Fan-Theobald-von Neumann systems

M. Seetharama Gowda
(Joint work with Juyoung Jeong)

Department of Mathematics
University of Maryland Baltimore County
Baltimore, Maryland 21250, USA

Contents

Motivation

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

$$\max\left\{\langle C, S\rangle : X \succeq 0,\ 1 \leq \lambda_{max}(X) \leq 2\right\},$$

where $\mathcal{S}^n$ is the space of all $n \times n$ real symmetric matrices,

$X \succeq 0$ means X is positive semidefinite, and

$\lambda_{max}(X) =$ maximum eigenvalue of X.

It turns out that this problem is equivalent to

$$\max\{\langle \lambda(C), q\rangle : 0 \leq q \in \mathcal{R}^n : q_1 \geq q_2 \geq \cdots \geq q_n,\ 1 \leq q_1 \leq 2\}$$

which is a (linear programming) problem in $\mathcal{R}^n$.

Moreover, as a consequence of a result due to Ramirez, Seeger, and Sossa, if X^* solves the former problem, then C and X^* commute.

So, we went from a problem in $\mathcal{S}^n$ to a problem in $\mathcal{R}^n$ and at the same time obtained a commutativity relation.

Can we address such a transfer of optimization problems from one space to another and also get commutativity in a general setting? The key features in the above example are: $\mathcal{S}^n$ and $\mathcal{R}^n$ are real inner product spaces, (the eigenvalue map) $\lambda : \mathcal{S}^n \to \mathcal{R}^n$ satisfies

$$\langle X, Y \rangle \leq \langle \lambda(X), \lambda(Y) \rangle$$

with a condition for equality. We formulate the definition of Fan-Theobald-von Neumann system based on these.

Fan-Theobald-von Neumann system

$\mathcal{V}$ and $\mathcal{W}$ are real inner product spaces,

$\lambda : \mathcal{V} \to \mathcal{W}$ is a map.

For $u \in \mathcal{V}$, its λ-orbit is $[u] = \{x \in \mathcal{V} : \lambda(x) = \lambda(u)\}$.

$(\mathcal{V}, \mathcal{W}, \lambda)$ is a **FTvN system** if

$$\max \Big\{ \langle c, x \rangle : x \in [u] \Big\} = \langle \lambda(c), \lambda(u) \rangle \quad (\forall c, u \in \mathcal{V}). \qquad (1)$$

From this we get **FTvN inequality**:

$\langle x, y \rangle \leq \langle \lambda(x), \lambda(y) \rangle \quad (x, y \in \mathcal{V})$.

If equality holds, we say x and y **commute** in the FTvN system.

Example 1

$\mathcal{V}$ - any real inner product space,

$\mathcal{W} = \mathcal{R}$, and $\lambda(x) = ||x||$.

For $u \in \mathcal{V}$, $[u] = \{x \in \mathcal{V} : ||x|| = ||u||\}$.

Then, $(\mathcal{V}, \mathcal{R}, \lambda)$ is a FTvN system.

Here, the FTvN inequality becomes

the Cauchy-Schwarz inequality.

Example 2

$\mathcal{V} = \mathcal{W} = \mathcal{R}^n$ with usual inner product.

$\lambda(x) = x^{\downarrow}$ (decreasing rearrangment of x).

For $u \in \mathcal{R}^n$, $[u] = \{Pu : P$ is a permutation matrix$\}$

Then, $(\mathcal{R}^n, \mathcal{R}^n, \lambda)$ is a FTvN system.

Here, the FTvN inequality becomes:

The Hardy-Littlewood-Pólya rearrangement inequality:

$\langle x, y \rangle \leq \langle x^{\downarrow}, y^{\downarrow} \rangle$.

Example 3

$\mathcal{V} = \mathcal{H}^n$ (space of all $n \times n$ Hermitian matrices), $\mathcal{W} = \mathcal{R}^n$.

$\lambda(X) =$ vector of eigenvalues of X written in the decreasing order.

For $X \in \mathcal{S}^n$, $[X] = \{UXU^* : U \text{ unitary}\}$,

Then, $(\mathcal{H}^n, \mathcal{R}^n, \lambda)$ is a FTvN system. Here, the FTvN inequality is:

Ky Fan/Richter inequality $\langle X, Y \rangle \leq \langle \lambda(X), \lambda(Y) \rangle$.

Equality case due to **Theobald**.

Similarly for $\mathcal{S}^n$ (space of all $n \times n$ real symmetric matrices).

Example 4

$\mathcal{V} = M_n$ (all $n \times n$ complex matrices) with $\langle X, Y \rangle := Re\, tr(X^*Y)$, $\mathcal{W} = \mathcal{R}^n$, and $\lambda(X) = s(X)$ (vector of singular values of X

written in the decreasing order).

For $X \in M_n$, $[X] = \{UXV : U, V \text{ unitary}\}$,

Then, $(M_n, \mathcal{R}^n, \lambda)$ is a FTvN system.

Here, the FTvN inequality is:

von Neumann's inequality $\langle X, Y \rangle \leq \langle s(X), s(Y) \rangle$.

Other Examples

$(V, \langle\cdot,\cdot\rangle, \circ)$ is a **Euclidean Jordan algebra** (in short, EJA) if

V is a finite dimensional real inner product space

and the bilinear Jordan product $x \circ y$ satisfies:

- $x \circ y = y \circ x$
- $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$
- $\langle x \circ y, z\rangle = \langle x, y \circ z\rangle$

$K = \{x^2 : x \in V\}$ is the symmetric cone in V.
(It is a self-dual, homogeneous cone.)

e $(\neq 0)$ is a unit element if $x \circ e = x$ for all x.

Any EJA is a product of the following:

- $\mathcal{S}^n = \text{Herm}(\mathcal{R}^{n\times n})$ - $n \times n$ real symmetric matrices.
- $\mathcal{H}^n = \text{Herm}(\mathcal{C}^{n\times n})$ - $n \times n$ complex Hermitian matrices.
- $\mathcal{Q}^n = \text{Herm}(\mathcal{Q}^{n\times n})$ - $n \times n$ quaternion Hermitian matrices.
- $\mathcal{O}^3 = \text{Herm}(\mathcal{O}^{3\times 3})$ - 3×3 octonion Hermitian matrices.
- $\mathcal{L}^n$ - Jordan spin algebra.

In a EJA, for $a \in \mathcal{V}$, $L_a(x) := a \circ x : \mathcal{V} \to \mathcal{V}$.

We say a and b *operator commute* in $\mathcal{V}$ if

$L_a L_b = L_b L_a$.

(On $\mathcal{H}^n$, this is the same as matrix commutativity.)

In a EJA, a nonzero element c is an idempotent

if $c^2 = c$. It is a primitive idempotent

if it is not the sum of two other idempotents.

Jordan frame: A finite set $\{e_1, e_2, \ldots, e_n\}$ of mutually orthogonal

primitive idempotents summing to (the unit) e.

As all Jordan frames have the same cardinality,

we let rank of $\mathcal{V} = n$.

In a EJA of rank n, every element x has a *spectral decomposition*

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n,$$

where $\{e_1, e_2, \ldots, e_n\}$ is a Jordan frame,

$x_1, x_2, \ldots, x_n$ are real numbers (called the eigenvalues of x).

- $\lambda(x) := (x_1, x_2, \ldots, x_n)^{\downarrow}$ is the eigenvalue vector of x.
- $\lambda : \mathcal{V} \to \mathcal{R}^n$ is the eigenvalue map.
- trace$(x) := x_1 + x_2 + \cdots + x_n$.

From now on, we replace the given inner product on $\mathcal{V}$ by

the *canonical inner product* $\langle x, y\rangle :=$ trace$(x \circ y)$.

Theorem*: In a EJA carrying the canonical inner product,

$$\langle x, y\rangle \leq \langle \lambda(x), \lambda(y)\rangle$$

with equality if and only if x and y have their spectral decompositions

$x = x_1 e_1 + x_2 e_2 + +x_n e_n$ and $y = y_1 e_1 + y_2 e_2 + \cdots + y_n e_n$,

where $x_1 \geq x_2 \geq \cdots \geq x_n$ and $y_1 \geq y_2 \geq \cdots \geq y_n$.

(A consequence of equality: x and y operator commute.)

- *If $\mathcal{V}$ is a EJA of rank n, then $(\mathcal{V}, \mathcal{R}^n, \lambda)$ is a FTvN system.*

* Lim-Kim-Faybusovich, Baes, Lewis, Gowda-Tao

Hyperbolic polynomials

$\mathcal{V}$ is a real finite dimensional space, $e \in \mathcal{V}$, p is a real homogeneous polynomial of degree n on $\mathcal{V}$. p is **hyperbolic with respect to** e if $p(e) \neq 0$ and for every $x \in \mathcal{V}$, roots of $p(te - x) = 0$ are all real. For any $x \in \mathcal{V}$, let $\lambda(x)$ denote the vector of roots of $p(te - x) = 0$ written in the decreasing order. Assuming p is complete (which means that $\lambda(x) = 0 \Rightarrow x = 0$), $\mathcal{V}$ can be made into an inner product space:

$$\langle x, y \rangle := \frac{1}{4}\Big[||\lambda(x+y)||^2 - ||\lambda(x-y)||^2 \Big].$$

Then λ is norm-preserving and $\langle x, y \rangle \leq \langle \lambda(x), \lambda(y) \rangle$. Under an 'isometric' condition (Bauschke et al., 2001), $(\mathcal{V}, \mathcal{R}^n, \lambda)$ becomes a FTvN system.

Normal decomposition systems (NDS)

Let $\mathcal{V}$ be a real inner product space, $\mathcal{G}$ be a closed subgroup of the orthogonal group of $\mathcal{V}$, and $\gamma : \mathcal{V} \to \mathcal{V}$ be a map satisfying the following conditions:

(a) γ is $\mathcal{G}$-invariant, that is, $\gamma(Ax) = \gamma(x)$ for all $x \in \mathcal{V}$ and $A \in \mathcal{G}$.

(b) For each $x \in \mathcal{V}$, there exists $A \in \mathcal{G}$ such that $x = A\gamma(x)$.

(c) For all $x, y \in \mathcal{V}$, we have $\langle x, y\rangle \leq \langle \gamma(x), \gamma(y)\rangle$.

Then, $(\mathcal{V}, \mathcal{G}, \gamma)$ is called a *normal decomposition system*.

• Lewis 1996, Group invariance and convex matrix analysis, SIAM J. Matrix Anal.

In a normal decomposition system,

(i) For any two elements x and y in $\mathcal{V}$, we have

$$\max_{A \in \mathcal{G}} \langle Ax, y \rangle = \langle \gamma(x), \gamma(y) \rangle .$$

Also, $\langle x, y \rangle = \langle \gamma(x), \gamma(y) \rangle$ holds for two elements x and y if and only if there exists an $A \in \mathcal{G}$ such that $x = A\gamma(x)$ and $y = A\gamma(y)$.

(ii) The range of γ, denoted by F, is a closed convex cone in $\mathcal{V}$.

Hence: If $(\mathcal{V}, \mathcal{G}, \gamma)$ is a NDS, then $(\mathcal{V}, \mathcal{V}, \gamma)$ is a FTvN system.

Eaton triple

Let $\mathcal{V}$ be a finite dimensional real inner product space, $\mathcal{G}$ be a closed subgroup of the orthogonal group of $\mathcal{V}$, and F be a closed convex cone in $\mathcal{V}$ satisfying the following conditions:

(a) $Orb(x) \cap F \neq \emptyset$ for all $x \in \mathcal{V}$, where $Orb(x) := \{Ax : A \in \mathcal{G}\}$.

(b) $\langle x, Ay \rangle \leq \langle x, y \rangle$ for all $x, y \in F$ and $A \in \mathcal{G}$.

Then, $(\mathcal{V}, \mathcal{G}, F)$ is called an *Eaton triple*.

• Eaton-Perlman 1977, Reflection groups, generalized Schur functions, and the geometry of majorization, Ann. Probab.
• Eaton 1987, Group induced orderings with some applications in statistics, CWI Newsletter.

Normal decomposition system = Eaton triple

It is known that in an Eaton triple $(\mathcal{V}, \mathcal{G}, F)$,

$Orb(x) \cap F$ consists of exactly one element for each $x \in \mathcal{V}$.

Defining $\gamma : \mathcal{V} \to \mathcal{V}$ with $Orb(x) \cap F = \{\gamma(x)\}$, $(\mathcal{V}, \mathcal{G}, \gamma)$ becomes

a normal decomposition system. Also, given a

finite dimensional normal decomposition system $(\mathcal{V}, \mathcal{G}, \gamma)$,

with $F := \gamma(\mathcal{V})$, $(\mathcal{V}, \mathcal{G}, F)$ becomes an Eaton triple.

Thus, *finite dimensional normal decomposition systems are equivalent to Eaton triples.*

Elementary properties of λ

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system.

Then,

- *λ is norm-preserving, positively homogeneous, and Lipschitz,*
- *For every $c \in \mathcal{V}$, $f(x) := \langle \lambda(c), \lambda(x) \rangle$ is sublinear,*
- *$\langle x, y \rangle = \langle \lambda(x), \lambda(y) \rangle$ iff $\lambda(x+y) = \lambda(x) + \lambda(y)$.*

Spectral set

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system.

A **spectral set** in $\mathcal{V}$ is of the form $E = \lambda^{-1}(Q)$ for some $Q \subseteq \mathcal{W}$.

Theorem: *Let E be spectral in $\mathcal{V}$. Then,*

- *closure/interior/boundary of E is spectral.*
- *If $\mathcal{V}$ is finite dimensional, then convex hull of E is spectral and sum of two convex spectral sets is spectral.*
- *If $\mathcal{V}$ is a Hilbert space, then the closed convex hull of E is spectral and sum of two compact convex spectral sets is spectral.*

Spectral function

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system. A **spectral function** on $\mathcal{V}$ is of the form $\Phi = \phi \circ \lambda$ for some $\phi : \mathcal{W} \to \mathcal{R}$.

Then,

- *A set E is spectral in $\mathcal{V}$ iff its indicator/characteristic function is spectral.*
- *A real-valued function Φ on $\mathcal{V}$ is spectral iff its epigraph is spectral in the product FTvN space $(\mathcal{R} \times \mathcal{V}, \mathcal{R} \times \mathcal{W}, \Lambda)$, where $\Lambda(t, x) = (t, \lambda(x))$.*

Center, unit element

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system. Then,

- x and y **commute** in $\mathcal{V}$ if $\langle x, y\rangle = \langle \lambda(x), \lambda(y)\rangle$.
- **Center** of FTvN system is $\mathcal{C} := \{x \in \mathcal{V} : x \,\text{commutes with every}\, y \in \mathcal{V}\}$.
- A nonzero e in $\mathcal{V}$ is a **unit element** if center $\mathcal{C} = \mathcal{R}\, e$.

Theorem:
*$\mathcal{C}$ is a closed subspace of $\mathcal{V}$ and λ is linear on it.
If $\mathcal{V}$ is a Hilbert space, then $\mathcal{V} = \mathcal{C} + \mathcal{C}^{\perp}$. Moreover, in the FTvN space $(\mathcal{C}, \mathcal{W}, \lambda)$, the center is $\mathcal{C}$ and in the FTvN space $(\mathcal{C}^{\perp}, \mathcal{W}, \lambda)$, the center is $\{0\}$.*

An example

In the FTvN system $(\mathcal{H}^n, \mathcal{R}^n, \lambda)$,

- spectrality $=$ unitary invariance.
- X and Y commute in the FTvN system iff

$$X = U diag(\lambda(X))U^* \quad \text{and} \quad Y = U diag(\lambda(Y))U^*$$

  for some unitary matrix U.

  (Note: $\lambda(X)$ and $\lambda(Y)$ have decreasing components.)
- Center $=$ All multiples of the identity matrix.
- The Identity matrix is a unit.

Majorization

In a FTvN system $(\mathcal{V}, \mathcal{W}, \lambda)$,

- x is **majorized** by y if $x \in \operatorname{conv}[y]$. We write $x \prec y$.
- A linear transformation $D : \mathcal{V} \to \mathcal{V}$ is **doubly stochastic** if $Dx \prec x$ for all $x \in \mathcal{V}$.
- An invertible linear transformation $A : \mathcal{V} \to \mathcal{V}$ is an **automorphism** if $\lambda(Ax) = \lambda(x)$ for all $x \in \mathcal{V}$.

Theorem: *D is a doubly stochastic transformation on $(\mathcal{V}, \mathcal{W}, \lambda)$ with adjoint D^* (whenever defined). Then,*

- *$D(E) \subseteq E$ for any convex spectral set E.*
- *$u \in \mathcal{C} \Rightarrow Du = u$, $D^*u = u$.*
  If e is a unit then, $De = e$, $D^*e = e$.
- *If A is linear and invertible, then A is an automorphism iff A and A^{-1} are doubly stochastic.*
- *If $\mathcal{V}$ is finite dimensional, then any convex combination of automorphisms is doubly stochastic.*

Reduced system of a FTVN system

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system. Then

a FTvN system $(\mathcal{W}, \mathcal{W}, \mu)$ is a **reduced system** of $(\mathcal{V}, \mathcal{W}, \lambda)$ if

- $range(\mu) \subseteq range(\lambda)$,
- $\mu \circ \lambda = \lambda$.

In this setting,

- $range(\mu) = range(\lambda)$,
- $\mu \circ \mu = \mu$.

Examples: $(\mathcal{R}^n, \mathcal{R}^n, \mu)$ with $\mu(x) = x^{\downarrow}$ is a reduced system of $(\mathcal{H}^n, \mathcal{R}^n, \lambda)$.

Every normal decomposition system is a reduced system of itself.

Theorem:

Suppose $(\mathcal{W}, \mathcal{W}, \mu)$ *is a reduced system of* $(\mathcal{V}, \mathcal{W}, \lambda)$ *with* $\mathcal{W}$ *finite dimensional. Then,*

- $\lambda(x+y) \prec \lambda(x) + \lambda(y)$.
- $x \prec y$ *implies* $\lambda(x) \prec \lambda(y)$. *The converse holds if* $\mathcal{V}$ *is finite dimensional.*

Question: When do we have **Lidskii type inequality**

$$\lambda(x) - \lambda(y) \prec \lambda(x-y).$$

True for hyperbolic polynomials: Gurvitz via Lax conjecture.

Transfer principles

Theorem: *Let $(\mathcal{W}, \mathcal{W}, \mu)$ be a reduced system of $(\mathcal{V}, \mathcal{W}, \lambda)$, Q be spectral in $\mathcal{W}$, and $E = \lambda^{-1}(Q)$. Then,*

- $\lambda^{-1}(Q^{\lozenge}) = (\lambda^{-1}(Q))^{\lozenge}$, *where $\lozenge$ denotes closure/interior/boundary operation.*
- *Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional. Then,*
  - *E is convex iff Q is convex.*
  - $\overline{\mathrm{conv}}\, \lambda^{-1}(Q) = \lambda^{-1}(\overline{\mathrm{conv}}\, Q)$.
  - *For convex spectral sets Q_1 and Q_2 in $\mathcal{W}$,*

$$\lambda^{-1}(Q_1 + Q_2) = \lambda^{-1}(Q_1) + \lambda^{-1}(Q_2).$$

Theorem: *Let $(\mathcal{W}, \mathcal{W}, \mu)$ be a reduced system of $(\mathcal{V}, \mathcal{W}, \lambda)$, $\phi : \mathcal{W} \to \mathcal{R}$ be spectral, $\Phi := \phi \circ \lambda$. Then*

- *When $\mathcal{W}$ is finite dimensional, convexity of ϕ implies that of Φ.*
- *When $\mathcal{V}$ and $\mathcal{W}$ are finite dimensional, Φ is convex iff ϕ is convex.*

This extends a celebrated result of Davis which says that a unitarily invariant function on $\mathcal{H}^n$ (the space of all n$\times$n complex Hermitian matrices) is convex if and only if its restriction to diagonal matrices is convex.

An optimization result

Theorem: In the FTvN system $(\mathcal{V}, \mathcal{W}, \lambda)$, for any spectral set E,

$c \in \mathcal{V}$, and $\phi : \mathcal{W} \to \mathcal{R}$,

$$\sup_{x \in E} \big\{ \langle c, x \rangle + (\phi \circ \lambda)(x) \big\} = \sup_{y \in \lambda(E)} \big\{ \langle \lambda(c), y \rangle + \phi(y) \big\}.$$

Attainment of one supremum implies that of the other. Moreover, if the supremum on the left is attained at x^*, then c commutes with x^*.

Note: Putting $\phi = 0$ and $E = [u]$, the above equality leads to our definition of FTvN system.

Fenchel and subdifferential formulas

Let X be a real inner product space, $f : X \to \mathcal{R} \cup \{\infty\}$.

Given $a \in S \subseteq X$ with $f(a) < \infty$,

recall the **subdifferential** of f at a relative to S:

$$\partial_S f(a) = \{d \in X : f(x) - f(a) \geq \langle d, x - a \rangle \, \forall x \in S\}$$

and the **Fenchel conjugate** of f relative to S:

$$f_S^*(z) = \sup\{\langle z, x \rangle - f(x) : \, x \in S\} \quad (z \in X).$$

Theorem:

Let $(\mathcal{V}, \mathcal{W}, \lambda)$ be a FTvN system, S be spectral in $\mathcal{V}$, $\phi : \mathcal{W} \to \mathcal{R} \cup \{\infty\}$, and $\Phi := \phi \circ \lambda$. Then,

- $\Phi_S^*(z) = \phi_{\lambda(S)}^*\big(\lambda(z)\big) \quad (z \in \mathcal{V})$.
- $y \in \partial_S \Phi(a) \Leftrightarrow \lambda(y) \in \partial_{\lambda(S)} \phi(\lambda(a)),\ y$ and a commute.

Moreover, when $(\mathcal{W}, \mathcal{W}, \mu)$ is a reduced system of $(\mathcal{V}, \mathcal{W}, \lambda)$ and ϕ is spectral on $\mathcal{W}$, we can replace $\lambda(S)$ by $[\lambda(S)]$.

In particular, with $S = \mathcal{V}$, get $(\phi \circ \lambda)^* = \phi^* \circ \lambda$, etc.

Note: The above two items are equivalent to the defining condition of FTvN system.

A geometric commutation principle

Let E be a spectral set in a FTvN system.

For any $a \in E$, let

$$N_E(a) := \{d \in \mathcal{V} : \langle d, x - a \rangle \leq 0 \ \forall\, x \in E\}$$

denote the **normal cone** *of E at a. Then every d in $N_E(a)$ commutes with a.*

Example: If x^* solves the **variational inequality problem** *VI(f, E), then x^* commutes with $-f(x^*)$.*

References

(1) M.S. Gowda, *Optimizing certain combinations of linear/distance functions over spectral sets*, arXiv:1902.06640v2 (2019).

(2) M.S. Gowda, *Commutation principles for optimization problems on spectral sets in Euclidean Jordan algebras*, Opt. Lett., 16, 1119–1128 (2022).

(3) M.S. Gowda and J. Jeong, *Commutativity, majorization, and reduction in FTvN systems*, Results Math., 78:72 (2023).

(4) J. Jeong and M.S. Gowda, *Transfer principles, Fenchel and subdifferential formulas in FTvN systems*, arXiv:2307.08478v1 (2023).