A Space-decoupling Framework for Optimization on Bounded-rank Matrices with Orthogonally Invariant Constraints

Yan Yang
第八届北京大学计算与应用数学拔尖博士生研讨会
November 19, 2025

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

Low-rank optimization

Low-rank problems: a geometric view

Optimization on the fixed-rank manifold

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

- $\mathbb{R}_r^{m\times n}$ is a smooth manifold [Helmke and Shayman, 1995]
- Riemannian optimization algorithms [Vandereycken, 2013]
- $\mathbb{R}_r^{m\times n}$ is not closed

Optimization on the set of bounded-rank matrices

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

- $\mathbb{R}_{\le r}^{m\times n}$ is a real-algebraic variety [Harris, 1992]
- Geometric methods [Schneider and Uschmajew, 2015; Hosseini and Uschmajew, 2019; Gao and Absil, 2022; Olikier et al., 2022; Levin et al., 2023; Rebjock and Boumal, 2024]
- $\mathbb{R}_{\le r}^{m\times n}$ is closed

Bounded-rank matrices + orthogonally invariant constraints

$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & \operatorname{rank}(X) \leq r \\ & h(X) = 0 \end{aligned}$$

Low rank + additional constraints

Bounded-rank matrices + orthogonally invariant constraints

$$\begin{array}{ll} \min & f(X) \\ \text{s.t.} & \operatorname{rank}(X) \leq r \\ & h(X) = 0 \end{array}$$

Blanket assumptions

- Objective: $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is twice continuously differentiable
- Smooth and orthogonally invariant constraints

$$h(XQ) = h(X), \text{ for } Q \in \mathcal{O}(n)$$

- The differential $\mathrm{D}h$ has full rank in

$$\mathcal{H} := \left\{X \in \mathbb{R}^{m\times n} : \ h(X) = 0\right\}$$

Bounded-rank matrices + orthogonally invariant constraints

$$\begin{array}{ll} \min & f(X) \\ \text{s.t.} & \operatorname{rank}(X) \leq r \\ & h(X) = 0 \end{array}$$

Blanket assumptions

- Objective: $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is twice continuously differentiable
- Smooth and orthogonally invariant constraints

$$h(XQ) = h(X), \text{ for } Q \in \mathcal{O}(n)$$

- The differential $\mathrm{D}h$ has full rank in

$$\mathcal{H} := \left\{X \in \mathbb{R}^{m\times n} :\ h(X) = 0\right\}$$

$\mathcal{H}$ is a smooth manifold embedded in $\mathbb{R}^{m\times n}$

Applications

Coupled feasible region

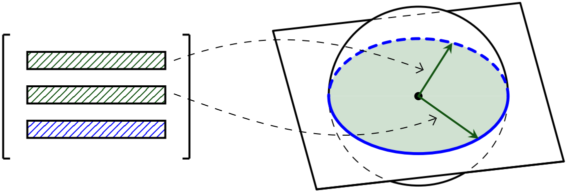
$$\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \;\text{ with }\; \mathcal{H} = \{X :\ h(X) = 0\}$$

Low rank + oblique manifold

- Low-rank data fitting on sphere [Chu et al., 2005 SIMAX]

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

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



Applications (cont'd)

Low rank + Frobenius sphere

- Low-rank approximation of graph similarity matrices [Cason et al., 2013 LAA]

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

Applications (cont'd)

Low rank + Frobenius sphere

- Low-rank approximation of graph similarity matrices [Cason et al., 2013 LAA]

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

Low rank + stacked Stiefel manifold

- Low-rank solutions of synchronization problems [Boumal, 2015]

$$\mathcal{H} = \left\{X = (X_1; X_2; \ldots; X_k) \in \mathbb{R}^{kp\times n} :\ X_j^\top \in \mathrm{St}(n, p) \text{ for } j = 1, 2, \ldots, k\right\}$$

$$\begin{bmatrix} \boxed{X_1 :\ X_1 X_1^\top = I} \\ \boxed{X_2 :\ X_2 X_2^\top = I} \\ \boxed{X_3 :\ X_3 X_3^\top = I} \end{bmatrix} \in \mathcal{H}$$

Low rank + Frobenius sphere

- Low-rank approximation of graph similarity matrices [Cason et al., 2013 LAA]

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

Low rank + stacked Stiefel manifold

- Low-rank solutions of synchronization problems [Boumal, 2015]

$$\mathcal{H} = \left\{X = (X_1; X_2; \ldots; X_k) \in \mathbb{R}^{kp\times n} : X_j^\top \in \mathrm{St}(n, p) \text{ for } j = 1, 2, \ldots, k\right\}$$

$$\begin{bmatrix} \boxed{X_1 : X_1 X_1^\top = I} \\ \boxed{X_2 : X_2 X_2^\top = I} \\ \boxed{X_3 : X_3 X_3^\top = I} \end{bmatrix} \in \mathcal{H}$$

Low rank + SDPs

- Low-rank semidefinite programs [Journée et al., 2010; Tang and Toh, 2023]

$$\min \ \langle C, Y\rangle \qquad \mathrm{s.t.}\ \ Y \in \mathbb{R}^{n\times n}_{\le r} \cap \mathbb{S}^n_+,\ \ \mathcal{A}(Y) = b$$

$$\min \ \left\langle C, XX^\top \right\rangle \quad \mathrm{s.t.}\ \ X \in \mathbb{R}^{n\times n}_{\le r},\ \ \mathcal{A}(XX^\top) = b$$

Challenges

Constraint-coupled optimization

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

Challenges

- Unknown **optimality conditions**

  Bouligand-tangent cone of the coupled feasible region

- Inherently **non-smooth structure** of $\mathbb{R}^{m\times n}_{\leq r}$ [Levin et al., 2023 MP]
- Unclear **strategy to preserve feasibility** on $\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}$

Constraint-coupled optimization

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

Challenges

- Unknown **optimality conditions**

  Bouligand-tangent cone of the coupled feasible region

- Inherently **non-smooth structure** of $\mathbb{R}^{m\times n}_{\le r}$ [Levin et al., 2023 MP]
- Unclear **strategy to preserve feasibility** on $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H}$

How to tackle the coupled and non-smooth constraints?

Bounded-rank optimization: optimality analysis

Variational geometry of $\mathbb{R}_{\leq r}^{m \times n}$

- Mordukhovich normal cone [Luke, 2013]
- Bouligand tangent cone and Frèchet normal cone [Schneider and Uschmajew, 2015]
- Clarke tangent cone and the corresponding normal cone [Hosseini et al., 2019; Li et al., 2019]

Bounded-rank optimization: optimality analysis

Variational geometry of $\mathbb{R}^{m\times n}_{\le r}$

- Mordukhovich normal cone [Luke, 2013]
- Bouligand tangent cone and Frèchet normal cone [Schneider and Uschmajew, 2015]
- Clarke tangent cone and the corresponding normal cone [Hosseini et al., 2019; Li et al., 2019]

Variational geometry of $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{C}$

- Mordukhovich normal cone when $\mathcal{C} = \mathrm{Sym}(n)$ [Tam, 2017]
- Frèchet normal cone when $\mathcal{C} = \mathrm{Sym}(n) \cap \mathcal{B}_i$ [Li et al., 2020]

  $\mathcal{B}_1 = \{X : \|X\|_\mathrm{F} \le 1\}$, $\mathcal{B}_2 = \{X : -tI_n \preceq X \preceq tI_n\}$, $\mathcal{B}_3 = \{X : X \succeq 0,\ \mathrm{Tr}(X) = 1\}$
- When $\mathcal{C}$ is a spectral set [Lewis, 1996; Pan, 2017; Li and Luo, 2023]

  $m \neq n$ **breaks the symmetry!**
- Bouligand tangent cone when $\mathcal{C} = \{X \in \mathbb{R}^{m\times n} : \|X\|_\mathrm{F} = 1\}$ [Cason et al., 2013]
- Frèchet normal cone when $\mathcal{C} = \{X \in \mathbb{R}^{m\times n} : \mathcal{A}(X) = b\}$ [Li and Luo, 2023]

Variational geometry of $\mathbb{R}^{m\times n}_{\le r}$

- Mordukhovich normal cone [Luke, 2013]
- Bouligand tangent cone and Frèchet normal cone [Schneider and Uschmajew, 2015]
- Clarke tangent cone and the corresponding normal cone [Hosseini et al., 2019; Li et al., 2019]

Variational geometry of $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{C}$

- Mordukhovich normal cone when $\mathcal{C} = \mathrm{Sym}(n)$ [Tam, 2017]
- Frèchet normal cone when $\mathcal{C} = \mathrm{Sym}(n) \cap \mathcal{B}_i$ [Li et al., 2020]

  $\mathcal{B}_1 = \{X : \|X\|_{\mathrm{F}} \le 1\}$, $\mathcal{B}_2 = \{X : -tI_n \preceq X \preceq tI_n\}$, $\mathcal{B}_3 = \{X : X \succeq 0,\ \mathrm{Tr}(X) = 1\}$
- When $\mathcal{C}$ is a spectral set [Lewis, 1996; Pan, 2017; Li and Luo, 2023]

  $m \neq n$ **breaks the symmetry!**
- Bouligand tangent cone when $\mathcal{C} = \{X \in \mathbb{R}^{m\times n} : \|X\|_{\mathrm{F}} = 1\}$ [Cason et al., 2013]
- Frèchet normal cone when $\mathcal{C} = \{X \in \mathbb{R}^{m\times n} : \mathcal{A}(X) = b\}$ [Li and Luo, 2023]

Not available for $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H}$

Bounded-rank optimization: algorithms

- Projected gradient descent framework [Jain et al., 2014; Schneider and Uschmajew, 2015; Olikier et al., 2022; 2024]
- Retraction-free methods [Schneider and Uschmajew, 2015; Olikier et al., 2023; 2024]
- Optimizing over a smooth parameterization

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

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

- LR factorization [Mishra et al. 2012; Levin et al. 2022]

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

- SVD-type lift [Mishra et al., 2014; Levin et al., 2023]

$$\mathcal{M} = \mathrm{St}(m, r) \times \mathrm{Sym}(r) \times \mathrm{St}(n, r), \;\phi(U, S, V) = USV^\top$$

- Desingularization [Khrulkov and Oseledets, 2018; Rebjock and Boumal, 2024]

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

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

How to tackle the coupled and non-smooth constraints?

A space-decoupling framework?

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

How to tackle the coupled and non-smooth constraints?

$$
\begin{tikzcd}
 & \mathcal{M}_h \arrow[dl] \arrow[d] \arrow[r, hook, "\phi"] & \mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H} \\
\boxed{\text{rank information?}} & \boxed{h(X) = 0?} &
\end{tikzcd}
$$

$$\begin{array}{ll} \min\limits_{X \in \mathbb{R}^{m\times n}} & f(X) \\ \text{s.t.} & X \in \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \end{array}$$

How to tackle the coupled and non-smooth constraints?

$$\begin{tikzcd} & \mathcal{M}_h \arrow[dl] \arrow[d] \arrow[r, hook, "\phi"] & \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \\ \boxed{\text{rank information?}} & \boxed{h(X) = 0?} & \end{tikzcd}$$

Goal

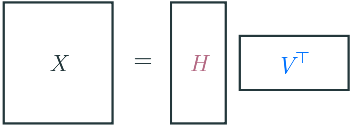
- optimality analysis
- smooth parametrization
- equivalence of two problems
- convergence analysis

Orthogonally invariant functions and optimality analysis

Orthogonal invariance “meets” low rank

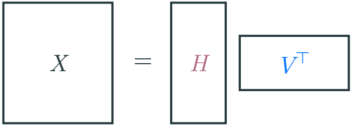
Low-rank decomposition

- $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} :\ h(X) = 0\}$ with orthogonally invariant h
- $X \in \mathbb{R}_s^{m\times n} \cap \mathcal{H}$ extracts low-rank (H) and orthogonally invariant (V)



Low-rank decomposition

- $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} : \ h(X) = 0\}$ with orthogonally invariant h
- $X \in \mathbb{R}_s^{m\times n} \cap \mathcal{H}$ extracts low-rank (H) and orthogonally invariant (V)



Submanifold inherites orthogonal invariance

$$X = HV^\top, \quad V^\top V = I$$

- natural embeddings $i^s(\cdot) : \ \mathbb{R}^{m\times s} \longrightarrow \mathbb{R}^{m\times n}, \ H \longmapsto [H \, 0^{m\times(n-s)}]$
- $h^s := h \circ i^s$ is orthogonally invariant
- $\mathcal{H}^s := \{H \in \mathbb{R}^{m\times s} : h^s(H) = 0\}$ is an embedded submanifold of $\mathbb{R}^{m\times s}$

Low-rank decomposition

- $\mathcal{H} = \{X \in \mathbb{R}^{m\times n} : \ h(X) = 0\}$ with orthogonally invariant h
- $X \in \mathbb{R}_s^{m\times n} \cap \mathcal{H}$ extracts low-rank (H) and orthogonally invariant (V)

$$X = H \; V^\top$$

Submanifold inherites orthogonal invariance

$$X = HV^\top, \quad V^\top V = I$$

- natural embeddings $i^s(\cdot) : \ \mathbb{R}^{m\times s} \longrightarrow \mathbb{R}^{m\times n}, \ H \longmapsto [H \, 0^{m\times(n-s)}]$
- $h^s := h \circ i^s$ is orthogonally invariant
- $\mathcal{H}^s := \{H \in \mathbb{R}^{m\times s} : h^s(H) = 0\}$ is an embedded submanifold of $\mathbb{R}^{m\times s}$

$$\boxed{X \in \mathcal{H} \iff H \in \mathcal{H}^s}$$

Space decoupling - level 1: tangent space

$$\mathcal{H} = \{X \in \mathbb{R}^{m\times n} :\ h(X) = 0\} \quad \text{and} \quad h(XQ) = h(X) \text{ for } Q \in \mathcal{O}(n)$$

Definition

- *Bouligand tangent cone*:

$$\mathrm{T}_X\mathcal{X} := \left\{\eta = \lim_{k\to\infty} \frac{(X_k - X)}{t_k} :\ X_k \in \mathcal{X}, t_k > 0 \textit{ for all } k, \lim_{k\to\infty} t_k = 0\right\}.$$

- *Fréchet normal cone*: $\mathrm{N}_X\mathcal{X} := (\mathrm{T}_X\mathcal{X})^\circ$

Tangent space decoupling at $X = HV^\top \in \mathcal{H}$

$$\begin{aligned}\mathrm{T}_X\mathcal{H} &= \left\{KV^\top + BV_\perp^\top :\ K \in \mathrm{T}_H\mathcal{H}^s,\ B \in \mathbb{R}^{m\times(n-s)}\right\}\\ &= (\mathrm{T}_H\mathcal{H}^s)\ V^\top \oplus \left(\mathbb{R}^{m\times(n-s)}\right)\ V_\perp^\top\end{aligned}$$

$$\mathcal{H} = \{X \in \mathbb{R}^{m\times n} : \ h(X) = 0\} \quad \text{and} \quad h(XQ) = h(X) \text{ for } Q \in \mathcal{O}(n)$$

Definition

- *Bouligand tangent cone*:

$$\mathrm{T}_X\mathcal{X} := \left\{ \eta = \lim_{k\to\infty} \frac{(X_k - X)}{t_k} : \ X_k \in \mathcal{X}, t_k > 0 \textit{ for all } k, \lim_{k\to\infty} t_k = 0 \right\}.$$

- *Fréchet normal cone*: $\mathrm{N}_X\mathcal{X} := (\mathrm{T}_X\mathcal{X})^\circ$

Tangent space decoupling at $X = HV^\top \in \mathcal{H}$

$$\begin{aligned} \mathrm{T}_X\mathcal{H} &= \left\{ KV^\top + BV_\perp^\top : \ K \in \mathrm{T}_H\mathcal{H}^s, \ B \in \mathbb{R}^{m\times(n-s)} \right\} \\ &= (\mathrm{T}_H\mathcal{H}^s)\ V^\top \oplus \left(\mathbb{R}^{m\times(n-s)}\right) V_\perp^\top \end{aligned}$$

$$\mathrm{T}_X\left(\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H}\right)?$$

Smooth layers

$$\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} = \bigcup_{s=0}^{s=r} (\mathbb{R}^{m\times n}_{s} \cap \mathcal{H}):\;\; X = HV^\top = U\Sigma V^\top$$

- $\mathbb{R}^{m\times n}_{s} \cap \mathcal{H}$ is a smooth manifold with

$$\begin{aligned} \mathrm{T}_X\left(\mathbb{R}^{m\times n}_{s} \cap \mathcal{H}\right) &= \mathrm{T}_X\mathbb{R}^{m\times n}_{s} \cap \mathrm{T}_X\mathcal{H} \\ &= \left\{KV^\top + UJV_\perp^\top :\; K \in \mathrm{T}_H\mathcal{H}^s,\; J \in \mathbb{R}^{s\times(n-s)}\right\} \end{aligned}$$

Closed-form characterization

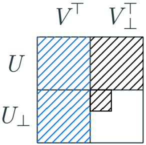
$$\mathrm{T}_X\left(\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}\right) = \left\{ \begin{array}{c} KV^\top + UJV_\perp^\top + U_\perp RV_\perp^\top :\; K \in \mathrm{T}_H\mathcal{H}^s \\ J \in \mathbb{R}^{s\times(n-s)},\; R \in \mathbb{R}^{(m-s)\times(n-s)},\; \mathrm{rank}(R) \leq r-s \end{array} \right\}$$

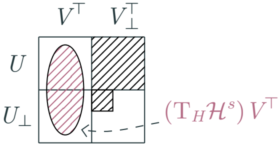
Projection onto the tangent cone

$$\mathcal{P}_{\mathrm{T}_X\left(\mathbb{R}^{m\times n}_{\leq r}\cap\mathcal{H}\right)}(E) = \left(\mathcal{P}_{\mathrm{T}_H\mathcal{H}^s}(EV)\right)V^\top + P_U E P_{V_\perp} + \mathcal{P}_{\mathbb{R}^{m\times n}_{r-s}}\left(P_{U_\perp} E P_{V_\perp}\right)$$

where $P_W := WW^\top$

How orthogonally invariant affects the tangent cone $X = U\Sigma V^\top$





$$\mathrm{T}_X\mathbb{R}^{m\times n}_{\leq r} = \mathbb{R}^{m\times s}\,V^\top + U\mathbb{R}^{s\times(n-s)}\,V_\perp^\top + U_\perp RV_\perp^\top$$

$$\mathrm{T}_X(\mathbb{R}^{m\times n}_{\leq r}\cap\mathcal{H}) = (\mathrm{T}_H\mathcal{H}^s)\,V^\top + U\mathbb{R}^{r\times(n-r)}\,V_\perp^\top + U_\perp RV_\perp^\top$$

Intersection rules for tangent and normal cones

$$\mathrm{T}_X\left(\mathbb{R}^{m\times n}_{\leq r}\cap\mathcal{H}\right) = \mathrm{T}_X\mathbb{R}^{m\times n}_{\leq r}\cap\mathrm{T}_X\mathcal{H}$$

$$\mathrm{N}_X\left(\mathbb{R}^{m\times n}_{\leq r}\cap\mathcal{H}\right) = \mathrm{N}_X\mathbb{R}^{m\times n}_{\leq r}\oplus\mathrm{N}_X\mathcal{H}$$

Definition

A point $X \in \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}$ is *stationary* if $-\nabla f(X) \in \mathrm{N}_X(\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H})$, or equivalently, $\mathcal{P}_{\mathrm{T}_X(\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H})}(-\nabla f(X)) = 0$

First-order optimality conditions

- Decompose $X = U\Sigma V^\top = HV^\top$ with $\mathrm{rank}(X) = s$ and $H = U\Sigma$
- Projected anti-gradient vanishes

$$\left(\mathcal{P}_{\mathrm{T}_H\mathcal{H}^s}\left(\nabla f(X)\, V\right)\right) V^\top + P_U \nabla f(X) P_{V_\perp} + \mathcal{P}_{\mathbb{R}^{m\times n}_{r-s}}\left(P_{U_\perp} \nabla f(X) P_{V_\perp}\right) = 0$$

- Stationarity at rank-deficient point X with $s < r$

$$-\nabla f(X) \in \left(\mathrm{N}_H \mathcal{H}^s\right) V^\top$$

which reduces to $\nabla f(X) = 0$ when $\mathcal{H} = \mathbb{R}^{m\times n}$

A space-decoupling parameterization

$$\begin{tikzcd}[column sep=small] & \mathcal{M}_h \arrow[dl] \arrow[d] \arrow[rr, hook, "\phi"] & & \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \\ \boxed{\text{rank information?}} & \boxed{h(X) = 0?} & & \end{tikzcd}$$

$$\begin{tikzcd} & \mathcal{M}_h \arrow[dl] \arrow[d] \arrow[rr, hook, "\phi"] & & \mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H} \\ \boxed{\text{rank information?}} & \boxed{h(X) = 0?} & & \end{tikzcd}$$

A space-decoupling parameterization $\mathcal{M}_h$

- Parameterize the coupled and non-smooth $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H}$

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

where $\mathrm{Gr}(n, n-r) = \{G \in \mathrm{Sym}(n) : G^2 = G, \operatorname{rank}(G) = n - r\}$

Space decoupling - level 3: a smooth parametrization

$$\begin{tikzcd}[column sep=small]
 & \mathcal{M}_h \arrow[dl] \arrow[dr] \arrow[rr, hook, "\phi(X,G):=X"] & & \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \\
\boxed{G \in \mathrm{Gr}(n, n-r)} & & \boxed{h(X) = 0} &
\end{tikzcd}$$

A space-decoupling parameterization $\mathcal{M}_h$

- Parameterize the coupled and non-smooth $\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}$

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

where $\mathrm{Gr}(n, n-r) = \{G \in \mathrm{Sym}(n) : G^2 = G,\ \mathrm{rank}(G) = n - r\}$

$\mathcal{M}_h$ is a smooth manifold embedded in $\mathbb{R}^{m\times n} \times \mathrm{Sym}(n)$

Represent $(X, G) \in \mathcal{M}_h$ by $X = HV^\top$ and $G = I - VV^\top$ with

$$H \in \mathcal{H}^r \quad \text{and} \quad V \in \mathrm{St}(n, r)$$

Tangent space at $(X, G) \in \mathcal{M}_h$

$$\mathrm{T}_{(X,G)}\mathcal{M}_h = \left\{ \left(KV^\top + HV_p^\top, -V_p V^\top - VV_p^\top \right) : K \in \mathrm{T}_H\mathcal{H}^r,\ V^\top V_p = 0 \right\}$$

Riemannian metric inherited from $\mathcal{E} := \mathbb{R}^{m\times n} \times \mathrm{Sym}(n)$

Represent $(\eta_i, \zeta_i) \in \mathrm{T}_{(X,G)}\mathcal{M}_h$ with $(K_2, V_{p,2})$ $(i = 1, 2)$

$$\langle (\eta_1, \zeta_1), (\eta_2, \zeta_2) \rangle_\omega = \langle K_1, K_2 \rangle + \langle V_{p,1}, V_{p,2} M_{H,\omega} \rangle$$

where $M_{H,\omega} := 2\omega I + H^\top H$

Projection of $(E, Z) \in \mathcal{E}$ onto $\mathrm{T}_{(X,G)}\mathcal{M}_h$

$$\bar{K} = \mathcal{P}_{\mathrm{T}_H\mathcal{H}^r}(EV), \qquad \bar{V}_p = G(E^\top H - 2\omega ZV) M_{H,\omega}^{-1}$$

Represent $(X, G) \in \mathcal{M}_h$ by $X = HV^\top$ and $G = I - VV^\top$ with

$$H \in \mathcal{H}^r \quad \text{and} \quad V \in \mathrm{St}(n, r)$$

Tangent space at $(X, G) \in \mathcal{M}_h$

$$\mathrm{T}_{(X,G)}\mathcal{M}_h = \left\{ \left(KV^\top + HV_p^\top, -V_pV^\top - VV_p^\top \right) : K \in \mathrm{T}_H\mathcal{H}^r,\ V^\top V_p = 0 \right\}$$

Riemannian metric inherited from $\mathcal{E} := \mathbb{R}^{m\times n} \times \mathrm{Sym}(n)$

Represent $(\eta_i, \zeta_i) \in \mathrm{T}_{(X,G)}\mathcal{M}_h$ with $(K_2, V_{p,2})$ $(i = 1, 2)$

$$\langle (\eta_1, \zeta_1), (\eta_2, \zeta_2) \rangle_\omega = \langle K_1, K_2 \rangle + \langle V_{p,1}, V_{p,2} M_{H,\omega} \rangle$$

where $M_{H,\omega} := 2\omega I + H^\top H$

Projection of $(E, Z) \in \mathcal{E}$ onto $\mathrm{T}_{(X,G)}\mathcal{M}_h$

$$\bar{K} = \mathcal{P}_{\mathrm{T}_H\mathcal{H}^r}(EV), \qquad \bar{V}_p = G(E^\top H - 2\omega ZV)M_{H,\omega}^{-1}$$

Complexity: $\mathcal{O}((m+n)r)$

Optimization via the parametrization

$$\begin{tikzcd} \mathcal{M}_h \arrow[d, "\phi"'] \arrow[dr, "\bar{f} := f \circ \phi"] & \\ \mathbb{R}^{m \times n}_{\leq r} \cap \mathcal{H} \arrow[r, "f"'] & \mathbb{R} \end{tikzcd}$$

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

$$\min_{X \in \mathbb{R}^{m \times n}_{\leq r} \cap \mathcal{H}} f(X)$$

Smooth parameterization

- $\phi : \mathcal{M}_h \to \mathbb{R}^{m \times n}_{\leq r} \cap \mathcal{H} : (X, G) \mapsto X$ is a surjection
- Smooth Riemannian optimization problem over $\mathcal{M}_h$
  - Riemannian derivatives
  - Riemannian retractions
  - Vector transports

Riemannian gradient $\nabla_{\mathcal{M}_h}\bar{f}(X, G)$

$$\bar{K} = \mathcal{P}_{\mathrm{T}_H\mathcal{H}^r}\left(\nabla f(X)\,V\right), \qquad \bar{V}_p = G\nabla f(X)^\top H M_{H,\omega}^{-1}.$$

Riemannian Hessian $\nabla^2_{\mathcal{M}_h}\bar{f}(X, G)[\eta, \zeta]$

$$\begin{cases} \bar{K} = \nabla^2_{\mathcal{H}} f(X)[\eta]\,V + \left(I - H M_{H,\omega}^{-1} H^\top\right)\nabla f(X)\,V_p \\ \bar{V}_p = G\left(-V_p[\mathcal{P}_{\mathrm{N}_H\mathcal{H}^r}(\nabla f(X)\,V)]^\top H + (\nabla^2 f(X)[\eta])^\top H + \nabla f(X)^\top W_{H,\omega} K\right) M_{H,\omega}^{-1} \end{cases}$$

First-order retraction

$$\mathrm{R}_{(X,G)}(\eta, \zeta) := (\mathrm{R}_H^{\mathcal{H}^r}(K)(\mathrm{R}_V^{\mathrm{St}}(V_p))^\top,\ I - \mathrm{R}_V^{\mathrm{St}}(V_p)(\mathrm{R}_V^{\mathrm{St}}(V_p))^\top),$$

Second-order retraction

$$\mathrm{R}_{(X,G)}(\eta, \zeta) := ([\mathcal{P}_{\mathcal{H}^r}((X + \eta)\,W)]\,W^\top, I - WW^\top)$$

$$W = L(L^\top L)^{-\frac{1}{2}}, \quad L = V + V_p(I - K^\top H M_{H,\omega}^{-1})$$

Vector transport $\bar{K} = \mathcal{T}_{\tilde{K}}^{\mathcal{H}}(K)$ and $\bar{V}_p = (I - \tilde{V}\tilde{V}^\top)\mathcal{T}_{\tilde{V}_p}^{\mathrm{St}}(V_p)$

Riemannian gradient $\nabla_{\mathcal{M}_h}\bar{f}(X, G)$

$$\bar{K} = \mathcal{P}_{\mathrm{T}_H\mathcal{H}^r}\left(\nabla f(X)\, V\right), \qquad \bar{V}_p = G\nabla f(X)^\top HM_{H,\omega}^{-1}.$$

Riemannian Hessian $\nabla^2_{\mathcal{M}_h}\bar{f}(X, G)[\eta, \zeta]$

$$\begin{cases} \bar{K} = \nabla^2_{\mathcal{H}} f(X)[\eta]\, V + \left(I - HM_{H,\omega}^{-1}H^\top\right)\nabla f(X)\, V_p \\ \bar{V}_p = G\left(-V_p[\mathcal{P}_{\mathrm{N}_H\mathcal{H}^r}(\nabla f(X)\, V)]^\top H + (\nabla^2 f(X)[\eta])^\top H + \nabla f(X)^\top W_{H,\omega}K\right) M_{H,\omega}^{-1} \end{cases}$$

First-order retraction

$$\mathrm{R}_{(X,G)}(\eta, \zeta) := (\mathrm{R}_H^{\mathcal{H}^r}(K)(\mathrm{R}_V^{\mathrm{St}}(V_p))^\top,\ I - \mathrm{R}_V^{\mathrm{St}}(V_p)(\mathrm{R}_V^{\mathrm{St}}(V_p))^\top),$$

Second-order retraction

$$\mathrm{R}_{(X,G)}(\eta, \zeta) := ([\mathcal{P}_{\mathcal{H}^r}((X+\eta)\, W)]\, W^\top, I - WW^\top)$$

$$W = L(L^\top L)^{-\frac{1}{2}}, \quad L = V + V_p(I - K^\top HM_{H,\omega}^{-1})$$

Vector transport $\bar{K} = \mathcal{T}_{\tilde{K}}^{\mathcal{H}}(K)$ and $\bar{V}_p = (I - \tilde{V}\tilde{V}^\top)\mathcal{T}_{\tilde{V}_p}^{\mathrm{St}}(V_p)$

Optimization on $\mathcal{M}_h$ is painless if the geometry of $\mathcal{H}$ is understood

Equivalence of problems

$$\begin{tikzcd}
\mathcal{M}_h \arrow[d, "\phi"'] \arrow[dr, "\bar{f} := f \circ \phi"] & \\
\mathbb{R}^{m \times n}_{\leq r} \cap \mathcal{H} \arrow[r, "f"'] & \mathbb{R}
\end{tikzcd}$$

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

$$\min_{X \in \mathbb{R}^{m \times n}_{\leq r} \cap \mathcal{H}} f(X)$$

$$\begin{tikzcd} \mathcal{M}_h \arrow[d, "\phi"'] \arrow[dr, "\bar{f} := f \circ \phi"] & \\ \mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H} \arrow[r, "f"'] & \mathbb{R} \end{tikzcd}$$

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

$$\min_{X\in\mathbb{R}^{m\times n}_{\leq r}\cap\mathcal{H}} f(X)$$

Proposition

The space-decoupling parametrization $\mathcal{M}_h$ of $\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}$ satisfies

(i) “$1 \Rightarrow 1$” and “$2 \Rightarrow 2$” hold at (X, G) if and only if $\operatorname{rank}(X) = r$.

(ii) “$2 \Rightarrow 1$” holds everywhere on $\mathcal{M}_h$.

Important ingredients

- $\mathcal{M}_h$ is a complete manifold
- Compact sublevel sets of $f \Longrightarrow$ compact sublevel sets of $\bar{f}$

General convergence theorem

- The sequence $\{(X_k, G_k)\}$ generated by monotone algorithms has at least one accumulation point.
- If the accumulation point (X^*, G^*) is second-order stationary on $\mathcal{M}_h$, then X^* is first-order stationary on $\mathbb{R}^{m \times n}_{\leq r}$.

Specific algorithms

- RGD accumulates to first-order critical points
- RTR accumulates to second-order critical points

Numerical experiments

Approximate $A \in \mathcal{H} = \mathrm{Ob}(m, n)$ with low-rank matrix [Chu et al., 2005 SIAMX]

$$\begin{aligned} \min_{X \in \mathbb{R}^{m \times n}} \quad & \frac{1}{2} \left\| \mathcal{P}_\Omega(X - A) \right\|^2 \\ \mathrm{s.\,t.} \quad & \mathrm{rank}(X) \leq r \\ & \mathrm{diag}(XX^\top) - \mathbf{1} = 0 \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & \bar{f}(x) := \frac{1}{2} \left\| \mathcal{P}_\Omega(\phi(x) - A) \right\|_\mathrm{F}^2 \\ \mathrm{s.\,t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

Approximate $A \in \mathcal{H} = \mathrm{Ob}(m, n)$ with low-rank matrix [Chu et al., 2005 SIAMX]

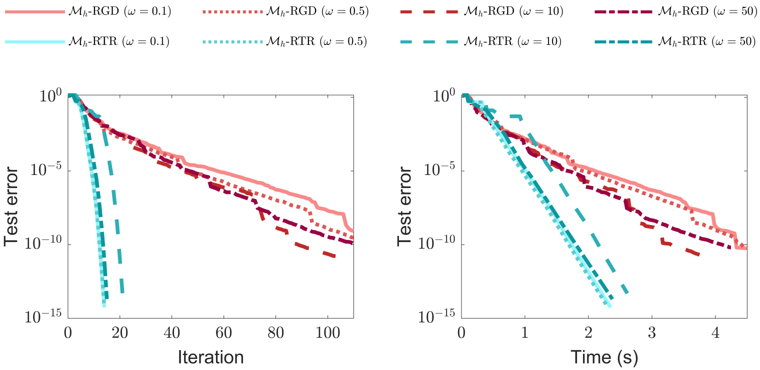
$$\begin{aligned} \min_{X \in \mathbb{R}^{m \times n}} \quad & \frac{1}{2} \|\mathcal{P}_\Omega(X - A)\|^2 \\ \mathrm{s.\,t.} \quad & \mathrm{rank}(X) \leq r \\ & \mathrm{diag}(XX^\top) - \mathbf{1} = 0 \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & \bar{f}(x) := \frac{1}{2} \|\mathcal{P}_\Omega(\phi(x) - A)\|_\mathrm{F}^2 \\ \mathrm{s.\,t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

Methods based on Manopt v8.0

- Riemannian gradient descent method (RGD)
- Riemannian trust region method (RTR)

Test with the unbiased rank parameter

- Synthetic $A \in \mathbb{R}^{m\times n}_{\leq r} \cap \mathrm{Ob}(m,n)$, $(m,n) = (5000, 6000)$
  rank parameter $r = r^* = 6$, oversampling factor $\mathrm{OS} = 5$



Test with over-estimated rank parameters

- Synthetic $A \in \mathbb{R}^{m\times n}_{\leq r} \cap \mathrm{Ob}(m,n)$, $(m,n) = (5000, 6000)$
  true rank $r^* = 6$, oversampling factor $\mathrm{OS} = 5$
  over-estimated rank parameters $r = 7, 8, 9, 10 > r^*$

Algorithm	$r=7$		$r=8$		$r=9$		$r=10$	
	Test err.	Time	Test err.	Time	Test err.	Time	Test err.	Time
$\mathcal{M}_h$-RGD	$1.97\text{e}{-11}$	10.23	$1.69\text{e}{-10}$	11.22	$1.49\text{e}{-11}$	9.56	$1.08\text{e}{-9}$	7.30
$\mathcal{M}_h$-RTR	$2.81\text{e}{-15}$	6.03	$3.97\text{e}{-15}$	15.34	$2.41\text{e}{-14}$	4.51	$6.77\text{e}{-14}$	17.08

Measuring problem [Cason et al., 2013 LAA]

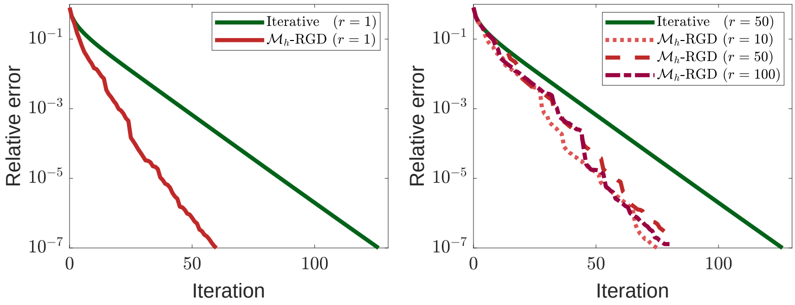
$X \in \mathcal{H} = \mathrm{S_F}(m, n)$ measures the node-to-node similarity between graphs

$$\begin{aligned} \min_{X \in \mathbb{R}^{m\times n}} \quad & -\operatorname{tr}(X^\top \mathcal{L} \circ \mathcal{L}(X)) \\ \text{s.t.} \quad & \operatorname{rank}(X) \le r \\ & \|X\|_\mathrm{F}^2 - 1 = 0 \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & -\operatorname{tr}\left(\phi(x)^\top \mathcal{L} \circ \mathcal{L}(\phi(x))\right) \\ \text{s.t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

Test on low-rank solution

Problem setting

- G_A of $m = 2000$ vertices: a single cycle
- G_B of $n = 3000$ vertices: binomial random graph
- Solution of low rank: $r^* = \mathrm{rank}(X^*) = 1$



Left: test with the unbiased rank parameter $r = r^* = 1$

Right: test with over-estimated rank parameters $r = 10, 50, 100 > r^*$

Problem setting

- $m = n$ to generate G_A and G_B: binomial random graphs
- Solution of full rank: $r^* = \mathrm{rank}(X^*) = m$

Numerical results of "$\mathcal{M}_h$-RTR" for different m and r

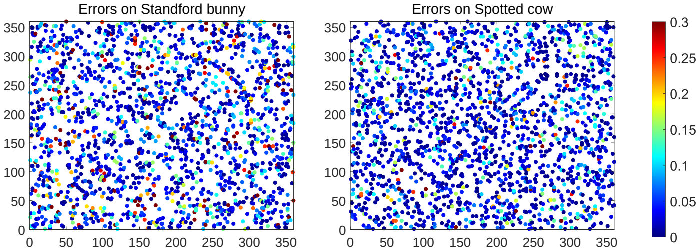
Dimension m	$r = 10\% \times m$		$r = 50\% \times m$		$r = 75\% \times m$		$r = 100\% \times m$	
	Rel. err.	Iter.	Rel. err.	Iter.	Rel. err.	Iter.	Rel. err.	Iter.
200	7.82e−7	38	1.25e−7	31	4.76e−8	26	1.43e−11	4
400	8.23e−7	32	1.85e−7	30	5.92e−8	26	1.07e−11	5
600	7.61e−8	39	5.73e−7	27	3.12e−7	28	8.71e−12	5
800	4.82e−7	37	5.79e−7	28	1.45e−7	21	9.31e−12	5
1000	9.98e−7	37	7.13e−7	31	3.85e−7	28	8.02e−12	5

Synchronization problem

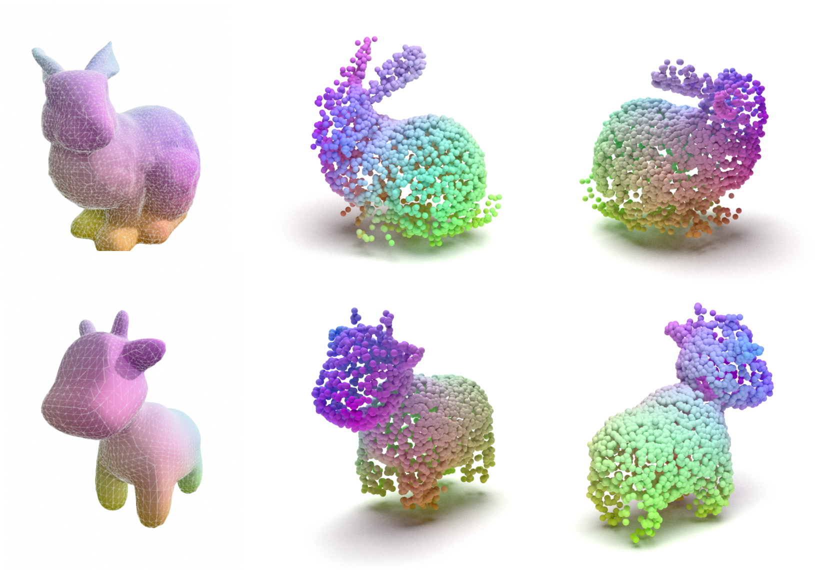
- n cameras with a set of relative rotations $\{\hat{R}_{ij} : (i,j) \in \mathscr{E}\}$
- reconstruct the absolute rotations $\{R_i\}_{i=1}^n$ such that $\hat{R}_{ij} \approx R_j R_i^\top$

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \quad & \langle C, XX^\top \rangle \\ \text{s.t.} \quad & \operatorname{rank}(X) \le r \\ & X \in \mathcal{H} = \mathrm{St}(3n,3)^n \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & \langle C, \phi(x)\phi(x)^\top \rangle \\ \text{s.t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

Reconstructed errors evaluated by $\|R_i R_j^\top - \hat{R}_{ij}\|_{\mathrm{F}}$



Visualization of the reconstruction



Problem formulation

- State space $\mathcal{S} = \{s_1, s_2, \ldots, s_n\}$ and Markov model $\mathrm{P} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$
- Hadamard parameterization, $P = X \odot X,\ X \in \mathrm{Ob}(|\mathcal{S}|, |\mathcal{S}|)$

Problem formulation

- State space $\mathcal{S} = \{s_1, s_2, \ldots, s_n\}$ and Markov model $\mathrm{P} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$
- Hadamard parameterization, $P = X \odot X$, $X \in \mathrm{Ob}(|\mathcal{S}|, |\mathcal{S}|)$
- Proposed approach

$$\begin{aligned} \min_{X \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}} \quad & \frac{1}{2} \|X \odot X - \hat{\mathrm{P}}\|_{\mathrm{F}}^2 \\ \mathrm{s.\,t.} \quad & \mathrm{rank}(X) \leq r \\ & \mathrm{diag}(XX^\top) - \mathbf{1} = 0 \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & \frac{1}{2} \|\phi(x) \odot \phi(x) - \hat{\mathrm{P}}\|_{\mathrm{F}}^2 \\ \mathrm{s.\,t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

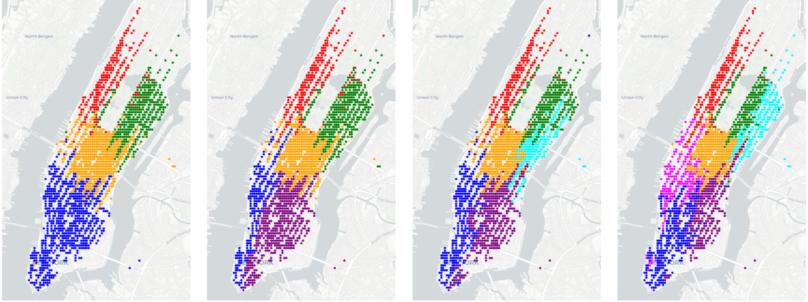
Test on the Manhattan transportation network

Problem setting

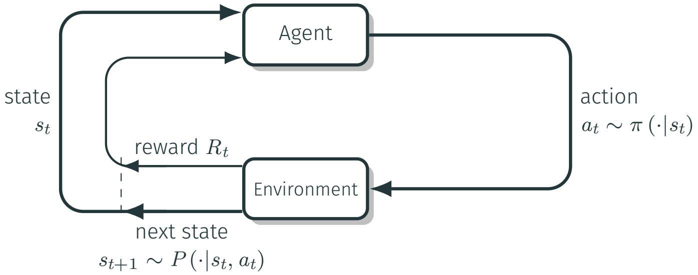
- Dataset of 1.1×10^7 NYC Yellow cab trips
- Model the transportation dynamics as a Markov proces

$$\hat{\mathrm{P}}(i,j) = \frac{\sum_{t=1}^{T} \mathbb{I}\left(s_t^{\text{pickup}} = i,\, s_t^{\text{dropoff}} = j\right)}{\sum_t^T \mathbb{I}\left(s_t^{\text{pickup}} = i\right)}, \quad \text{for} \quad i, j \in \mathcal{S},$$

State compression



$r = 4$ $\qquad$ $r = 5$ $\qquad$ $r = 6$ $\qquad$ $r = 7$



Markov decision process

- State space $\mathcal{S}$, action space $\mathcal{A}$, transition dynamics $\mathrm{P_d} \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{A}|\times|\mathcal{S}|}$
- Reward $R \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{A}|}$, policy $\pi \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{A}|}$
- Maximize $J(\pi) := \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \mid s_0 \sim \mu, \pi]$

Low-rank strategy

- Hadamard parameterization $\pi(X) := X \odot X,\ X \in \mathrm{Ob}(|\mathcal{S}|, |\mathcal{A}|)$
- Impose the rank constraint $\mathrm{rank}(X) \leq r$

$$\begin{aligned} \min_{X \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{A}|}} \quad & -J(\pi(X)) \\ \mathrm{s.\,t.} \quad & \mathrm{rank}(X) \leq r, \\ & \mathrm{diag}(XX^\top) - \mathbf{1} = 0. \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & -J(\pi(\phi(x))) \\ \mathrm{s.\,t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

Low-rank strategy

- Hadamard parameterization $\pi(X) := X \odot X,\ X \in \mathrm{Ob}(|\mathcal{S}|, |\mathcal{A}|)$
- Impose the rank constraint $\mathrm{rank}(X) \leq r$

$$\begin{aligned} \min_{X \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{A}|}} \quad & -J(\pi(X)) \\ \mathrm{s.\,t.} \quad & \mathrm{rank}(X) \leq r, \\ & \mathrm{diag}(XX^\top) - \mathbf{1} = 0. \end{aligned} \longrightarrow \begin{aligned} \min_{x} \quad & -J(\pi(\phi(x))) \\ \mathrm{s.\,t.} \quad & x \in \mathcal{M}_h \end{aligned}$$

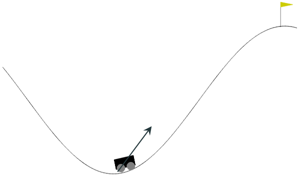
Compared methods

- Low-rank Riemannian policy gradient (LRRPG)
- Q-learning
- REINFORCE

Test in two environments

Pendulum

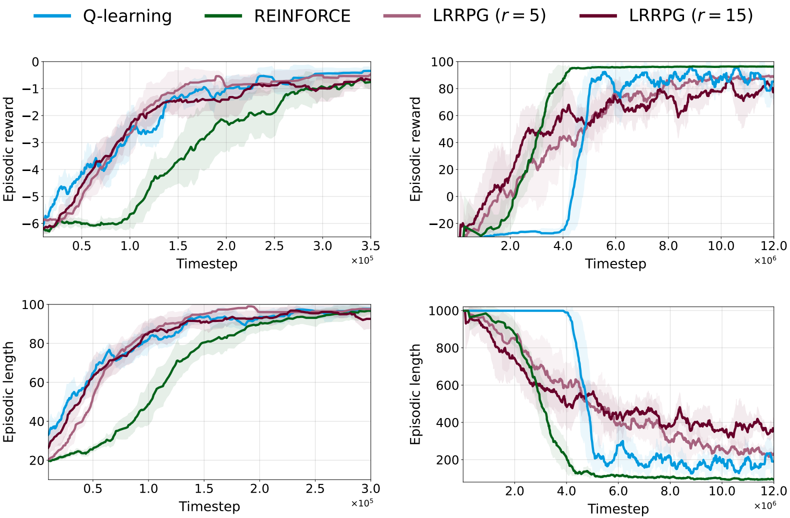
Mountain car



Parameter efficiency

Algorithm	Pendulum			Mountain car		
	Parameters	ρ_{storage}	Reward	Parameters	ρ_{storage}	Reward
REINFORCE	$86,961$	1.00	-0.94	$582,096$	1.00	96.99
Q-learning	$86,961$	1.00	-0.75	$582,096$	1.00	83.95
LRRPG ($r=5$)	$10,810$	0.12	-0.73	$15,485$	0.03	88.31
LRRPG ($r=10$)	$21,620$	0.25	-0.58	$30,970$	0.05	86.44
LRRPG ($r=15$)	$32,430$	0.37	-0.83	$46,455$	0.08	78.23

Convergence curves



Test in “pendulum”

Test in “mountain car”

A space-decoupling framework: four levels

- Relation between rank and orthogonal invariance
- Tangent and normal cones to $\mathbb{R}^{m\times n}_{\leq r} \cap \mathcal{H}$
- Space-decoupling parametrization $\mathcal{M}_h$ and its geometry
- Riemannian optimization on $\mathcal{M}_h$

Future work

- Local convergence results on $\mathcal{M}_h$
- Bounded-rank optimization with additional constraints

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
$\mathcal{S}_{\leq r}(n)$	matrix	[Li-Xiu-Zhou'20 JOTA]	Proposition 3
$\mathcal{S}^{+}_{\leq r}(n)$	matrix	[Levin-Kileel-Boumal'25 MP]	Proposition 4
Intersection of sets	**Structured set**	**First-order**	**Second-order**
$\mathcal{M} \cap \mathcal{K}$	general $(\mathcal{M}, \mathcal{K})$	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.1	Appendix D.1
$\mathcal{S}^{+}_{\leq r}(n) \cap \mathcal{U}$	$\mathcal{U} = \{X \mid \mathcal{A}(X) = b\}$	[Levin-Kileel-Boumal'25 MP]	Appendix D.2

Conclusion and perspectives

A space-decoupling framework: four levels

- Relation between rank and orthogonal invariance
- Tangent and normal cones to $\mathbb{R}^{m\times n}_{\le r} \cap \mathcal{H}$
- Space-decoupling parametrization $\mathcal{M}_h$ and its geometry
- Riemannian optimization on $\mathcal{M}_h$

Future work

- Local convergence results on $\mathcal{M}_h$
- Bounded-rank optimization with additional constraints

References

- **Yan Yang**, Bin Gao, Ya-xiang Yuan. *A space-decoupling framework for optimization on bounded-rank matrices with orthogonally invariant constraints.* Under major revisions at Mathematical Programming
- **Yan Yang**, Bin Gao, Ya-xiang Yuan. *Variational analysis of determinantal varieties.*

- Code is publicly available from https://github.com/UCAS-YanYang

Thanks for your attention!

Email: yangyan@amss.ac.cn