# Jordan Algebras, Symmetric Cones and General Semidefinite Programming

Farid Alizadeh
Rutgers University and
Stefan Schmieta
Dash Software

Presented at
Workshop on
Novel Approaches to Hard Discrete Optimization
University of Waterloo
Waterloo, ON, Canada
April 27, 2001

# Nonlinear Relaxations

**Question:** Can a combinatorial optimization problem be relaxed to a convex optimization problem?

**Known Answers:**

- Linear Programming
- semidefinite programming

**Glaring Omission:** Lorentz cone programming: Convex quadratic programming with convex quadratic constraints

In this talk we explore generalization of Semidefinite programming

## “Word-by-Word Extensions”

**Question:**

1. How much of analysis, algorithms, proofs, etc. extend from LP to SDP to symmetric cone optimization?
2. Which ones extend from SDP to symmetric cone optimization?

# Some Algebra

- $(\mathcal{A}, *)$ is an Algebra if

  1. $\mathcal{A}$ is a vector space over $\Re$,
  2. "*" distributes over "+": $x * (y + z) = x * y + x * z$, $(y + z) * x = y * x + z * x$

  **Observe:**

  1. If $x * y = z$ then $z_i$ are bilinear in $x$ and $y$:

  $$z_i = x^T A_i y \qquad A_i \in \Re^{n \times n}$$

  2. For every $x \in \mathcal{A}$ there is a matrix $L(x)$:

  $$L(x) y = x * y$$

- $(\mathcal{A}, .)$ is *associative* if $x(yz) = (xy)z$ always
  1. in associative algebras $x \to L(x)$ defines an endomorphism into algebra of $n \times n$ matrices
  2. Thus every associative algebra is isomorphic to some subalgebra of $n \times n$ matrices.
- $(\mathcal{V}, \cdot)$ is *power-associative algebra* if the subalgebra generated by an element $\mathbf{x}$ is associative
  1. We assume identity $\mathbf{e}$: $\mathbf{ex} = \mathbf{xe} = \mathbf{x}$
  2. $x^p$ is unambiguous
- Power associative algebras have many of formal properties of $n \times n$ matrices

In power associative algebra $(\mathcal{V}, \cdot)$:

- for each $\mathbf{x} \in \mathcal{V}$ there is a least integer $r$ such that

$$\{\mathbf{e}, \mathbf{x}, \mathbf{x}^2, \ldots, \mathbf{x}^r\}$$

  is linearly dependent: $\mathbf{x}^r - a_1(\mathbf{x})\mathbf{x}^{r-1} + \cdots + (-1)^r a_r(\mathbf{x})\mathbf{e} = \mathbf{0}$

- *The minimum polynomial* of $\mathbf{x}$:
  $p(\lambda; \mathbf{x}) \stackrel{\text{def}}{=} \lambda^r - a_1(\mathbf{x})\lambda^{r-1} + \cdots + (-1)^r a_r(\mathbf{x})$ and its roots are *eigenvalues* of $\mathbf{x}$ is $\mathcal{V}$.

- $a_i(\mathbf{x})$ are homogeneous of degree $i$: $a_i(t\mathbf{x}) = t^i a_i(\mathbf{x})$

- $$\mathrm{tr}(\mathbf{x}) \stackrel{\text{def}}{=} a_1(\mathbf{x}) \qquad \text{is linear in } \mathbf{x}$$

  $$\det(\mathbf{x}) \stackrel{\text{def}}{=} a_r(\mathbf{x})$$

  $$\mathrm{cork}(\mathbf{x}) \stackrel{\text{def}}{=} \text{multiplicity of zero eigenvalue}$$

  $$\deg(\mathbf{x}) \stackrel{\text{def}}{=} r$$

  $$r \stackrel{\text{def}}{=} \mathrm{rank}(\mathcal{V}) = \max_{\mathbf{x}}\{\mathbf{x} \in \mathcal{V}\}$$

  $$\mathrm{rk}(\mathbf{x}) \stackrel{\text{def}}{=} r - \mathrm{cork}(x)$$

- An element $\mathbf{x}$ is *regular* if $\deg(\mathbf{x}) = \mathrm{rank}(\mathcal{V})$
- The set of regular elements of $\mathcal{V}$ are dense and generic
- **Characteristic polynomial**:
  1. For regular $\mathbf{x}$: $P_{\mathrm{Char}}(\mathbf{x}) \stackrel{\text{def}}{=} P_{\mathrm{Min}}(\mathbf{x})$
  2. Analytically continue characteristic polynomials all $\mathbf{x} \in \mathcal{V}$

## Jordan Algebras

- $(\mathcal{J}, \circ)$ is *Jordan algebra* if

  1. $\mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$
  2. $\mathbf{x} \circ (\mathbf{x}^2 \circ \mathbf{y}) = \mathbf{x}^2 \circ (\mathbf{x} \circ \mathbf{y})$ that is $\mathsf{L}(\mathbf{x})\mathsf{L}(\mathbf{x}^2) = \mathsf{L}(\mathbf{x}^2)\mathsf{L}(\mathbf{x})$.

- Jordan algebras are power-associative

## Three Examples

1. Every associative algebra $\mathcal{A}$ induces a Jordan algebra: $\mathbf{x} \circ \mathbf{y} \stackrel{\text{def}}{=} \frac{\mathbf{xy}+\mathbf{yx}}{2}$ for example $n \times n$ matrices $(\mathcal{M}_n, \circ)$
2. Symmetric matrices under $\circ$: $(\mathcal{S}_n, \circ)$ Here $L(X) = \frac{X \otimes I + I \otimes X}{2}$
3. $(\Re^{n+1}, \circ)$ where:

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \circ \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} x_0 y_0 + \cdots + x_n y_n \\ x_0 y_1 + y_0 x_1 \\ \vdots \\ x_0 y_n + y_0 x_n \end{pmatrix}$$

and

$$\mathrm{L}(\mathbf{x}) = \mathrm{Arw}(\mathbf{x}) = \begin{pmatrix} x_0 & \bar{\mathbf{x}}^\top \\ \bar{\mathbf{x}} & x_0 \mathrm{I} \end{pmatrix}$$

where

$$\bar{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{x}} = \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \mathbf{e} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

## Euclidean Jordan Algebras

- if $(\mathcal{E}, \circ)$ is Jordan and $\mathrm{tr}(\mathbf{x}^2) > 0$ for all $\mathbf{x} \neq \mathbf{0}$ then $\mathcal{E}$ is a *Euclidean Jordan Algebra*
- **Jordan Frame:**
  1. $\mathbf{c}$ is idempotent if $\mathbf{c}^2 = \mathbf{c}$
  2. An orthogonal system of idempotents $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$:

$$\mathbf{c}_i \circ \mathbf{c}_j = \mathbf{0} \quad \text{for } i \neq j \qquad \sum \mathbf{c}_i = \mathbf{e}$$

- An idempotent is *primitive* if it is not sum of two other idempotents
- Definition: A *Jordan frame* is a system of primitive orthogonal idempotents.
- It can be shown that if $\{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$ is a Jordan frame then $r = \mathrm{rank}(\mathcal{J})$

- FACT: Spectral decomposition theorem
  1. All eigenvalues $\lambda_i$ of $\mathbf{x} \in \mathcal{E}$ are real numbers
  2. For each $\mathbf{x} \in \mathcal{E}$ there is a Jordan frame $\{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$ such that

$$\mathbf{x} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_r \mathbf{c}_r$$

  3. The cone of squares (or the *semidefinite cone:*) $\mathcal{K}(\mathcal{E}) = \{\mathbf{x}^2 : \mathbf{x} \in \mathcal{E}\}$
  4. FACT: A cone is symmetric (homogeneous and self-dual) iff it is the cone of squares of some Euclidean Jordan Algebra

## Examples

- $n \times n$ **Symmetric matrices** $(\mathcal{S}, \circ)$**:** Here concepts of eigenvalues, rank, determinant and norms are the usual ones

$$X = Q \Lambda Q^T = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

- If $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ an orthonormal set $\{\mathbf{q}_1 \mathbf{q}_1^T, \ldots, \mathbf{q}_n \mathbf{q}_n^T\}$ is a Jordan frame

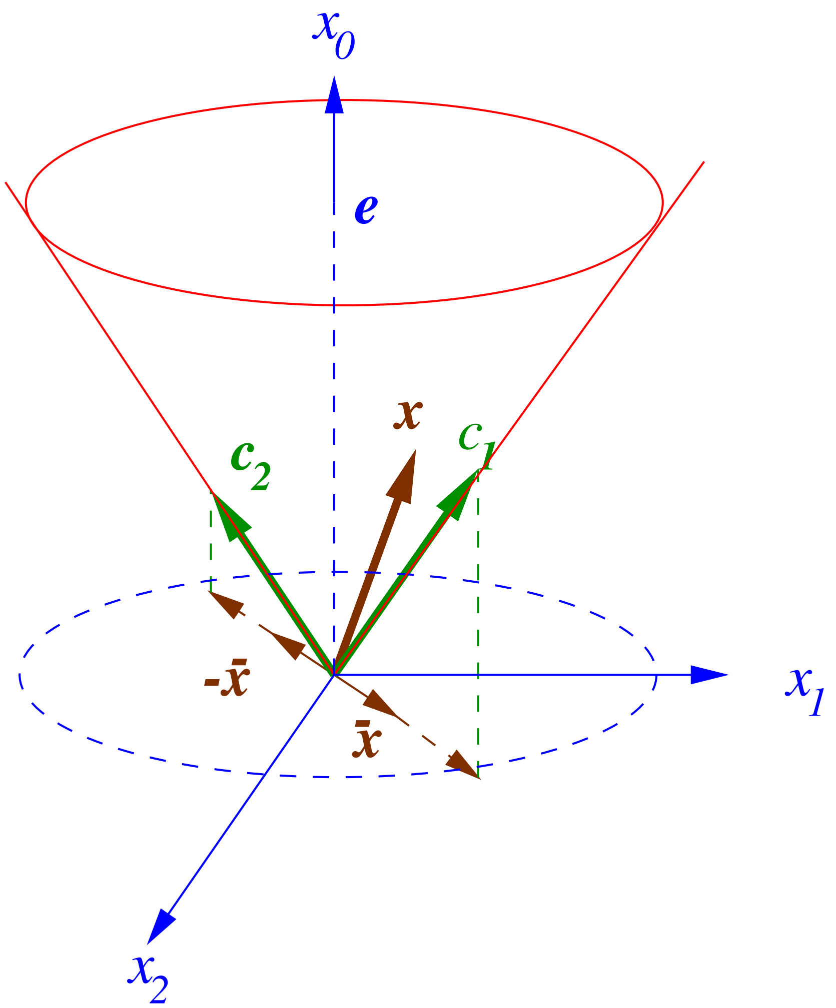
- Cone of squares is the positive semidefinite cone

- **Example:** $(\Re^{n+1}, \circ)$ Here one can see that for any $(x_0; \bar{\mathbf{x}})$

1. Cone of squares: Lorentz cone: $\mathcal{L} = \{(x_0; \bar{\mathbf{x}}) : x_0 \geq \|\bar{\mathbf{x}}\|\}$
2. Since $\mathbf{x}^2 - 2x_0\mathbf{x} + (x_0^2 - \|\bar{\mathbf{x}}\|^2)\mathbf{e} = \mathbf{0}$, the characteristic polynomial is $\lambda^2 - 2x_0\lambda + (x_0^2 - \|\bar{x}\|^2)$
3. Eigenvalues are $x_0 + \|\bar{x}\|$ and $x_0 - \|\bar{x}\|$
4. $\operatorname{tr}_{\Re^{n+1}}(x) = 2x_0$, $\det_{\Re^{n+1}}(x) = x_0^2 - \|\bar{x}\|^2$, and $\operatorname{rk}(\Re^{n+1}) = 2$, independent of $n$
5. $\operatorname{tr}_{\Re^{n+1}}(x^2) = 2\|x\|^2 \geq 0$ and thus $(\Re^{n+1}, \circ)$ is Euclidean
6. For each $(x_0; \bar{\mathbf{x}}) \in \Re^{n+1}$

$$\mathbf{x} = \frac{1}{2}(x_0 + \|\bar{\mathbf{x}}\|)\begin{pmatrix} 1 \\ \frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix} + \frac{1}{2}(x_0 - \|\bar{\mathbf{x}}\|)\begin{pmatrix} 1 \\ -\frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix}$$

- The Jordan frame $\left\{ \frac{1}{2} \begin{pmatrix} 1 \\ \frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 \\ -\frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|} \end{pmatrix} \right\}$



- Since all $\lambda_i$ are real we can define functions and norms:
  1. $\|\mathbf{x}\|_{F/\mathcal{E}} \stackrel{\text{def}}{=} \sqrt{\lambda_1^2 + \cdots + \lambda_r^2}$
  2. $\|\mathbf{x}\|_2 \stackrel{\text{def}}{=} \max_i |\lambda_i|$
  3. $f(\mathbf{x}) \stackrel{\text{def}}{=} f(\lambda_1)\mathbf{c}_1 + \cdots + f(\lambda_r)\mathbf{c}_r$
  4. $x^{-1} \stackrel{\text{def}}{=} \lambda_1^{-1}\mathbf{c}_1 + \cdots + \lambda_r^{-1}\mathbf{c}_r$
- $\mathbf{x} \succcurlyeq \mathbf{0}$ if $\lambda_i(\mathbf{x}) \geq 0$,
- **Quadratic Representation:** For each $\mathbf{x}, \mathbf{y} \in \mathcal{V}$:

$$Q_{\mathbf{x},\mathbf{y}} \stackrel{\text{def}}{=} L(\mathbf{x})L(\mathbf{y}) - L(\mathbf{y})L(\mathbf{x}) - L(\mathbf{x} \circ \mathbf{y})$$

$$Q_{\mathbf{x}} \stackrel{\text{def}}{=} 2L^2(\mathbf{x}) - L(\mathbf{x}^2)$$

## Examples

- **Symmetric matrices:** Concepts of norms are the usual ones
- $Q_{X,Y}(Z) = \frac{XZY+YZX}{2}$ and $Q_X(Y) = XYX$, $Q = X \otimes X$
- **Lorentz cone algebra:**
- $\|x\|_{F/\mathfrak{R}^{n+1}} = \sqrt{2}\|x\|$ and $\|x\|_2 = \max\Big\{\big|x_0 + \|\bar{x}\|\big|, \big|x_0 - \|\bar{x}\|\big|\Big\}$
- 

$$\mathbf{Q_x} = \begin{pmatrix} \|\mathbf{x}\|^2 & 2x_0\bar{\mathbf{x}}^T \\ 2x_0\bar{\mathbf{x}} & (x_0^2 - \|\bar{\mathbf{x}}\|^2)I + 2\bar{\mathbf{x}}\bar{\mathbf{x}}^T \end{pmatrix}$$

## Simple Jordan Algebras and Cones

- If $(\mathcal{E}_1, \circ_1)$ and $\mathcal{E}_2, \circ_2)$ are two Euclidean Algebras then $(\mathcal{E}_1 \oplus \mathcal{E}_2, \circ)$ is also a Euclidean Jordan algebra

$$(\mathbf{x}_1; \mathbf{x}_2) \circ (\mathbf{y}_; \mathbf{y}_2) \stackrel{\text{def}}{=} (\mathbf{x}_1 \circ_1 \mathbf{y}_1; \mathbf{x}_2 \circ_2 \mathbf{y}_2)$$

- An algebra is *simple* if it cannot be decomposed in this way
- There are only 5 Simple Jordan Algebras:
  1. $(\Re^{n+1}, \circ)$ of Lorentz cone
  2. $(\mathcal{S}_n, \circ)$ of $n \times n$ real symmetric matrices
  3. $(\mathcal{H}_n, \circ)$ of $n \times n$ complex Hermitian matrices
  4. $(\mathcal{A}_n, \circ)$ of $n \times n$ quaternion Hermitian matrices
  5. $(\mathcal{O}_n, \circ)$ of $3 \times 3$ octonion Hermitian matrices (Albert)
- All of these except the last one are “representable”

## Representable Jordan algebras

- An Euclidean Jordan Algebra $(\mathcal{E}, \circ)$ is

  **representable:** if it is isomorphic to a subalgebra of $(\mathcal{S}_, \circ)$

  **exceptional** if it is not representable

- Among simple Euclidean Jordan algebras only the *Albert algebra* is exceptional

## Euclidean-Jordan-Associative System

Definition: $(\mathcal{A}, \mathcal{S}, \mathcal{E}, \prime)$ is an EAJ system if

1. $(\mathcal{A}, \cdot)$ is associative
2. $\prime$ is the adjoint operation:
   (a) $\prime : \mathcal{A} \to \mathcal{A}$ is linear and $(\mathbf{x}')' = \mathbf{x}$
   (b) $(xy)' = y'x'$
   (c) $\mathrm{tr}(xx') \geq 0$
3. $\mathcal{S}$ is the set of self-adjoint elements in $\mathcal{A}$ which makes a *Euclidean* Jordan algebra under the $\mathcal{A}$-induced $\circ$
4. $(\mathcal{E}, \circ)$ is a *simple* Jordan subalgebra of $\mathcal{S}$ which generates $\mathcal{A}$.

Note: $\mathcal{E}$ is a Jordan subalgebra of $\mathcal{S}$ but they may have different ranks

In an EAJ system

- $\operatorname{rank}(\mathcal{S})/\operatorname{rank}(\mathcal{E}) = k$ an integer
- $\lambda_i$ is an eigenvalue of $\mathbf{x}$ in $\mathcal{E}$ with multiplicity $t_i$ iff it has multiplicity $kt_i$ in $\mathcal{S}$
- Since $xx' \in \mathcal{S}$ we can define norms and an inner product on $\mathcal{A}$:
  1. $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{A}} \stackrel{\text{def}}{=} \operatorname{tr}_{\mathcal{A}}(\mathbf{xy})$, $\|\mathbf{x}\|_{F/\mathcal{A}} \stackrel{\text{def}}{=} \sqrt{\sum_i \lambda_i(xx')}$,
     $\|\mathbf{x}\|_2 \stackrel{\text{def}}{=} \sqrt{\max_i \lambda_i(\mathbf{xx}')}$
  2. We can define norms in terms of $\mathcal{E}$: $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{E}} \stackrel{\text{def}}{=} \frac{\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{A}}}{k}$
     $\|\mathbf{x}\|_{F/\mathcal{E}} \stackrel{\text{def}}{=} \frac{\|\mathbf{x}\|_{F/\mathcal{S}}}{\sqrt{k}}$
- Fact: $\|\mathbf{xy}\|_{F/\mathcal{E}} \leq \|\mathbf{y}\|_2 \, \|\mathbf{x}\|_{F/\mathcal{E}}$
- **Question:** Is it true that $\|\mathbf{xy}\|_{F/\mathcal{E}} \leq \|\mathbf{y}\|_{F/\mathcal{E}} \, \|\mathbf{x}\|_{F/\mathcal{E}}$?

- **Answer:** Only sometimes!

  1. DEFINITION: $\mathbf{x} \in \mathcal{A}$ is *simple* if $\mathbf{x}\mathbf{x}' \in \mathcal{E}$
  2. FACT: if $\mathbf{y} \in \mathcal{A}$ is simple then $\|\mathbf{x}\mathbf{y}\|_{\mathrm{F}/\mathcal{E}} \leq \|\mathbf{y}\|_{\mathrm{F}/\mathcal{E}} \|\mathbf{x}\|_{\mathrm{F}/\mathcal{E}}$
  3. How do we identify simple elements in $\mathcal{A}$?

     FACT: If $\alpha \in \Re$, $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in \mathcal{E}$ and $\mathbf{a} \in \mathcal{A}$ is simple then the following are simple:

     (a) $\mathbf{xy}$

     (b) $\mathbf{xa}$

     (c) $\mathbf{u}^{-1}\mathbf{xyu}$

     (d) $\mathbf{uv} + \alpha\mathbf{u}^{-1}\mathbf{v}^{-1}$

     (e) $\mathbf{xy} + \mathbf{u}$

## Peirce Decomposition

Let $(\mathcal{E}, \circ)$ be a Euclidean Jordan algebra of dim $n$ and $\mathrm{rank}(\mathcal{E}) = r$

- for a primitive idempotent $c$ $L(c)$ is symmetric and
  1. has some eigenvalue equal to $0$
  2. has some eigenvalue equal to $1$
  3. has some eigenvalues equal to $1/2$.
- Thus $\mathcal{E} = \mathcal{E}_0 \oplus \mathcal{E}_1 \oplus \mathcal{E}_{1/2}$

$$
\begin{aligned}
\mathcal{E}_0 &= \{\mathbf{x} : \mathbf{c} \circ \mathbf{x} = \mathbf{0}\} \\
\mathcal{E}_1 &= \{\mathbf{x} : \mathbf{c} \circ \mathbf{x} = \mathbf{x}\} \\
\mathcal{E}_{\frac{1}{2}} &= \{\mathbf{x} : \mathbf{c} \circ \mathbf{x} = \mathbf{x}/2\}
\end{aligned}
$$

- If $\{\mathbf{c}_1, \dots, \mathbf{c}_r\}$ and $\{\mathbf{d}_1, \dots, \mathbf{d}_r\}$ are Jordan frames then there is $\mathbf{w} \in \mathcal{E}$:
  1. $\mathbf{w}^2 = \mathbf{e}$
  2. $Q_{\mathbf{w}}(\mathbf{c}_i) = \mathbf{d}_i$ for $i = 1, \dots r$

- The Jordan frame $\{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$ induces the following decomposition of $\mathcal{E}$

$$\mathcal{E}_{ii} \stackrel{\text{def}}{=} \mathcal{E}_1(\mathbf{c}_i) \qquad \text{with dimension } 1$$

$$\mathcal{E}_{ij} \stackrel{\text{def}}{=} \mathcal{E}_{\frac{1}{2}}(\mathbf{c}_i) \cap \mathcal{E}_{\frac{1}{2}}(\mathbf{c}_j) \qquad \text{with dimension } \mathbf{d}$$

$$\mathcal{E} = \oplus_{i \leq j} \mathcal{E}_{ij}$$

- Projection of $\mathbf{x}$ to $\mathcal{E}_{i,j}$ are given by
  1. Onto $\mathcal{E}_{ii}$ : $\qquad Q_{\mathbf{c}_i}$
  2. Onto $\mathcal{E}_{ij}$ : $\qquad 4L(\mathbf{c}_i)(L(\mathbf{c}_j))$

One can show

- $\mathcal{E}_{ij} \circ \mathcal{E}_{ij} \subset \mathcal{E}_i + \mathcal{E}_j$
- $\mathcal{E}_{ij} \circ \mathcal{E}_{jk} \subset \mathcal{E}_{ik}$ for $i \neq k$
- $\mathcal{E}_{ij} \circ \mathcal{E}_{kl} = \{\mathbf{0}\}$ for $i, j, k, l$ all distinct

**Conclusion:** Every $\mathbf{x} \in \mathcal{E}$ can be written as

$$\mathbf{x} = \sum_i x_i \mathbf{c}_i + \sum_{i<j} \mathbf{x}_{ij}$$

This is formally similar to an $r \times r$ matrix

# Partial Answer to the Question:

1. **From LP to $\mathcal{K}$-LP:** Primal-only or dual-only algorithms extend "Word-by-Word" to all symmetric cone optimization problems (**Alizadeh-Schmieta** Handbook)
   They include
   - Karmarkar's original algorithms,
   - Ye's primal and dual and projective algorithms
   - Gonzaga's $\mathcal{O}(n^3)$ algorithm etc.

   **Reason:** If a proof *does not* rely on $L(\mathbf{x})L(\mathbf{s}) = L(\mathbf{s})L(\mathbf{x})$ it extends from LP (in LP $L(\mathbf{x}) = \mathrm{Diag}(\mathbf{x})$)

2. Primal-dual algorithms (**Kojima-Mizuno-Yoshise** and **Monteiro-Adler**) don't extend, *however*

3. **From SDP to $\mathcal{K}$-LP:** Those algorithms that need not assume existence of inducing associative algebra (**Faybusovich, Tsuchiya, Schmieta-Alizadeh**):
   (a) Commutative class of Monteiro-Zhang family (includes Nesterov-Todd, $\mathbf{xs}$ and $\mathbf{sx}$ methods (**HRVW/KSH/M**) and
   (b) $\mathbf{x}^{1/2}\mathbf{s}\mathbf{x}^{1/2}$ and variants (**Monteiro-Zanjacomo, Tseng**)
4. **From SDP to *representable* $\mathcal{K}$-LP:** Those algorithms that rely on existence of and inducing associative algebra (**Schmieta Thesis, Schmieta-Alizadeh**):
   The entire Monteiro-Zhang family and most importantly the $\mathbf{xs}+\mathbf{sx}$ method (**AHO**)
5. The last group does not extend to the Albert Algebra of $3 \times 3$ octonions, this algebra is *exceptional* (not representable)

## Generalization of Monteiro-Zhang family

Primal-dual Methods apply Newton's method to the following system:

$$\textbf{Primal}: \qquad \min\{\langle \mathbf{c}, \mathbf{x}\rangle_{\mathcal{E}} : \langle \mathbf{a}_i, \mathbf{x}\rangle_{\mathcal{E}} = b_i, \quad \mathbf{x} \in \mathcal{K}(\mathcal{E})\}$$

$$\textbf{Dual}: \qquad \max\{\mathbf{b}^{\mathsf{T}}\mathbf{y} : \mathcal{A}^{\mathsf{T}}\mathbf{y} + s = \mathbf{c}, \quad \mathbf{s} \in \mathcal{K}(\mathcal{E})\}$$

**Observation:** Polynomiality proofs of primal-dual methods in LP assume $L(\mathbf{x})L(\mathbf{s}) = L(\mathbf{s})L(\mathbf{x})$.

**Monteiro-Zhang Scaling:** Let $\mathbf{p} \succ \mathbf{0}$, and $\mathbf{Q_p} = 2\mathsf{L}^2(\mathbf{p}) - \mathsf{L}(\mathbf{p}^2)$, then

$$\mathbf{x} \to \widetilde{\mathbf{x}} = \mathbf{Q_p x} = \mathbf{pxp}$$

$$\mathbf{s} \to \utilde{\mathbf{s}} \stackrel{\text{def}}{=} \mathbf{Q}_{\mathbf{p}^{-1}} \mathbf{s} = \mathbf{p}^{-1}\mathbf{s}\mathbf{p}^{-1}$$

**Scaled Optimization problem:**

$$\underline{\textbf{Primal:}} \qquad \min \quad \{\langle \utilde{\mathbf{c}}, \widetilde{\mathbf{x}} \rangle_{\mathcal{E}} : \langle \utilde{\mathbf{a}_i}, \widetilde{\mathbf{x}} \rangle_{\mathcal{E}} = b_i, \widetilde{\mathbf{x}} \in \mathcal{K}(\mathcal{E})\}$$

$$\underline{\textbf{Dual:}} \qquad \max \quad \{\mathbf{b}^\mathsf{T}\mathbf{y} : \utilde{A}^\mathsf{T}\mathbf{y} + \utilde{\mathbf{s}} = \utilde{\mathbf{c}}, \utilde{\mathbf{s}} \in \mathcal{K}(\mathcal{E})\}$$

The Newton iteration is applied to

$$\utilde{A}\widetilde{\mathbf{x}} = \mathbf{b}$$

$$\utilde{A}^\mathsf{T}\mathbf{y} + \utilde{\mathbf{s}} = \utilde{\mathbf{c}}$$

$$\widetilde{\mathbf{x}} \circ \utilde{\mathbf{s}} = \mathsf{L}(\widetilde{\mathbf{x}})\mathsf{L}(\utilde{\mathbf{s}})\mathbf{e} = \mu\sigma\mathbf{e}$$

- If $\mathbf{p}$ is chosen such that $\mathsf{L}(\widetilde{\mathbf{x}})$ and $\mathsf{L}(\underset{\sim}{\mathbf{s}})$ commute then we get the Monteiro-Zhang *commutative* class
- Choosing $\mathbf{p} = \mathbf{s}^{1/2}$ yields $\mathbf{xs}$ method, $\mathbf{p} = \mathbf{x}^{-1/2}$ the $\mathbf{sx}$ method, Nesterov-Todd by $\mathbf{p} = \mathbf{w}^{1/2}$ where $\mathbf{w} = \mathbf{Q}_{x^{1/2}} (\mathbf{Q}_{x^{1/2}}\mathbf{s})^{-1/2}$
- The $\mathbf{xs} + \mathbf{sx}$ method is not in commutative class
- Optimization over scaled problem is entirely equivalent to optimization over the original problem

# Centrality Measure

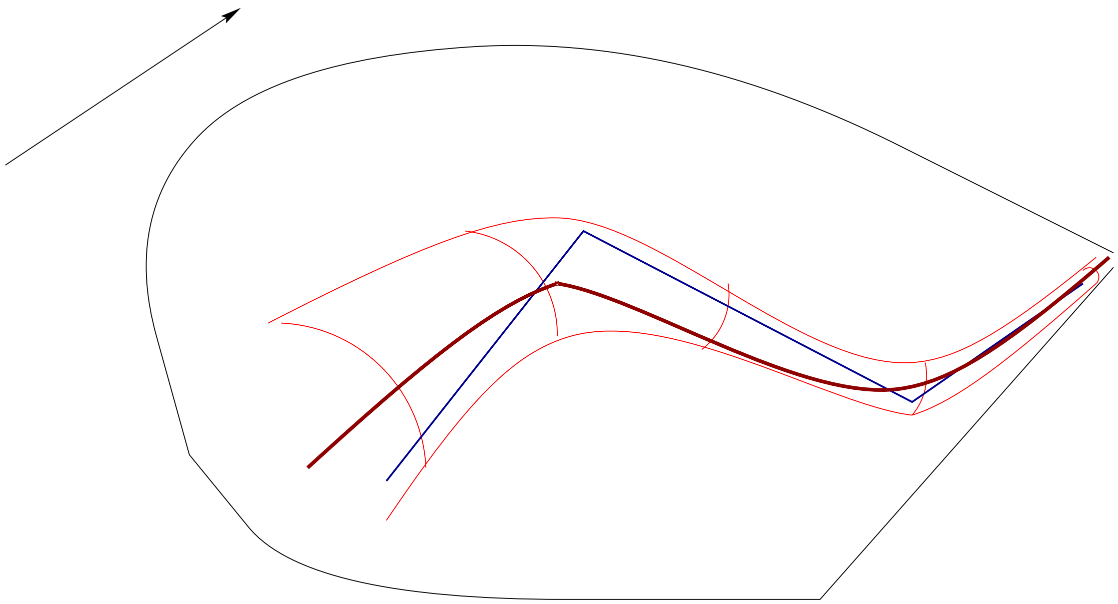
There are several centrality measures for $(\mathbf{x}, \mathbf{s}) \in \mathcal{K} \times \mathcal{K}$

$$
\begin{aligned}
d_F(\mathbf{x}, \mathbf{s}) &\stackrel{\text{def}}{=} \|\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s} - \mu \mathbf{e}\|_{F/\mathcal{E}} \\
&= \sqrt{\frac{1}{k} \sum_{i=1}^{n} (\lambda_i(\mathbf{xs}) - \mu)^2} \\
d_2(\mathbf{x}, \mathbf{s}) &\stackrel{\text{def}}{=} \|\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s} - \mu \mathbf{e}\|_2 = \max_{i=1,\dots,r} |\lambda_i(\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s}) - \mu| \\
&= \max\{\lambda_{\max}(\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s}) - \mu, \mu - \lambda_{\min}(\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s})\} \\
d_{-\infty}(\mathbf{x}, \mathbf{s}) &\stackrel{\text{def}}{=} \mu - \lambda_{\min}(\mathbf{Q}_{\mathbf{x}^{1/2}} \mathbf{s})
\end{aligned}
$$

For a given constant $\gamma \in (0, 1)$, we denote by $\mathcal{N}_\bullet(\gamma)$ the neighborhood of the central path

$$
\mathcal{N}_\bullet(\gamma) = \{(\mathbf{x}, \mathbf{s}, y) \mid d_\bullet(\mathbf{x}, \mathbf{s}) \le \gamma\mu, (\mathbf{x}, \mathbf{s}) \text{ interior feasible}\}
$$

- Let $r = \langle \mathbf{e}, \mathbf{e} \rangle_{\mathcal{E}}$. Then technical lemmas go through when $n$ is replaced by $r$.
- In particular one get $\mathcal{O}(\sqrt{r})$ iteration complexity for short-step path-following method



Here is the what we can prove

- Assume that $(\mathbf{x}, \mathbf{s}, y) \in \mathcal{N}_\bullet(\gamma)$ and let $(\Delta\mathbf{x}, \Delta\mathbf{s}, \Delta\mathbf{y})$ be the solution of the perturbed Newton system.
- By choosing $\gamma, \Gamma$, and $\delta$ appropriately, we can assure that

  $$(\mathbf{x} + \Delta\mathbf{x}, \mathbf{s} + \Delta\mathbf{s}, y + \Delta\mathbf{y}) \in \mathcal{N}_\bullet(\Gamma),$$

  and

  $$\langle \mathbf{x} + \Delta\mathbf{x}, \mathbf{s} + \Delta\mathbf{s} \rangle = (1 - \frac{\delta \kappa_{\mathbf{p}}}{\sqrt{r}}) \langle \mathbf{x}, \mathbf{s} \rangle$$

- **Scaling invariance** $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{N}_\bullet(\gamma)$ iff $(\widetilde{\mathbf{x}}, \mathbf{y}, \underset{\sim}{\mathbf{s}}) \in \mathcal{N}_\bullet(\gamma)$
- Therefore any proof for $\mathbf{xs} + \mathbf{sx}$ method extends to all scalings

- $\kappa_{\mathbf{p}}$ depends on our choice of scaling parameter $\mathbf{p}$
- $\kappa_{\mathbf{p}} \stackrel{\text{def}}{=} \operatorname{cond} \mathsf{L}(\widetilde{\mathbf{x}})\mathsf{L}(\utilde{\mathbf{s}})$ the condition number of $\mathsf{L}(\widetilde{\mathbf{x}})\mathsf{L}(\utilde{\mathbf{s}})$

EXAMPLE: LEMMA 3.6 (**Monteiro**): For every $\alpha \in \Re$, we have

$$\left\| \mathbf{x}^{-1/2} \left[ \mathbf{x}(\alpha)\mathbf{s}(\alpha) - \mu(\alpha)\mathbf{e} \right] \mathbf{x}^{1/2} \right\|_F \leq (1-\alpha)d(\mathbf{x},\mathbf{s}) + \alpha^2 \delta_x \delta_y + \alpha \left\| \mathbf{w_x} \right\|_F ,$$

where

$$\delta_x = \left\| \mathbf{x}^{-1/2} \Delta \mathbf{x} \mathbf{s}^{1/2} \right\|_F , \qquad \delta_s = \left\| \mathbf{s}^{-1/2} \Delta \mathbf{s} \mathbf{x}^{1/2} \right\|_F .$$

and

$$\mathbf{w_x} = \mathbf{x}^{-1/2} \left[ \Delta\mathbf{x}\mathbf{s} + \mathbf{x}\Delta\mathbf{s} + \mathbf{x}\mathbf{s} - \sigma\mu\mathbf{e} \right] \mathbf{x}^{1/2},$$

**Proof:**

$$\mathbf{x}^{-1/2}\left[\mathbf{x}(\alpha)\ \mathbf{s}(\alpha)-\mu(\alpha)\mathbf{e}\right]\mathbf{x}^{1/2}$$

$$= \mathbf{x}^{-1/2}\left[\mathbf{x}\mathbf{s}+\alpha(\mathbf{x}\Delta\mathbf{s}+\Delta\mathbf{x}\mathbf{s})+\alpha^2\Delta\mathbf{x}\Delta\mathbf{s}-\mu(\alpha)\mathbf{e}\right]\mathbf{x}^{1/2}$$

$$= (1-\alpha)(\mathbf{x}^{1/2}\mathbf{s}\mathbf{x}^{1/2}-\mu\mathbf{e})$$

$$+\ \alpha\mathbf{x}^{-1/2}\left[\mathbf{x}\mathbf{s}+\mathbf{x}\Delta\mathbf{s}+\Delta\mathbf{x}\mathbf{s}-\sigma\mu\mathbf{e}\right]\mathbf{x}^{1/2}$$

$$+\ \alpha^2\mathbf{x}^{-1/2}\Delta\mathbf{x}\Delta\mathbf{s}\mathbf{x}^{1/2}$$

$$\Rightarrow$$

$$\left\|\mathbf{x}^{-1/2}\left[\mathbf{x}(\alpha)\ \mathbf{s}(\alpha)-\mu(\alpha)\mathbf{e}\right]\mathbf{x}^{1/2}\right\|_F$$

$$= (1-\alpha)d(\mathbf{x},\mathbf{s})+\|w_x\|_F+\alpha^2\delta_x\delta_s$$

## Clifford algebras

**Definition 1** *1. The Clifford algebra $\mathcal{C}_n$ is a $2^n$-dimensional algebra which is generated by the set of elements $\{\mathbf{e}, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ having the property*

$$\mathbf{e}^2 = \mathbf{e}, \qquad \mathbf{e}\mathbf{e}_i = \mathbf{e}_i\mathbf{e} = \mathbf{e}_i, \qquad \mathbf{e}_i^2 = \mathbf{e},$$

$$\mathbf{e}_i\mathbf{e}_j = -\mathbf{e}_j\mathbf{e}_i \quad \textit{for } i \neq j,\ i, j \in \{1, \dots, n\}.$$

*2. Its basis consists of the $2^n$ products*

$$\mathbf{e}_I = \prod_{i \in I} \mathbf{e}_i \quad I \subseteq N = \{1, \dots, n\}$$

- Elements of $\mathcal{C}_n$ and their product:

$$x = \sum_{I \subseteq N} x_I e_I$$

$$y = \sum_{I \subseteq N} y_I e_I$$

$$xy = \left(\sum_{I \subseteq N} x_I e_I\right)\left(\sum_{I \subseteq N} y_I e_I\right) = \sum_{I \subseteq N}\left(\sum_{J} \operatorname{sgn}(I, J) x_J y_{I \Delta J}\right) e_I$$

- A base element $\mathbf{e}_I$ is called *simple* if $|I| \leq 1$.
- It is called *even* iff $\mathbf{e}_I^2 = \mathbf{e}$ and *odd* iff $\mathbf{e}_I^2 = -\mathbf{e}$
- $\mathbf{x}$ is called *symmetric* if $x_I \neq 0 \Rightarrow \mathbf{e}_I$ even
- Similarly $\mathbf{x}$ is called *simple* if $x_I \neq 0 \Rightarrow \mathbf{e}_I$ simple

Now We have our EAJ system:

- $\mathcal{C}$ is the associative algebra
- $\mathbf{x}'$ is defined as

$$x'_I \stackrel{\text{def}}{=} \begin{cases} -x_I & \text{if } \mathbf{e}_I \text{ odd} \\ x_I & \text{if } \mathbf{e}_I \text{ even} \end{cases}$$

- $S_{2^n}$ is the set of all symmetric elements in $\mathcal{C}_n$
- $\mathfrak{R}^{n+1}$ is the set of all simple elements in $S_{2^n}$
- The EAJ system: $(\mathcal{C}_n, S_{2^n}, \mathfrak{R}^{n+1}, \prime)$.