Optimization over Orbit Closures of Real Reductive Lie Group Actions

Zhan Zhiyuan
Associate Prof. Hiroshi Hirai

August 1, 2022

Matrix Scaling [Sinkhorn 1964]

Given $A = (a_{ij}) \in M_n(\mathbb{R})$ with $a_{ij} \geqslant 0$,

- A is *almost doubly stochastic scalable* if $\forall\ \varepsilon > 0$, $\exists$ positive diagonal X, Y s.t. $B = XAY$ satisfies

$$\|B\mathbb{1} - \mathbb{1}\|_2 < \varepsilon, \ \ \left\|B^T\mathbb{1} - \mathbb{1}\right\|_2 < \varepsilon$$

  where $\mathbb{1} = (1, \cdots, 1)^T \in \mathbb{R}^n$.

Remark: In short, we say A is almost scalable.

Theorem (Rothblum and Schneider 1989)

A is almost scalable if and only if for every zero minor $I \times J$ of A,

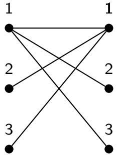
$$|I| + |J| \leqslant n$$

Connection with Graph Theory

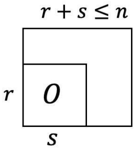
Given $A = (a_{ij}) \in M_n(\mathbb{R})$:

- Bipartite graph $G_A = ([n] \cup [n], E)$: $(i,j) \in E \;\Leftrightarrow\; a_{ij} \neq 0$;

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \Rightarrow$$



- Hall's theorem: G_A has a perfect matching $\Leftrightarrow$



- [RS89]: G_A has a perfect matching $\Leftrightarrow$ A is almost scalable

From Invariant Theory

Let $G = ST(n) \times ST(n)$ act on $V = M(n, \mathbb{R})$, that is

$$\pi: \quad G \quad \longrightarrow \quad GL(V)$$

where $ST(n) = \left\{ \mathrm{diag}\,(t_1, \cdots, t_n) : t_j \in \mathbb{R} \backslash \{0\}, \ \prod_{j=1}^n t_j = 1 \right\}$.

$$\pi(X, Y)A := XAY, \ \forall \ (X, Y) \in G$$

Problem: Null Cone Membership

$$0 \in \overline{\pi(G)A}?$$

By Hilbert-Mumford Criterion [Hil93; MFK94] and [RS89],

Invariant Theoretic View

A is almost scalable $\Leftrightarrow$ $0 \notin \overline{\pi(G)A}$.

Optimization over Orbit Closure

The *Kempf-Ness function* of $A = (a_{ij}) \in M_n(\mathbb{R})$,

$$f_A(x,y) = \log\left(\frac{(x^T Ay)^n}{\prod_{i=1}^n x_i \prod_{j=1}^n y_j}\right) : (\mathbb{R}^+)^n \times (\mathbb{R}^+)^n \to \mathbb{R}$$

- $0 \notin \overline{\pi(G)A} \;\Leftrightarrow\; \inf f_A > -\infty$
- $x_i = e^{s_i},\; y_j = e^{t_j}$,

$$\nabla f_A(s,t) \sim \text{const} \times (\text{Row sums} - \mathbb{1}, \text{Column sums} - \mathbb{1})$$

Observation

A is almost scalable $\;\Leftrightarrow\; \inf f_A > -\infty \;\Leftrightarrow\; 0 \in \overline{\nabla f\left(\mathbb{R}^n \times \mathbb{R}^n\right)}$

Operator Scaling [Gurvits 2004]

Let $G = SL(n, \mathbb{C}) \times SL(n, \mathbb{C})$ act $V = M(n, \mathbb{C})^{\oplus m}$ by

$$\pi(g, h)(A_1, \cdots, A_m) := (gA_1h^\dagger, \cdots, gA_mh^\dagger)$$

$A = (A_1, \cdots, A_m)$ is *almost scalable* for if $\forall\ \varepsilon > 0,\ \exists\ g, h$ s.t.

$$\left\| \sum_{i=1}^m gA_ih^\dagger hA_i^\dagger g^\dagger - I_n \right\|_{\mathrm{F}} < \varepsilon, \quad \left\| \sum_{i=1}^m hA_i^\dagger g^\dagger gX_ih^\dagger - I_n \right\|_{\mathrm{F}} < \varepsilon$$

Also by Hilbert-Mumford Criterion,

Theorem (Garg et al. 2020)

A *is almost scalable* $\Leftrightarrow 0 \notin \overline{\pi(G)A}.$

Optimization: Given $A = (A_1, \cdots, A_m)$,

$$f_A(g,h) = \log \|\pi(g,h)A\|, \ \ g,h \in SL(n,\mathbb{C})$$

where $\|A\|^2 = \sum_i \|A_i\|_F^2$.

Theorem ([GGOW20])

$$A \textit{ is almost scalable} \ \Leftrightarrow\ 0 \in \overline{\nabla f_A(G)} \ \Leftrightarrow\ \inf f_A > -\infty$$

Applications:

- Noncommutative PIT [GGOW20]:

$$p(x_1, \cdots, x_n) : x_i x_j \neq x_j x_i \ \Rightarrow\ p(x_1, \cdots, x_n) = 0?$$

- Brascamp–Lieb inequalities [GGOW18]:

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i\left(B_i(x)\right) dx \leqslant C \prod_{i=1}^m \|f_i\|_{1/p_i}$$

Generalization [Bürgisser, Franks, Garg, Oliveira, Walter and Wigderson 2019]

- $G \subset GL(n, \mathbb{C})$: complex reductive group
- $V \simeq \mathbb{C}^m$ and $\pi : V \to GL(V)$: Given $0 \neq v \in V$

Problems

- **Null cone membership:** determine if $0 \in \overline{\pi(G)v}$
- **Norm minimization:** find g s.t. $\|\pi(g)v\| \to \inf \|\pi(g)v\|$
- **Scaling problem:** Kempf-Ness function $f_v \rightsquigarrow$ gradient ∇f_v

$$\text{find } g \text{ s.t. } \nabla f_v(g) \to 0$$

Applications

- Horn's problem: A, B Hermitians $\Rightarrow$ eigenvalues of $A + B$
- Quantum marginal, Orbit problems, invariant theory, $\cdots\cdots$

Previous Research [Bür+19]

Optimizing the Kempf-Ness function $f_v(g) := \log \|\pi(g)v\|$,

$$\inf_{g\in G} f_v(g)$$

- Smoothness of f_v: first order algorithm $\rightsquigarrow \nabla f_v(g_s) \to 0$
- Convexity of f_v: second order algorithm
- Norm minimization: quantizing the relationship

$$\nabla f_v(g) \to 0 \;\Leftrightarrow\; f_v(g) \to \inf_{g\in G} f_v(g)$$

- Nullcone membership, analyzing parameters, $\cdots\cdots$

This Thesis

Problem Settings

$G \subset GL(n, \mathbb{R}) \curvearrowright V \simeq \mathbb{R}^m$: real reductive Lie group actions

- Without considering the complex structure
- More general geometric structure of $P \simeq G/K$

Contributions: Extending some of results in [Bür+19] to G

- $f_v : P \to \mathbb{R}$ on a Riemannian manifold P.
- Smoothness and convexity of f_v.
- Riemannian optimization on f_v: RGD algorithm.
- Quantizing $\nabla f_v(g) \to 0 \;\Leftrightarrow\; f_v(g) \to \inf f_v(g)$ for real case.
- Analyzing a general scaling problem on P.

Real Reductive Lie Group

Definition (Real Reductive Lie Group)

$G \subset GL(n, \mathbb{R})$ subgroup satisfies:

- $\exists\ f_1, \cdots, f_l$ polynomials on $GL(n, \mathbb{R})$ s.t.

$$G = \{g \in GL(n, \mathbb{R}) \colon f_i(g) = 0,\ \forall\ i = 1, \cdots, l\}$$

- $g \in G \ \Rightarrow\ g^T \in G$

Examples:

1. $GL(n, \mathbb{R})$, $O(n)$, $SL(n, \mathbb{R})$, $SO(n, \mathbb{R})$, $ST(n)$, $\cdots$
2. $GL(n, \mathbb{C})$, $U(n)$, $SL(n, \mathbb{C})$, $SU(n, \mathbb{C})$, $\cdots$

Norm Minimization

Rational representation: (π, V)

- V: finite-dimesional $\mathbb{R}$-vector space with $\langle \cdot, \cdot \rangle$
- $\pi : G \to GL(V)$: group homomorphism

$$\pi(g)v = (\pi(g)_{ij})\, v$$

s.t. $\pi(g)_{ij}$ is a polynomial in g and $\det g^{-1}$.

Norm Minization Problem

Given $\varepsilon > 0, (\pi, V)$ of G, $v \in V$ with $0 \notin \overline{\pi(G)v}$, minimizing

$$\|\pi(g)v\|\,, \;\; g \in G$$

Optimization

For $v \in V \setminus \{0\}$, define $\tilde{f}_v : G \to \mathbb{R}$

$$\tilde{f}_v(g) := \log \|\pi(g)v\|^2 , \ \forall \ g \in G$$

- Norm minimization $\rightsquigarrow$ If $\tilde{f}_{v,\inf} := \inf_{g \in G} \tilde{f}_v(g) > -\infty$,

$$g \in G \ \Rightarrow \ \tilde{f}_v(g) - \tilde{f}_{v,\inf} < \varepsilon$$

- $K = G \cap O(n)$: maximal compact subgroup

  $\exists \ \langle \cdot, \cdot \rangle$ on V s.t. $\langle \pi(k)v, \pi(k)w \rangle = \langle v, w \rangle , \ \forall \ k \in K$

$$\Rightarrow \ \tilde{f}_v : G/K \to \mathbb{R}$$

- Optimizing $\tilde{f}_v$ on G/K: Riemannian optimization

Structure of G/K

G: real reductive Lie group with Lie algebra

$$\mathfrak{g} = \mathrm{Lie}(G) := \left\{ X \in M_n(\mathbb{R}) \colon e^{tX} \in G,\ \forall\, t \in \mathbb{R} \right\}$$

- $\mathfrak{p} := \mathfrak{g} \cap S_n = \left\{ X \in \mathfrak{g} \colon X = X^T \right\}$: Cartan Decomosition

$$G \ \ni \ g = ke^X \ (\exists\, k \in K,\ X \in \mathfrak{p})$$

$$\Rightarrow \ \mathfrak{g} = \mathfrak{k} + \mathfrak{p},\ \mathfrak{k} = \mathrm{Lie}(K)$$

- $P := G \cap P(n) = \{g \in P(n) \colon g \in G\}$:

$$G/K \simeq P \ \left(\left[ke^X \right] \mapsto e^X \right)$$

Settings in [Bür+19]: $G \subset GL(n, \mathbb{C})$, then $K = G \cap U(n)$, $\mathfrak{p} := i\mathfrak{k}$

Kempf-Ness Function

For any $v \in V \backslash \{0\}$, the Kempf-Ness Function f_v

$$\begin{array}{cccc} f_v: & P & \longrightarrow & \mathbb{R} \cup \{\infty\} \\ & x & \longmapsto & \log \langle v, \pi(x) v \rangle \end{array}$$

- [RS90] $\exists\ \langle \cdot, \cdot \rangle$ on V s.t. $\pi(g)^T = \pi\left(g^T\right),\ \forall\ g \in G$

$$f_v(g^T g) = \log \|\pi(g) v\|^2 = \tilde{f}_v(g)$$

- $\forall\ x \in P \subset G,\ x^{\frac{1}{2}} \in G$,

$$f_v(x) = \log \left\| \pi(x^{\frac{1}{2}}) v \right\|^2 = \tilde{f}_v(x^{\frac{1}{2}})$$

Optimization

$$0 \notin \overline{\pi\left(G\right) v} \ \Leftrightarrow\ \inf_{x \in P} f_v(x) > -\infty$$

Riemannian Structure on P

$P = G \cap P(n) \simeq G/K$: Riemannian manifold

- $\langle \cdot, \cdot \rangle_{\mathfrak{p}}$ on $\mathfrak{p}$: $\langle X, Y \rangle_{\mathfrak{p}} = \operatorname{tr}(XY) \rightsquigarrow \|\cdot\|_{\mathfrak{p}}$
- For $x \in P$, the tangent space at x is

$$T_x P := \left\{ x^{\frac{1}{2}} X x^{\frac{1}{2}} : X \in \mathfrak{p} \right\} \Rightarrow T_I P = \mathfrak{p}$$

- $\langle \cdot, \cdot \rangle_x$ on $T_x P$: $\rightsquigarrow \|\cdot\|_x$

$$\langle H_1, H_2 \rangle_x = \operatorname{tr}\left(x^{-1} H_1 x^{-1} H_2 \right), \ \forall\, H_1, H_2 \in T_x P$$

Geodesic: starting at $x \in P$ with the direction $x^{\frac{1}{2}} X x^{\frac{1}{2}} \in T_x P$

$$\gamma_X(t) = x^{\frac{1}{2}} e^{tX} x^{\frac{1}{2}}, \ t \in \mathbb{R}$$

Scaling Problem

For $x \in P$, the gradient $\nabla f_v(x) \in T_x P$ is

$$\left\langle \nabla f_v(x), x^{\frac{1}{2}} X x^{\frac{1}{2}} \right\rangle_x = \left.\frac{d}{dt}\right|_{t=0} f_v(x^{\frac{1}{2}} e^{tX} x^{\frac{1}{2}})$$

Theorem (Kempf-Ness Theorem)

$$f_{v,\inf} := \inf_{x \in P} f_v(x) > -\infty \;\Leftrightarrow\; \inf_{x \in P} \|\nabla f_v(x)\|_x = 0$$

Scaling Problem

Given $\varepsilon > 0, (\pi, V)$ of G, $v \in V$ with $0 \in \overline{\nabla f_v(P)}$,

find $x \in P$ s.t. $\|\nabla f_v(x)\|_x < \varepsilon$

Norm Minimization and Scaling Problem

Given (π, V) of G and $\varepsilon > 0$, $v \in V$ s.t. $f_{v,\inf} > -\infty$

- **Scaling Problem:** find $x \in P$ s.t.

$$\|\nabla f_v(x)\|_x < \varepsilon$$

- **Norm Minimization Problem:** find $x \in P$ s.t.

$$f_v(x) - f_{v,\inf} < \varepsilon$$

RGD Algorithm

Algorithm 1: Riemannian Gradient Descent

Input: Target funtion: $f\colon P \to \mathbb{R} \cup \{\infty\}$;
Step size: η;
Number of iterations: T

Output: $x \in P$

```
x_0 = I
for t = 1 to T do
    x_{t+1} = x_t^{1/2} e^{-η x_t^{-1/2} ∇f(x_t) x_t^{-1/2}} x_t^{1/2}
end
x = argmin { x_t : ‖∇f(x_t)‖_{x_t} }
return x_s
```

$x_0 = I$

for $t = 1$ **to** T **do**

$$x_{t+1} = x_t^{\frac{1}{2}} e^{-\eta x_t^{-\frac{1}{2}} \nabla f(x_t) x_t^{-\frac{1}{2}}} x_t^{\frac{1}{2}}$$

end

$x = \operatorname{argmin} \left\{ x_t \colon \; \|\nabla f(x_t)\|_{x_t} \right\}$

return x_s

c.f. N. Boumal, *An Introduction to Optimization on Smooth Manifolds*, 2022

Validity of RGD

Theorem ([Bou22] Corollary 4.9)

If $f_{\inf} := \inf_x f(x) > -\infty$ *and* f *is* L*-smooth, then* $\forall\, \varepsilon > 0$, *by setting*

$$\eta = \frac{1}{L},\ T > \frac{2L}{\varepsilon^2}\left(f(I) - f_{\inf}\right)$$

in RGD, it can get

$$\|\nabla f(x)\|_x < \varepsilon$$

Remark: $f : P \to \mathbb{R} \cup \{\infty\}$ is called L-smooth if for any $x \in P$,

$$\left| \frac{d^2}{dt^2} f(x^{\frac{1}{2}} e^{tX} x^{\frac{1}{2}}) \right| \leqslant L \|X\|_{\mathfrak{p}}^2,\ \forall\ X \in \mathfrak{p}$$

Smoothness and Convexity

Theorem (This Thesis)

- *f_v is $N(\pi)^2$-smooth;*
- *f_v is geodesically convex, that is $g(t) = f_v(\gamma_X(t))$ is convex for any geodesic $\gamma_X(t)$.*

Remark: The weight norm [Bür+19] $N(\pi)$ is

$$N(\pi) = \max_{X \in \mathfrak{p}} \frac{\|\Pi(X)\|_{\mathrm{op}}}{\|X\|_{\mathfrak{p}}}$$

where $\|\Pi(X)\|_{\mathrm{op}}$ is the operator norm of $\Pi(X) := \frac{d}{dt}\Big|_{t=0} \pi(e^{tX})$.

RGD for Scaling Problem

Theorem (This Thesis)

If $f_{v,\inf} > -\infty$, then for any $\varepsilon > 0$, by setting

$$\eta = \frac{1}{N(\pi)^2}, \quad T > \frac{2N(\pi)^2}{\varepsilon^2}\left(\log \|v\|^2 - f_{v,\inf}\right)$$

in RGD, it returns $x \in P$ s.t. $\|\nabla f_v(x)\|_x < \varepsilon$.

Strengthening Kempf-Ness Theorem

Quantizing the relation: $\nabla f_v(g) \to 0 \;\Leftrightarrow\; f_v(g) \to \inf f_v(g)$

Theorem (This Thesis)

Given a real reductive Lie group G and a rational representation (π, V) of G, for any $v \in V \setminus \{0\}$ then

$$\log\left(1 - \frac{\|\nabla f_v(x)\|_x}{\gamma(\pi)}\right) \leqslant f_{v,\inf} - f_v(x) \leqslant \log\left(1 - \frac{\|\nabla f_v(x)\|_x^2}{4N(\pi)^2}\right)$$

Remark:

- $\gamma(\pi)$ is the weight margin [Bür+19], characterized by π;
- Complex case $\mathfrak{g} = \mathfrak{k} + i\mathfrak{k} \Rightarrow$ Real case $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$.

RGD for Norm Minimization Problem

For any $0 < \varepsilon < \log 2$,

$$\|\nabla f_v(x)\|_x < \frac{1}{2}\gamma(\pi)\varepsilon \;\Rightarrow\; f_v(x) - f_{v,\inf} < \varepsilon$$

Theorem (This Thesis)

For $v \in V$ with $f_{v,\inf} > -\infty$ and $\varepsilon > 0$, by setting

$$\eta = \frac{1}{N(\pi)^2}, \quad T > \frac{8N(\pi)^2}{\gamma(\pi)^2\varepsilon^2}\left(\log\|v\|^2 - f_{v,\inf}\right)$$

in RGD, it returns $x \in P$ s.t. $f_v(x) - f_{v,\inf} < \varepsilon$.

Conclusions

- Optimization over an orbit of real reductive Lie group
  $\rightsquigarrow$ Optimization on $P \subset P(n)$, Riemannian submanifold
- Null cone, scaling and norm minimization problem
  $\rightsquigarrow$ Optimizing the Kempf-Ness function f_v on P
- Smoothness of the Kempf-Ness function f_v
  $\rightsquigarrow$ Applying RGD method to f_v
  $\rightsquigarrow$ Solving the scaling problem
- Strengthening the Kempf-Ness theorem
  $\rightsquigarrow$ RGD method applied for the norm minimization problem

Future work:

- Boundedness of $N(\pi)$ and $\gamma(\pi)$ for some cases;
- Second order algorithm for these problems.

N. Boumal. *An Introduction to Optimization on Smooth Manifolds*. Cambridge: Cambridge University Press, 2022.

P. Bürgisser, C. Franks, A. Garg, R. Oliveira, M. Walter, and A. Wigderson. "Towards a theory of non-commutative optimization: Geodesic 1st and 2nd order methods for moment maps and polytopes". In: *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*. 2019, pp. 845–861.

A. Garg, L. Gurvits, R. Oliveira, and A. Wigderson. "Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via operator scaling". In: *Geometric and Functional Analysis* 28.1 (2018), pp. 100–145.

A. Garg, L. Gurvits, R. Oliveira, and A. Wigderson. "Operator scaling: theory and applications". In: *Foundations of Computational Mathematics* 20.2 (2020), pp. 223–290.

L. Gurvits. "Classical complexity and quantum entanglement". In: *Journal of Computer and System Sciences* 69.3 (2004), pp. 448–484.

D. Hilbert. "Über die vollen Invariantensysteme". In: *Mathematische Annalen* 42.3 (1893), pp. 313–373.

H. Hirai. "Convex analysis on Hadamard spaces and scaling problems". In: *arXiv preprint arXiv:2203.03193* (2022).

D. Mumford, J. Fogarty, and F. Kirwan. *Geometric Invariant Theory*. Berlin: Springer, 1994.

P. Paradan. "Moment polytopes in real symplectic geometry I". In: *arXiv:2012.08837* (2020).

R. Richardson and P. Slodowy. "Minimum vectors for real reductive algebraic groups". In: *Journal of the London Mathematical Society* 2.3 (1990), pp. 409–429.

U. Rothblum and H. Schneider. "Scalings of matrices which have prespecified row sums and column sums via optimization". In: *Linear Algebra and its Applications* 114 (1989), pp. 737–764.

R. Sinkhorn. "A relationship between arbitrary positive matrices and doubly stochastic matrices". In: *The Annals of Mathematical Statistics* 35.2 (1964), pp. 876–879.

T. Uzawa. "Real Moment Maps". In: *RIMS Kôkyûroku* 778 (1992), pp. 80–83.

Moment Polytope

- Given (π, V) of $G \subset GL(n, \mathbb{R})$ and $v \in V \setminus \{0\}$
- $K \curvearrowright \mathfrak{p}$: $k \cdot X := kXk^T \rightsquigarrow \mathfrak{p}/K$
- Moment map $\mu \colon V \setminus \{0\} \to \mathfrak{p}$: $\mu(v) := \nabla f_v(I)$

Definition (Moment Polytope)

The moment polytope is defined as

$$\Delta(v) := \left\{ s\left(\mu(w)\right) : w \in \overline{\pi(G) \cdot v} \right\} \subset \mathfrak{p}/K$$

where $s\left(\mu(w)\right) = \operatorname{diag}\left(\lambda_1, \cdots, \lambda_n\right)$ with $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ and all λ_i are eigenvalues of $\mu(w)$.

Remark

$\Delta(v)$ is a convex polytope in $\mathfrak{p}/K$ by [Par20; Uza92].

p-Scaling Problem

Matrix Scaling: positive vectors $\mathbf{r}, \mathbf{c}$ with $\sum_{i=1}^{n} r_i = \sum_{j=1}^{n} c_j = R$

- A almost $(\mathbf{r}, \mathbf{c})$-scalable: $\forall\, \varepsilon > 0$, $\exists$ positive diagonal X, Y s.t. $B = XAY$ satisfies

$$\|B\mathbb{1} - \mathbf{r}\|_2 < \varepsilon, \quad \left\|B^T\mathbb{1} - \mathbf{c}\right\|_2 < \varepsilon$$

- A almost $(\mathbf{r}, \mathbf{c})$-scalable $\Leftrightarrow$ $(\mathbf{r}, \mathbf{c}) - R/n(\mathbb{1}, \mathbb{1}) \in \Delta(A)$

p-Scaling Problem

Given $\varepsilon > 0$ and (π, V) of G and $v \in V$ such that $p \in \Delta(v)$, find a $g \in G$ s.t.

$$\|s\left(\mu(\pi(g)v)\right) - p\|_F < \varepsilon$$

Busemann Function

For any $p \in \Delta(v)$, the Busemann function b_p corresponding to p

$$\begin{array}{cccc} b_p : & P & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \lim_{t\to\infty} d\left(x, e^{tp}\right) - t \end{array}$$

where $d\left(x, e^{tp}\right)$ is the geodesic distance from x to e^{tp}.

Theorem ([Hir22] Theorem 2.27)

1. *Consider the Kempf-Ness function f_v on P and $p \in \Delta(v)$,*

$$\exists \ \{x_i\} \subset P \text{ s.t. } \lim_{i\to\infty} \left\|\nabla\left(f_v + b_p\right)(x_i)\right\|_{x_i} = 0$$

$$\Leftrightarrow \inf_{x\in P} \left(f_v + b_p\right) > -\infty$$

2. *If $p \in \Delta(v)$, then $\inf_{x\in P}(f_v + b_p)(x) > -\infty$.*

Application to p-scaling Problem

Proposition ([Hir22] Proposition 2.35)

For any $p \in \Delta(v)$*,*

$$\|\nabla(f_v + b_p)(x)\|_x = \left\| \mu(\pi(x^{\frac{1}{2}})v) - ks(X)k^T \right\|_F$$

where $x^{\frac{1}{2}} = bk$ *for upper triangular matrix* b*, orthogonal matrix* k*.*

Problem

Whether is the Busemann function b_p L-smooth? i.e.

$$\left| \frac{d^2}{dt^2} b_p \left(x^{\frac{1}{2}} e^{tY} x^{\frac{1}{2}} \right) \right| \leqslant L \|Y\|_{\mathfrak{p}}^2, \ \forall \ Y \in \mathfrak{p}$$

$\rightsquigarrow$ RGD algorithm for p-scaling