Variational analysis of determinantal varieties

Yan Yang
ICOTA 2026, Shanghai

April 11, 2026

State Key Laboratory of Mathematical Sciences
Academy of Mathematics and Systems Science
Chinese Academy of Sciences, China

Joint work with **Dr. Bin Gao** and **Prof. Ya-xiang Yuan**

Outline

1. Background: low-rank optimization
2. Variational analysis of low-rank sets
3. First- and second-order tangent sets
4. Tangent sets bridge optimization landscapes
5. Second-order optimality on bounded-rank matrices

Background: low-rank optimization

Low-rank problems: a geometric view

Optimization on the fixed-rank manifold $r < \min\{m, n\}$

$$\min \ f(X) \quad \text{s.t. } X \in \mathcal{M}_r := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank}(X) = r\}$$

- $\mathcal{M}_r$ is a smooth manifold [Helmke-Shayman'95]
- Riemannian optimization algorithms [Vandereycken'13; He-Li-Jiang-Ma-Zhang'26]
- $\mathcal{M}_r$ is not closed

Low-rank problems: a geometric view

Optimization on the fixed-rank manifold $r < \min\{m, n\}$

$$\min \; f(X) \quad \text{s.t. } X \in \mathcal{M}_r := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank}(X) = r\}$$

- $\mathcal{M}_r$ is a smooth manifold [Helmke-Shayman'95]
- Riemannian optimization algorithms [Vandereycken'13; He-Li-Jiang-Ma-Zhang'26]
- $\mathcal{M}_r$ is not closed

Optimization on the set of bounded-rank matrices $r < \min\{m, n\}$

$$\min \; f(X) \quad \text{s.t. } X \in \mathcal{M}_{\le r} := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank}(X) \le r\}$$

- $\mathcal{M}_{\le r}$ is closed
- $\mathcal{M}_{\le r}$ is a real-algebraic variety: determinantal variety [Harris'92]
- Finding first-order stationary points [Schneider-Uschmajew'15; Levin-Kileel-Boumal'22; '24; Olikier-Absil'22; '23; '24]

Low-rank problems: a geometric view

Optimization on the fixed-rank manifold $r < \min\{m, n\}$

$$\min\ f(X) \quad \text{s.t. } X \in \mathcal{M}_r := \{X \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(X) = r\}$$

- $\mathcal{M}_r$ is a smooth manifold [Helmke-Shayman'95]
- Riemannian optimization algorithms [Vandereycken'13; He-Li-Jiang-Ma-Zhang'26]
- $\mathcal{M}_r$ is not closed

Optimization on the set of bounded-rank matrices $r < \min\{m, n\}$

$$\min\ f(X) \quad \text{s.t. } X \in \mathcal{M}_{\leq r} := \{X \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(X) \leq r\}$$

- $\mathcal{M}_{\leq r}$ is closed
- $\mathcal{M}_{\leq r}$ is a real-algebraic variety: determinantal variety [Harris'92]
- Finding first-order stationary points [Schneider-Uschmajew'15; Levin-Kileel-Boumal'22; '24; Olikier-Absil'22; '23; '24]

A unified variational analysis for low-rank sets ?

Applications of low-rank optimization

$$\boxed{\begin{array}{ll} \min & f(X) \\ \text{s.t.} & X \in \mathcal{M}_{\leq r} \end{array}}$$

- Recommendation system: movie ratings [Frolov-Oseledets'17]
- Hyperspectral Images [Zhang-He-Zhang-Shen-Yuan'13; Zhuang-Fu-Ng'21]
- Image and video inpainting [Bertalmio-Sapiro-Caselles-Ballester'00; Fu-Ruan-Luo-An-Jin'21; Luo-Zhao-Li-Ng-Meng'23]
- EEG (brain signals) data [Mørup-Hansen-Herrmann-Parnas-Arnfred'06; Kong-Kong-Fan-Zhao-Cichoki'17]
- Magnetic resonance imaging (MRI) [Banco-Aeron-Hoge'16; Choi-Bao-Zhang'18; Fessler'20]
- Data analysis, e.g., weather forecast [Loucheur-Absil-Journée'23] and exoplanet detection [Daglayan-Vary-Absil-Cantalloube-Christiaens-Gillis-Jacques-Leplat-Absil'24]

Low rank + additional constraints

Bounded-rank matrices + an additional constraint

$$\begin{array}{ll} \min & f(X) \\ \text{s.t.} & X \in \mathcal{M}_{\leq r} \cap \mathcal{H} \end{array}$$

Low rank + additional constraints

Bounded-rank matrices + an additional constraint

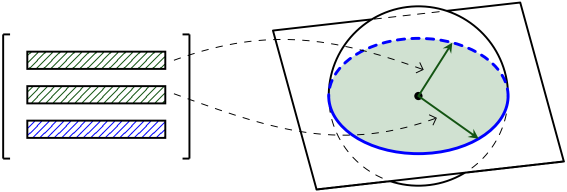
$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & X \in \mathcal{M}_{\leq r} \cap \mathcal{H} \end{aligned}$$

Low rank + oblique manifold

- Low-rank data fitting on sphere [Chu-Del Buono-Lopez-Politi'05 SIMAX]

$$\mathcal{H} = \mathrm{Ob}(m, n) := \{X \in \mathbb{R}^{m\times n} \mid \operatorname{diag}\left(XX^\top\right) - \mathbf{1} = \mathbf{0}\}$$

$\mathbb{R}^{3\times 4}_{\leq 2} \cap \mathrm{Ob}(3, 4) \ni$



Low rank + additional constraints (cont'd)

Low rank + Frobenius sphere

- Approximation of graph similarity matrices [Cason-Absil-Van Dooren'13 LAA]

$$\mathcal{H} = \mathrm{S}_{\mathrm{F}}(m,n) := \{X \in \mathbb{R}^{m \times n} \mid \|X\|_{\mathrm{F}}^2 - 1 = 0\}$$

Low rank + hyperbolic manifold

- Approximation of hyperbolic embeddings [Jawanpuria-Meghwanshi-Mishra'19 CDC]

$$\mathcal{H} = \mathbb{H}^m_{n-1} := \{X \in \mathbb{R}^{m \times n} \mid \mathrm{diag}(X\,\mathrm{Diag}(-1,1,\dots,1)X^\top) - \mathbf{1} = \mathbf{0}\}$$

Low rank + additional constraints (cont'd)

Low rank + Frobenius sphere

- Approximation of graph similarity matrices [Cason-Absil-Van Dooren'13 LAA]

$$\mathcal{H} = \mathrm{S}_{\mathrm{F}}(m, n) := \{X \in \mathbb{R}^{m\times n} \mid \|X\|_{\mathrm{F}}^2 - 1 = 0\}$$

Low rank + hyperbolic manifold

- Approximation of hyperbolic embeddings [Jawanpuria-Meghwanshi-Mishra'19 CDC]

$$\mathcal{H} = \mathbb{H}_{n-1}^{m} := \{X \in \mathbb{R}^{m\times n} \mid \mathrm{diag}(X\,\mathrm{Diag}(-1, 1, \dots, 1)X^\top) - \mathbf{1} = \mathbf{0}\}$$

Low rank + SDPs

- Low-rank semidefinite programming [Journée-Bach-Absil-Sepulchre'10 SIOPT]

$$\begin{aligned} \min \quad & \langle C, X\rangle \\ \text{s.t.} \quad & X \in \mathcal{M}_{\leq r} \cap \mathcal{S}^{+}(n) \cap \mathcal{U} \end{aligned}$$

where $\mathcal{U} = \{X \in \mathbb{R}^{n\times n} \mid \mathcal{A}(X) = b\}$

Low rank + additional constraints (cont'd)

Low rank + Frobenius sphere

- Approximation of graph similarity matrices [Cason-Absil-Van Dooren'13 LAA]

$$\mathcal{H} = \mathrm{S}_{\mathrm{F}}(m, n) := \{X \in \mathbb{R}^{m \times n} \mid \|X\|_{\mathrm{F}}^2 - 1 = 0\}$$

Low rank + hyperbolic manifold

- Approximation of hyperbolic embeddings [Jawanpuria-Meghwanshi-Mishra'19 CDC]

$$\mathcal{H} = \mathbb{H}_{n-1}^{m} := \{X \in \mathbb{R}^{m \times n} \mid \operatorname{diag}(X \operatorname{Diag}(-1, 1, \ldots, 1) X^\top) - \mathbf{1} = \mathbf{0}\}$$

Low rank + SDPs

- Low-rank semidefinite programming [Journée-Bach-Absil-Sepulchre'10 SIOPT]

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & X \in \mathcal{M}_{\leq r} \cap \mathcal{S}^{+}(n) \cap \mathcal{U} \end{aligned}$$

where $\mathcal{U} = \{X \in \mathbb{R}^{n \times n} \mid \mathcal{A}(X) = b\}$

☆ First- and second-order optimality conditions

Variational analysis of low-rank sets

Challenges

Constraint-coupled optimization

$$\begin{aligned} \min \quad & f(X) \\ \text{s. t.} \quad & X \in \mathcal{M}_{\leq r} \cap \mathcal{H} \end{aligned}$$

Challenges

- Inherent non-smoothness of $\mathcal{M}_{\leq r}$ [Olikier-Absil'22 SVVA; Levin-Kileel-Boumal'23; '25 MP]
- Geometry of the coupled region $\mathcal{M}_{\leq r} \cap \mathcal{H}$
  - Complicated Bouligand-tangent cone
  - Unknown second-order tangent set

Geometry: first-order analysis of determinantal varieties

Tangent and normal cones to $\mathcal{M}_{\leq r}$

- Mordukhovich normal cone [Luke'13]
- Bouligand tangent cone and Frèchet normal cone [Cason-Absil-Van Dooren'13]
- Clarke tangent cone and Clarke normal cone [Hosseini-Uschmajew'19; Li-Song-Xiu'19]
- Continuity of tangent cones [Olikier-Absil'22]

Geometry: first-order analysis of determinantal varieties

Tangent and normal cones to $\mathcal{M}_{\le r}$

- Mordukhovich normal cone [Luke'13]
- Bouligand tangent cone and Frèchet normal cone [Cason-Absil-Van Dooren'13]
- Clarke tangent cone and Clarke normal cone [Hosseini-Uschmajew'19; Li-Song-Xiu'19]
- Continuity of tangent cones [Olikier-Absil'22]

First-order geometry of $\mathcal{M}_{\le r} \cap \mathcal{H}$

- Mordukhovich normal cone: $\mathcal{H} = \mathrm{Sym}(n)$ [Tam'17]
- Frèchet normal cone: $\mathcal{H} = \mathrm{Sym}(n) \cap \mathcal{B}_i$ [Li-Xiu-Zhou'20 JOTA]

$$\mathcal{B}_1 = \{X \mid \|X\|_{\mathrm{F}}^2 \le 1\},\ \mathcal{B}_2 = \{X \mid -tI_n \preceq X \preceq tI_n\},\ \mathcal{B}_3 = \{X \mid X \succeq 0,\ \mathrm{tr}(X) = 1\}$$

- Bouligand tangent cone: $\mathcal{H} = \mathcal{S}^+(n)$ [Levin-Kileel-Boumal'25 MP]

Geometry: first-order analysis of determinantal varieties

Tangent and normal cones to $\mathcal{M}_{\leq r}$

- Mordukhovich normal cone [Luke'13]
- Bouligand tangent cone and Frèchet normal cone [Cason-Absil-Van Dooren'13]
- Clarke tangent cone and Clarke normal cone [Hosseini-Uschmajew'19; Li-Song-Xiu'19]
- Continuity of tangent cones [Olikier-Absil'22]

First-order geometry of $\mathcal{M}_{\leq r} \cap \mathcal{H}$

- Mordukhovich normal cone: $\mathcal{H} = \mathrm{Sym}(n)$ [Tam'17]
- Frèchet normal cone: $\mathcal{H} = \mathrm{Sym}(n) \cap \mathcal{B}_i$ [Li-Xiu-Zhou'20 JOTA]

  $$\mathcal{B}_1 = \{X \mid \|X\|_{\mathrm{F}}^2 \leq 1\},\ \mathcal{B}_2 = \{X \mid -tI_n \preceq X \preceq tI_n\},\ \mathcal{B}_3 = \{X \mid X \succeq 0,\ \mathrm{tr}(X) = 1\}$$

- Bouligand tangent cone: $\mathcal{H} = \mathcal{S}^+(n)$ [Levin-Kileel-Boumal'25 MP]
- Bouligand tangent cone: $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} \mid \|X\|_{\mathrm{F}}^2 = 1\}$ [Cason-Absil-Van Dooren'13 LAA]
- Frèchet normal cone: $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} \mid \mathcal{A}(X) = b\}$ [Li-Luo'23 SIOPT]

★ Bouligand tangent cone and Frèchet normal cone:
$\mathcal{H} = \{X \in \mathbb{R}^{m\times n} \mid h(X) = 0\}$ with orthogonally invariant h [Yang-Gao-Yuan'26 MP]

Geometry: first-order analysis of determinantal varieties

Tangent and normal cones to $\mathcal{M}_{\leq r}$

- Mordukhovich normal cone [Luke'13]
- Bouligand tangent cone and Frèchet normal cone [Cason-Absil-Van Dooren'13]
- Clarke tangent cone and Clarke normal cone [Hosseini-Uschmajew'19; Li-Song-Xiu'19]
- Continuity of tangent cones [Olikier-Absil'22]

First-order geometry of $\mathcal{M}_{\leq r} \cap \mathcal{H}$

- Mordukhovich normal cone: $\mathcal{H} = \mathrm{Sym}(n)$ [Tam'17]
- Frèchet normal cone: $\mathcal{H} = \mathrm{Sym}(n) \cap \mathcal{B}_i$ [Li-Xiu-Zhou'20 JOTA]

$$\mathcal{B}_1 = \{X \mid \|X\|_{\mathrm{F}}^2 \leq 1\},\ \mathcal{B}_2 = \{X \mid -tI_n \preceq X \preceq tI_n\},\ \mathcal{B}_3 = \{X \mid X \succeq 0,\ \mathrm{tr}(X) = 1\}$$

- Bouligand tangent cone: $\mathcal{H} = \mathcal{S}^{+}(n)$ [Levin-Kileel-Boumal'25 MP]
- Bouligand tangent cone: $\mathcal{H}=\{X \in \mathbb{R}^{m\times n} \mid \|X\|_{\mathrm{F}}^2 = 1\}$ [Cason-Absil-Van Dooren'13 LAA]
- Frèchet normal cone: $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} \mid \mathcal{A}(X) = b\}$ [Li-Luo'23 SIOPT]

★ Bouligand tangent cone and Frèchet normal cone:
$\mathcal{H} = \{X \in \mathbb{R}^{m\times n} \mid h(X) = 0\}$ with orthogonally invariant h [Yang-Gao-Yuan'26 MP]

A unified analysis remains unclear

Algorithm: finding first-order stationary points

- Projected gradient descent framework [Schneider-Uschmajew'15; Olikier-Absil'22;'24]
- Rank-adaptive methods [Olikier-Absil'23;'24; Gao-Peng-Yuan'25]
- Optimizing via a smooth parameterization

$$\begin{tikzcd} \overline{\mathcal{M}} \arrow[d, "\phi"'] \arrow[dr, "\bar{f} := f \circ \phi"] & \\ \mathcal{M}_{\leq r} \arrow[r, "f"'] & \mathbb{R} \end{tikzcd}$$

$$\min \bar{f}(Y) \quad \text{s. t. } Y \in \overline{\mathcal{M}}$$

  - LR factorization [Park-Kyrillidis-Caramanis-Sanghavi'18; Levin-Kileel-Boumal'22]

$$\overline{\mathcal{M}} = \mathbb{R}^{m \times r} \times \mathbb{R}^{n \times r}, \; \phi(L, R) = LR^\top$$

  - Desingularization [Khrulkov-Oseledets'18; Rebjock-Boumal'24]

$$\overline{\mathcal{M}} = \left\{ (X, G) \in \mathbb{R}^{m \times n} \times \mathrm{Gr}(n, n-r) : XG = 0 \right\}, \; \phi(X, G) = X$$

  - Conditions for "$1 \Rightarrow 1$" and "$2 \Rightarrow 1$" [Levin-Kileel-Boumal'25 MP]

Algorithm: finding first-order stationary points

- Projected gradient descent framework [Schneider-Uschmajew'15; Olikier-Absil'22;'24]
- Rank-adaptive methods [Olikier-Absil'23;'24; Gao-Peng-Yuan'25]
- Optimizing via a smooth parameterization

$$\begin{tikzcd}
\overline{\mathcal{M}} \arrow[d, "\phi"'] \arrow[dr, "\bar{f} := f \circ \phi"] & \\
\mathcal{M}_{\leq r} \arrow[r, "f"'] & \mathbb{R}
\end{tikzcd}$$

$$\min \bar{f}(Y) \quad \text{s.t. } Y \in \overline{\mathcal{M}}$$

  - LR factorization [Park-Kyrillidis-Caramanis-Sanghavi'18; Levin-Kileel-Boumal'22]

$$\overline{\mathcal{M}} = \mathbb{R}^{m\times r} \times \mathbb{R}^{n\times r},\ \phi(L,R) = LR^\top$$

  - Desingularization [Khrulkov-Oseledets'18; Rebjock-Boumal'24]

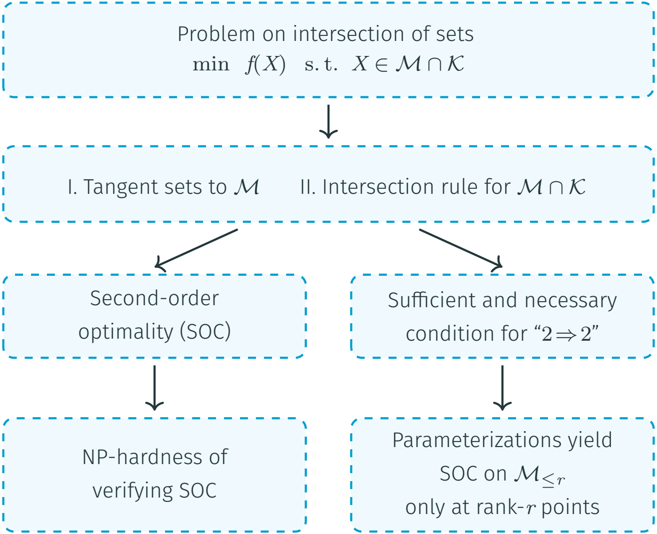
$$\overline{\mathcal{M}} = \left\{(X,G) \in \mathbb{R}^{m\times n} \times \mathrm{Gr}(n, n-r) : XG = 0\right\},\ \phi(X,G) = X$$

  - Conditions for "$1 \Rightarrow 1$" and "$2 \Rightarrow 1$" [Levin-Kileel-Boumal'25 MP]

An open question: characterization of "$2 \Rightarrow 2$"

Overview of our contributions

Tangent sets and optimization



Problem on intersection of sets
$\min \quad f(X) \quad \mathrm{s.t.} \quad X \in \mathcal{M} \cap \mathcal{K}$

I. Tangent sets to $\mathcal{M}$

II. Intersection rule for $\mathcal{M} \cap \mathcal{K}$

Second-order optimality (SOC)

Sufficient and necessary condition for “$2 \Rightarrow 2$”

NP-hardness of verifying SOC

Parameterizations yield SOC on $\mathcal{M}_{\leq r}$ only at rank-r points

A unified analysis of tangent sets

Set	Format	First-order	Second-order
$\mathcal{M}$	Assumption 1	Theorem 1	Theorem 1
$\mathcal{M}_{\leq r}$	matrix	[Schneider-Uschmajew'15 SIOPT]	Proposition 1
$\mathcal{M}^{\mathrm{ht}}_{\leq \mathbf{r}}$	hierarchical Tucker	Proposition 2	Proposition 2
$\mathcal{M}^{\mathrm{tc}}_{\leq \mathbf{r}}$	Tucker	[Gao-Peng-Yuan'25 MP]	Proposition 2
$\mathcal{M}^{\mathrm{tt}}_{\leq \mathbf{r}}$	tensor train	[Kutschan'18 LAA]	Proposition 2
$\mathcal{S}_{\leq r}(n)$	symmetric matrix	[Li-Xiu-Zhou'20 JOTA]	Proposition 3
$\mathcal{S}^{+}_{\leq r}(n)$	PSD matrix	[Levin-Kileel-Boumal'25 MP]	Proposition 4
Intersection of sets	Structured set	First-order	Second-order
$\mathcal{M} \cap \mathcal{K}$	Assumption 2	Theorem 2	Theorem 2
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is an affine manifold	[Li-Luo'23 SIOPT]	Appendix C.1
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is orthogonally invariant	[Yang-Gao-Yuan'25]	Appendix C.2
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is hyperbolic	Appendix C.3	Appendix C.3
$\mathcal{S}_{\leq r}(n) \cap \mathcal{U}$	$\mathcal{U} = \{X \mid \|X\|_{\mathrm{F}}^2 = 1\}$	Appendix D.2	Appendix D.2
$\mathcal{S}^{+}_{\leq r}(n) \cap \mathcal{U}$	$\mathcal{U} = \{X \mid \mathcal{A}(X) = b\}$	[Levin-Kileel-Boumal'25 MP]	Appendix D.3

Preliminaries

Bouligand tangent cone

$$\mathrm{T}_{\mathcal{X}}(X) = \{\eta \in \mathcal{E} : \exists t_i \to 0, \text{ s. t. } \operatorname{dist}(X + t_i\eta, \mathcal{X}) = o(t_i)\}$$

Fréchet and Mordukhovich normal cones

$$\begin{tikzcd}\mathrm{T}_{\mathcal{X}}(X) \arrow[r, "\text{polar}"] & \hat{\mathrm{N}}_{\mathcal{X}}(X) \arrow[r, "\lim"] & \mathrm{N}_{\mathcal{X}}(X)\end{tikzcd}$$

Second-order tangent set in the direction η

$$\mathrm{T}^2_{\mathcal{X}}(X, \eta) = \left\{\zeta \in \mathcal{E} : \exists t_i \to 0, \text{ s. t. } \operatorname{dist}(X + t_i\eta + \frac{1}{2}t_i^2\zeta, \mathcal{X}) = o(t_i^2)\right\}$$

Directionally derivative in the direction η

$$h'(X; \eta) := \lim_{t \to 0} \frac{h(X + t\eta) - h(X)}{t}$$

Parabolic second-order directional derivative in the direction (η, ζ)

$$h''(X; \eta, \zeta) := \lim_{t \to 0} \frac{h(X + t\eta + \frac{1}{2}t^2\zeta) - h(X) - th'(X; \eta)}{\frac{1}{2}t^2}$$

First-order geometry of low-rank sets

$\mathcal{M}_s$ as a smooth manifold based on the SVD $X = U\Sigma V^\top$

$$\mathrm{T}_{\mathcal{M}_s}(X) = \left\{ [U\ U_\perp] \begin{bmatrix} \mathbb{R}^{s\times s} & \mathbb{R}^{s\times(n-s)} \\ \mathbb{R}^{(m-s)\times s} & 0^{(m-s)\times(n-s)} \end{bmatrix} [V\ V_\perp]^\top \right\}$$

$$\mathrm{N}_{\mathcal{M}_s}(X) = \left\{ [U\ U_\perp] \begin{bmatrix} 0^{s\times s} & 0^{s\times(n-s)} \\ 0^{(m-s)\times s} & \mathbb{R}^{(m-s)\times(n-s)} \end{bmatrix} [V\ V_\perp]^\top \right\}$$

Geometry of $\mathcal{M}_{\le r}$ at $X = U\Sigma V^\top \in \mathcal{M}_s$ [Luke'13; Cason-Absil-Van Dooren'13; Schneider-Uschmajew'15]

$$\mathrm{T}_{\mathcal{M}_{\le r}}(X) = \mathrm{T}_{\mathcal{M}_s}(X) + \{R \in \mathrm{N}_X\mathcal{M}_s :\ \mathrm{rank}\,(R) \le r - s\}$$

$$\hat{\mathrm{N}}_{\mathcal{M}_{\le r}}(X) = \begin{cases} \mathrm{N}_{\mathcal{M}_s}(X), & \textit{if } s = r \\ \{0\}\,, & \textit{if } s < r \end{cases}$$

$$\mathrm{N}_{\mathcal{M}_{\le r}}(X) = \{Y \in \mathrm{N}_{\mathcal{M}_s}(X) : \mathrm{rank}(Y) \le \min\{m, n\} - r\}$$

All the cone mappings are not continuous at **singular** points [Olikier-Absil'22]

First- and second-order tangent sets

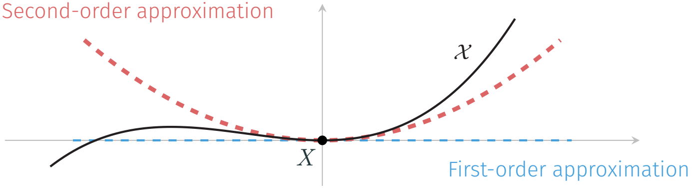
Motivation and challenge

First- and second-order tangent sets

$$\mathrm{T}_{\mathcal{X}}(X) = \{\eta \in \mathcal{E} : \exists t_i \to 0, \text{ s. t. } \operatorname{dist}(X + t_i\eta, \mathcal{X}) = o(t_i)\}$$

$$\mathrm{T}^2_{\mathcal{X}}(X, \eta) = \left\{\zeta \in \mathcal{E} : \exists t_i \to 0, \text{ s. t. } \operatorname{dist}(X + t_i\eta + \frac{1}{2}t_i^2\zeta, \mathcal{X}) = o(t_i^2)\right\}$$

Approximations to the set $\mathcal{X} = \{(x, y) \in \mathbb{R}^2 \mid y = x^2 + x^3\}$



Challenge when $\mathcal{X} = \mathcal{M}_{\leq r} = \left\{X \in \mathbb{R}^{m\times n} \mid \operatorname{rank}(X) \leq r\right\}$

- The geometry of $\mathcal{M}_{\leq r}$ is **complicated** [Olikier-Absil'22]
- The mapping $\operatorname{rank}(\cdot)$ is **not continuous**

Characterize low-rank sets via singular value functions

A new perspective

$$\mathcal{M}_{\le r} = \{X \in \mathbb{R}^{m\times n} \mid \sigma_{r+1}(X) = 0\}$$

- ☹ σ_{r+1} is both non-differentiable and non-convex
- ☺ σ_{r+1} is Lipschitz continuous [Weyl 1912]

$$|\sigma_{r+1}(X) - \sigma_{r+1}(X+\Delta)| \le \|\Delta\|_2$$

- ☺ σ_{r+1} controls the distance to $\mathcal{M}_{\le r}$

$$\mathrm{dist}(\tilde{X}, \mathcal{M}_{\le r}) \le (\min\{m,n\} - r)^{1/2}\sigma_{r+1}(\tilde{X})$$

Error bound condition at X

Let $\mathcal{M} := \{\tilde{X} \in \mathbb{R}^q \mid c_1(\tilde{X}) = 0, c_2(\tilde{X}) \le 0\}$. There exists a neighborhood $\mathcal{B}$ of $X \in \mathcal{M}$ and a constant $\rho > 0$ such that

$$\mathrm{dist}(\tilde{X}, \mathcal{M}) \le \rho\|(c_1(\tilde{X}),\ [c_2(\tilde{X})]_+)\|_2, \quad \text{for all } \tilde{X} \in \mathcal{B}$$

Bridge directional derivatives and tangent sets

Theorem

Let $\mathcal{M} = \{\tilde{X} \in \mathbb{R}^q \mid c_1(\tilde{X}) = 0, c_2(\tilde{X}) \leq 0\}$ with c_1 and c_2 satisfying

- error bound condition
- local Lipschitz property

Define the active index set $I_0(X) := \{j \in \{1, \dots, n_2\} \mid c_2(X)_j = 0\}$

(i) *(First-order)* If c_1 and c_2 are directionally differentiable at X, then

$$\mathrm{T}_{\mathcal{M}}(X) = \left\{\eta \in \mathbb{R}^q \mid c_1'(X;\eta) = 0,\ c_2'(X;\eta)_j \leq 0 \text{ for all } j \in I_0(X)\right\}$$

(ii) *(Second-order)* If, in addition, c_1 and c_2 admit parabolic second-order directional derivatives at X, then for any $\eta \in \mathrm{T}_{\mathcal{M}}(X)$

$$\mathrm{T}^2_{\mathcal{M}}(X;\eta) = \left\{\zeta \in \mathbb{R}^q \mid c_1''(X;\eta,\zeta) = 0,\ c_2''(X;\eta,\zeta)_j \leq 0 \text{ for all } j \in I_1(X;\eta)\right\}$$

where $I_1(X;\eta) := \{j \in I_0(X) \mid c_2'(X;\eta)_j = 0\}$

Application to low-rank matrices and tensors

Low-rank matrices

$$\mathcal{M}_{\leq r} = \{X \in \mathbb{R}^{m \times n} \mid \sigma_{r+1}(X) = 0\}$$

- $\| Y - \mathcal{P}_{\mathcal{M}_{\leq r}}(Y) \|_{\mathrm{F}} = \sqrt{\sum_{i=r+1}^{\min\{m,n\}} \sigma_i^2} \leq (\min\{m, n\} - r)^{1/2} \sigma_{r+1}$
- $\mathrm{T}_{\mathcal{M}_{\leq r}}(X) = \{\eta \in \mathbb{R}^{m \times n} \mid \sigma'_{r+1}(X; \eta) = 0\}$

Application to low-rank matrices and tensors

Low-rank matrices

$$\mathcal{M}_{\le r} = \{X \in \mathbb{R}^{m\times n} \mid \sigma_{r+1}(X) = 0\}$$

- $\| Y - \mathcal{P}_{\mathcal{M}_{\le r}}(Y)\|_{\mathrm{F}} = \sqrt{\sum_{i=r+1}^{\min\{m,n\}} \sigma_i^2} \le (\min\{m,n\} - r)^{1/2}\sigma_{r+1}$
- $\mathrm{T}_{\mathcal{M}_{\le r}}(X) = \{\eta \in \mathbb{R}^{m\times n} \mid \sigma'_{r+1}(X;\eta) = 0\}$

Low-rank tensors hierarchical Tucker, Tucker, and tensor train

$$\mathcal{M}^{\mathrm{ht}}_{\le \mathbf{r}} = \bigcap_{t\in T} \mathrm{ten}^{\mathrm{ht}}_{(t)}\left(\mathbb{R}^{n_t\times n_{t\_}}_{\le r_t}\right) = \left\{\mathbf{X} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_d} \mid \sigma_{r_t+1}(X^{\mathrm{ht}}_{(t)}) = 0 \text{ for } t \in T\right\}$$

- $\|\mathbf{Y} - \mathcal{P}^{\mathrm{HOSVD}}_{\le \mathbf{r}}(\mathbf{Y})\|_{\mathrm{F}} \le \rho' \sqrt{\sum_{t\in T} \sigma^2_{r_t+1}(Y^{\mathrm{ht}}_{(t)})}$
- $\mathrm{T}_{\mathcal{M}^{\mathrm{ht}}_{\le \mathbf{r}}}(\mathbf{X}) = \bigcap_{t\in T} \mathrm{ten}^{\mathrm{ht}}_{(t)}\left(\mathrm{T}_{\mathcal{R}_t}(X^{\mathrm{ht}}_{(t)})\right)$ with $\mathcal{R}_t := \mathbb{R}^{n_t\times n_{t\_}}_{\le r_t}$

Application to symmetric matrices

Low-rank symmetric matrices

$$\mathcal{S}_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \mathrm{rank}(X) \leq r\}$$

Application to symmetric matrices

Low-rank symmetric matrices

$$\mathcal{S}_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \mathrm{rank}(X) \leq r\} \; = \; \boxed{?}$$

- A decomposition

$$\mathcal{S}_{\leq r}(n) = \bigcup_{j=1}^{r+1} \mathcal{S}_j, \text{ with } \mathcal{S}_j := \{X \in \mathrm{Sym}(n) \mid \lambda_j(X) = 0,\ \lambda_{j+n-r-1}(X) = 0\}$$

- Verify the error bound condition of $(\lambda_j,\ \lambda_{j+n-r-1})$
- Derivation of tangent sets $\mathrm{T}_{\bigcup_{j=1}^{r+1} \mathcal{S}_j}(X) = \bigcup_{j=1}^{r+1} \mathrm{T}_{\mathcal{S}_j}(X)$

Application to symmetric matrices

Low-rank symmetric matrices

$$\mathcal{S}_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \mathrm{rank}(X) \leq r\} \ = \ ?$$

- A decomposition

$$\mathcal{S}_{\leq r}(n) = \bigcup_{j=1}^{r+1} \mathcal{S}_j, \text{ with } \mathcal{S}_j := \{X \in \mathrm{Sym}(n) \mid \lambda_j(X) = 0,\ \lambda_{j+n-r-1}(X) = 0\}$$

- Verify the error bound condition of $(\lambda_j,\ \lambda_{j+n-r-1})$
- Derivation of tangent sets $\mathrm{T}_{\bigcup_{j=1}^{r+1} \mathcal{S}_j}(X) = \bigcup_{j=1}^{r+1} \mathrm{T}_{\mathcal{S}_j}(X)$

Low-rank PSD matrices

$$\mathcal{S}^{+}_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \lambda_{r+1}(X) = 0, \lambda_n(X) = 0\}$$

Application to symmetric matrices

Low-rank symmetric matrices

$$\mathcal{S}_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \mathrm{rank}(X) \leq r\} \;=\; \boxed{?}$$

- A decomposition

$$\mathcal{S}_{\leq r}(n) = \bigcup_{j=1}^{r+1} \mathcal{S}_j, \text{ with } \mathcal{S}_j := \{X \in \mathrm{Sym}(n) \mid \lambda_j(X) = 0,\ \lambda_{j+n-r-1}(X) = 0\}$$

- Verify the error bound condition of $(\lambda_j,\ \lambda_{j+n-r-1})$
- Derivation of tangent sets $\mathrm{T}_{\bigcup_{j=1}^{r+1} \mathcal{S}_j}(X) = \bigcup_{j=1}^{r+1} \mathrm{T}_{\mathcal{S}_j}(X)$

Low-rank PSD matrices

$$\mathcal{S}^+_{\leq r}(n) = \{X \in \mathrm{Sym}(n) \mid \lambda_{r+1}(X) = 0, \lambda_n(X) = 0\}$$

$\mathcal{S}^+_{\leq r}(n)$ coincides with $\mathcal{S}_{r+1}$

Directional derivatives of singular values

Necessary notation [Zhang-Zhang-Xiao'13 SVVA]

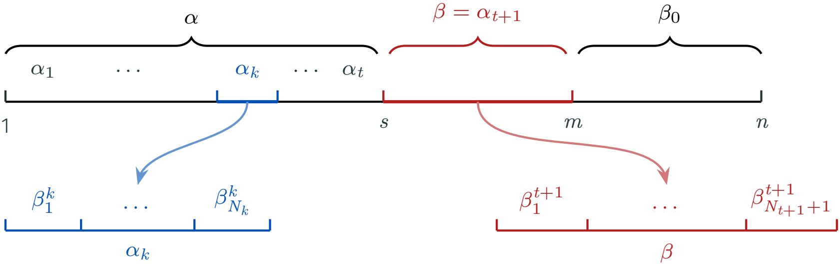
- the SVD $X = \bar{U}[\bar{\Sigma}(X)\ 0]\,\bar{V}^\top$ with $\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_m(X)$

$$\alpha = \{i : \sigma_i(X) > 0, 1 \leq i \leq m\}, \quad \beta = \{i : \sigma_i(X) = 0, 1 \leq i \leq m\}$$

$$\beta_0 = \{m+1, \ldots, n\}$$

- the distinct singular values $\mu_1 > \mu_2 > \ldots > \mu_p > \mu_{p+1} = 0$

$$\alpha_k = \{i : \sigma_i(X) = \mu_k, 1 \leq i \leq m\}, \ k = 1, \ldots, p, \text{ and } \widehat{\beta} = \beta \cup \beta_0$$

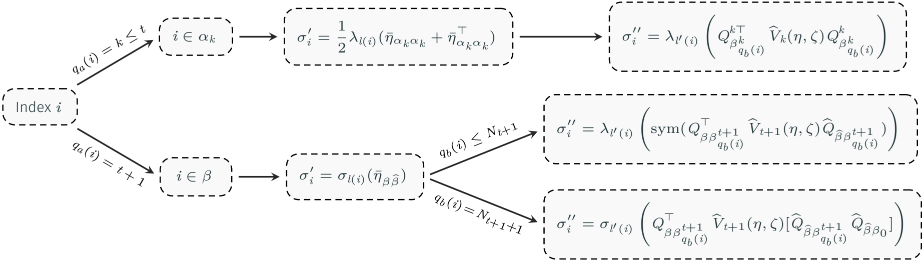


Directional derivatives of singular values (cont'd)

Four mappings

$$
\begin{aligned}
q_a &: \{1,\ldots,m\} \to \{1,\ldots,p+1\},\, q_a(i) = k, \text{ if } i \in \alpha_k \\
l &: \{1,\ldots,m\} \to \mathbb{N},\, l(i) = i - \kappa_{q_a(i)-1} \\
q_b &: \{1,\ldots,m\} \to \mathbb{N},\, q_b(i) = e, \text{ if } l(i) \in \beta_e^{q_a(i)} \\
l' &: \{1,\ldots,m\} \to \mathbb{N},\, l'(i) = l(i) - \kappa^{(q_a(i))}_{q_b(i)-1}
\end{aligned}
$$

Formula of $\sigma_i'(X;\eta)$ and $\sigma_i''(X;\eta,\zeta)$ at $X = \bar{U}[\bar{\Sigma}(X)\ 0]\,\bar{V}^\top$ for $i = 1, 2, \ldots, m$



Index i

$q_a(i) = k \le t$

$i \in \alpha_k$

$\sigma_i' = \frac{1}{2}\lambda_{l(i)}(\bar{\eta}_{\alpha_k\alpha_k} + \bar{\eta}_{\alpha_k\alpha_k}^\top)$

$\sigma_i'' = \lambda_{l'(i)}\left(Q^{k\top}_{\beta^k_{q_b(i)}}\widehat{V}_k(\eta,\zeta)\,Q^k_{\beta^k_{q_b(i)}}\right)$

$q_a(i) = t+1$

$i \in \beta$

$\sigma_i' = \sigma_{l(i)}(\bar{\eta}_{\beta\widehat{\beta}})$

$q_b(i) \le N_{t+1}$

$\sigma_i'' = \lambda_{l'(i)}\left(\mathrm{sym}(Q^\top_{\beta\beta^{t+1}_{q_b(i)}}\widehat{V}_{t+1}(\eta,\zeta)\widehat{Q}_{\widehat{\beta}\beta^{t+1}_{q_b(i)}})\right)$

$q_b(i) = N_{t+1}+1$

$\sigma_i'' = \sigma_{l'(i)}\left(Q^\top_{\beta\beta^{t+1}_{q_b(i)}}\widehat{V}_{t+1}(\eta,\zeta)[\widehat{Q}_{\widehat{\beta}\beta^{t+1}_{q_b(i)}}\ \widehat{Q}_{\widehat{\beta}\beta_0}]\right)$

Compute second-order tangent set

Important geometry of $\mathcal{M}_{\leq r}$

- consider $X = \bar{U}[\Sigma\ 0]\bar{V}^\top \in \mathcal{M}_s$ with $\bar{U} = [U\ U_\perp]$ and $\bar{V} = [V\ V_\perp]$
- $\eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X)$ if and only if

$$\eta = UAV^\top + U_\perp BV^\top + UCV_\perp^\top + U_\perp RV_\perp^\top$$

where $\operatorname{rank}(R) = \ell - s \leq r - s$

Index mappings associated with $\sigma_{r+1}(X) = 0$

- $r+1 \in \beta$, $q_a(r+1) = p+1$, $l(r+1) = r+1-s$
- $\bar{\eta}_{\beta\widehat{\beta}} = U_\perp^\top \eta V_\perp = R$ and $\operatorname{rank}(R) = \ell - s \Rightarrow q_b(r+1) = N_{p+1}+1$
- $l'(r+1) = \ell + 1 - s$

$$\begin{tikzcd}
 & & \sigma_i'' = \lambda_{l'(i)}(\operatorname{sym}(Q^T_{\beta\beta^{p+1}_{q_b(i)}}\widehat{V}_{p+1}(\eta,\zeta)\widehat{Q}_{\widehat{\beta}\beta^{p+1}_{q_b(i)}})) \\
i \in \beta \arrow[r] & \sigma_i' = \sigma_{l(i)}(\bar{\eta}_{\beta\widehat{\beta}}) \arrow[ur, "q_b(i) \leq N_{p+1}"] \arrow[dr, "q_b(i) = N_{p+1+1}"'] & \\
 & & \sigma_i'' = \sigma_{l'(i)}(Q^T_{\beta\beta^{p+1}_{q_b(i)}}\widehat{V}_{p+1}(\eta,\zeta)[\widehat{Q}_{\widehat{\beta}\beta^{p+1}_{q_b(i)}}\ \widehat{Q}_{\widehat{\beta}\beta_0}])
\end{tikzcd}$$

Second-order tangent set to $\mathcal{M}_{\leq r}$

Theorem

Given $X \in \mathcal{M}_{\leq r}$ and $\eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X)$ with $\operatorname{rank}(X) = s$ and the SVD $X = U\Sigma V^\top$. Let $\mathcal{P}_{\mathrm{N}_{\mathcal{M}_s}(X)}(\eta) = U_\eta \Sigma_\eta V_\eta^\top$, with $\operatorname{rank}(\Sigma_\eta) = \ell - s$. Take $[U\ U_\eta\ \ U_{\eta\perp}] \in \mathcal{O}(m)$ and $[V\ V_\eta\ \ V_{\eta\perp}] \in \mathcal{O}(n)$. It holds that

$$\mathrm{T}^2_{\mathcal{M}_{\leq r}}(X;\eta) = \left\{ 2\eta X^\dagger \eta + [U^+\ \ U_{\eta\perp}] \begin{bmatrix} W_1 & W_2 \\ W_3 & J \end{bmatrix} [V^+\ \ V_{\eta\perp}]^\top \;\middle|\; \begin{array}{l} W_1 \in \mathbb{R}^{\ell\times\ell} \\ W_2 \in \mathbb{R}^{\ell\times(n-\ell)} \\ W_3 \in \mathbb{R}^{(m-\ell)\times\ell} \\ J \in \mathbb{R}^{(m-\ell)\times(n-\ell)} \\ \operatorname{rank}(J) \leq r-\ell \end{array} \right\}$$

where $U^+ = [U\ U_\eta]$, $V^+ = [V\ V_\eta]$

A unified analysis of tangent sets to low-rank sets

Set	Format	First-order	Second-order
$\mathcal{M}$	general	Theorem 1	Theorem 1
$\mathcal{M}_{\leq r}$	matrix	[Schneider-Uschmajew'15 SIOPT]	Proposition 1
$\mathcal{M}^{\mathrm{ht}}_{\leq \mathbf{r}}$	hierarchical Tucker	Proposition 2	Proposition 2
$\mathcal{M}^{\mathrm{tc}}_{\leq \mathbf{r}}$	Tucker	[Gao-Peng-Yuan'25 MP]	Proposition 2
$\mathcal{M}^{\mathrm{tt}}_{\leq \mathbf{r}}$	tensor train	[Kutschan'18 LAA]	Proposition 2
$\mathcal{S}_{\leq r}(n)$	symmetric matrix	[Li-Xiu-Zhou'20 JOTA]	Proposition 3
$\mathcal{S}^{+}_{\leq r}(n)$	PSD matrix	[Levin-Kileel-Boumal'25 MP]	Proposition 4

A unified analysis of tangent sets to low-rank sets

Set	Format	First-order	Second-order
$\mathcal{M}$	general	Theorem 1	Theorem 1
$\mathcal{M}_{\leq r}$	matrix	[Schneider-Uschmajew'15 SIOPT]	Proposition 1
$\mathcal{M}^{\mathrm{ht}}_{\leq \mathbf{r}}$	hierarchical Tucker	Proposition 2	Proposition 2
$\mathcal{M}^{\mathrm{tc}}_{\leq \mathbf{r}}$	Tucker	[Gao-Peng-Yuan'25 MP]	Proposition 2
$\mathcal{M}^{\mathrm{tt}}_{\leq \mathbf{r}}$	tensor train	[Kutschan'18 LAA]	Proposition 2
$\mathcal{S}_{\leq r}(n)$	symmetric matrix	[Li-Xiu-Zhou'20 JOTA]	Proposition 3
$\mathcal{S}^{+}_{\leq r}(n)$	PSD matrix	[Levin-Kileel-Boumal'25 MP]	Proposition 4

Extension to intersection of sets $\mathcal{M} \cap \mathcal{K}$?

Intersection of sets

Theorem

- $\mathcal{M} = \{\tilde{X} \in \mathbb{R}^q \mid c(\tilde{X}) = 0\}$ with $c : \mathbb{R}^q \to \mathbb{R}^{n_1}$,
- $\mathcal{K} = \{\tilde{X} \in \mathbb{R}^q \mid h(\tilde{X}) = 0\}$ with smooth $h : \mathbb{R}^q \to \mathbb{R}^k$

Intersection of sets

Theorem

- $\mathcal{M} = \{\tilde{X} \in \mathbb{R}^q \mid c(\tilde{X}) = 0\}$ with $c : \mathbb{R}^q \to \mathbb{R}^{n_1}$,
- $\mathcal{K} = \{\tilde{X} \in \mathbb{R}^q \mid h(\tilde{X}) = 0\}$ with smooth $h : \mathbb{R}^q \to \mathbb{R}^k$
- Suppose that $(\mathcal{M}, \mathcal{K})$ satisfy

$$\begin{tikzcd}
\text{(transversal)}\ \overline{\mathcal{M}} \cap \overline{\mathcal{K}} \subseteq \mathbb{R}^p \arrow[d, "\text{(open) } \phi"'] \arrow[dr, "h \circ \phi\ \text{(rank-constant)}"] & \\
\mathcal{M} \cap \mathcal{K} \subseteq \mathbb{R}^q \arrow[r, "h"'] & \mathbb{R}^k
\end{tikzcd}$$

Intersection of sets

Theorem

- $\mathcal{M} = \{\tilde{X} \in \mathbb{R}^q \mid c(\tilde{X}) = 0\}$ with $c : \mathbb{R}^q \to \mathbb{R}^{n_1}$,
- $\mathcal{K} = \{\tilde{X} \in \mathbb{R}^q \mid h(\tilde{X}) = 0\}$ with smooth $h : \mathbb{R}^q \to \mathbb{R}^k$
- Suppose that $(\mathcal{M}, \mathcal{K})$ satisfy

$$\begin{tikzcd}
\text{(transversal)}\ \overline{\mathcal{M}} \cap \overline{\mathcal{K}} \subseteq \mathbb{R}^p \arrow[d, "\text{(open) } \phi"'] \arrow[rd, "h \circ \phi \text{ (rank-constant)}"] & \\
\mathcal{M} \cap \mathcal{K} \subseteq \mathbb{R}^q \arrow[r, "h"'] & \mathbb{R}^k
\end{tikzcd}$$

We have the following intersection rules for the tangent sets

$$\mathrm{T}_{\mathcal{M}\cap\mathcal{K}}(X) = \mathrm{T}_{\mathcal{M}}(X) \cap \mathrm{T}_{\mathcal{K}}(X)$$

$$\mathrm{T}^2_{\mathcal{M}\cap\mathcal{K}}(X;\eta) = \mathrm{T}^2_{\mathcal{M}}(X;\eta) \cap \mathrm{T}^2_{\mathcal{K}}(X;\eta)$$

Applications

Low-rank matrix sets $\mathcal{M}_{\le r} \cap \mathcal{H}$

- $\mathcal{H} = \mathrm{Aff}(m,n) = \{X \in \mathbb{R}^{m\times n} \mid \mathcal{A}(X) - b = 0\}$ [Li-Luo'23 SIOPT]
- $\mathcal{H} = \mathrm{S}_{\mathrm{F}}(m,n) = \{X \in \mathbb{R}^{m\times n} \mid \|X\|_{\mathrm{F}}^2 - 1 = 0\}$ [Cason-Absil-Van Dooren'13 LAA]
- $\mathcal{H} = \mathrm{Ob}(m,n) = \{X \in \mathbb{R}^{m\times n} \mid \mathrm{diag}(XX^\top) - \mathbf{1} = \mathbf{0}\}$ [Yang-Gao-Yuan'25]
- $\mathcal{H} = \mathbb{H}^m_{n-1} = \{X \in \mathbb{R}^{m\times n} \mid \mathrm{diag}(X\,\mathrm{Diag}(-1,1,\ldots,1)X^\top) + \mathbf{1} = \mathbf{0}\}$ ☺ new

A smooth parameterization

$$(\mathcal{M}_{\mathrm{LR}}, \phi_{\mathrm{LR}}) = (\mathbb{R}^{m\times r} \times \mathbb{R}^{n\times r},\ (L,R) \mapsto LR^\top)$$

☺ $\mathcal{M}_{\mathrm{LR}}$ coincides with the ambient space

☺ ϕ_{LR} is open around "balanced" (L,R)

★ It remains to prove $\overline{\mathcal{H}} = \phi_{\mathrm{LR}}^{-1}(\mathcal{H})$ is manifold

$$\begin{tikzcd}
\text{(transversal)}\ \mathcal{M}_{\mathrm{LR}} \cap \overline{\mathcal{H}} \arrow[d, "\text{(open)}\ \phi_{\mathrm{LR}}"'] \arrow[dr, "h \circ \phi\ \text{(rank-constant)}"] & \\
\mathcal{M}_{\le r} \cap \mathcal{H} \arrow[r, "h"'] & \mathbb{R}^k
\end{tikzcd}$$

A unified analysis of tangent sets

Set	Format	First-order	Second-order
$\mathcal{M}$	Assumption 1	Theorem 1	Theorem 1
$\mathcal{M}_{\leq r}$	matrix	[Schneider-Uschmajew'15 SIOPT]	Proposition 1
$\mathcal{M}^{\mathrm{ht}}_{\leq \mathbf{r}}$	hierarchical Tucker	Proposition 2	Proposition 2
$\mathcal{M}^{\mathrm{tc}}_{\leq \mathbf{r}}$	Tucker	[Gao-Peng-Yuan'25 MP]	Proposition 2
$\mathcal{M}^{\mathrm{tt}}_{\leq \mathbf{r}}$	tensor train	[Kutschan'18 LAA]	Proposition 2
$\mathcal{S}_{\leq r}(n)$	symmetric matrix	[Li-Xiu-Zhou'20 JOTA]	Proposition 3
$\mathcal{S}^{+}_{\leq r}(n)$	PSD matrix	[Levin-Kileel-Boumal'25 MP]	Proposition 4
Intersection of sets	**Structured set**	**First-order**	**Second-order**
$\mathcal{M} \cap \mathcal{K}$	Assumption 2	Theorem 2	Theorem 2
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is an affine manifold	[Li-Luo'23 SIOPT]	Appendix C.1
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is orthogonally invariant	[Yang-Gao-Yuan'25]	Appendix C.2
$\mathcal{M}_{\leq r} \cap \mathcal{H}$	$\mathcal{H}$ is hyperbolic	Appendix C.3	Appendix C.3
$\mathcal{S}_{\leq r}(n) \cap \mathcal{U}$	$\mathcal{U} = \{X \mid \|X\|_{\mathrm{F}}^2 = 1\}$	Appendix D.2	Appendix D.2
$\mathcal{S}^{+}_{\leq r}(n) \cap \mathcal{U}$	$\mathcal{U} = \{X \mid \mathcal{A}(X) = b\}$	[Levin-Kileel-Boumal'25 MP]	Appendix D.3

Tangent sets bridge optimization landscapes

Second-order optimality conditions

Second-order stationarity [Ruszczynski'06]

$$\min_{X \in \mathcal{M}} f(X),$$

- $-\nabla f(X^*) \in \hat{\mathrm{N}}_{\mathcal{M}}(X^*)$
- $\langle \nabla f(X^*), \zeta \rangle + \big\langle \eta, \nabla^2 f(X^*)[\eta] \big\rangle \geq 0$, for every $\eta \in \mathrm{T}_{\mathcal{M}}(X)$ such that $\langle \nabla f(X^*), \eta \rangle = 0$ and all $\zeta \in \mathrm{T}^2_{\mathcal{M}}(X^*; \eta)$

Smooth parameterization finds first-order stationary points [Levin-Kileel-Boumal'25 MP]

$$\begin{tikzcd}
\overline{\mathcal{M}} \subseteq \mathcal{W} \arrow[d, "\phi"'] \arrow[dr, "\bar{f} = f \circ \phi"] & \\
\mathcal{M} \subseteq \mathcal{E} \arrow[r, "f"'] & \mathbb{R}
\end{tikzcd}$$

$$\min_{Y \in \overline{\mathcal{M}}} \bar{f}(Y) = f(\phi(Y))$$

$$\min_{X \in \mathcal{M}} f(X)$$

When does a parameterization satisfy "$2 \Rightarrow 2$"?

A sufficient and necessary condition for "$2 \Rightarrow 2$"

Mappings connecting the geometry

$$\mathbf{L}_Y : \mathrm{T}_{\overline{\mathcal{M}}}(Y) \to \mathcal{E} : \; v \mapsto \mathrm{D}\phi_Y[v], \quad \text{for } Y \in \overline{\mathcal{M}}$$

$$\mathbf{Q}_{Y,v} : \mathrm{T}^2_{\overline{\mathcal{M}}}(Y; v) \to \mathcal{E} : \; u_v \mapsto \mathrm{D}\phi_Y[u_v] + \mathrm{D}^2\phi_Y[v, v], \quad \text{for } Y \in \overline{\mathcal{M}} \text{ and } v \in \mathrm{T}_{\overline{\mathcal{M}}}(Y)$$

Two important sets

$$\mathcal{Q}_Y(v) := \bigcup_{\{v_i\}_{i\in\mathbb{N}} :\, \mathbf{L}_Y(v_i) \to \mathbf{L}_Y(v)} \lim_{i\to\infty} \left(\mathbf{Q}_{Y,v_i}(u_{v_i}) + \mathrm{im}(\mathbf{L}_Y)\right), \quad \text{for } v \in \mathrm{T}_{\overline{\mathcal{M}}}(Y)$$

$$\mathit{\Gamma}_Y := \left\{\mathcal{P}_{\hat{\mathrm{N}}_{\mathcal{M}}(\phi(Y))} \left(\mathrm{D}\phi_Y[\mathrm{II}(v_0, v_1)] + \mathrm{D}^2\phi_Y[v_0, v_1]\right) \mid v_0 \in \ker(\mathbf{L}_Y),\; v_1 \in \mathrm{T}_{\overline{\mathcal{M}}}(Y)\right\}$$

Theorem

The parameterization $(\overline{\mathcal{M}}, \phi)$ satisfies "$2 \Rightarrow 2$" at $Y \in \overline{\mathcal{M}}$ if and only if $\mathrm{im}(\mathbf{L}_Y) = \mathrm{T}_{\mathcal{M}}(X)$ where $X = \phi(Y)$, and for all $v \in \mathrm{T}_{\overline{\mathcal{M}}}(Y)$

$$\mathrm{T}^2_{\mathcal{M}}(X; \mathbf{L}_Y(v)) \subseteq \overline{\mathrm{conv}}\left(\mathcal{Q}_Y(v) + \mathcal{Q}_Y(0) + \mathit{\Gamma}_Y\right)$$

Second-order optimality on bounded-rank matrices

Second-order stationarity on low-rank optimization

Proposition

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \quad & f(X) \\ \mathrm{s.\,t.} \quad & X \in \mathcal{M}_{\leq r} \cap \mathcal{H} \end{aligned}$$

The point X^* is second-order stationary if

$$\langle \nabla f(X^*), \eta\rangle = 0, \text{ for all } \eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X^*) \cap \mathrm{T}_{\mathcal{H}}(X^*)$$

and for all $\eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X^*) \cap \mathrm{T}_{\mathcal{H}}(X^*)$, it holds that

$$\langle \nabla f(X^*), \zeta\rangle + \left\langle \eta, \nabla^2 f(X^*)[\eta]\right\rangle \geq 0, \quad \text{for all } \zeta \in \mathrm{T}^2_{\mathcal{M}_{\leq r}}(X^*;\eta) \cap \mathrm{T}^2_{\mathcal{H}}(X^*;\eta)$$

Corollary $\mathcal{H} = \mathbb{R}^{m\times n}$

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \quad & f(X) \\ \mathrm{s.\,t.} \quad & X \in \mathcal{M}_{\leq r} \end{aligned}$$

The point $X^* \in \mathcal{M}_{\leq r}$ with $\mathrm{rank}(X^*) = s$ is second-order stationary if

$$\begin{cases} \nabla_{\mathcal{M}_r} f(X^*) = 0 \text{ and } \nabla^2_{\mathcal{M}_r} f(X^*) \succeq 0, & \text{if } s = r \\ \nabla f(X^*) = 0 \text{ and } \left\langle \eta, \nabla^2 f(X^*)[\eta]\right\rangle \geq 0 \text{ for all } \eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X^*), & \text{if } s < r \end{cases}$$

Complexity of verifying second-order conditions (SOC)

Find a negative curvature direction at X with $\operatorname{rank}(X) < r$

$$\begin{aligned} \min_{\eta \in \mathbb{R}^{m \times n}} \quad & \langle \eta, \mathcal{A}(\eta) \rangle \\ \text{s.t.} \quad & \|\eta\|_{\mathrm{F}} = 1, \\ & \eta \in \mathrm{T}_{\mathcal{M}_{\le r}}(X). \end{aligned} \tag{Q}$$

Problem: VERSOC
Input: Parameters m, n, r; point $X \in \mathcal{M}_{\le r}$; symmetric operator $\mathcal{A}$
Question: Does the optimal value of (Q) $\lambda^\star < 0$?

Complexity of verifying second-order conditions (SOC)

Find a negative curvature direction at X with $\mathrm{rank}(X) < r$

$$\begin{aligned} \min_{\eta \in \mathbb{R}^{m \times n}} \quad & \langle \eta, \mathcal{A}(\eta) \rangle \\ \text{s. t.} \quad & \|\eta\|_{\mathrm{F}} = 1, \\ & \eta \in \mathrm{T}_{\mathcal{M}_{\leq r}}(X). \end{aligned} \tag{Q}$$

Problem: VERSOC
Input: Parameters m, n, r; point $X \in \mathcal{M}_{\leq r}$; symmetric operator $\mathcal{A}$
Question: Does the optimal value of (Q) $\lambda^\star < 0$?

An NP-complete problem [Karp'72]

Problem: CLIQUE
Input: Undirected graph $G = (\mathcal{V}, E)$; clique size K
Question: Does there exist a clique of size K in G?

NP-hardness of verifying SOC

Theorem (NP-hardness)

The problem CLIQUE is polynomially reducible to VERSOC, and thus verifying second-order optimality conditions is NP-hard.

NP-hardness of verifying SOC

Theorem (NP-hardness)

The problem CLIQUE is polynomially reducible to VERSOC, and thus verifying second-order optimality conditions is NP-hard.

Theorem (No FPTAS)

Unless P=NP, there is no fully polynomial-time approximation scheme for verifying whether a point is second-order stationary for optimization problems over bounded-rank matrices.

Equivalence of problems

$$\begin{tikzcd}
\overline{\mathcal{M}} \arrow[d, "\phi"'] \arrow[dr, "\bar{f} = f \circ \phi"] & \\
\mathcal{M}_{\leq r} \arrow[r, "f"'] & \mathbb{R}
\end{tikzcd}$$

$$\min_{Y \in \overline{\mathcal{M}}} \quad \bar{f}(Y) = f(\phi(Y))$$

$$\min_{X \in \mathcal{M}_{\leq r}} \quad f(X)$$

Smooth parameterization of $\mathcal{M}_{\leq r}$

- LR factorization: $\overline{\mathcal{M}} = \mathbb{R}^{m \times r} \times \mathbb{R}^{n \times r}, \ \phi(L, R) = LR^\top$
- Desingularization: $\overline{\mathcal{M}} = \left\{(X, G) \in \mathbb{R}^{m \times n} \times \mathrm{Gr}(n, n-r) \mid XG = 0\right\}, \ \phi(X, G) = X$

$\mathrm{T}_{\mathcal{M}_{\leq r}}(X) = \mathrm{im}(\mathbf{L}_Y)$ and $\mathrm{T}^2_{\mathcal{M}_{\leq r}}(X; \mathbf{L}_Y(v)) = \mathrm{im}(\mathbf{Q}_{Y,v})$ when $\mathrm{rank}(X) = r$

Proposition

The parameterizations of $\mathcal{M}_{\leq r}$ satisfy "$2 \Rightarrow 2$" if and only if $\mathrm{rank}(X) = r$.

Conclusion and perspective

A unified analysis to derive tangent sets

★ Second-order optimality conditions [Bonnans-Cominetti-Shapiro'99 SIOPT]

- Metric subregularity [Gfrerer'11 SIOPT]
- System stability [Gfrerer-Mordukhovich'17 SIOPT]

Future work

☺ Efficient algorithm exploiting the geometry

$$\mathrm{T}_{\mathcal{H}\cap\mathcal{M}}(X) = \mathrm{T}_{\mathcal{H}}(X) \cap \mathrm{T}_{\mathcal{M}}(X)$$

? Second-order development beneficial from the property

$$\mathrm{T}^2_{\mathcal{H}\cap\mathcal{M}}(X;\eta) = \mathrm{T}^2_{\mathcal{H}}(X;\eta) \cap \mathrm{T}^2_{\mathcal{M}}(X;\eta)$$

Conclusion and perspective

A unified analysis to derive tangent sets

★ Second-order optimality conditions [Bonnans-Cominetti-Shapiro'99 SIOPT]

- Metric subregularity [Gfrerer'11 SIOPT]
- System stability [Gfrerer-Mordukhovich'17 SIOPT]

Future work

☺ Efficient algorithm exploiting the geometry

$$\mathrm{T}_{\mathcal{H}\cap\mathcal{M}}(X)=\mathrm{T}_{\mathcal{H}}(X)\cap\mathrm{T}_{\mathcal{M}}(X)$$

? Second-order development beneficial from the property

$$\mathrm{T}^2_{\mathcal{H}\cap\mathcal{M}}(X;\eta)=\mathrm{T}^2_{\mathcal{H}}(X;\eta)\cap\mathrm{T}^2_{\mathcal{M}}(X;\eta)$$

References

- **Yan Yang**, Bin Gao, Ya-xiang Yuan. Variational analysis of determinantal varieties. arXiv:2511.22613 2025
- **Yan Yang**, Bin Gao, Ya-xiang Yuan. *A space-decoupling framework for optimization on bounded-rank matrices with orthogonally invariant constraints.* Mathematical Programming (2026)

Thanks for your attention!

✉ Email: yangyan@amss.ac.cn
Homepage: https://ucas-yanyang.github.io/