial\phi_1(\overline{\mathbf{x}}) = \operatorname{conv}\{\mathbf{a}^i,\ i \in \iota_1(\overline{\mathbf{x}})\}, \partial\phi_2(\overline{\mathbf{x}}) = \mathcal{N}_{\operatorname{dom}\phi}(\overline{\mathbf{x}}) = \operatorname{cone}\{\mathbf{b}^i,\ i \in \iota_2(\overline{\mathbf{x}})\}$$

$\phi_1(\mathbf{x}) = \max_{1\le i\le p} \left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\}$ is finite everywhere,

- ϕ_1 is directionally differentiable
- the **directional derivate**:

$$\phi_1'(\overline{\mathbf{x}};\mathbf{h}) = \max_{i\in\iota_1(\overline{\mathbf{x}})} \langle \mathbf{a}^i, \mathbf{h}\rangle, \quad \mathbf{h} \in \mathbb{R}^n.$$

Let $\psi(\mathbf{x}) := \max_{1\le i\le q} \left\{\langle \mathbf{b}^i, \mathbf{x}\rangle - d^i\right\}$. Then, $\operatorname{dom}\phi = \left\{\mathbf{x} \in \mathbb{R}^n \mid \psi(\mathbf{x}) \le 0\right\}$

Convex piecewise linear + Symmetric (cont'd)

Rockafellar & Wets, 1998, **Mordukhovich & Sarabi, 2016**:

- the **subgradients**:

$$\partial\phi_1(\overline{\mathbf{x}}) = \mathrm{conv}\{\mathbf{a}^i,\ i \in \iota_1(\overline{\mathbf{x}})\}, \partial\phi_2(\overline{\mathbf{x}}) = \mathcal{N}_{\mathrm{dom}\,\phi}(\overline{\mathbf{x}}) = \mathrm{cone}\{\mathbf{b}^i,\ i \in \iota_2(\overline{\mathbf{x}})\}$$

$\phi_1(\mathbf{x}) = \max\limits_{1\le i\le p} \left\{\langle \mathbf{a}^i, \mathbf{x}\rangle - c^i\right\}$ is finite everywhere,

- ϕ_1 is directionally differentiable
- the **directional derivate**:

$$\phi_1'(\overline{\mathbf{x}};\mathbf{h}) = \max_{i\in\iota_1(\overline{\mathbf{x}})} \langle \mathbf{a}^i, \mathbf{h}\rangle, \quad \mathbf{h} \in \mathbb{R}^n.$$

Let $\psi(\mathbf{x}) := \max\limits_{1\le i\le q} \left\{\langle \mathbf{b}^i, \mathbf{x}\rangle - d^i\right\}$. Then, $\mathrm{dom}\,\phi = \left\{\mathbf{x} \in \mathbb{R}^n \mid \psi(\mathbf{x}) \le 0\right\}$

- ψ is directionally differentiable
- the **directional derivate**:

$$\psi'(\overline{\mathbf{x}};\mathbf{h}) = \max_{i\in\iota_2(\overline{\mathbf{x}})} \langle \mathbf{b}^i, \mathbf{h}\rangle, \quad \mathbf{h} \in \mathbb{R}^n.$$

Tangent sets

Tangent sets

For $\theta_1 = \phi_1 \circ \lambda$:

Tangent sets

For $\theta_1 = \phi_1 \circ \lambda$:

- **Tangent set** of epigraph:

$$\mathcal{T}_{\operatorname{epi}\theta_1}(\overline{X}, \theta(\overline{X})) = \operatorname{epi}\theta_1'(\overline{X};\cdot) := \left\{(H,y) \in \mathbb{S}^n \times \mathbb{R} \mid \theta_1'(\overline{X};H) \le y\right\}$$

Tangent sets

For $\theta_1 = \phi_1 \circ \lambda$:

- **Tangent set** of epigraph:

$$\mathcal{T}_{\mathrm{epi}\,\theta_1}(\overline{X}, \theta(\overline{X})) = \mathrm{epi}\,\theta_1'(\overline{X}; \cdot) := \left\{(H, y) \in \mathbb{S}^n \times \mathbb{R} \mid \theta_1'(\overline{X}; H) \leq y\right\}$$

- The **lineality space**:

$$\mathcal{T}_{\theta_1}^{\mathrm{lin}}(\overline{X}) := \left\{H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) = -\theta_1'(\overline{X}; -H)\right\}$$

Tangent sets

For $\theta_1 = \phi_1 \circ \lambda$:

- **Tangent set** of epigraph:

$$\mathcal{T}_{\mathrm{epi}\,\theta_1}(\overline{X}, \theta(\overline{X})) = \mathrm{epi}\,\theta_1'(\overline{X}; \cdot) := \left\{(H, y) \in \mathbb{S}^n \times \mathbb{R} \mid \theta_1'(\overline{X}; H) \le y\right\}$$

- The **lineality space**:

$$\mathcal{T}_{\theta_1}^{\mathrm{lin}}(\overline{X}) := \left\{H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) = -\theta_1'(\overline{X}; -H)\right\}$$

Proposition

$H \in \mathcal{T}_{\theta_1}^{\mathrm{lin}}(\overline{X})$ if and only if $\langle \mathbf{z}, \lambda'(\overline{X}; H)\rangle$ is a constant for any $\mathbf{z} \in \partial\phi_1(\lambda(\overline{X}))$, i.e.,

$$\langle \lambda'(\overline{X}; H), \mathbf{a}^i - \mathbf{a}^j\rangle = 0 \quad \forall\, i, j \in \iota_1(\lambda(\overline{X})).$$

Tangent sets (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

Tangent sets (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

- $\theta_2 = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom} \phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \leq 0\},$$

  where $\zeta = \psi \circ \lambda$

Tangent sets (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

- $\theta_2 = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom}\phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \leq 0\},$$

  where $\zeta = \psi \circ \lambda$

- **Tangent set** of $\mathcal{K}$:

$$\begin{aligned} \mathcal{T}_{\mathcal{K}}(\overline{X}) &= \left\{H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) \leq 0\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{b}^i, \lambda'(\overline{X}; H)\rangle \leq 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\} \end{aligned}$$

Tangent sets (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

- $\theta_2 = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom}\phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \leq 0\},$$

  where $\zeta = \psi \circ \lambda$

- **Tangent set** of $\mathcal{K}$:

$$\begin{aligned} \mathcal{T}_{\mathcal{K}}(\overline{X}) &= \left\{H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) \leq 0\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{b}^i, \lambda'(\overline{X}; H)\rangle \leq 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\} \end{aligned}$$

- The **lineality space**:

$$\operatorname{lin}(\mathcal{T}_{\mathcal{K}}(\overline{X})) = \left\{H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) = -\zeta'(\overline{X}; -H) = 0\right\}$$

Tangent sets (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

- $\theta_2 = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom}\phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \le 0\},$$

  where $\zeta = \psi \circ \lambda$

- **Tangent set** of $\mathcal{K}$:

$$\begin{aligned}\mathcal{T}_{\mathcal{K}}(\overline{X}) &= \left\{H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) \le 0\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{b}^i, \lambda'(\overline{X}; H)\rangle \le 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\}\end{aligned}$$

- The **lineality space**:

$$\operatorname{lin}(\mathcal{T}_{\mathcal{K}}(\overline{X})) = \left\{H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) = -\zeta'(\overline{X}; -H) = 0\right\}$$

Proposition

$H \in \operatorname{lin}(\mathcal{T}_{\mathcal{K}}(\overline{X}))$ *if and only if* $\langle \mathbf{b}^i, \lambda'(\overline{X}; H)\rangle = 0$ *for any* $i \in \iota_2(\lambda(\overline{X}))$.

Tangent sets: SDP

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1\leq i\leq n}\{\langle \mathbf{e}^i, \lambda(X)\rangle\} \leq 0\}$$

Tangent sets: SDP

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n}\{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$$

$$\overline{X} = \overline{V}\begin{bmatrix} 0_\alpha & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix}\overline{V}^T, \quad \iota_2(\lambda(\overline{X})) = \alpha \cup \beta = \overline{\gamma}$$

Tangent sets: SDP

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$$

$$\overline{X} = \overline{V} \begin{bmatrix} 0_\alpha & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \iota_2(\lambda(\overline{X})) = \alpha \cup \beta = \overline{\gamma}$$

$$\begin{aligned} \mathcal{T}_{\mathbb{S}^n_-}(\overline{X}) &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{e}^i, \lambda'(\overline{X}; H)\rangle \le 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \overline{V}_{\overline{\gamma}}^T H \overline{V}_{\overline{\gamma}} \preceq 0\right\} \end{aligned}$$

Tangent sets: SDP

$$\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n}\{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$$

$$\overline{X} = \overline{V}\begin{bmatrix} 0_\alpha & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix}\overline{V}^T, \quad \iota_2(\lambda(\overline{X})) = \alpha \cup \beta = \overline{\gamma}$$

$$\begin{aligned} \mathcal{T}_{\mathbb{S}^n_-}(\overline{X}) &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{e}^i, \lambda'(\overline{X}; H)\rangle \le 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \overline{V}_{\overline{\gamma}}^T H \overline{V}_{\overline{\gamma}} \preceq 0\right\} \end{aligned}$$

$$\begin{aligned} \operatorname{lin}(\mathcal{T}_{\mathbb{S}^n_-}(\overline{X})) &= \left\{H \in \mathbb{S}^n \mid \langle \mathbf{e}^i, \lambda'(\overline{X}; H)\rangle = 0 \quad \forall\, i \in \iota_2(\lambda(\overline{X}))\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \overline{V}_{\overline{\gamma}}^T H \overline{V}_{\overline{\gamma}} = 0\right\} \end{aligned}$$

Critical cone

For $\theta_1 = \phi_1 \circ \lambda$:

Critical cone

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$.
- **Critical cone**:

$$\begin{aligned}\mathcal{C}(A; \partial\theta_1(\overline{X})) &:= \left\{H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) \le \langle \overline{Y}, H\rangle\right\} \\ &= \left\{H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) = \langle \overline{Y}, H\rangle\right\}\end{aligned}$$

Critical cone

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$.
- **Critical cone**:

$$\begin{aligned} \mathcal{C}(A; \partial\theta_1(\overline{X})) &:= \left\{ H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) \leq \langle \overline{Y}, H \rangle \right\} \\ &= \left\{ H \in \mathbb{S}^n \mid \theta_1'(\overline{X}; H) = \langle \overline{Y}, H \rangle \right\} \end{aligned}$$

Proposition

$H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$ if and only if $H \in \mathbb{S}^n$ satisfies for any $i, j \in \eta_1(\overline{\mathbf{x}}, \overline{\mathbf{y}})$,

$$\langle \mathrm{diag}(\overline{U}^T H \overline{U}), \mathbf{a}^i \rangle = \langle \mathrm{diag}(\overline{U}^T H \overline{U}), \mathbf{a}^j \rangle = \max_{i \in \iota_1(\overline{\mathbf{x}})} \langle \lambda'(\overline{X}; H), \mathbf{a}^i \rangle,$$

where the index set $\eta_1(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \subseteq \iota_1(\overline{\mathbf{x}})$:

$$\eta_1(\overline{\mathbf{x}}, \overline{\mathbf{y}}) := \Big\{ i \in \iota_1(\overline{\mathbf{x}}) \mid \sum_{i \in \iota_1(\overline{\mathbf{x}})} u^i \mathbf{a}^i = \overline{\mathbf{y}}, \ \sum_{i \in \iota_1(\overline{\mathbf{x}})} u^i = 1, \ 0 < u^i \leq 1 \Big\}$$

with $\overline{\mathbf{x}} := \lambda(\overline{X})$ and $\overline{\mathbf{y}} := \lambda(\overline{Y})$.

Critical cone (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

Critical cone (cont'd)

For $\theta_2 = \phi_2 \circ \lambda$:

- Let $\overline{Z} \in \mathcal{N}_{\mathcal{K}}(\overline{X})$. Denote $B = \overline{X} + \overline{Z}$.
- **Critical cone**:

$$\mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(\overline{X})) := \mathcal{T}_{\mathcal{K}}(\overline{X}) \cap \overline{Z}^{\perp} = \left\{ H \in \mathbb{S}^n \mid \zeta'(\overline{X}; H) \leq 0,\ \langle \overline{Z}, H \rangle = 0 \right\}$$

Proposition

$H \in \mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(\overline{X}))$ if and only if $H \in \mathbb{S}^n$ satisfies for any $i \in \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}})$,

$$0 = \langle \mathrm{diag}(\overline{V}^T H \overline{V}), \mathbf{b}^i \rangle = \max_{i \in \iota_2(\overline{\mathbf{x}})} \langle \lambda'(\overline{X}; H), \mathbf{b}^i \rangle,$$

where the index set $\eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) \subseteq \iota_2(\overline{\mathbf{x}})$:

$$\eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) := \left\{ i \in \iota_2(\overline{\mathbf{x}}) \mid \sum_{i \in \iota_2(\overline{\mathbf{x}})} u^i \mathbf{b}^i = \overline{\mathbf{z}},\ u^i > 0 \right\}$$

with $\overline{\mathbf{x}} := \lambda(\overline{X})$ and $\overline{\mathbf{z}} := \lambda(\overline{Z})$.

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n}\{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1\leq i\leq n}\{\langle \mathbf{e}^i, \lambda(X)\rangle\} \leq 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \left\{ \begin{array}{l} \iota_2(\overline{\mathbf{x}}) = \alpha \cup \beta = \overline{\gamma} \\ \\ \end{array} \right.$$

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\} = \{X \in \mathbb{S}^n \mid \max_{1\leq i\leq n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \leq 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \begin{cases} \iota_2(\overline{\mathbf{x}}) = \alpha \cup \beta = \overline{\gamma} \\ \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) = \alpha \end{cases}$$

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \begin{cases} \iota_2(\overline{\mathbf{x}}) = \alpha \cup \beta = \overline{\gamma} \\ \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) = \alpha \end{cases}$$

$H \in \mathcal{C}(B; \mathcal{N}_{\mathbb{S}^n_-}(\overline{X}))$ if and only if for any $i \in \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}})$,

$$0 = \langle \mathrm{diag}(\overline{V}^T H \overline{V}), \mathbf{b}^i \rangle = \max_{i \in \iota_2(\overline{\mathbf{x}})} \langle \lambda'(\overline{X}; H), \mathbf{b}^i \rangle$$

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max\limits_{1\le i\le n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \begin{cases} \iota_2(\overline{\mathbf{x}}) = \alpha \cup \beta = \overline{\gamma} \\ \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) = \alpha \end{cases}$$

$H \in \mathcal{C}(B; \mathcal{N}_{\mathbb{S}^n_-}(\overline{X}))$ if and only if for any $i \in \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}})$,

$$0 = \langle \mathrm{diag}(\overline{V}^T H \overline{V}), \mathbf{b}^i\rangle = \max_{i \in \iota_2(\overline{\mathbf{x}})} \langle \lambda'(\overline{X}; H), \mathbf{b}^i\rangle$$

i.e.,

$$\overline{V}^T H \overline{V} = \begin{bmatrix} \mathrm{diag} = 0 & \preceq 0 & \times \\ \preceq 0 & \preceq \vdots & \times \\ \times & \times & \times \end{bmatrix}$$

Critical cone: SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\} = \{X \in \mathbb{S}^n \mid \max_{1\le i\le n} \{\langle \mathbf{e}^i, \lambda(X)\rangle\} \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & \vdots \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T, \quad \begin{cases} \iota_2(\overline{\mathbf{x}}) = \alpha \cup \beta = \overline{\gamma} \\ \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}}) = \alpha \end{cases}$$

$H \in \mathcal{C}(B; \mathcal{N}_{\mathbb{S}^n_-}(\overline{X}))$ if and only if for any $i \in \eta_2(\overline{\mathbf{x}}, \overline{\mathbf{z}})$,

$$0 = \langle \mathrm{diag}(\overline{V}^T H \overline{V}), \mathbf{b}^i \rangle = \max_{i \in \iota_2(\overline{\mathbf{x}})} \langle \lambda'(\overline{X}; H), \mathbf{b}^i \rangle$$

i.e.,

$$\overline{V}^T H \overline{V} = \begin{bmatrix} \mathrm{diag} = 0 & \preceq 0 & \times \\ \preceq 0 & \preceq \vdots & \times \\ \times & \times & \times \end{bmatrix} \quad \Longleftrightarrow \quad \overline{V}^T H \overline{V} = \begin{bmatrix} 0 & 0 & \times \\ 0 & \preceq 0 & \times \\ \times & \times & \times \end{bmatrix}$$

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$ and $H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$.
- θ_1 is (parabolic) second-order directionally differentiable:

$$F(W) := \theta_1''(\overline{X}; H, W) = \phi_1''(\lambda(\overline{X}); \lambda'(\overline{X}; H), \lambda''(\overline{X}; H, W))$$

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$ and $H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$.
- θ_1 is (parabolic) second-order directionally differentiable:

$$F(W) := \theta_1''(\overline{X}; H, W) = \phi_1''(\lambda(\overline{X}); \lambda'(\overline{X}; H), \lambda''(\overline{X}; H, W))$$

The σ-term of θ_1 $\triangleq$ the conjugate function $F^*(\overline{Y})$

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$ and $H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$.
- θ_1 is (parabolic) second-order directionally differentiable:

$$\digamma(W) := \theta_1''(\overline{X}; H, W) = \phi_1''(\lambda(\overline{X}); \lambda'(\overline{X}; H), \lambda''(\overline{X}; H, W))$$

The σ-term of θ_1 $\quad \triangleq \quad$ the conjugate function $\digamma^*(\overline{Y})$

Moreover,

$$\digamma^*(\overline{Y}) = 2\sum_{l=1}^{r} \langle \Lambda(\overline{Y})_{\alpha^l \alpha^l}, \overline{U}_{\alpha^l}^T H(\overline{X} - \bar{v}^l I)^\dagger H \overline{U}_{\alpha^l} \rangle$$

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$ and $H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$.
- θ_1 is (parabolic) second-order directionally differentiable:

$$\digamma(W) := \theta_1''(\overline{X}; H, W) = \phi_1''(\lambda(\overline{X}); \lambda'(\overline{X}; H), \lambda''(\overline{X}; H, W))$$

The σ-term of θ_1 $\quad \triangleq \quad$ the conjugate function $\digamma^*(\overline{Y})$

Moreover,

$$\digamma^*(\overline{Y}) = 2\sum_{l=1}^{r} \langle \Lambda(\overline{Y})_{\alpha^l \alpha^l}, \overline{U}_{\alpha^l}^T H(\overline{X} - \bar{v}^l I)^\dagger H \overline{U}_{\alpha^l} \rangle := \Upsilon_{\overline{X}}^1(\overline{Y}, H)$$

The "σ-term"

For $\theta_1 = \phi_1 \circ \lambda$:

- Let $\overline{Y} \in \partial\theta_1(\overline{X})$. Denote $A = \overline{X} + \overline{Y}$ and $H \in \mathcal{C}(A; \partial\theta_1(\overline{X}))$.
- θ_1 is (parabolic) second-order directionally differentiable:

$$F(W) := \theta_1''(\overline{X}; H, W) = \phi_1''(\lambda(\overline{X}); \lambda'(\overline{X}; H), \lambda''(\overline{X}; H, W))$$

The σ-term of θ_1 $\quad\triangleq\quad$ the conjugate function $F^*(\overline{Y})$

Moreover,

$$F^*(\overline{Y}) = 2\sum_{l=1}^{r} \langle \Lambda(\overline{Y})_{\alpha^l \alpha^l}, \overline{U}_{\alpha^l}^T H(\overline{X} - \bar{v}^l I)^\dagger H \overline{U}_{\alpha^l} \rangle := \Upsilon_{\overline{X}}^1(\overline{Y}, H)$$

$$\Upsilon_{\overline{X}}^1(\overline{Y}, H) = -2 \sum_{1 \le l < l' \le r} \sum_{i \in \alpha^l} \sum_{j \in \alpha^{l'}} \frac{\lambda_i(\overline{Y}) - \lambda_j(\overline{Y})}{\lambda_i(\overline{X}) - \lambda_j(\overline{X})} (\overline{U}_{\alpha^l}^T H \overline{U}_{\alpha^{l'}})_{ij}^2$$

The "σ-term" (cont'd)

For $\theta_2 = \phi_2 \circ \lambda = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom} \phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \leq 0\},$$

where $\zeta = \psi \circ \lambda$

The "σ-term" (cont'd)

For $\theta_2 = \phi_2 \circ \lambda = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom} \phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \le 0\},$$

where $\zeta = \psi \circ \lambda$

Let $\overline{Z} \in \mathcal{N}_{\mathcal{K}}(\overline{X})$. Denote $B = \overline{X} + \overline{Z}$ and $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

the "σ-term" of $\mathcal{K}$ $\triangleq$ the support function of $\mathcal{T}^2_{\mathcal{K}}(\overline{X}, H)$

$$\delta^*_{\mathcal{T}^2_{\mathcal{K}}(\overline{X},H)}(\overline{Z}) = 2\sum_{l=1}^{r} \langle \Lambda(\overline{Z})_{\alpha^l\alpha^l}, \overline{V}_{\alpha^l}^T H(\overline{X} - \bar{v}^l I)^\dagger H \overline{V}_{\alpha^l} \rangle$$

The "σ-term" (cont'd)

For $\theta_2 = \phi_2 \circ \lambda = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom}\phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \le 0\},$$

where $\zeta = \psi \circ \lambda$

Let $\overline{Z} \in \mathcal{N}_{\mathcal{K}}(\overline{X})$. Denote $B = \overline{X} + \overline{Z}$ and $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

the "σ-term" of $\mathcal{K}$ $\quad\triangleq\quad$ the support function of $\mathcal{T}^2_{\mathcal{K}}(\overline{X}, H)$

$$\delta^*_{\mathcal{T}^2_{\mathcal{K}}(\overline{X},H)}(\overline{Z}) = 2\sum_{l=1}^{r}\langle \Lambda(\overline{Z})_{\alpha^l\alpha^l}, \overline{V}^T_{\alpha^l} H(\overline{X} - \bar{v}^l I)^\dagger H \overline{V}_{\alpha^l}\rangle := \Upsilon^2_{\overline{X}}\left(\overline{Z}, H\right)$$

The "σ-term" (cont'd)

For $\theta_2 = \phi_2 \circ \lambda = \delta_{\mathcal{K}}$ with

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \lambda(X) \in \operatorname{dom}\phi\} = \{X \in \mathbb{S}^n \mid \zeta(X) \leq 0\},$$

where $\zeta = \psi \circ \lambda$

Let $\overline{Z} \in \mathcal{N}_{\mathcal{K}}(\overline{X})$. Denote $B = \overline{X} + \overline{Z}$ and $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

the "σ-term" of $\mathcal{K}$ $\triangleq$ the support function of $\mathcal{T}^2_{\mathcal{K}}(\overline{X}, H)$

$$\delta^*_{\mathcal{T}^2_{\mathcal{K}}(\overline{X},H)}(\overline{Z}) = 2\sum_{l=1}^{r}\langle \Lambda(\overline{Z})_{\alpha^l\alpha^l}, \overline{V}^T_{\alpha^l} H(\overline{X} - \bar{v}^l I)^\dagger H\overline{V}_{\alpha^l}\rangle := \Upsilon^2_{\overline{X}}(\overline{Z}, H)$$

$$\Upsilon^2_{\overline{X}}(\overline{Z}, H) = -2\sum_{1\leq l<l'\leq r}\sum_{i\in\alpha^l}\sum_{j\in\alpha^{l'}}\frac{\lambda_i(\overline{Z}) - \lambda_j(\overline{Z})}{\lambda_i(\overline{X}) - \lambda_j(\overline{X})}(\overline{V}^T_{\alpha^l} H\overline{V}_{\alpha^{l'}})^2_{ij}$$

The "σ-term": SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \leq 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$, $B = \overline{X} + \overline{Z}$, $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

The "σ-term": SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$, $B = \overline{X} + \overline{Z}$, $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & 0 \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T$$

The "σ-term" of $\mathbb{S}^n_-$:

$$\Upsilon^2_{\overline{X}}(\overline{Z}, H) = 2 \sum_{i \in \gamma, j \in \alpha} \frac{\lambda_j(\overline{Z})}{\lambda_i(\overline{X})} (\widetilde{H})^2_{ij},$$

The "σ-term": SDP

- $\mathbb{S}^n_- = \{X \in \mathbb{S}^n \mid \lambda_{\max}(X) \le 0\}$
- $\overline{Z} \in \mathcal{N}_{\mathbb{S}^n_-}(\overline{X})$, $B = \overline{X} + \overline{Z}$, $H \in \mathcal{C}(B, \mathcal{N}_{\mathcal{K}}(\overline{X}))$

$$\overline{X} + \overline{Z} = \overline{V} \begin{bmatrix} \Lambda_\alpha(\overline{Z}) & \cdots & 0 \\ \vdots & 0_\beta & 0 \\ 0 & \cdots & \Lambda_\gamma(\overline{X}) \end{bmatrix} \overline{V}^T$$

The "σ-term" of $\mathbb{S}^n_-$:

$$\Upsilon^2_{\overline{X}}(\overline{Z}, H) = 2 \sum_{i\in\gamma, j\in\alpha} \frac{\lambda_j(\overline{Z})}{\lambda_i(\overline{X})} (\widetilde{H})^2_{ij}, \quad \text{cf. } \textbf{(Sun, 2006)}$$

where $\widetilde{H} = \overline{V}^T H \overline{V}$.

Robinson CQ

CMatOP:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} \quad & f(\mathbf{x}) + \theta_1(g(\mathbf{x})) \\ \text{subject to} \quad & h(\mathbf{x}) = 0, \\ & g(\mathbf{x}) \in \mathcal{K} \end{aligned}$$

Robinson CQ

CMatOP:

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{X}}{\text{minimize}} & f(\mathbf{x}) + \theta_1(g(\mathbf{x})) \\ \text{subject to} & h(\mathbf{x}) = 0, \\ & g(\mathbf{x}) \in \mathcal{K} \end{array}$$

Proposition

Let $\overline{\mathbf{x}} \in \mathbb{X}$ be a feasible point of the CMatOP. We say that the Robinson CQ (RCQ) holds at $\overline{\mathbf{x}}$ if

$$\begin{bmatrix} h'(\overline{\mathbf{x}}) \\ g'(\overline{\mathbf{x}}) \end{bmatrix} \mathbb{X} + \begin{bmatrix} \{0\} \\ \mathcal{T}_{\mathcal{K}}(g(\overline{\mathbf{x}})) \end{bmatrix} = \begin{bmatrix} \mathbb{Y} \\ \mathbb{S}^n \end{bmatrix}.$$

Thus, the set of Lagrange multipliers $\mathcal{M}(\overline{\mathbf{x}})$ is a non-empty, convex, bounded and compact subset if and only if the RCQ holds at $\overline{\mathbf{x}}$.

Second-order optimality conditions

Critical cone of CMatOP:

$\mathcal{C}(\overline{\mathbf{x}}) := \{\mathbf{d} \in \mathbb{X} \mid h'(\overline{\mathbf{x}})\mathbf{d} = 0, g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(A; \partial\theta_1(g(\overline{\mathbf{x}}))), g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(g(\overline{\mathbf{x}})))\}$

Second-order optimality conditions

Critical cone of CMatOP:

$\mathcal{C}(\overline{\mathbf{x}}) := \{\mathbf{d} \in \mathbb{X} \mid h'(\overline{\mathbf{x}})\mathbf{d} = 0, g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(A; \partial\theta_1(g(\overline{\mathbf{x}}))), g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(g(\overline{\mathbf{x}})))\}$

Theorem ("no gap" second-order optimality conditions)

Suppose that $\overline{\mathbf{x}} \in \mathbb{X}$ is a locally optimal solution of the CMatOP and the RCQ holds. Then, the following inequality holds: for any $\mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}})$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{\langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}}, \mathbf{y}, Y, Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right)\right\} \geq 0.$$

Second-order optimality conditions

Critical cone of CMatOP:

$\mathcal{C}(\overline{\mathbf{x}}) := \{\mathbf{d} \in \mathbb{X} \mid h'(\overline{\mathbf{x}})\mathbf{d} = 0, g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(A; \partial\theta_1(g(\overline{\mathbf{x}}))), g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(g(\overline{\mathbf{x}})))\}$

Theorem ("no gap" second-order optimality conditions)

Suppose that $\overline{\mathbf{x}} \in \mathbb{X}$ is a locally optimal solution of the CMatOP and the RCQ holds. Then, the following inequality holds: for any $\mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}})$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{\langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}},\mathbf{y},Y,Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right)\right\} \geq 0.$$

Conversely, let $\overline{\mathbf{x}}$ be a feasible point of the CMatOP such that $\mathcal{M}(\overline{\mathbf{x}})$ is nonempty. Suppose that the RCQ holds at $\overline{\mathbf{x}}$. Then the following condition: for any $\mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}}) \setminus \{0\}$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{\langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}},\mathbf{y},Y,Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right)\right\} > 0$$

Second-order optimality conditions

Critical cone of CMatOP:

$\mathcal{C}(\overline{\mathbf{x}}) := \{\mathbf{d} \in \mathbb{X} \mid h'(\overline{\mathbf{x}})\mathbf{d} = 0, g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(A; \partial\theta_1(g(\overline{\mathbf{x}}))), g'(\overline{\mathbf{x}})\mathbf{d} \in \mathcal{C}(B; \mathcal{N}_{\mathcal{K}}(g(\overline{\mathbf{x}})))\}$

Theorem ("no gap" second-order optimality conditions)

Suppose that $\overline{\mathbf{x}} \in \mathbb{X}$ is a locally optimal solution of the CMatOP and the RCQ holds. Then, the following inequality holds: for any $\mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}})$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{ \langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}},\mathbf{y},Y,Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right) \right\} \geq 0.$$

Conversely, let $\overline{\mathbf{x}}$ be a feasible point of the CMatOP such that $\mathcal{M}(\overline{\mathbf{x}})$ is nonempty. Suppose that the RCQ holds at $\overline{\mathbf{x}}$. Then the following condition: for any $\mathbf{d} \in \mathcal{C}(\overline{\mathbf{x}}) \setminus \{0\}$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{ \langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}},\mathbf{y},Y,Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right) \right\} > 0$$

is necessary and sufficient for the quadratic growth condition at the point $\overline{\mathbf{x}}$: for any $\mathbf{x} \in N$ such that $h(\mathbf{x}) = 0$ and $g(\mathbf{x}) \in \mathcal{K}$,

$$f(\mathbf{x}) + \phi_1 \circ \lambda(g(\mathbf{x})) \geq f(\overline{\mathbf{x}}) + \phi_1 \circ \lambda(g(\overline{\mathbf{x}})) + \rho\|\mathbf{x} - \overline{\mathbf{x}}\|^2.$$

Strong second-order sufficient condition

Definition

Let $\overline{\mathbf{x}} \in \mathbb{X}$ be a stationary point of the CMatOP. We say the strong second-order sufficient condition holds at $\overline{\mathbf{x}}$ if for any $\mathbf{d} \in \widehat{\mathcal{C}}(\overline{\mathbf{x}}) \setminus \{0\}$,

$$\sup_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \left\{ \langle \mathbf{d}, \mathcal{L}''_{\mathbf{xx}}(\overline{\mathbf{x}},\mathbf{y},Y,Z)\mathbf{d}\rangle - \Upsilon^1_{g(\overline{\mathbf{x}})}\left(Y, g'(\overline{\mathbf{x}})\mathbf{d}\right) - \Upsilon^2_{g(\overline{\mathbf{x}})}\left(Z, g'(\overline{\mathbf{x}})\mathbf{d}\right) \right\} > 0$$

with

$$\widehat{\mathcal{C}}(\overline{\mathbf{x}}) := \bigcap_{(\mathbf{y},Y,Z)\in\mathcal{M}(\overline{\mathbf{x}})} \mathrm{app}(\mathbf{y},Y,Z),$$

where for any $(\mathbf{y},Y,Z) \in \mathcal{M}(\overline{\mathbf{x}})$, the set $\mathrm{app}(\mathbf{y},Y,Z)$ is given by

$$\begin{aligned}\mathrm{app}(\overline{\mathbf{y}},\overline{Y},\overline{Z}) := \{\mathbf{d} \in \mathbb{X} \quad | \quad & h'(\overline{\mathbf{x}})\mathbf{d} = 0,\ g'(\overline{\mathbf{x}})\mathbf{d} \in \mathrm{aff}(\mathcal{C}(A;\partial\theta_1(g(\overline{\mathbf{x}})))),\\ & g'(\overline{\mathbf{x}})\mathbf{d} \in \mathrm{aff}(\mathcal{C}(B;\mathcal{N}_{\mathcal{K}}(g(\overline{\mathbf{x}}))))\}.\end{aligned}$$

Constraint nondegeneracy (LICQ)

The **constraint nondegeneracy** for the CMatOP is defined as follows

$$\begin{bmatrix} h'(\overline{\mathbf{x}}) \\ g'(\overline{\mathbf{x}}) \\ g'(\overline{\mathbf{x}}) \end{bmatrix} \mathbb{X} + \begin{bmatrix} \{0\} \\ \mathcal{T}^{\mathrm{lin}}_{\theta_1}(g(\overline{\mathbf{x}})) \\ \mathrm{lin}\left(\mathcal{T}_{\mathcal{K}}(g(\overline{\mathbf{x}}))\right) \end{bmatrix} = \begin{bmatrix} \mathbb{Y} \\ \mathbb{S}^n \\ \mathbb{S}^n \end{bmatrix}.$$

Strong regularity of CMatOPs

Theorem

Let $\overline{\mathbf{x}} \in \mathbb{X}$ *be a stationary point of CMatOP with multipliers* $(\overline{\mathbf{y}}, \overline{Y}, \overline{Z})$*:*

(i) the strong second order sufficient condition and constraint nondegeneracy hold at $\overline{\mathbf{x}}$*;*

(ii) every element in $\partial F(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{Y}, \overline{Z})$ *is nonsingular;*

(iii) $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{Y}, \overline{Z})$ *is a strongly regular solution of the KKT system.*

It holds that (i) $\Longrightarrow$ *(ii)* $\Longrightarrow$ *(iii).*

Strong regularity of CMatOPs

Theorem

Let $\overline{\mathbf{x}} \in \mathbb{X}$ *be a stationary point of CMatOP with multipliers* $(\overline{\mathbf{y}}, \overline{Y}, \overline{Z})$*:*

(i) the strong second order sufficient condition and constraint nondegeneracy hold at $\overline{\mathbf{x}}$*;*

(ii) every element in $\partial F(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{Y}, \overline{Z})$ *is nonsingular;*

(iii) $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{Y}, \overline{Z})$ *is a strongly regular solution of the KKT system.*

It holds that (i) $\Longrightarrow$ *(ii)* $\Longrightarrow$ *(iii).*

(iii) $\Longrightarrow$ (i) can be established for particular CMatOPs:

- NLSDP (**Sun, MOR 2006**)
- CMatOPs with the sum of k-largest eigenvalues, etc (in our work)

Thank you!