

An Introduction to Transfer and Commutation Principles in Euclidean Jordan Algebras

Juyoung Jeong

Department of Mathematics
Soongsil University

NUU – SSU G-LAMP International Workshop
October 02 - 04, 2026

1. Introduction

Why Spectral Optimization in Machine Learning?

- Learning distances and kernels leads to optimization over positive semidefinite matrices.
- Spectral regularizers control eigenvalues without selecting a preferred eigenbasis.

A typical model is

$$\min_{X \succeq O} \underbrace{\mathcal{L}(X)}_{\text{data-fitting loss}} + \mu \underbrace{f(\lambda(X))}_{\text{spectral regularizer}}, \quad \mu > 0.$$

- **Transfer principles** ask which analytic/convexity properties of f can be transferred to the matrix function $f \circ \lambda$.
- **Commutation principles** seek spectral alignment conditions an optimizer must satisfy.

Why Spectral Optimization in Machine Learning?

- A data-fitting loss function aggregates discrepancies between squared distances and target values r_{ij} . For example, with data points $z_i \in \mathcal{R}^n$ and a matrix $X \succ 0$

$$\mathcal{L}(X) = \sum_{(i,j) \in \mathcal{P}} (d_X(z_i, z_j)^2 - r_{ij})^2, \quad d_X(z_i, z_j)^2 = (z_i - z_j)^\top X (z_i - z_j).$$

where $\mathcal{P}$ is a set of observed pairs and $r_{ij} \geq 0$ are target squared distances.

- For a spectral regularizer, one can choose the LogDet divergence from I :

$$G(X) = D_{\text{ld}}(X, I) = \text{tr}(X) - \log \det X - n.$$

It penalizes eigenvalues near zero or very large, with its minimum at $X = I$.

- The resulting learning model becomes

$$\min_{X \in \mathcal{S}_{++}^n} \sum_{(i,j) \in \mathcal{P}} (d_X(z_i, z_j)^2 - r_{ij})^2 + \mu (\text{tr}(X) - \log \det X - n).$$

Why Spectral Optimization in Machine Learning?

- Note that the regularizer above depends only on the eigenvalues, that is,

$$G(X) = \sum_{i=1}^n (\lambda_i(X) - \log \lambda_i(X) - 1)$$

- Let $\lambda(X) \in \mathcal{R}^n$ list the eigenvalues of $X \in \mathcal{S}^n$ in decreasing order. Then G can be written as $G = f \circ \lambda$, where

$$f(u) = \sum_{i=1}^n (u_i - \log u_i - 1).$$

- To analyze G , focus on its nonlinear term

$$-\log \det X = -\sum_{i=1}^n \log \lambda_i(X).$$

Strict convexity transfers from the eigenvalue function to $-\log \det X$; adding the affine term $\text{tr}(X) - n$ preserves it.

Motivating Example

Question. Show that $G(X) = -\log \det(X)$ is a strictly convex function on $\mathcal{S}_{++}^n$.

Answer 1.

- For $A \in \mathcal{S}_{++}^n$ and $H \in \mathcal{S}^n \setminus \{O\}$, define

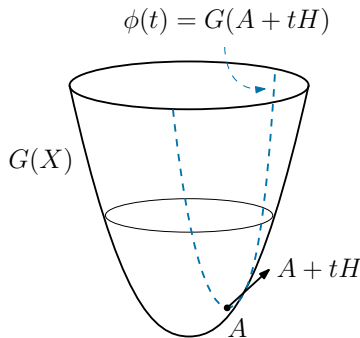
$$\phi(t) = G(A + tH), \quad t \in I_{A,H},$$

where $I_{A,H} = \{t \in \mathcal{R} : A + tH \in \mathcal{S}_{++}^n\}$.

- Using a line characterization of strict convexity, it is enough to show that

$$\phi''(0) = \frac{d^2}{dt^2} G(A + tH) \Big|_{t=0} > 0$$

for all $A \in \mathcal{S}_{++}^n$ and $H \in \mathcal{S}^n \setminus \{O\}$.



Motivating Example

Answer 1. (continued...)

- By Jacobi's formula and the derivative of the inverse,

$$\begin{aligned}\frac{d}{dt}G(A + tH) &= -\operatorname{tr}((A + tH)^{-1}H), \\ \frac{d}{dt}(A + tH)^{-1} &= -(A + tH)^{-1}H(A + tH)^{-1}.\end{aligned}$$

- Hence, we see that

$$\left.\frac{d^2}{dt^2}G(A + tH)\right|_{t=0} = \operatorname{tr}(A^{-1}HA^{-1}H).$$

- Since A is positive definite and $H \neq O$, the matrix $A^{-1/2}HA^{-1/2}$ is symmetric and nonzero. Hence,

$$\operatorname{tr}(A^{-1}HA^{-1}H) = \operatorname{tr}((A^{-1/2}HA^{-1/2})^2) = \|A^{-1/2}HA^{-1/2}\|_F^2 > 0.$$

- This verifies that G is a strictly convex function.

Motivating Example

Answer 2.

- We realize that $G(A) = -\log \det(A)$ can be written as $G = f \circ \lambda$, where the symmetric function $f : \mathcal{R}_{++}^n \rightarrow \mathcal{R}$ is defined by

$$f(u_1, \dots, u_n) = -\log \left(\prod_{i=1}^n u_i \right) = -\sum_{i=1}^n \log u_i.$$

- The Hessian $\nabla^2 f(u_1, \dots, u_n)$ is simply

$$\nabla^2 f(u_1, \dots, u_n) = \text{Diag} \left(\frac{1}{u_1^2}, \dots, \frac{1}{u_n^2} \right).$$

Thus, $\nabla^2 f(u_1, \dots, u_n)$ is positive definite on $\mathcal{R}_{++}^n$, implying strict convexity of f .

- Thanks to the **transfer principle for strict convexity**, we conclude that $G = f \circ \lambda$ is strictly convex as well.

2. Euclidean Jordan algebras

A **Euclidean Jordan algebra (EJA)** is a triple $(\mathcal{V}, \circ, \langle \cdot, \cdot \rangle)$, where $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ is a finite-dimensional real inner product space with the bilinear Jordan product

$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \circ \mathbf{y} : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$$

satisfying

- $\mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.
- $\mathbf{x}^2 \circ (\mathbf{x} \circ \mathbf{y}) = \mathbf{x} \circ (\mathbf{x}^2 \circ \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ where $\mathbf{x}^2 = \mathbf{x} \circ \mathbf{x}$.
- $\langle \mathbf{x} \circ \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \circ \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$.
- There is a unit $\mathbf{e} \in \mathcal{V}$ with $\mathbf{e} \circ \mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathcal{V}$.

Example. For $A, B \in \mathcal{S}^n$, define

$$A \circ B = \frac{1}{2}(AB + BA) \quad \text{and} \quad \langle A, B \rangle = \text{tr}(AB).$$

Then $(\mathcal{S}^n, \circ, \langle \cdot, \cdot \rangle)$ becomes a Euclidean Jordan algebra.

■ Spectral decomposition theorem:

Let n be the rank of $\mathcal{V}$. Any $\mathbf{x} \in \mathcal{V}$ can be written as

$$\mathbf{x} = \lambda_1(\mathbf{x})\mathbf{e}_1 + \lambda_2(\mathbf{x})\mathbf{e}_2 + \cdots + \lambda_n(\mathbf{x})\mathbf{e}_n,$$

where $\lambda_1(\mathbf{x}) \geq \cdots \geq \lambda_n(\mathbf{x})$ and $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a Jordan frame. Its elements are primitive idempotents with $\mathbf{e}_i \circ \mathbf{e}_j = \mathbf{0}$ for $i \neq j$ and $\sum_i \mathbf{e}_i = \mathbf{e}$.

- Using the spectral decomposition theorem, define the **eigenvalue map** $\lambda : \mathcal{V} \rightarrow \mathcal{R}^n$, which takes $\mathbf{x}$ to $\lambda(\mathbf{x})$ defined by

$$\lambda(\mathbf{x}) = (\lambda_1(\mathbf{x}), \lambda_2(\mathbf{x}), \dots, \lambda_n(\mathbf{x}))^\top.$$

- Using the eigenvalue map $\lambda : \mathcal{V} \rightarrow \mathcal{R}^n$, we can now define a spectral function $G : \mathcal{V} \rightarrow \mathcal{R}$ in the same way.

- Throughout, $\mathcal{V}$ carries the trace inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \text{tr}(\mathbf{x} \circ \mathbf{y}), \quad \|\mathbf{x}\|^2 = \text{tr}(\mathbf{x}^2) = \sum_{i=1}^n \lambda_i(\mathbf{x})^2.$$

- A set $Q \subseteq \mathcal{R}^n$ is **permutation invariant** if $PQ = Q$ for every permutation matrix P .
A function $f : Q \rightarrow \mathcal{R}$ is **symmetric** (or **permutation invariant**) if $f(Pu) = f(u)$.
- We use the representations

$$C = \lambda^{-1}(Q), \quad G = f \circ \lambda,$$

with permutation invariant Q and symmetric f . These are called spectral sets and spectral functions, respectively.

- For $\mathbf{a} \in \mathcal{V}$, let $L_{\mathbf{a}}(\mathbf{x}) = \mathbf{a} \circ \mathbf{x}$. Elements $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ **operator commute** if $L_{\mathbf{a}}L_{\mathbf{b}} = L_{\mathbf{b}}L_{\mathbf{a}}$, equivalently, they share a Jordan frame:

$$\mathbf{a} = \sum_{i=1}^n \alpha_i \mathbf{e}_i, \quad \mathbf{b} = \sum_{i=1}^n \beta_i \mathbf{e}_i,$$

where $\{\alpha_1, \dots, \alpha_n\} = \{\lambda_1(\mathbf{a}), \dots, \lambda_n(\mathbf{a})\}$ and $\{\beta_1, \dots, \beta_n\} = \{\lambda_1(\mathbf{b}), \dots, \lambda_n(\mathbf{b})\}$.

- They **strongly operator commute** if a common Jordan frame can be chosen with both eigenvalue lists in decreasing order:

$$\mathbf{a} = \sum_{i=1}^n \lambda_i(\mathbf{a}) \mathbf{e}_i, \quad \mathbf{b} = \sum_{i=1}^n \lambda_i(\mathbf{b}) \mathbf{e}_i.$$

- In $\mathcal{S}^n$, operator commutativity is reduced to $AB = BA$. Strong operator commutativity also requires the two eigenvalue lists to have the same order in a common eigenbasis.

- It is well-known for $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ that

$$\langle \mathbf{x}, \mathbf{y} \rangle \leq \langle \lambda(\mathbf{x}), \lambda(\mathbf{y}) \rangle.$$

- Equivalently, the eigenvalue map λ is Lipschitz continuous, i.e.

$$\|\lambda(\mathbf{x}) - \lambda(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$$

Hence, we see that a spectral function $G = f \circ \lambda$ is continuous if f is continuous.

- What is more interesting is that the converse holds as well for symmetric f !
- Moreover, we have a transfer principle for Fréchet differentiability.

$G = f \circ \lambda$ is differentiable at $\mathbf{x}$ if (and only if) f is differentiable at $\lambda(\mathbf{x})$.
In this case, $\mathbf{x}$ and $\nabla G(\mathbf{x})$ operator commute, and

$$\lambda(\nabla G(\mathbf{x})) = (\nabla f(\lambda(\mathbf{x})))^\downarrow.$$

Here $u^\downarrow$ denotes the decreasing rearrangement of u .

3. Transfer Principles

Definition

Let C be a convex set in $\mathcal{V}$ and consider a function $G : C \rightarrow \mathcal{R}$.

(a) G is said to be **convex** on C if

$$G((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)G(\mathbf{x}) + tG(\mathbf{y})$$

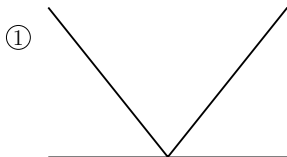
for all $\mathbf{x}, \mathbf{y} \in C$ and $t \in [0, 1]$.

(b) G is said to be **strictly convex** on C if

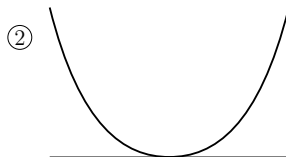
$$G((1-t)\mathbf{x} + t\mathbf{y}) < (1-t)G(\mathbf{x}) + tG(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in C$ with $\mathbf{x} \neq \mathbf{y}$ and $t \in (0, 1)$.

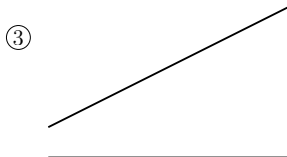
Convex functions – Definitions



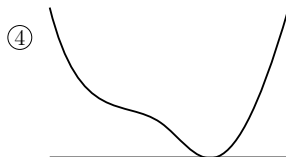
convex
not strictly convex



strictly convex



convex
not strictly convex

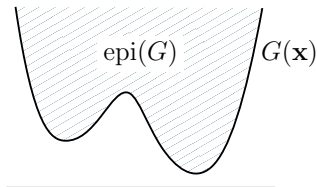


not convex

- G is convex if and only if the epigraph

$$\text{epi}(G) = \{(\mathbf{x}, t) \in C \times \mathcal{R} : G(\mathbf{x}) \leq t\}$$

is a convex set in $\mathcal{V} \times \mathcal{R}$.



- Assume that C is open and G is Fréchet differentiable on C . Then G is convex on C if and only if

$$\langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq G(\mathbf{y}) - G(\mathbf{x})$$

for all $\mathbf{x}, \mathbf{y} \in C$.

- If G is convex, then every local minimum of G is also global. When G is also Fréchet differentiable, then every stationary point is a global minimum.

Definition

Let C be a convex set in $\mathcal{V}$ and consider a function $G : C \rightarrow \mathcal{R}$.

(a) G is said to be **quasiconvex** on C if the implication

$$G(\mathbf{y}) \leq G(\mathbf{x}) \implies G((1-t)\mathbf{x} + t\mathbf{y}) \leq G(\mathbf{x})$$

holds for all $\mathbf{x}, \mathbf{y} \in C$ and $t \in [0, 1]$.

(b) G is said to be **strictly quasiconvex** on C if the implication

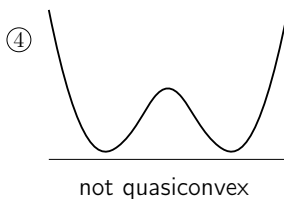
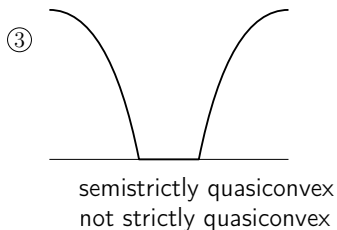
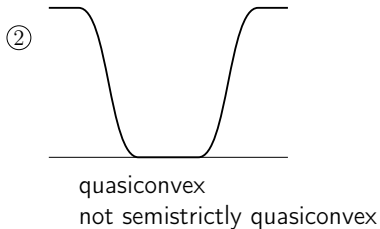
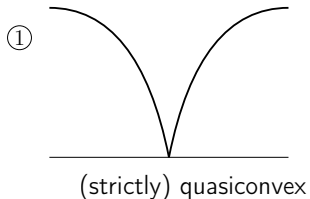
$$G(\mathbf{y}) \leq G(\mathbf{x}) \implies G((1-t)\mathbf{x} + t\mathbf{y}) < G(\mathbf{x})$$

holds for all $\mathbf{x}, \mathbf{y} \in C$ with $\mathbf{x} \neq \mathbf{y}$ and $t \in (0, 1)$.

(c) G is said to be **semistrictly quasiconvex** on C if the implication

$$G(\mathbf{y}) < G(\mathbf{x}) \implies G((1-t)\mathbf{x} + t\mathbf{y}) < G(\mathbf{x})$$

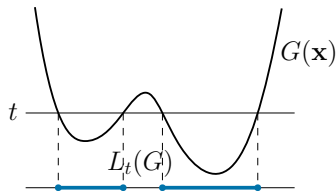
holds for all $\mathbf{x}, \mathbf{y} \in C$ and $t \in (0, 1)$.



- G is quasiconvex if and only if the t -sublevel set

$$L_t(G) = \{\mathbf{x} \in C : G(\mathbf{x}) \leq t\}$$

is a convex set in $\mathcal{V}$ for every $t \in \mathcal{R}$.



- Assuming C is open and G is Fréchet differentiable on C , G is quasiconvex on C if and only if the implication

$$G(\mathbf{y}) \leq G(\mathbf{x}) \implies \langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq 0$$

holds for all $\mathbf{x}, \mathbf{y} \in C$.

- Let G be a continuous quasiconvex function on C . Then, G is semistrictly quasiconvex if and only if every local minimum is also global.

Definition

Let C be an open convex set in $\mathcal{V}$ and consider Fréchet differentiable $G : C \rightarrow \mathcal{R}$.

(a) G is said to be **pseudoconvex** on C if the implication

$$G(\mathbf{y}) < G(\mathbf{x}) \implies \langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle < 0$$

$$[\text{equivalently, } \langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \implies G(\mathbf{y}) \geq G(\mathbf{x})]$$

holds for all $\mathbf{x}, \mathbf{y} \in C$.

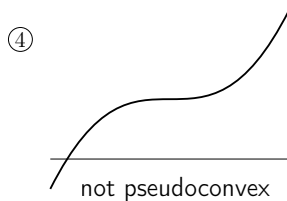
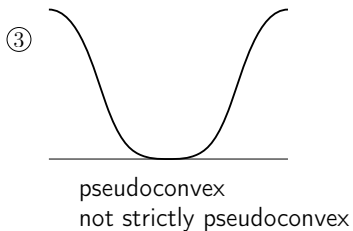
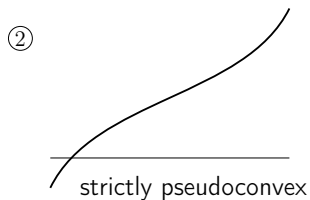
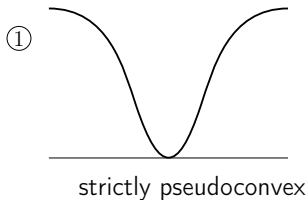
(b) G is said to be **strictly pseudoconvex** on C if the implication

$$G(\mathbf{y}) \leq G(\mathbf{x}) \implies \langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle < 0$$

$$[\text{equivalently, } \langle \nabla G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0 \implies G(\mathbf{y}) > G(\mathbf{x})]$$

holds for all $\mathbf{x}, \mathbf{y} \in C$ with $\mathbf{x} \neq \mathbf{y}$.

Pseudoconvex functions – Definitions



- A derivative-free characterization: G is pseudoconvex on C if, for $\mathbf{x}, \mathbf{y} \in C$ satisfying $G(\mathbf{y}) < G(\mathbf{x})$, there exists $\gamma > 0$ and $\delta \in (0, 1)$ such that

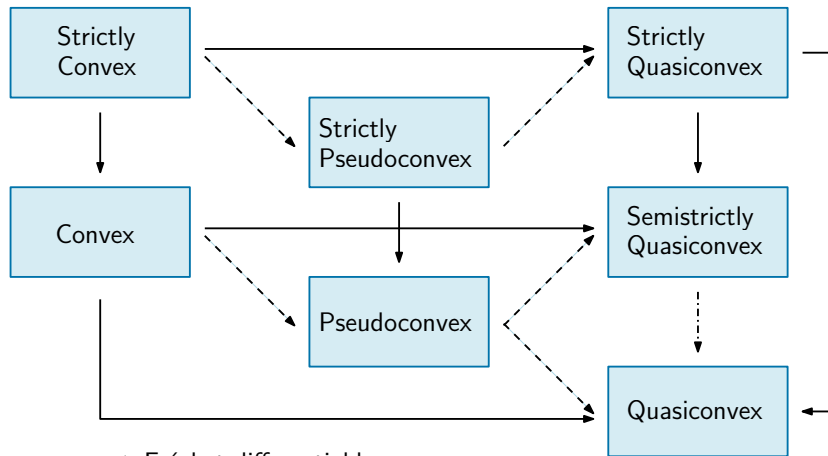
$$G((1 - t)\mathbf{x} + t\mathbf{y}) < G(\mathbf{x}) - \gamma t(1 - t)$$

whenever $t \in (0, \delta)$.

Two definitions of pseudoconvexity coincide when G is Fréchet differentiable on an open convex set C .

- If G is (Fréchet differentiable and) pseudoconvex, then every stationary point is a global minimum. Furthermore, if G is strictly pseudoconvex, then there exists at most one stationary point.
- Indeed, when G is Fréchet differentiable and quasiconvex, G is pseudoconvex if and only if every stationary point is a global minimum.

Hierarchy between Generalized Convexity Concepts



Theorem (Baes, 2007; Jeong-Gowda, 2017)

Let $C = \lambda^{-1}(Q)$ be a spectral convex set in $\mathcal{V}$ and $G : C \rightarrow \mathcal{R}$ be a spectral function with the corresponding symmetric function $f : Q \rightarrow \mathcal{R}$. Then G is 'convex' on C if and only if f is 'convex' on Q .

- Baes proved that convexity of f implies that of G .
- Later, Jeong and Gowda proved that convexity can be transferred in both directions with a (possibly) simpler proof, which relies on the following fact.

Let $f : Q \subseteq \mathcal{R}^n \rightarrow \mathcal{R}$ be symmetric and convex. If $u, v \in Q$ satisfy $u \prec v$, then $f(u) \leq f(v)$.

M. Baes
Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras
Linear Algebra and its Applications, 422.2 (2007): 664–700

J. Jeong and M. S. Gowda
Spectral sets and functions on Euclidean Jordan algebras
Linear Algebra and its Applications, 518 (2017): 31–56

Theorem (Jeong-Sossa, 2025)

Let $C = \lambda^{-1}(Q)$ be a spectral convex set in $\mathcal{V}$ and $G : C \rightarrow \mathcal{R}$ be a spectral function with the corresponding symmetric function $f : Q \rightarrow \mathcal{R}$. Then G is 'strictly convex' on C if and only if f is 'strictly convex' on Q .

- Starting with the assumption $\mathbf{x} \neq \mathbf{y}$, one needs to consider two possible cases: When $\lambda(\mathbf{x}) \neq \lambda(\mathbf{y})$, one can apply Schur-convexity of and strict convexity of f .
- The case $\lambda(\mathbf{x}) = \lambda(\mathbf{y})$ involves the following lemma

Let $f : Q \subseteq \mathcal{R}^n \rightarrow \mathcal{R}$ be symmetric and strictly convex. If $u, v \in Q$ satisfy $u \prec v$ and $u^\downarrow \neq v^\downarrow$, then $f(u) < f(v)$.

Theorem (Jeong, submitted)

Let $C = \lambda^{-1}(Q)$ be a spectral convex set in $\mathcal{V}$ and $G : C \rightarrow \mathcal{R}$ be a spectral function with the corresponding symmetric function $f : Q \rightarrow \mathcal{R}$. Then G is '(strictly/semistrictly) quasiconvex' on C if and only if f is '(strictly/semistrictly) quasiconvex' on Q .

- Luckily enough, symmetric quasiconvex functions are Schur-convex!

Let $f : Q \subseteq \mathcal{R}^n \rightarrow \mathcal{R}$ be symmetric and quasiconvex. If $u, v \in Q$ satisfy $u \prec v$, then $f(u) \leq f(v)$. Moreover, when f is strictly quasiconvex and $u^\downarrow \neq v^\downarrow$, then $f(u) < f(v)$.

Theorem (Jeong, submitted)

Let $C = \lambda^{-1}(Q)$ be a spectral open convex set in $\mathcal{V}$ and $G : C \rightarrow \mathcal{R}$ be a Fréchet differentiable spectral function with the corresponding symmetric function $f : Q \rightarrow \mathcal{R}$. Then G is ‘(strictly) pseudoconvex’ on C if and only if f is ‘(strictly) pseudoconvex’ on Q .

- Here, we may quote a transfer principle for Fréchet differentiability.

$G = f \circ \lambda$ is differentiable at $\mathbf{x}$ if (and only if) f is differentiable at $\lambda(\mathbf{x})$.
In this case, $\mathbf{x}$ and $\nabla G(\mathbf{x})$ operator commute, and

$$\lambda(\nabla G(\mathbf{x})) = (\nabla f(\lambda(\mathbf{x})))^\downarrow.$$

Here $u^\downarrow$ denotes the decreasing rearrangement of u .

- Transfer principle for strict pseudoconvexity is more involved, though.

4. Commutation Principles

- Consider the following semidefinite optimization problem over $\mathcal{S}^n$.

$$\min \left\{ \|B - X\|^2 : X \in \mathcal{S}_+^n, 1 \leq \lambda_{\max}(X) \leq 2 \right\}$$

- It turns out the problem is equivalent to

$$\min \left\{ \|\lambda(B) - u\|^2 : u \in \mathcal{R}_+^n, u_1 \geq \cdots \geq u_n, 1 \leq u_1 \leq 2 \right\},$$

which is a linearly constrained least-squares problem over $\mathcal{R}^n$.

- Moreover, if A solves the original problem, then A and B commute, i.e., $AB = BA$. Equivalently, A and B are simultaneously diagonalizable.
- Hence, if v is a solution to the second problem, then

$$A = U \operatorname{Diag}(v) U^T, \quad \text{where } B = U \operatorname{Diag}(\lambda(B)) U^T.$$

Theorem (Iusem-Seeger, 2007)

Suppose that E in $\mathcal{S}^n$ and $F : E \rightarrow \mathcal{R}$ are spectral. Let $B \in \mathcal{S}^n$ be given. If $A \in E$ is a local minimizer/maximizer of the map

$$X \in E \mapsto \langle B, X \rangle + F(X),$$

then A and B commute, i.e., $AB = BA$.

Remark.

- Note that A and B commute if and only if they are simultaneously diagonalizable, i.e., there exists $U \in O^n$ such that

$$U^T A U = D_A \quad \text{and} \quad U^T B U = D_B,$$

where D_A, D_B are diagonal matrices.

Remark.

- Moreover, the global optimal value (minimum or maximum, when attained) equals that of the following relatively simpler problem defined on $\mathcal{R}^n$

$$u \in Q \mapsto \langle \lambda(B), u \rangle + f(u),$$

where $Q \subseteq \mathcal{R}^n$ and $f : \mathcal{R}^n \rightarrow \mathcal{R}$ so that $E = \lambda^{-1}(Q)$ and $F = f \circ \lambda$.

- Hence, if $v \in Q$ is a global solution of this problem, then

$$A = U \operatorname{Diag}(v) U^T,$$

where $U \in O^n$ is chosen so that $U^T B U = \operatorname{Diag}(\lambda(B))$, solves the original problem.

Commutation Principle in EJAs

Ramirez-Seeger-Sossa generalize the commutation principle in $\mathcal{S}^n$ to EJAs.

Theorem (Ramírez-Seeger-Sossa, 2013)

Suppose that E in $\mathcal{V}$ and $F : E \rightarrow \mathcal{R}$ are 'spectral'. Let $\Theta : \mathcal{V} \rightarrow \mathcal{R}$ be 'Fréchet differentiable'. If $\mathbf{a} \in E$ is a 'local minimizer/maximizer' of the map

$$\mathbf{x} \in E \mapsto \Theta(\mathbf{x}) + F(\mathbf{x}),$$

then $\mathbf{a}$ and $\nabla\Theta(\mathbf{a})$ 'operator commute' in $\mathcal{V}$.

Suppose additionally that E is convex. Then Gowda-Jeong observed that

- If F is convex and $\mathbf{a} \in E$ is a local minimizer, then $\mathbf{a}$ and $-\nabla\Theta(\mathbf{a})$ strongly operator commute.
- If F is concave and $\mathbf{a} \in E$ is a local maximizer, then $\mathbf{a}$ and $\nabla\Theta(\mathbf{a})$ strongly operator commute.

H. Ramírez, A. Seeger, and D. Sossa
Commutation principle for variational problems on Euclidean Jordan algebras
SIAM Journal on Optimization, 23 (2013): 687–694

Commutation Principle with Weak Spectrality in EJAs

- An **algebra automorphism** is an invertible linear map $T : \mathcal{V} \rightarrow \mathcal{V}$ satisfying $T(\mathbf{x} \circ \mathbf{y}) = T(\mathbf{x}) \circ T(\mathbf{y})$. The group of these maps is denoted by $\text{Aut}(\mathcal{V})$.
- A set $E \subseteq \mathcal{V}$ is **weakly spectral** if $T(E) = E$ for all $T \in \text{Aut}(\mathcal{V})$. A function $F : E \rightarrow \mathcal{R}$ is **weakly spectral** if $F(T(\mathbf{x})) = F(\mathbf{x})$ for all $\mathbf{x} \in E$ and $T \in \text{Aut}(\mathcal{V})$.
- Every spectral set/function is weakly spectral, but the converse need not hold in general.

Theorem (Gowda-Jeong, 2017)

Suppose that E in $\mathcal{V}$ and $F : E \rightarrow \mathcal{R}$ are ‘weakly spectral’. Let $\Theta : \mathcal{V} \rightarrow \mathcal{R}$ be ‘Fréchet differentiable’. If $\mathbf{a} \in E$ is a ‘local minimizer/maximizer’ of the map

$$\mathbf{x} \in E \mapsto \Theta(\mathbf{x}) + F(\mathbf{x}),$$

then $\mathbf{a}$ and $\nabla\Theta(\mathbf{a})$ ‘operator commute’ in $\mathcal{V}$.

Nonsmooth Commutation Principle with Spectrality in EJAs

Theorem (Gowda, 2022)

Suppose that E in $\mathcal{V}$ and $F : \mathcal{V} \rightarrow \mathcal{R}$ are ‘spectral’. Let $\Theta : \mathcal{V} \rightarrow \mathcal{R}$ be ‘any’ function. Consider the map $\mathbf{x} \in E \mapsto \Theta(\mathbf{x}) + F(\mathbf{x})$.

- If $\mathbf{a}$ is a ‘global maximizer’ of the map and $\partial_E \Theta(\mathbf{a}) \neq \emptyset$, then $\mathbf{a}$ ‘strongly operator commutes’ with every element in $\partial_E \Theta(\mathbf{a})$.
- Assume E and Θ are convex. If $\mathbf{a}$ is a global minimizer of Θ over E , then $\mathbf{a}$ ‘strongly operator commutes’ with some $\mathbf{b} \in -\partial \Theta(\mathbf{a})$.

- Here, the **subdifferential** of Θ at $\mathbf{a}$ relative to E is defined by

$$\partial_E \Theta(\mathbf{a}) = \{\mathbf{d} \in \mathcal{V} : \Theta(\mathbf{x}) - \Theta(\mathbf{a}) \geq \langle \mathbf{d}, \mathbf{x} - \mathbf{a} \rangle \text{ for all } \mathbf{x} \in E\}.$$

Nonsmooth Commutation Principle with Weak Spectrality in EJAs

Theorem (Jeong-Sossa, 2025; Jeong, 2026)

Suppose that E in $\mathcal{V}$ and $F : \mathcal{V} \rightarrow \mathcal{R}$ are ‘weakly spectral’. Let $\Theta : \mathcal{V} \rightarrow \mathbb{R}$ be ‘any’ function. Let $\mathbf{a} \in \text{dom } \Theta$ be such that $\partial_E \Theta(\mathbf{a}) \neq \emptyset$. If $\mathbf{a}$ is a ‘local maximizer’ of the map

$$\mathbf{x} \in E \mapsto \Theta(\mathbf{x}) + F(\mathbf{x}),$$

then $\mathbf{a}$ ‘operator commutes’ with every element of $\partial_E \Theta(\mathbf{a})$.

- Note that the theorem above only requires weak spectrality and locality of the maximizer in comparison to those in Gowda’s result.
- However, the price we pay is operator commutativity rather than strong operator commutativity.

J. Jeong and D. Sossa
Commutation principles for nonsmooth variational problems on Euclidean Jordan algebras
SIAM Journal on Optimization, 35.3 (2025): 1551–1569

J. Jeong
Geometric commutation principles for weakly spectral sets in Euclidean Jordan algebras
Optimization Letters 20 (2026): 401–410

Nonsmooth Commutation Principle with Weak Spectrality in EJAs

Theorem (Jeong-Sossa, 2025)

Suppose that E in $\mathcal{V}$ and $F : \mathcal{V} \rightarrow \mathcal{R}$ are ‘weakly spectral’. Let $\Theta : \mathcal{V} \rightarrow \mathcal{R}$ be ‘any’ function. Let $\mathbf{a} \in \text{dom } \Theta$ be such that Θ is subdifferentially regular at $\mathbf{a}$. If $\mathbf{a}$ is a ‘local minimizer’ of the map

$$\mathbf{x} \in E \mapsto \Theta(\mathbf{x}) + F(\mathbf{x})$$

and if the following condition is satisfied:

If $\mathbf{a}$ operator commutes with $\mathbf{v} \in \partial^\infty \Theta(\mathbf{a})$, then $\mathbf{v} = \mathbf{0}$,

then $\mathbf{a}$ ‘operator commutes’ with some element in $\partial\Theta(\mathbf{a})$.

Here $\partial\Theta$ denotes the limiting (Mordukhovich) subdifferential and $\partial^\infty \Theta$ denotes the horizon subdifferential.

5. Conclusions

Conclusions

- In this talk, we reviewed the classical notions of generalized convexity, including convexity, quasiconvexity, and pseudoconvexity, and their key properties in optimization.
- It is always a good exercise to convert an original problem to a simple one as much as you can. The transfer principle and commutation principle are such techniques.
- We also briefly talked about how the commutation principle has been developed in the setting of Euclidean Jordan algebras.
- These principles provide mathematical tools for analyzing spectral regularization and designing structured matrix updates in learning.

Thank you for your attention!

For Transfer principle:

- M. Baes, “Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras”, *Linear Algebra and its Applications*, 422.2 (2007): 664–700.
- J. Jeong and M. S. Gowda “Spectral sets and functions on Euclidean Jordan algebras” *Linear Algebra and its Applications* 518 (2017): 31–56.
- J. Jeong, “Generalized convexity for spectral functions on Euclidean Jordan algebras”, Submitted (2026)
- B. F. Lourenço, and A. Takeda, “Generalized subdifferentials of spectral functions over Euclidean Jordan algebras”, *SIAM Journal on Optimization*, 30.4 (2020): 3387–3414.
- D. Sun and J. Sun, “Löwner’s operator and spectral functions in Euclidean Jordan algebras”, *Mathematics of Operations Research*, 33.2 (2008): 421–445.

For Commutation principle:

- M. S. Gowda, “Commutation principles for optimization problems on spectral sets in Euclidean Jordan algebras”, Optimization Letters, 16 (2022): 1119–1128.
- M. S. Gowda and J. Jeong, “Commutation principles in Euclidean Jordan algebras and normal decomposition systems”, SIAM Journal on Optimization, 27 (2017): 1390–1402.
- J. Jeong and D. Sossa, “Commutation principles for nonsmooth variational problems on Euclidean Jordan algebras”, SIAM Journal on Optimization, 35.3 (2025): 1551–1569.
- J. Jeong, “Geometric commutation principles for weakly spectral sets in Euclidean Jordan algebras” Optimization Letters 20 (2026): 401–410.
- H. Ramírez, A. Seeger, and D. Sossa, “Commutation principle for variational problems on Euclidean Jordan algebras”, SIAM Journal on Optimization, 23 (2013) 687–694.