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

ILAS 2023, Madrid, 13 June 2023

Contents

- Motivation
- Definition
- Examples
- Concepts and results
- Reduced system
- Transfer principles
- Fenchel conjugate and subdifferential formulas

Motivation

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

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

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 $n \times n$ 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

- $(\mathcal{V}, \mathcal{R}^n, \lambda)$ where $\mathcal{V}$ is a Euclidean Jordan algebra of rank n,

  $\mathcal{W} = \mathcal{R}^n$, $\lambda(x)$ is the eigenvector of x.

- Systems induced by complete, isometric hyperbolic polynomials.

- Normal decomposition systems (includes Eaton triples).

- Infinite dimensional system (l_2, l_2, λ).

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

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 $(\mathcal{H}^n, \mathcal{H}^n, \lambda)$,

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

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

  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.

An optimization result

In the FTvN system $(\mathcal{V}, \mathcal{W}, \lambda)$, for any spectral set E, $c \in \mathcal{V}$, $\phi : \mathcal{W} \to \mathcal{R}$, and $\Phi = \phi \circ \lambda$,

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

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^*.

In terms of the Fenchel conjugate, above equality is equivalent to

$$(\phi \circ \lambda)^*_E(c) = \phi^*_{\lambda(E)}(\lambda(c)),$$

and, in a specialized form, to

$$(\phi \circ \lambda)^* = \phi^* \circ \lambda.$$

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**

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

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,*

- $\lambda^{-1}(Q^{\Diamond}) = (\lambda^{-1}(Q))^{\Diamond}$, *where* $\Diamond$ *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.

Fenchel conjugate 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*, Manuscript, May 2023.