LEHIGH UNIVERSITY

MOPTA 2017

On the identification of the optimal partition for semidefinite optimization

Ali Mohammad-Nezhad Tamás Terlaky

Department of Industrial and Systems Engineering
Lehigh University

August 2017

ISE Industrial and Systems Engineering

Outline

Outline

Semidefinite optimization

Optimal partition

Identification of the optimal partition

Rounding procedure

Standard form of semidefinite optimization (SDO)

Let C, A^i for $i = 1, \ldots, m$, and X: $n \times n$ symmetric matrices

$$\text{(P)} \qquad \min \left\{ \langle C, X \rangle \mid \langle A^i, X \rangle = b_i, \;\; i = 1, \ldots, m, \; X \succeq 0 \right\},$$

$$\text{(D)} \qquad \max \left\{ b^T y \;\middle|\; \sum_{i=1}^{m} y_i A^i + S = C, \; S \succeq 0, \; y \in \mathbb{R}^m \right\}.$$

Assumption

- *There exists a primal-dual feasible* (X, y, S) *so that* $X, S \succ 0$.
- A^i *for* $i = 1, \ldots, m$ *are linearly independent.*

- We have strong duality:

$$\begin{aligned} \langle A^i, X \rangle = b_i, \quad & i = 1, \ldots, m, \\ \sum_{i=1}^{m} A^i y_i + S = C, & \\ XS = 0, \quad & X, S \succeq 0. \end{aligned}$$

Strict and maximal complementarity

Let $\mathcal{P}^*$ and $\mathcal{D}^*$ denote the primal and dual optimal sets.

Definition

A primal-dual optimal solution (X^*, y^*, S^*) is maximally complementary if

$$(X^*, y^*, S^*) \in \mathrm{ri}(\mathcal{P}^* \times \mathcal{D}^*).$$

A maximally complementary optimal solution is strictly complementary if

$$X^* + S^* \succ 0.$$

Alternatively,

- $X^* \in \mathcal{P}^*$ and $S^* \in \mathcal{D}^*$ are maximally complementary optimal solutions if

$$\mathcal{R}(X) \subset \mathcal{R}(X^*), \ \ \forall X \in \mathcal{P}^*, \quad \mathcal{R}(S) \subset \mathcal{R}(S^*), \ \ \forall S \in \mathcal{D}^*.$$

- $X^* \in \mathcal{P}^*$ and $S^* \in \mathcal{D}^*$ are strictly complementary optimal solutions if

$$\mathcal{R}(X^*) + \mathcal{R}(S^*) = \mathbb{R}^n.$$

- A SDO problem may fail strict complementarity.

Outline

Optimal partition

Let (X^*, y^*, S^*) be a maximally complementary optimal solution, and

$$\mathcal{B} := \mathcal{R}(X^*), \qquad \mathcal{N} := \mathcal{R}(S^*).$$

Then it is immediate that

- $\mathcal{R}(X) \subseteq \mathcal{B}$ and $\mathcal{R}(S) \subseteq \mathcal{N}$ for all $(X, y, S) \in \mathcal{P}^* \times \mathcal{D}^*$.
- $\dim(\mathcal{B}) + \dim(\mathcal{N}) \leq n$.
- $\mathcal{B}$ and $\mathcal{N}$ are spanned by the eigenvectors of positive eigenvalues

$$\mathcal{B} = \mathcal{R}(Q\Lambda(X^*)), \qquad \mathcal{N} = \mathcal{R}(Q\Lambda(S^*)).$$

If $\dim(\mathcal{B}) + \dim(\mathcal{N}) < n$:

- The orthogonal complement to $\mathcal{B} + \mathcal{N}$, which we call $\mathcal{T}$, is nonzero.
- The strict complementarity fails.

Definition

The partition $(\mathcal{B}, \mathcal{N}, \mathcal{T})$ of $\mathbb{R}^n$ is called the optimal partition of an SDO problem.

Optimal partition and optimal solutions

Let $Q := [Q_{\mathcal{B}}, Q_{\mathcal{T}}, Q_{\mathcal{N}}]$ denote an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$.

Let $n_{\mathcal{B}} := \dim(\mathcal{B})$, $n_{\mathcal{N}} := \dim(\mathcal{N})$, and $n_{\mathcal{T}} := \dim(\mathcal{T})$.

Theorem (de Klerk et al.)

Every primal-dual optimal solution $(X, y, S) \in \mathcal{P}^* \times \mathcal{D}^*$ *can be represented as*

$$X = Q_{\mathcal{B}} U_X Q_{\mathcal{B}}^T, \qquad S = Q_{\mathcal{N}} U_S Q_{\mathcal{N}}^T,$$

where $U_X \in \mathbb{S}_+^{n_{\mathcal{B}}}$ *and* $U_S \in \mathbb{S}_+^{n_{\mathcal{N}}}$.

- *If* $n_{\mathcal{B}} > 0$ *and* $X^* \in \mathrm{ri}(\mathcal{P}^*)$, *then there exists* $U_{X^*} \succ 0$.
- *If* $n_{\mathcal{N}} > 0$ *and* $S^* \in \mathrm{ri}(\mathcal{D}^*)$, *then there exists* $U_{S^*} \succ 0$.

It can be deducted that

- $Q_{\mathcal{B}} \mathbb{S}_+^{n_{\mathcal{B}}} Q_{\mathcal{B}}^T$ is the minimal face of $\mathbb{S}_+^n$ which contains $\mathcal{P}^*$.
- $Q_{\mathcal{N}} \mathbb{S}_+^{n_{\mathcal{N}}} Q_{\mathcal{N}}^T$ is the minimal face of $\mathbb{S}_+^n$ which contains $\mathcal{D}^*$.

Optimal partition and parametric SDO

Consider a pair of SDO problems with perturbed objective vector:

$$(\mathrm{P}(\omega)) \qquad \min \left\{ \langle C + \omega \bar{C}, X \rangle \mid \langle A^i, X \rangle = b_i, \quad i = 1, \ldots, m, \ X \succeq 0 \right\},$$

$$(\mathrm{D}(\omega)) \qquad \max \left\{ b^T y \;\middle|\; \sum_{i=1}^{m} y_i A^i + S = C + \omega \bar{C}, \ S \succeq 0, \ y \in \mathbb{R}^m \right\}.$$

- We assume that the Slater condition holds for all ω in a closed interval.

The optimal value function is defined as

$$\phi(\omega) := \langle C + \omega \bar{C}, X(\omega) \rangle = b^T y(\omega),$$

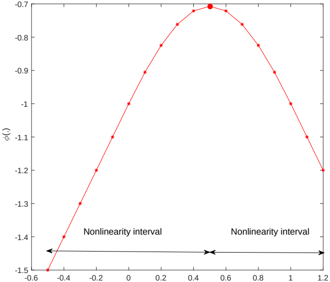
where $(X(\omega), y(\omega), S(\omega))$ is a primal-dual optimal solution.

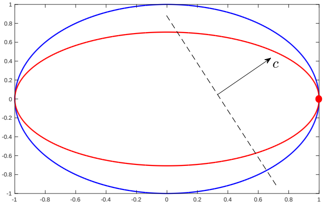
- The optimal value function is concave and piecewise algebraic (Nie et al.).
- Linearity and nonlinearity intervals are joined at the transition points.

Optimal partition and parametric SDO

We can describe $\phi(.)$ using the optimal partition:

- Differentiability of $\phi(.)$ at a given point ω:
  - Left and right derivatives.
- Constancy interval of the optimal partition $\Rightarrow$ linearity interval.
- Length of linearity interval.



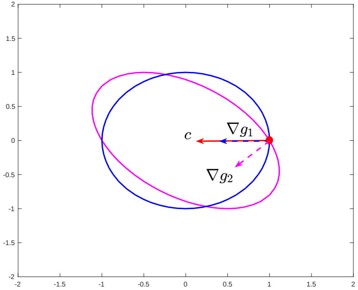


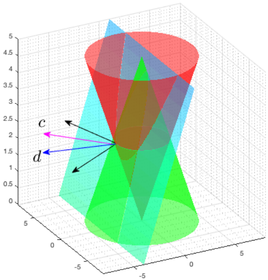
Optimal partition and parametric SDO

Identification of strongly unique optimal solutions ⇒ Linearity intervals:

$$\langle C, X^* \rangle \geq \langle C, \bar{X} \rangle + \alpha \| X^* - \bar{X} \|, \quad \forall \bar{X} \in \mathcal{P}^*,$$

or optimal solutions which are strongly unique in lower dimensions.





Outline

Semidefinite optimization

Optimal partition

Identification of the optimal partition

Rounding procedure

Central path

We aim to identify the sets of eigenvectors converging to an orthonormal bases.

- The central path equations are defined as

$$\begin{aligned} \langle A_i, X \rangle &= b_i, \quad i = 1, ..., m, \\ \sum_{i=1}^{m} A_i y_i + S &= C, \\ XS &= \mu I, \\ X, S &\succeq 0. \end{aligned}$$

- This system has a unique solution, the so called μ-center, $\forall \mu > 0$.
- The trajectory of the μ-centers is known as the central path.
- As $\mu \to 0$, the trajectory converges to a solution in the optimal set.

Condition number

We measure the magnitude of the eigenvalues using a condition number.

- The condition number σ is defined as

$$\sigma_{\mathcal{B}} := \begin{cases} \max_{X \in \mathcal{P}^*} \lambda_{\min}(Q_{\mathcal{B}}^T X Q_{\mathcal{B}}), & \mathcal{B} \neq \{0\}, \\ \infty, & \mathcal{B} = \{0\}, \end{cases}$$

$$\sigma_{\mathcal{N}} := \begin{cases} \max_{(y,S) \in \mathcal{D}^*} \lambda_{\min}(Q_{\mathcal{N}}^T S Q_{\mathcal{N}}), & \mathcal{N} \neq \{0\}, \\ \infty, & \mathcal{N} = \{0\}, \end{cases}$$

$$\sigma := \min\{\sigma_{\mathcal{B}}, \sigma_{\mathcal{N}}\}.$$

- By the Slater condition σ is well-defined and positive.
- For some instances σ can be doubly exponentially small.

Lower bound for the condition number

Lemma

Let L denote the binary length of the largest absolute value of the entries in b, C, and A^i for $i = 1, \ldots, m$. Then we have

$$\sigma \geq \min\left\{\frac{1}{r_{\mathcal{P}^*}\sum_{i=1}^m \|A^i\|}, \frac{1}{r_{\mathcal{D}^*}}\right\},$$

where

- *$R_{\mathcal{P}^*}$ is the radius of the ball which intersects $\mathcal{P}^*$,*
- *$R_{\mathcal{D}^*}$ is the radius of the ball which intersects $\mathcal{D}^*$.*

Regular system and degree of singularity

- The primal and dual feasible sets are regular systems:

$$\left\{X \in \mathbb{S}^n \mid \langle A^i, X\rangle = b_i,\ i = 1, \ldots, m\right\} \cap \operatorname{int}(\mathbb{S}^n_+) \neq \emptyset,$$

$$\left\{S \in \mathbb{S}^n \mid \sum_{i=1}^{m} y_i A^i + S = C,\ \text{for some } y \in \mathbb{R}^m\right\} \cap \operatorname{int}(\mathbb{S}^n_+) \neq \emptyset.$$

- The optimal set is not regular due to the complementarity condition:
  - Primal optimal face: $Q_{\mathcal{B}} \mathbb{S}^{n_{\mathcal{B}}}_+ Q_{\mathcal{B}}^T$
  - Dual optimal face: $Q_{\mathcal{N}} \mathbb{S}^{n_{\mathcal{N}}}_+ Q_{\mathcal{N}}^T$

- The number of regularization steps (Borwein and Wolkowicz) is called the degree of singularity.

Distance to the optimal set

To identify eigenvectors converging to an orthonormal basis of $\mathcal{T}$ we need

- The distance of a central solution to the optimal set,
- The degree of singularity of the subspace which contains the optimal set.

Lemma

Assume $n\mu \leq 1$. There exists $(X, y, S) \in \mathcal{P}^* \times \mathcal{D}^*$ so that

$$\left\|X(\mu) - X\right\| \leq \kappa(n\mu)^\gamma, \qquad \left\|S(\mu) - S\right\| \leq \kappa(n\mu)^\gamma,$$

where

- $\gamma = 2^{-d(\mathrm{lin}(\mathcal{P}^* \times \mathcal{D}^*),\ \mathbb{S}^n_+)}$,
- $d(\mathrm{lin}(\mathcal{P}^* \times \mathcal{D}^*), \mathbb{S}^n_+)$ *is the degree of singularity of the minimal subspace containing the optimal set,*
- κ *is a positive condition number.*

- In general, we have $\gamma \geq \frac{1}{2^{n-1}}$ for $n \geq 2$.
- If the strict complementarity holds, then $\gamma = \frac{1}{2}$.

Bounds for the eigenvalues

Theorem

For a central solution $(X(\mu), y(\mu), S(\mu))$ *with* $n\mu \leq 1$ *it holds that:*

1. *For* $i = 1, \dots, n_{\mathcal{B}}$ *we have*

$$\lambda_{[n-i+1]}(S(\mu)) \leq \frac{n\mu}{\sigma}, \qquad \lambda_{[i]}(X(\mu)) \geq \frac{\sigma}{n}.$$

2. *For* $i = 1, \dots, n_{\mathcal{N}}$ *we have*

$$\lambda_{[i]}(S(\mu)) \geq \frac{\sigma}{n}, \qquad \lambda_{[n-i+1]}(X(\mu)) \leq \frac{n\mu}{\sigma}.$$

Furthermore, we have

$$\lambda_{[n-i+1]}(X(\mu)) \leq \kappa\sqrt{n}(n\mu)^{\gamma}, \qquad \lambda_{[i]}(S(\mu)) \geq \frac{\mu}{\kappa\sqrt{n}(n\mu)^{\gamma}}, \; i = 1, \dots, n_{\mathcal{N}} + n_{\mathcal{T}},$$
$$\lambda_{[n-i+1]}(S(\mu)) \leq \kappa\sqrt{n}(n\mu)^{\gamma}, \qquad \lambda_{[i]}(X(\mu)) \geq \frac{\mu}{\kappa\sqrt{n}(n\mu)^{\gamma}}, \; i = 1, \dots, n_{\mathcal{B}} + n_{\mathcal{T}}.$$

- *If* $n_{\mathcal{T}} > 0$, *then we have* $\kappa \geq \frac{1}{n}$, *and* $\frac{1}{2^{n-1}} \leq \gamma \leq \frac{1}{2}$.

Identification of eigenvectors converging to $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$

In general, there exist three sets of eigenvectors $q_i(\mu)$ for which

- $\lambda_i(X(\mu))$ converges to a positive value and $\lambda_i(S(\mu))$ converges to 0;
- $\lambda_i(S(\mu))$ converges to a positive value and $\lambda_i(X(\mu))$ converges to 0;
- both $\lambda_i(X(\mu))$ and $\lambda_i(S(\mu))$ converge to 0,

where $\lambda_i(X(\mu))$ and $\lambda_i(S(\mu))$ correspond to the eigenvector $q_i(\mu)$.

- As $\mu \to 0$, the eigenvectors converge to an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$.

Theorem

If μ satisfies

$$\mu < \min\left\{\frac{1}{n}\left(\frac{\sigma}{\kappa n^{\frac{3}{2}}}\right)^{\frac{1}{\gamma}},\ \frac{\sigma^2}{n^2},\ \frac{1}{n}\right\},$$

then we can identify the sets of eigenvectors converging to an orthonormal bases for $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{T}$.

Outline

Approximate maximally complementary solution

- We do not have the exact orthonormal bases for $\mathcal{B}$, $\mathcal{N}$ and $\mathcal{T}$:
  - An exact solution cannot be obtained from a given central solution.
- If we project $(X(\mu), y(\mu), S(\mu))$ onto the boundary of the cone:
  - The solution has zero complementary gap.
  - The solution is ϵ-infeasible.
  - The solution is called approximate maximally complementary.

The eigenvectors of $X(\mu)$ and $S(\mu)$ can be rearranged so that

$$Q(\mu) := [Q_{\mathcal{B}}(\mu), Q_{\mathcal{T}}(\mu), Q_{\mathcal{N}}(\mu)],$$

if μ allows for the identification of eigenvectors.

$$Q(\mu)^T X(\mu) Q(\mu) = \begin{bmatrix} \Lambda_{\mathcal{B}}(X(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(X(\mu)) & 0 \\ 0 & 0 & \Lambda_{\mathcal{N}}(X(\mu)) \end{bmatrix}.$$

- $\Lambda_{\mathcal{T}}(X(\mu)) \to 0$ and $\Lambda_{\mathcal{N}}(X(\mu)) \to 0$ as $\mu \to 0$.
- We get a complementary solution if we discard $\Lambda_{\mathcal{T}}(X(\mu))$ and $\Lambda_{\mathcal{N}}(X(\mu))$.

Primal auxiliary problem

Let $\bar{A}^i := Q(\mu)^T A^i Q(\mu)$.

- For the primal problem we solve

$$\begin{aligned} \min \quad & \|\epsilon_p\|^2 + \|\Delta X\|^2 \\ \text{s.t.} \quad & \langle \bar{A}^i_{\mathcal{B}}, \Delta X \rangle - (\epsilon_p)_i = \langle \bar{A}^i_{\mathcal{T}}, \Lambda_{\mathcal{T}}(X(\mu)) \rangle + \langle \bar{A}^i_{\mathcal{N}}, \Lambda_{\mathcal{N}}(X(\mu)) \rangle, \quad i = 1, \ldots, m. \end{aligned}$$

- The optimal solution $(\epsilon_p^*, \Delta X^*)$ to the auxiliary problem yields

$$\tilde{X}_{\mathcal{B}} := \Lambda_{\mathcal{B}}(X(\mu)) + \Delta X^*$$

so that

$$\langle \bar{A}^i_{\mathcal{B}}, \tilde{X}_{\mathcal{B}} \rangle = b_i + (\epsilon_p^*)_i, \qquad i = 1, \ldots, m.$$

- Thus, $\tilde{X}_{\mathcal{B}}$ has $\|\epsilon_p^*\|$ infeasibility for the primal constraints.

Dual auxiliary problem

Let E denote a residual matrix as

$$E := \begin{bmatrix} E_{\mathcal{B}} & E_{\mathcal{B}\mathcal{T}} & E_{\mathcal{B}\mathcal{N}} \\ E_{\mathcal{T}\mathcal{B}} & E_{\mathcal{T}} & E_{\mathcal{T}\mathcal{N}} \\ E_{\mathcal{N}\mathcal{B}} & E_{\mathcal{N}\mathcal{T}} & 0 \end{bmatrix}.$$

For the dual problem we solve

$$\begin{aligned} \min \quad & \|E\|^2 + \|\Delta y\|^2 + \|\Delta S\|^2 \\ \text{s.t.} \quad & \sum_{i=1}^{m} \Delta y_i \bar{A}^i + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta S \end{bmatrix} - E = \begin{bmatrix} \Lambda_{\mathcal{B}}(S(\mu)) & 0 & 0 \\ 0 & \Lambda_{\mathcal{T}}(S(\mu)) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

- The optimal solution $(E^*, \Delta y^*, \Delta S^*)$ gives

$$\begin{aligned} \tilde{y}_i &:= y_i(\mu) + \Delta y_i^*, \qquad i = 1, \ldots, m, \\ \tilde{S}_{\mathcal{N}} &:= \Lambda_{\mathcal{N}}(S(\mu)) + \Delta S^*. \end{aligned}$$

Cone feasibility

Let π_p and π_d denote parameters dependent on linear mapping $\mathcal{A}$ and $Q(\mu)$, and

$$r(n) := \frac{n(n+1)}{2}.$$

- If μ is sufficiently small, then the rounded solution is cone feasible.

Theorem

Let $\vartheta_1 := 2n^2\|\mathcal{A}\|^2$, $\vartheta_2 := 2\kappa n\sqrt{nn_{\mathcal{T}}}\|\mathcal{A}\|^2$, *and let*

$$\tilde{\mu} := \min\left\{\frac{\sigma^2}{\vartheta_1 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})n_{\mathcal{N}}},\ \pi_d\sqrt{mn_{\mathcal{B}}}\}}, \frac{1}{n}\left(\frac{\sigma}{\vartheta_2 \max\{\pi_p\sqrt{r(n_{\mathcal{B}})},\ \pi_d\sqrt{m}\}}\right)^{\frac{1}{\gamma}}\right\}.$$

If $\mu \leq \tilde{\mu}$, *then we have* $\tilde{X}_{\mathcal{B}}, \tilde{S}_{\mathcal{N}} \succ 0$.

- Only $\mathcal{O}(\max\{n_{\mathcal{B}}^6, m^3\})$ arithmetic operations are needed.

Summary

- Introduction of the optimal partition.
- Application of optimal partition in sensitivity analysis.
- Identification of eigenvectors converging to an orthonormal bases.
- A rounding procedure for an approximate maximally complementary solution.

Thank you for your attention

Any questions?