Computing cone-constrained singular values of matrices

Giovanni Barbarino [1] Nicolas Gillis [1] David Sossa [2]

Hidden structures in dynamical systems
optimization and machine learning

22 May 2025

[1]Université de Mons, Belgium

[2]Universidad de O'Higgins, Rancagua, Chile

Class of Computational Complexity

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} \quad u^\top A v \qquad P, Q \quad \begin{array}{l}\text{closed convex cones} \\ \text{finitely generated}\end{array}$$

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} \quad u^\top A v \qquad P, Q \quad \text{closed convex cones} \atop \text{finitely generated}$$

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1,\\ v \geq 0,\ \|v\| = 1,}} \quad u^\top A v$$

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v \qquad P, Q \quad \begin{array}{l}\text{closed convex cones} \\ \text{finitely generated}\end{array}$$

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1, \\ v \geq 0,\ \|v\| = 1,}} u^\top A v$$

Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top v$$

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v \qquad P, Q \quad \text{closed convex cones} \atop \text{finitely generated}$$

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1,\\ v \geq 0,\ \|v\| = 1,}} u^\top A v$$

Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top v$$

Singular Values

$$\min_{\|u\|=\|v\|=1} u^\top A v$$

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v \qquad P, Q$$

closed convex cones
finitely generated

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1,\\ v \geq 0,\ \|v\| = 1,}} u^\top A v$$

Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top v$$

Singular Values

$$\min_{\|u\|=\|v\|=1} u^\top A v$$

Polynomial Time
$O(mn^2)$ to compute
all Singular Values

Conical Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} \quad u^\top A v \qquad P, Q \quad \text{closed convex cones} \atop \text{finitely generated}$$

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1, \\ v \geq 0,\ \|v\| = 1,}} \quad u^\top A v$$

?

Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} \quad u^\top v$$

?

Singular Values

$$\min_{\|u\|=\|v\|=1} u^\top A v$$

Polynomial Time
$O(mn^2)$ to compute
all Singular Values

Reductions

Lemma (B., G., S. 2024)

Any matrix $A \in \mathbb{R}^{m\times n}$ of spectral norm 1 and $m \geq n$ can be decomposed as $A = U^\top V$ where $U \in \mathbb{R}^{(m+n)\times m}$, $V \in \mathbb{R}^{(m+n)\times n}$ have orthonormal columns

$$\min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top A v = \min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top U^\top V v = \min_{\substack{\widetilde{u}\in UP,\ \|\widetilde{u}\|=1,\\ \widetilde{v}\in VQ,\ \|\widetilde{v}\|=1}} \widetilde{u}^\top \widetilde{v}$$

The minimum conical singular value problem

reduces polynomially

to the maximum angle between cones problem

Theorem (G., Glineur 2013)

Let $B \in \{0,1\}^{m\times n}$ be the bi-adjacency matrix of a bipartite graph with $d \geq \max\{m,n\}$.

$$\min_{x,y\geq 0} \|B - d(1-B) - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

is solved by binary vectors x, y that identify the **Maximum Edge Biclique**

Theorem (Seeger, S. 2023)

$$\sigma_0 = (u^*)^\top A v^* = \min_{u,v\geq 0} u^\top A v \quad : \quad \|u\| = \|v\| = 1 \qquad \textbf{(Pareto SV)}$$

If A has at least one negative entry then $(x^*, y^*) = \sqrt{-\sigma_0}(u^*, v^*)$ is optimal for

$$\min_{x,y\geq 0} \|-A - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

Reductions

Lemma (B., G., S. 2024)

Any matrix $A \in \mathbb{R}^{m\times n}$ of spectral norm 1 and $m \geq n$ can be decomposed as $A = U^\top V$ where $U \in \mathbb{R}^{(m+n)\times m}$, $V \in \mathbb{R}^{(m+n)\times n}$ have orthonormal columns

$$\min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top A v = \min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top U^\top V v = \min_{\substack{\widetilde{u}\in UP,\ \|\widetilde{u}\|=1,\\ \widetilde{v}\in VQ,\ \|\widetilde{v}\|=1}} \widetilde{u}^\top \widetilde{v}$$

The minimum conical singular value problem

reduces polynomially

to the maximum angle between cones problem

Theorem (G., Glineur 2013)

Let $B \in \{0,1\}^{m\times n}$ be the bi-adjacency matrix of a bipartite graph with $d \geq \max\{m,n\}$.

$$\min_{x,y\geq 0} \|B - d(1-B) - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

is solved by binary vectors x, y that identify the **Maximum Edge Biclique**

Theorem (Seeger, S. 2023)

$$\sigma_0 = (u^*)^\top A v^* = \min_{u,v\geq 0} u^\top A v \quad : \quad \|u\| = \|v\| = 1 \qquad \textbf{(Pareto SV)}$$

If A has at least one negative entry then $(x^*, y^*) = \sqrt{-\sigma_0}(u^*, v^*)$ is optimal for

$$\min_{x,y\geq 0} \|-A - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

Reductions

Lemma (B., G., S. 2024)

Any matrix $A \in \mathbb{R}^{m\times n}$ of spectral norm 1 and $m \geq n$ can be decomposed as $A = U^\top V$ where $U \in \mathbb{R}^{(m+n)\times m}$, $V \in \mathbb{R}^{(m+n)\times n}$ have orthonormal columns

$$\min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top A v = \min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top U^\top V v = \min_{\substack{\tilde u\in UP,\ \|\tilde u\|=1,\\ \tilde v\in VQ,\ \|\tilde v\|=1}} \tilde u^\top \tilde v$$

The minimum conical singular value problem

reduces polynomially

to the maximum angle between cones problem

Theorem (G., Glineur 2013)

Let $B \in \{0,1\}^{m\times n}$ be the bi-adjacency matrix of a bipartite graph with $d \geq \max\{m,n\}$.

$$\min_{x,y\geq 0} \|B - d(1-B) - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

is solved by binary vectors x, y that identify the **Maximum Edge Biclique**

Theorem (Seeger, S. 2023)

$$\sigma_0 = (u^*)^\top A v^* = \min_{u,v\geq 0} u^\top A v \quad : \quad \|u\| = \|v\| = 1 \qquad \textbf{(Pareto SV)}$$

If A has at least one negative entry then $(x^*, y^*) = \sqrt{-\sigma_0}(u^*, v^*)$ is optimal for

$$\min_{x,y\geq 0} \|-A - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

Reductions

Lemma (B., G., S. 2024)

Any matrix $A \in \mathbb{R}^{m\times n}$ of spectral norm 1 and $m \geq n$ can be decomposed as $A = U^\top V$ where $U \in \mathbb{R}^{(m+n)\times m}$, $V \in \mathbb{R}^{(m+n)\times n}$ have orthonormal columns

$$\min_{\substack{u \in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top A v = \min_{\substack{u \in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top U^\top V v = \min_{\substack{\tilde{u} \in UP,\ \|\tilde{u}\|=1,\\ \tilde{v}\in VQ,\ \|\tilde{v}\|=1}} \tilde{u}^\top \tilde{v}$$

The minimum conical singular value problem

reduces polynomially

to the maximum angle between cones problem

Theorem (G., Glineur 2013)

Let $B \in \{0,1\}^{m\times n}$ be the bi-adjacency matrix of a bipartite graph with $d \geq \max\{m, n\}$.

$$\min_{x,y\geq 0} \|B - d(1-B) - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

is solved by binary vectors x, y that identify the **Maximum Edge Biclique**

Theorem (Seeger, S. 2023)

$$\sigma_0 = (u^*)^\top A v^* = \min_{u,v\geq 0} u^\top A v \quad : \quad \|u\| = \|v\| = 1 \qquad \textbf{(Pareto SV)}$$

If A has at least one negative entry then $(x^*, y^*) = \sqrt{-\sigma_0}(u^*, v^*)$ is optimal for

$$\min_{x,y\geq 0} \|-A - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

Reductions

Lemma (B., G., S. 2024)

Any matrix $A \in \mathbb{R}^{m\times n}$ of spectral norm 1 and $m \geq n$ can be decomposed as $A = U^\top V$ where $U \in \mathbb{R}^{(m+n)\times m}$, $V \in \mathbb{R}^{(m+n)\times n}$ have orthonormal columns

$$\min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top A v = \min_{\substack{u\in P,\ \|u\|=1,\\ v\in Q,\ \|v\|=1}} u^\top U^\top V v = \min_{\substack{\widetilde{u}\in UP,\ \|\widetilde{u}\|=1,\\ \widetilde{v}\in VQ,\ \|\widetilde{v}\|=1}} \widetilde{u}^\top \widetilde{v}$$

The minimum conical singular value problem

reduces polynomially

to the maximum angle between cones problem

Theorem (G., Glineur 2013)

Let $B \in \{0,1\}^{m\times n}$ be the bi-adjacency matrix of a bipartite graph with $d \geq \max\{m,n\}$.

$$\min_{x,y\geq 0} \|B - d(1-B) - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

is solved by binary vectors x, y that identify the **Maximum Edge Biclique**

Theorem (Seeger, S. 2023)

$$\sigma_0 = (u^*)^\top A v^* = \min_{u,v\geq 0} u^\top A v \quad : \quad \|u\| = \|v\| = 1 \qquad \textbf{(Pareto SV)}$$

If A has at least one negative entry then $(x^*, y^*) = \sqrt{-\sigma_0}(u^*, v^*)$ is optimal for

$$\min_{x,y\geq 0} \|-A - xy^\top\|_F^2 \qquad \textbf{(Nonnegative Rank 1)}$$

Everything is Hard

Theorem (Peeters 2003)

The Maximal Edge Biclique problem is NP-hard

Maximal Edge Biclique

Maximum Number of Edges in a Bipartite Connected Subgraph

NP-hard

Everything is Hard

Theorem (Peeters 2003)

The Maximal Edge Biclique problem is NP-hard

Maximal Edge Biclique

Maximum Number of Edges in a Bipartite Connected Subgraph

NP-hard

Nonnegative Rank 1

$$\min_{x,y\geq 0} \|M - xy^\top\|$$

NP-hard

Everything is Hard

Theorem (Peeters 2003)

The Maximal Edge Biclique problem is NP-hard

Maximal Edge Biclique

Maximum Number of Edges in a Bipartite Connected Subgraph

NP-hard

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1, \\ v \geq 0,\ \|v\| = 1,}} u^\top A v$$

NP-hard

Nonnegative Rank 1

$$\min_{x,y \geq 0} \|M - xy^\top\|$$

NP-hard

Everything is Hard

Theorem (Peeters 2003)

The Maximal Edge Biclique problem is NP-hard

Maximal Edge Biclique

Maximum Number of Edges in a Bipartite Connected Subgraph

NP-hard

Conic Singular Values

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} \quad u^\top A v$$

NP-hard

Pareto Singular Values

$$\min_{\substack{u \geq 0,\ \|u\| = 1,\\ v \geq 0,\ \|v\| = 1,}} \quad u^\top A v$$

NP-hard

Nonnegative Rank 1

$$\min_{x,y \geq 0} \|M - xy^\top\|$$

NP-hard

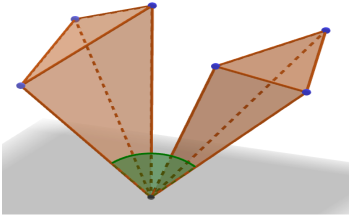
Conic Angles

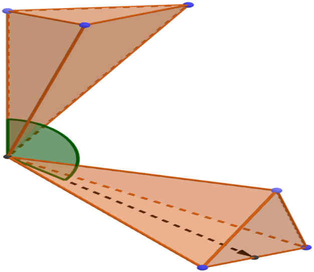
$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} \quad u^\top v$$

NP-hard

Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top v \qquad P, Q \subseteq \mathbb{R}^n \text{ non trivial (polyhedral) cones}$$



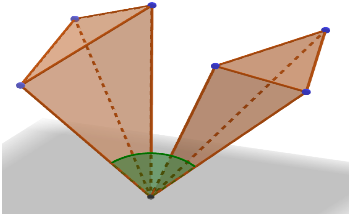


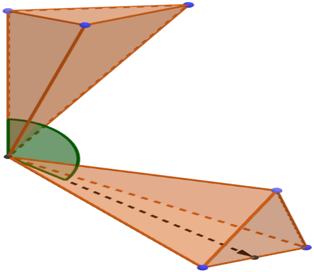
Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top v \qquad P, Q \subseteq \mathbb{R}^n \text{ non trivial (polyhedral) cones}$$

"Simple" Case:

$$\min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top v \geq 0 \implies u, v \text{ are vertices of } P, Q$$





Conic Angles

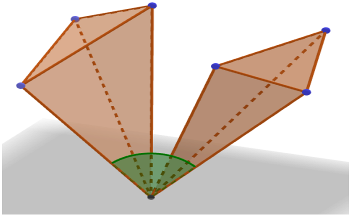
$$\min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} \quad u^\top v \qquad P, Q \subseteq \mathbb{R}^n \text{ non trivial (polyhedral) cones}$$

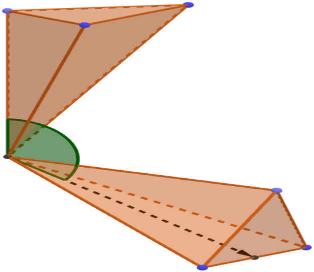
"Simple" Case:

If one of u, v in the antipodal pair is a vertex then the problem is **Polynomial** in n and the number of generators of P, Q

$$\min_{v \in Q,\ \|v\|=1} u^\top v = -\max_{v \in -Q,\ \|v\|=1} u^\top v \implies v = -\frac{Proj(u, -Q)}{\|Proj(u, -Q)\|}$$

$$Proj(u, -Q) \equiv \min_{y \geq 0} \|u - (-H)y\| \qquad \langle H \rangle = Q, \text{ NNLS, convex}$$





Conic Angles

$$\min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top v \qquad P, Q \subseteq \mathbb{R}^n \text{ non trivial (polyhedral) cones}$$

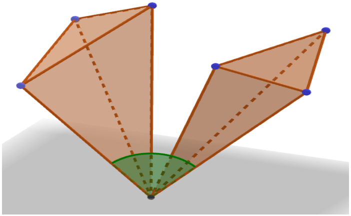
if $(u^*)^\top v^* < 0$, when is it that one among u, v is a vertex?

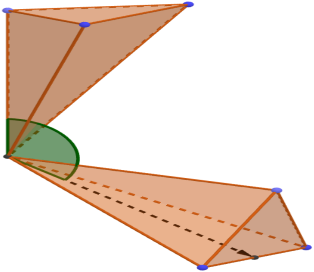
Theorem (B., G., S. 2024)

Let (u, v) be a stationary point and let $u \in int(F_u)$, $v \in int(F_v)$ where F_u, F_v are faces of P, Q. If $\dim(F_u) + \dim(F_v) > n$ and $v \neq \pm u$, then (u, v) is a saddle point

Corollary (B., G., S. 2024)

If (u, v) is a local minimum in dimension $n \leq 3$ with $u \neq -v$, then at least one among u, v is a vertex





Algorithms

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\|=1,\\ y \geq 0,\ \|Hy\|=1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\top A \overline{H}\overline{y} - \lambda^* \overline{G}^\top \overline{G}\overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\top A^\top \overline{G}\overline{x} - \lambda^* \overline{H}^\top \overline{H}\overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A\, v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\| = 1,\\ y \geq 0,\ \|Hy\| = 1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\top A \overline{H}\overline{y} - \lambda^* \overline{G}^\top \overline{G}\overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\top A^\top \overline{G}\overline{x} - \lambda^* \overline{H}^\top \overline{H}\overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A\, v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\|=1,\\ y \geq 0,\ \|Hy\|=1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\top A \overline{H}\overline{y} - \lambda^* \overline{G}^\top \overline{G}\overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\top A^\top \overline{G}\overline{x} - \lambda^* \overline{H}^\top \overline{H}\overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\| = 1,\\ y \geq 0,\ \|Hy\| = 1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\top A \overline{H}\overline{y} - \lambda^* \overline{G}^\top \overline{G}\overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\top A^\top \overline{G}\overline{x} - \lambda^* \overline{H}^\top \overline{H}\overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\| = 1,\\ y \geq 0,\ \|Hy\| = 1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\dagger A \overline{H}\overline{y} - \lambda^* \overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\dagger A^\top \overline{G}\overline{x} - \lambda^* \overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A\, v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\|=1,\\ y \geq 0,\ \|Hy\|=1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

KKT Conditions:

$$\begin{cases} 0 \leq x^* \perp G^\top A H y^* - \lambda^* G^\top G x^* \geq 0 \\ 0 \leq y^* \perp H^\top A^\top G x^* - \lambda^* H^\top H y^* \geq 0 \\ \|Gx^*\| = \|Hy^*\| = 1 \end{cases} \implies \begin{cases} 0 < \overline{x}, \quad \overline{G}^\dagger A \overline{H}\,\overline{y} - \lambda^* \overline{x} = 0 \\ 0 < \overline{y}, \quad \overline{H}^\dagger A^\top \overline{G}\,\overline{x} - \lambda^* \overline{y} = 0 \\ \overline{x} := x^*_{\mathcal{I}}, \overline{y} := y^*_{\mathcal{J}}, \overline{G} := G_{:,\mathcal{I}}, \overline{H} := H_{:,\mathcal{J}} \end{cases}$$

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\,\overline{x}, \overline{H}\,\overline{y})$ and $\lambda^* = (u^*)^\top A\, v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1,\\ v \in Q,\ \|v\| = 1,}} u^\top A v = \min_{\substack{x \geq 0,\ \|Gx\| = 1,\\ y \geq 0,\ \|Hy\| = 1,}} x^\top G^\top A H y$$

Idea: If we know the sets $\mathcal{I}$, $\mathcal{J}$ of indices for which $x_i^*, y_j^* > 0$, called **Active Sets**, then a direct gradient computation solves the problem

The Active Set algorithm cycles over all subsets of indices $\mathcal{I}, \mathcal{J}$ and tests if the least eigenvalue of M has a nonnegative eigenvector, giving us upper bounds on λ^*, and the exact solution when $\mathcal{I}, \mathcal{J}$ coincide with the active sets of (x^*, y^*)

Optimizations: $2 < |\mathcal{I}| + |\mathcal{J}| \leq m + n - \mathrm{Null}(A^\top A - \|A\|^2 I)$ and $\overline{G}$, $\overline{H}$ must be full rank

Theorem (B., G., S. 2024)

For the optimal solution $(u^*, v^*) = (Gx^*, Hy^*) = (\overline{G}\overline{x}, \overline{H}\overline{y})$ and $\lambda^* = (u^*)^\top A\, v^*$

$$M^* := \begin{pmatrix} 0 & \overline{H}^\dagger A^\top \overline{G} \\ \overline{G}^\dagger A \overline{H} & 0 \end{pmatrix} \implies M^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix} = \lambda^* \begin{pmatrix} \overline{y} \\ \overline{x} \end{pmatrix}$$

where λ^* is the least eigenvalue of M^* (from 2° order KKT)

For any $\mathcal{I}, \mathcal{J}$, if λ is the least eigenvalue λ of M and admits a nonnegative eigenvector then $\lambda \geq \lambda^*$

Brute Force Active Set

Input: $A \in \mathbb{R}^{m\times n}$, $G \in \mathbb{R}^{m\times p}$, $H \in \mathbb{R}^{n\times q}$, $P = \langle G\rangle$, $Q = \langle H\rangle$

Output: $\lambda = \min u^\top A v$ such that $\|u\| = \|v\| = 1$, $u \in P$, $v \in Q$

```
λ = g_i^T A h_j = min_{k,ℓ} (G^T A H)_{k,ℓ}, u = g_i, v = h_j, r = Null(A^T A − ‖A‖² I_n)
𝒥 := {(ℐ, 𝒥) : 2 < |ℐ| + |𝒥| ≤ m + n − r, Ḡ := G_{:,ℐ} and H̄ := H_{:,𝒥} full column rank}
for (ℐ, 𝒥) ∈ 𝒥, do
    A_x = Ḡ^† A^T H̄, A_y = H̄^† A Ḡ
    A_λ = A_y A_x, Ã_λ = A_x (or A_λ = A_x A_y, Ã_λ = A_y if |ℐ| > |𝒥|)
    if ρ(A_λ) ≤ λ² then Skip to the next (ℐ, 𝒥) ∈ 𝒥
    U right eigenspace of ρ(A_λ) in A_λ, μ = −√ρ(A_λ), W = ( Ã_λ U/μ ; U )
    Compute the reduced QR of W = VR
    if (VV^T − I)z = 0, z ≥ 0, e^T z = 1 admits a solution then
        λ = μ, z = [y^T x^T]^T (or z = [x^T y^T]^T if |ℐ| > |𝒥|)
        u = Ḡx/‖Ḡx‖, v = H̄y/‖H̄y‖
    end if
end for
```

1. $\lambda = g_i^\top A h_j = \min_{k,\ell} (G^\top A H)_{k,\ell}$, $u = g_i$, $v = h_j$, $r = \mathrm{Null}(A^\top A - \|A\|^2 I_n)$
2. $\mathscr{I} := \{(\mathcal{I}, \mathcal{J}) : 2 < |\mathcal{I}| + |\mathcal{J}| \leq m + n - r,\ \overline{G} := G_{:,\mathcal{I}}$ and $\overline{H} := H_{:,\mathcal{J}}$ full column rank$\}$
3. **for** $(\mathcal{I}, \mathcal{J}) \in \mathscr{I}$, **do**
    - $A_x = \overline{G}^\dagger A^\top \overline{H}$, $A_y = \overline{H}^\dagger A \overline{G}$
    - $A_\lambda = A_y A_x$, $\widetilde{A}_\lambda = A_x$ (or $A_\lambda = A_x A_y$, $\widetilde{A}_\lambda = A_y$ if $|\mathcal{I}| > |\mathcal{J}|$)
    - **if** $\rho(A_\lambda) \leq \lambda^2$ **then** Skip to the next $(\mathcal{I}, \mathcal{J}) \in \mathscr{I}$
    - U right eigenspace of $\rho(A_\lambda)$ in A_λ, $\mu = -\sqrt{\rho(A_\lambda)}$, $W = \begin{pmatrix} \widetilde{A}_\lambda U/\mu \\ U \end{pmatrix}$
    - Compute the reduced QR of $W = VR$
    - **if** $(VV^\top - I)z = 0$, $z \geq 0$, $e^\top z = 1$ admits a solution **then**
        - $\lambda = \mu$, $z = [y^\top\ x^\top]^\top$ (or $z = [x^\top\ y^\top]^\top$if $|\mathcal{I}| > |\mathcal{J}|$)
        - $u = \overline{G}x/\|\overline{G}x\|$, $v = \overline{H}y/\|\overline{H}y\|$
    - **end if**
4. **end for**

Alternating projection with extrapolation

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v$$

Idea: We have seen that if we know u^* or v^*, then finding the other is equivalent to solve an easy convex problem

Alternate Projection: starting from an initial feasible point (u_0, v_0) and $k = 0$

- $u_{k+1} = \arg\min_{x \in P} x^\top A v_k$ such that $\|x\|_2 = 1$
- $v_{k+1} = \arg\min_{y \in Q} u_{k+1}^\top A y$ such that $\|y\|_2 = 1$
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step** after each update

- $u_{k+1} = u_{k+1} + \beta(u_{k+1} - u_k)$
- $v_{k+1} = v_{k+1} + \beta(v_{k+1} - v_k)$
- If the objective increases then we decrease β and go back to (u_k, v_k), otherwise we increase β

The method **converges** to a stationary point, that **may not be optimal**

Alternating projection with extrapolation

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v$$

Idea: We have seen that if we know u^* or v^*, then finding the other is equivalent to solve an easy convex problem

Alternate Projection: starting from an initial feasible point (u_0, v_0) and $k = 0$

- $u_{k+1} \;=\; \arg\min_{x \in P} x^\top A v_k$ such that $\|x\|_2 = 1$
- $v_{k+1} \;=\; \arg\min_{y \in Q} u_{k+1}^\top A y$ such that $\|y\|_2 = 1$
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step** after each update

- $u_{k+1} = u_{k+1} + \beta(u_{k+1} - u_k)$
- $v_{k+1} = v_{k+1} + \beta(v_{k+1} - v_k)$
- If the objective increases then we decrease β and go back to (u_k, v_k), otherwise we increase β

The method **converges** to a stationary point, that **may not be optimal**

Alternating projection with extrapolation

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v$$

Idea: We have seen that if we know u^* or v^*, then finding the other is equivalent to solve an easy convex problem

Alternate Projection: starting from an initial feasible point (u_0, v_0) and $k = 0$

- $u_{k+1} = \arg\min_{x \in P} x^\top A v_k$ such that $\|x\|_2 = 1$
- $v_{k+1} = \arg\min_{y \in Q} u_{k+1}^\top A y$ such that $\|y\|_2 = 1$
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step** after each update

- $u_{k+1} = u_{k+1} + \beta(u_{k+1} - u_k)$
- $v_{k+1} = v_{k+1} + \beta(v_{k+1} - v_k)$
- If the objective increases then we decrease β and go back to (u_k, v_k), otherwise we increase β

The method **converges** to a stationary point, that **may not be optimal**

Alternating projection with extrapolation

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v$$

Idea: We have seen that if we know u^* or v^*, then finding the other is equivalent to solve an easy convex problem

Alternate Projection: starting from an initial feasible point (u_0, v_0) and $k = 0$

- $u_{k+1} \;=\; \arg\min_{x \in P} x^\top A v_k$ such that $\|x\|_2 = 1$
- $v_{k+1} \;=\; \arg\min_{y \in Q} u_{k+1}^\top A y$ such that $\|y\|_2 = 1$
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step** after each update

- $u_{k+1} = u_{k+1} + \beta(u_{k+1} - u_k)$
- $v_{k+1} = v_{k+1} + \beta(v_{k+1} - v_k)$
- If the objective increases then we decrease β and go back to (u_k, v_k), otherwise we increase β

The method **converges** to a stationary point, that **may not be optimal**

Alternating projection with extrapolation

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v$$

Idea: We have seen that if we know u^* or v^*, then finding the other is equivalent to solve an easy convex problem

Alternate Projection: starting from an initial feasible point (u_0, v_0) and $k = 0$

- $u_{k+1} \ = \ \arg\min_{x \in P} x^\top A v_k$ such that $\|x\|_2 = 1$
- $v_{k+1} \ = \ \arg\min_{y \in Q} u_{k+1}^\top A y$ such that $\|y\|_2 = 1$
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step** after each update

- $u_{k+1} = u_{k+1} + \beta(u_{k+1} - u_k)$
- $v_{k+1} = v_{k+1} + \beta(v_{k+1} - v_k)$
- If the objective increases then we decrease β and go back to (u_k, v_k), otherwise we increase β

The method **converges** to a stationary point, that **may not be optimal**

Alternating projection with extrapolation

Input: $A \in \mathbb{R}^{m \times n}$, cones $P \subseteq \mathbb{R}^m$ and $Q \subseteq \mathbb{R}^n$

Output: An approximate solution to $\min_{u \in P, v \in Q} u^\top A v$ such that $\|u\|_2 = \|v\|_2 = 1$.

$u = 0,\ v = 0,\ v_e = v_0,\ k = 1.$

while $k \leq K$ and $(\|u - u_p\|_2 \geq \delta$ or $\|v - v_p\|_2 \geq \delta)$ **do**

  $u_p = u$. *% Keep previous iterate in memory*

  $u = \arg\min_{x \in P} x^\top A v_e$ such that $\|x\|_2 = 1$.

  $u_e = u + \beta(u - u_p)$. *% Extrapolated point*

  $v_p = v$. *% Keep previous iterate in memory*

  $v = \arg\min_{y \in Q} u_e^\top A y$ such that $\|y\|_2 = 1$.

  $v_e = v + \beta(v - v_p)$. *% Extrapolated point*

  $e_k \leftarrow u^\top A v$.

  if $k \geq 2$ and $e_k > e_{k-1}$ **then**

    $u = u_p,\ v = v_p,\ \beta = \frac{\beta}{\eta}$.

  else

    $\beta \leftarrow \min(1, \gamma\beta)$.

  end if

  $k \leftarrow k + 1$.

end while

Sequential Regularized Partial Linearization

$$\lambda^* = \min_{\substack{u \in P,\ \|u\| = 1, \\ v \in Q,\ \|v\| = 1,}} u^\top A v = \min_{\substack{u \in P,\ u \neq 0, \\ v \in Q,\ v \neq 0,}} \frac{u^\top A v}{\|u\|\|v\|} = \min_{\substack{e^\top x = 1,\ x \geq 0, \\ e^\top y = 1,\ y \geq 0,}} \frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|}$$

Idea: If the minimum of $f_\delta(x, y) := x^\top G^\top A H y - \delta \|Gx\|\|Hy\|$ over $(x, y) \in \Delta_p \times \Delta_q$ is $\mu < 0$ then we get a decrease in the objective function

$$\frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|} = \delta + \frac{\mu}{\|Gx\|\|Hy\|} < \delta$$

Partial Linearization: starting from an initial feasible point (x_0, y_0) and $k = 0$,

- $\delta = \frac{x_k^\top G^\top A H y_k}{\|Gx_k\|\|Hy_k\|}$
- Linearize wrt x the function $f_\delta(x, y_k)$, penalize it with $\|x - x_k\|^2$ and minimize it to obtain x_{k+1}
- Linearize wrt y the function $f_\delta(x_{k+1}, y)$, penalize it with $\|y - y_k\|^2$ and minimize it to obtain y_{k+1}
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step**

Sequential Regularized Partial Linearization

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{u \in P,\ u \neq 0,\\ v \in Q,\ v \neq 0,}} \frac{u^\top A v}{\|u\|\|v\|} = \min_{\substack{e^\top x = 1,\ x \geq 0,\\ e^\top y = 1,\ y \geq 0,}} \frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|}$$

Idea: If the minimum of $f_\delta(x,y) := x^\top G^\top A H y - \delta \|Gx\|\|Hy\|$ over $(x,y) \in \Delta_p \times \Delta_q$ is $\mu < 0$ then we get a decrease in the objective function

$$\frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|} = \delta + \frac{\mu}{\|Gx\|\|Hy\|} < \delta$$

Partial Linearization: starting from an initial feasible point (x_0, y_0) and $k = 0$,

- $\delta = \frac{x_k^\top G^\top A H y_k}{\|Gx_k\|\|Hy_k\|}$
- Linearize wrt x the function $f_\delta(x, y_k)$, penalize it with $\|x - x_k\|^2$ and minimize it to obtain x_{k+1}
- Linearize wrt y the function $f_\delta(x_{k+1}, y)$, penalize it with $\|y - y_k\|^2$ and minimize it to obtain y_{k+1}
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step**

Sequential Regularized Partial Linearization

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{u \in P,\ u \neq 0,\\ v \in Q,\ v \neq 0,}} \frac{u^\top A v}{\|u\|\|v\|} = \min_{\substack{e^\top x = 1,\ x \geq 0,\\ e^\top y = 1,\ y \geq 0,}} \frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|}$$

Idea: If the minimum of $f_\delta(x,y) := x^\top G^\top A H y - \delta \|Gx\|\|Hy\|$ over $(x,y) \in \Delta_p \times \Delta_q$ is $\mu < 0$ then we get a decrease in the objective function

$$\frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|} = \delta + \frac{\mu}{\|Gx\|\|Hy\|} < \delta$$

Partial Linearization: starting from an initial feasible point (x_0, y_0) and $k = 0$,

- $\delta = \frac{x_k^\top G^\top A H y_k}{\|Gx_k\|\|Hy_k\|}$
- Linearize wrt x the function $f_\delta(x, y_k)$, penalize it with $\|x - x_k\|^2$ and minimize it to obtain x_{k+1}
- Linearize wrt y the function $f_\delta(x_{k+1}, y)$, penalize it with $\|y - y_k\|^2$ and minimize it to obtain y_{k+1}
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step**

Sequential Regularized Partial Linearization

$$\lambda^* = \min_{\substack{u \in P,\ \|u\|=1,\\ v \in Q,\ \|v\|=1,}} u^\top A v = \min_{\substack{u \in P,\ u \neq 0,\\ v \in Q,\ v \neq 0,}} \frac{u^\top A v}{\|u\|\|v\|} = \min_{\substack{e^\top x = 1,\ x \geq 0,\\ e^\top y = 1,\ y \geq 0,}} \frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|}$$

Idea: If the minimum of $f_\delta(x, y) := x^\top G^\top A H y - \delta \|Gx\|\|Hy\|$ over $(x, y) \in \Delta_p \times \Delta_q$ is $\mu < 0$ then we get a decrease in the objective function

$$\frac{x^\top G^\top A H y}{\|Gx\|\|Hy\|} = \delta + \frac{\mu}{\|Gx\|\|Hy\|} < \delta$$

Partial Linearization: starting from an initial feasible point (x_0, y_0) and $k = 0$,

- $\delta = \frac{x_k^\top G^\top A H y_k}{\|Gx_k\|\|Hy_k\|}$
- Linearize wrt x the function $f_\delta(x, y_k)$, penalize it with $\|x - x_k\|^2$ and minimize it to obtain x_{k+1}
- Linearize wrt y the function $f_\delta(x_{k+1}, y)$, penalize it with $\|y - y_k\|^2$ and minimize it to obtain y_{k+1}
- $k = k + 1$

To accelerate the convergence, we add an **Extrapolation step**

Input: $A \in \mathbb{R}^{m \times n}$, cones $P \subseteq \mathbb{R}^m$ and $Q \subseteq \mathbb{R}^n$

Output: An approximate solution to $\min_{u \in P, v \in Q} \langle u, Av \rangle$ such that $\|u\| = \|v\| = 1$

1: Set

$$\delta_k := \frac{\langle Gx^k, AHy^k \rangle}{\|Gx^k\| \|Hy^k\|}$$

2: Let $L_1^k(x) := \left\langle Gx, AHy^k - \delta_k \|Gx^k\|^{-1} \|Hy^k\| Gx^k \right\rangle$
Compute a solution $\tilde{x}^k$ to the convex program

$$\min L_1^k(x) + \tfrac{\mu_1}{2} \|x - x^k\|^2 \quad \text{such that } x \in \Delta_p$$

3: Let $L_2^k(y) := \left\langle Hy, A^\top Gx^k - \delta_k \|Gx^k\| \|Hy^k\|^{-1} Hy^k \right\rangle$
Compute a solution $\tilde{y}^k$ to the convex program

$$\min L_2^k(y) + \tfrac{\mu_2}{2} \|y - y^k\|^2 \quad \text{such that } y \in \Delta_q$$

4: Let $d_1^k := \tilde{x}^k - x^k$ and $d_2^k := \tilde{y}^k - y^k$

5: If $(|L_1^k(d_1^k)| < \delta$ and $|L_2^k(d_2^k)| < \delta)$ or $k \geq K$ terminate
Otherwise, let $t_k := \beta \rho^{\ell_k}$, where ℓ_k is the smallest nonnegative integer ℓ such that

$$\Phi(x^k + t^k d_1^k, y^k + t^k d_2^k) \leq \Phi(x^k, y^k) + \alpha t_k \frac{L_1^k(d_1^k) + L_2^k(d_2^k)}{\|Gx^k\| \|Hy^k\|}$$

Set $\left(x^{k+1}, y^{k+1}\right) := (x^k, y^k) + t_k(d_1^k, d_2^k)$ and $k = k + 1$. Go to step 1

Experiments

An Example: Schur Cone

We test and compare the following algorithms on several problems:

- Brute Force Active Set
- Alternating projection with extrapolation
- Sequential Regularized Partial Linearization
- Gurobi (exact nonconvex quadratic solver based on McCormick relaxation)

The **Schur Cone** is generated by the matrix

$$H = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & -1 \end{pmatrix} \in \mathbb{R}^{n \times n-1} \qquad \langle H \rangle \subseteq e^{\perp}$$

One can prove that the maximum angle between the Schur cone Q and $\mathbb{R}^n_+$ is achieved by

$$y = e_n \in P \quad x = (a\, a\, \dots\, a\, b) \in Q \quad a = \sqrt{\frac{1}{n(n-1)}} \quad b = -\sqrt{1 - \frac{1}{n}} = x^\top y$$

An Example: Schur Cone

We test and compare the following algorithms on several problems:

- Brute Force Active Set
- Alternating projection with extrapolation
- Sequential Regularized Partial Linearization
- Gurobi (exact nonconvex quadratic solver based on McCormick relaxation)

The **Schur Cone** is generated by the matrix

$$H = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & -1 \end{pmatrix} \in \mathbb{R}^{n\times n-1} \qquad \langle H \rangle \subseteq e^{\perp}$$

One can prove that the maximum angle between the Schur cone Q and $\mathbb{R}^n_+$ is achieved by

$$y = e_n \in P \quad x = (a\, a \dots a\, b) \in Q \quad a = \sqrt{\frac{1}{n(n-1)}} \quad b = -\sqrt{1 - \frac{1}{n}} = x^\top y$$

Schur Cone and Positive Orthant

Table 1: Numerical comparison for Gur and BFAS for different dimensions for the problem of finding the maximum angle between the Schur cone and $\mathbb{R}^n_+$. The table reports the optimal objective functions values found in the timelimit (60 seconds) and the actual elapsed time. We also report the exact value for each problem.

n	5	10	20	50
exact	0.852416π	0.897584π	0.928217π	0.954833π
Gur	0.852416π	0.897584π	0.928218π	0.954833π
	0.1134 s	0.2016 s	20.1493 s	60^* s
BFAS	0.852416π	0.897584π	0.750000π	0.750000π
	0.3310 s	48.3153 s	60^* s	60^* s

n	100	200	500
exact	0.968116π	0.977473π	0.985760π
Gur	0.968116π	0.977473π	0.985756π
	60^* s	60^* s	60^* s
BFAS	0.750000π	0.750000π	0.750000π
	60^* s	60^* s	60^* s

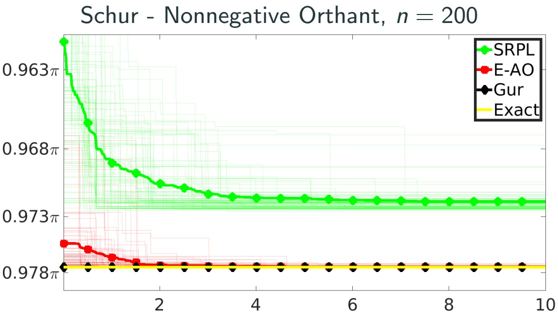
Schur Cone and Positive Orthant

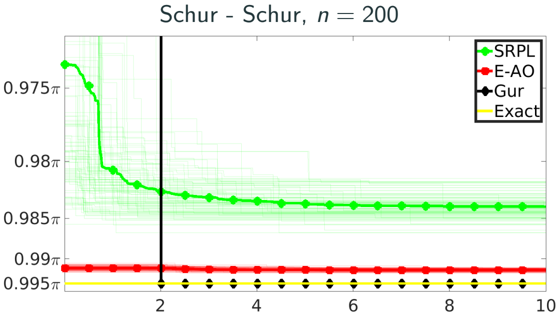
Table 1: Numerical comparison for Gurobi and BFAS for different dimensions for the problem of finding the maximum angle between the Schur cone and itself. The table reports the optimal objective functions values found in the timelimit (60 seconds) and the actual elapsed time. We also report the exact value for each problem.

n	5	10	20	50
exact	0.800000π	0.900000π	0.950000π	0.980000π
Gur	0.800001π	0.900000π	0.950000π	0.980000π
	0.2508 s	60^* s	60^* s	60^* s
BFAS	0.800000π	0.900000π	0.859157π	0.804087π
	0.3856 s	60^* s	60^* s	60^* s

n	100	200	500
exact	0.990000π	0.995000π	0.998000π
Gur	0.936315π	0.994996π	0.998011π
	60^* s	60^* s	60^* s
BFAS	0.750000π	0.750000π	0.750000π
	60^* s	60^* s	60^* s

Schur Cone and Positive Orthant





Maximum Edge Biclique Problem

Recall that solving the Pareto singular value problem is equivalent to solve the maximum edge biclique problem.

Here we thus test all four algorithms on four bipartite graphs taken from a benchmark dataset[1]. All graphs have been randomly generated with a fixed edge density, and then a biclique has been added to them. In particular,

- the first graph is a 100×100 graph with density 0.2 and planted biclique of size $50 \times 50 = 2500$,
- the second graph is a 300×300 graph with density 0.3 and planted biclique of size $2 \times 55 = 110$,
- the third graph is a 100×100 graph with density 0.71 and planted biclique of size $80 \times 80 = 6400$,
- the fourth graph is a 10000×300 graph with density 0.03 and planted biclique of size $22 \times 2 = 44$.

[1] Shaham, E.: maximum biclique benchmark. https://github.com/shahamer/maximum-biclique-benchmark (2019)

Maximum Edge Biclique Problem

Table 1: Numerical comparison for Gurobi, BFAS, E-AO and SRPL for the problem of finding the maximum edge biclique in four different bipartite graphs. The table reports the maximum edgee biclique found in the timelimit (10 seconds) for Gurobi and BFAS. The reported number for E-AO and SRPL are instead the average value found at 10 seconds for 100 runs, and in parentheses the best value found throughout all 100 runs when it differs from the average one. Gurobi cannot be executed on the last graph due to its excessive size.

n	100×100	300×300	100×100	10000×300
Gur	2500	0	310	NA
BFAS	3	2	2	2
E-AO	66	114	87	12
SRPL	2500	114	6400	46(358)

Maximum Angle between PSD and Nonnegative Symmetric Matrices

Given $\langle A, B \rangle = Tr(A^\top B)$ an open question is the maximum angle between the cone of PSD matrices $\mathcal{P}^n$ and the cone of nonnegative symmetric matrices $\mathcal{N}^n$ for $n \geq 5$

Maximum Angle between PSD and Nonnegative Symmetric Matrices

Given $\langle A, B \rangle = Tr(A^\top B)$ an open question is the maximum angle between the cone of PSD matrices $\mathcal{P}^n$ and the cone of nonnegative symmetric matrices $\mathcal{N}^n$ for $n \geq 5$

$$n = 2, 3, 4 \implies \gamma_n = \frac{3}{4}\pi \qquad \lim_{n \to \infty} \gamma_n \uparrow \pi$$

All antipodal couples (and the best known for $n = 5$) are circulant matrices

Maximum Angle between PSD and Nonnegative Symmetric Matrices

Given $\langle A, B\rangle = Tr(A^\top B)$ an open question is the maximum angle between the cone of PSD matrices $\mathcal{P}^n$ and the cone of nonnegative symmetric matrices $\mathcal{N}^n$ for $n \geq 5$

$$n = 2, 3, 4 \implies \gamma_n = \frac{3}{4}\pi \qquad \lim_{n\to\infty} \gamma_n \uparrow \pi$$

All antipodal couples (and the best known for $n = 5$) are circulant matrices

If $\mathcal{SC}^n$ is the algebra of circulant matrices, $\mathcal{SC}^n \cap \mathcal{P}^n$ and $\mathcal{SC}^n \cap \mathcal{N}^n$ are both polyhedral cones and an alternating algorithm using projections to maximize γ_n will converge to a stationary point of the problem that is still circulant

Maximum Angle between PSD and Nonnegative Symmetric Matrices

Given $\langle A, B\rangle = Tr(A^\top B)$ an open question is the maximum angle between the cone of PSD matrices $\mathcal{P}^n$ and the cone of nonnegative symmetric matrices $\mathcal{N}^n$ for $n \geq 5$

$$n = 2, 3, 4 \implies \gamma_n = \frac{3}{4}\pi \qquad \lim_{n\to\infty} \gamma_n \uparrow \pi$$

All antipodal couples (and the best known for $n = 5$) are circulant matrices

If $\mathcal{SC}^n$ is the algebra of circulant matrices, $\mathcal{SC}^n \cap \mathcal{P}^n$ and $\mathcal{SC}^n \cap \mathcal{N}^n$ are both polyhedral cones and an alternating algorithm using projections to maximize γ_n will converge to a stationary point of the problem that is still circulant

5	$0.7575\ \pi$	$0.7575\ \pi$	18	$0.7699\ \pi$	$0.7670\ \pi$
6	$0.7575\ \pi$	$0.7575\ \pi$	19	$0.7703\ \pi$	$0.7681\ \pi$
7	$0.7575\ \pi$	$0.7575\ \pi$	20	$0.7719\ \pi$	$0.7719\ \pi$
8	$0.7608\ \pi$	$0.7608\ \pi$	21	$0.7719\ \pi$	$0.7719\ \pi$
9	$0.7608\ \pi$	$0.7608\ \pi$	22	$0.7719\ \pi$	$0.7719\ \pi$
10	$0.7609\ \pi$	$0.7608\ \pi$	23	$0.7722\ \pi$	$0.7719\ \pi$
11	$0.7627\ \pi$	$0.7627\ \pi$	24	$0.7735\ \pi$	$0.7730\ \pi$
12	$0.7649\ \pi$	$0.7649\ \pi$	25	$0.7735\ \pi$	$0.7730\ \pi$
13	$0.7649\ \pi$	$0.7649\ \pi$	26	$0.7735\ \pi$	$0.7730\ \pi$
14	$0.7659\ \pi$	$0.7649\ \pi$	27	$0.7739\ \pi$	$0.7730\ \pi$
15	$0.7678\ \pi$	$0.7649\ \pi$	28	$0.7750\ \pi$	$0.7730\ \pi$
16	$0.7699\ \pi$	$0.7670\ \pi$	29	$0.7750\ \pi$	$0.7741\ \pi$
17	$0.7699\ \pi$	$0.7670\ \pi$	30	$0.7757\ \pi$	$0.7741\ \pi$

Left: Best known lower bounds on γ_n

Right: Gurobi solutions

- In **black** the exact angle $\mathcal{SC}^n \cap \mathcal{P}^n \angle \mathcal{SC}^n \cap \mathcal{N}^n$
- In blue if a previous angle was bigger then the exact solution
- In red if it is a lower bound

Table 2: Numerical comparison of Gur and BFAS for different dimensions for the problem of finding the maximum angle between the PSD cone and the nonnegative symmetric cone, both restricted to the subalgebra of circulant matrices. Timelimit: 60 seconds

n	13	15	17	19	21	23
exact	0.762950π	0.757765π	0.764971π	0.768062π	0.768769π	0.766370π
Gur	0.762950π	0.757765π	0.764971π	0.767876π	0.765409π	0.766370π
	0.854 s	25.061 s	60^* s	60^* s	60^* s	60^* s
BFAS	0.762950π	0.757765π	0.764971π	0.768062π	0.768768π	0.766370π
	0.333 s	0.356 s	1.114 s	4.418 s	19.953 s	60^* s

Table 3: Numerical comparison of Gur, BFAS, E-AO and SRPL for the same problem. Timelimit: 10 seconds. When the exact value is not available, the best known lower bound is reported with an asterisk

n	17	19	21	23	25	27
exact	0.764971π	0.768062π	0.768769π	0.766370π	$0.767385\pi^*$	$0.768258\pi^*$
Gur	0.764971π	0.759309π	0.765409π	0.766370π	0.767385π	0.760879π
BFAS	0.764971π	0.768062π	0.768768π	0.766370π	0.762620π	0.756841π
E-AO	0.764971π	0.768062π	0.768768π	0.766370π	0.767385π	0.768258π
SRPL	0.764970π	0.768062π	0.768768π	0.766369π	0.767384π	0.768257π

PSD and SNN matrices

Since E-AO and SRPL main steps are projections, they can be adapted to the case of NON-polyhedral cones, as long as we know how to compute the projection on such cones

We can thus test them on the task to find the maximum angle between the cone of Positive Semi-Definite matrices and the cone of Symmetric Nonnegative matrices

Table 4: Numerical comparison for E-AO and SRPL for different dimensions for the problem of finding the maximum angle between the PSD cone and the nonnegative symmetric cone. The table reports the best and average value found over 10000 random initializations, together with the average elapsed time. We also report the best known value for each dimension.

n	30	40	50	60
best known	0.7757π	0.7789π	0.7812π	0.7837π
EAO_b	0.7757π	0.7789π	0.7812π	0.7837π
EAO_a	0.7741π	0.7768π	0.7790π	0.7805π
	0.111 ± 0.054 s	0.701 ± 0.235 s	1.263 ± 0.273 s	2.852 ± 0.321 s
SRPL_b	0.7757π	0.7789π	0.7812π	0.7837π
SRPL_a	0.7739π	0.7766π	0.7787π	0.7802π
	0.062 ± 0.025 s	0.155 ± 0.060 s	0.319 ± 0.130 s	0.565 ± 0.229 s

PSD and SNN matrices

Since E-AO and SRPL main steps are projections, they can be adapted to the case of NON-polyhedral cones, as long as we know how to compute the projection on such cones

We can thus test them on the task to find the maximum angle between the cone of Positive Semi-Definite matrices and the cone of Symmetric Nonnegative matrices

Table 4: Numerical comparison for E-AO and SRPL for different dimensions for the problem of finding the maximum angle between the PSD cone and the nonnegative symmetric cone. The table reports the best and average value found over 10000 random initializations, together with the average elapsed time. We also report the best known value for each dimension.

n	30	40	50	60
best known	0.7757π	0.7789π	0.7812π	0.7837π
EAO_b	0.7757π	0.7789π	0.7812π	0.7837π
EAO_a	0.7741π	0.7768π	0.7790π	0.7805π
	0.111 ± 0.054 s	0.701 ± 0.235 s	1.263 ± 0.273 s	2.852 ± 0.321 s
SRPL_b	0.7757π	0.7789π	0.7812π	0.7837π
SRPL_a	0.7739π	0.7766π	0.7787π	0.7802π
	0.062 ± 0.025 s	0.155 ± 0.060 s	0.319 ± 0.130 s	0.565 ± 0.229 s

Giovanni Barbarino, Nicolas Gillis, and David Sossa. **Computing cone-constrained singular values of matrices.** *Arxiv 2504.04069*, 2025.

Welington de Oliveira, Valentina Sessa, and David Sossa. **Computing critical angles between two convex cones.** *Journal of Optimization Theory and Applications*, 201(2):866–898, 2024.

Alberto Seeger and David Sossa. **Cone-constrained singular value problems.** *J. Convex Anal.*, 30:1285–1306, 2023.

Alberto Seeger and David Sossa. **Singular value problems under nonnegativity constraints.** *Positivity*, 27, 2023.

Nicolas Gillis and Francois Glineur. **A continuous characterization of the maximum-edge biclique problem.** *J. Glob. Optim.*, 58:439–464, 2014.

Felix Goldberg and Naomi Shaked-Monderer. **On the maximum angle between copositive matrices.** *The Electronic Journal of Linear Algebra*, 27(1), 2014.

Alberto Seeger and Alfredo Iusem. **Axiomatization of the index of pointedness for closed convex cones.** *Computational and Applied Mathematics*, 24(2):245–283, 2005.

R. Peeters. **The maximum edge biclique problem is np-complete.** *Discret. Appl. Math.*, 131(3):651–654, 2003.

Thank You!