Continuous Optimisation, Chpt 9: Semidefinite Problems

Peter J.C. Dickinson

DMMP, University of Twente

p.j.c.dickinson@utwente.nl

http://dickinson.website/Teaching/2016CO.html

version: 21/11/16

Monday 21st November 2016

Book

Semidefinite Optimization – M.J. Todd
http://people.orie.cornell.edu/miketodd/soa5.ps

Primal Problem: Chpt 2: Problems

Dual problems: Chpt 4: Duality

Examples: Chpt 3: Examples

Table of Contents

Inner Product

For a vector space $\mathbb{V}$, we say that $\langle \bullet, \bullet \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ is an inner product if the following properties hold:

1. Symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$,
2. Linearity 1: $\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$, $\lambda \in \mathbb{R}$,
3. Linearity 2: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}$,
4. Positive definiteness: $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ for all $\mathbf{x} \in \mathbb{V} \setminus \{\mathbf{0}\}$.

Example

For vector space $\mathbb{R}^n$ and a matrix $\mathsf{A} \in \mathcal{S}^n$, $\mathsf{A} \succ \mathsf{O}$ have an (induced) inner product $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathsf{A}} = \mathbf{x}^{\mathsf{T}} \mathsf{A} \mathbf{y}$.

Usually consider the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathsf{T}} \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathsf{I}}$.

Example

For $\mathcal{B} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq 1\}$, we can consider the space of continuous functions from $\mathcal{B}$ to $\mathbb{R}$, which has an inner product $\langle f, g \rangle = \int_{\mathcal{B}} f(\mathbf{x}) g(\mathbf{x}) \mathrm{d}\mathbf{x}$.

Primal and Dual Problem

Consider convex cone $\mathcal{K} \subseteq \mathbb{V}$ and $\mathbf{c}, \mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{V}$ and $\mathbf{b} \in \mathbb{R}^m$:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \langle \mathbf{c}, \mathbf{x} \rangle \\ \text{s. t.} \quad & \langle \mathbf{a}_i, \mathbf{x} \rangle = b_i \qquad \text{for all } i = 1, \ldots, m \\ & \mathbf{x} \in \mathcal{K}, \end{aligned} \tag{P}$$

$$\begin{aligned} \max_{\mathbf{y}} \quad & \mathbf{b}^\mathsf{T} \mathbf{y} \\ \text{s. t.} \quad & \mathbf{c} - \sum_{i=1}^m y_i \mathbf{a}_i \in \mathcal{K}^* \\ & \mathbf{y} \in \mathbb{R}^m. \end{aligned} \tag{D}$$

$$\mathcal{K}^* := \{ \mathbf{z} \in \mathbb{V} \mid \langle \mathbf{x}, \mathbf{z} \rangle \geq 0 \text{ for all } \mathbf{x} \in \mathcal{K} \}.$$

<u>Ex. 9.1</u> Show that $\mathbf{x} \in \text{feas}(P),\ \mathbf{y} \in \text{feas}(D) \Rightarrow \langle \mathbf{c}, \mathbf{x} \rangle \geq \mathbf{b}^\mathsf{T} \mathbf{y}$.

Symmetric Matrices

Definition 9.1

For space of symmetric matrices, $\mathcal{S}^n$, define inner product $\langle \mathrm{A}, \mathrm{B} \rangle := \mathrm{trace}(\mathrm{AB}) = \sum_{i,j=1}^{n} a_{ij} b_{ij}$.

The definitions/results from the previous two lectures can be naturally extended for this, noting that $\mathcal{S}^n$ is a space of dimension $\frac{1}{2}n(n+1)$.

Lemma 9.2

For $\mathrm{A} \in \mathbb{R}^{n \times m}$ and $\mathrm{B} \in \mathbb{R}^{m \times n}$ have $\mathrm{trace}(\mathrm{AB}) = \mathrm{trace}(\mathrm{BA})$.

Lemma 9.3

Have $\langle \mathbf{a}\mathbf{b}^\mathsf{T} + \mathbf{b}\mathbf{a}^\mathsf{T}, \mathrm{X} \rangle = 2\mathbf{a}^\mathsf{T}\mathrm{X}\mathbf{b}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathrm{X} \in \mathcal{S}^n$.

Nonnegative symmetric matrices

Cone of **N**onnegative symmetric matrices:

$$\mathcal{N}^n := \{\mathrm{X} \in \mathcal{S}^n \mid x_{ij} \geq 0 \text{ for all } i,j\} = \mathrm{conic}\{\mathbf{e}_i\mathbf{e}_j^\mathsf{T} + \mathbf{e}_j\mathbf{e}_i^\mathsf{T} \mid i,j\}.$$

$$\begin{aligned}(\mathcal{N}^n)^* &= \{\mathbf{e}_i\mathbf{e}_j^\mathsf{T} + \mathbf{e}_j\mathbf{e}_i^\mathsf{T} \mid i,j\}^* \\ &= \{\mathrm{Y} \in \mathcal{S}^n \mid \langle \mathbf{e}_i\mathbf{e}_j^\mathsf{T} + \mathbf{e}_j\mathbf{e}_i^\mathsf{T}, \mathrm{Y}\rangle \geq 0 \text{ for all } i,j\} \\ &= \{\mathrm{Y} \in \mathcal{S}^n \mid y_{ji} + y_{ij} \geq 0 \text{ for all } i,j\} = \mathcal{N}^n.\end{aligned}$$

- Closed convex cone as intersection of closed convex cones.
- Pointed as $\pm\mathrm{X} \in \mathcal{N}^n \Rightarrow \pm x_{ij} \geq 0$ for all $i,j \Rightarrow \mathrm{X} = \mathrm{O}$.
- Full-dimensional as dual to a pointed convex cone.

Denote $\mathrm{A} \geq \mathrm{B}$ iff $\mathrm{A} - \mathrm{B} \in \mathcal{N}^n$.

Primal and Dual Problem

Proper cone $\mathcal{K} \subseteq \mathcal{S}^n$ and $\mathsf{C}, \mathsf{A}_1, \ldots, \mathsf{A}_m \in \mathcal{S}^n$ and $\mathbf{b} \in \mathbb{R}^m$:

$$\begin{aligned} \min_{\mathsf{X}} \quad & \langle \mathsf{C}, \mathsf{X} \rangle \\ \text{s. t.} \quad & \langle \mathsf{A}_i, \mathsf{X} \rangle = b_i \qquad \text{for all } i = 1, \ldots, m \\ & \mathsf{X} \in \mathcal{K}, \end{aligned} \tag{P}$$

$$\begin{aligned} \max_{\mathbf{y}} \quad & \mathbf{b}^{\mathsf{T}} \mathbf{y} \\ \text{s. t.} \quad & \mathsf{C} - \textstyle\sum_{i=1}^m y_i \mathsf{A}_i \in \mathcal{K}^* \\ & \mathbf{y} \in \mathbb{R}^m. \end{aligned} \tag{D}$$

$$\mathcal{K}^* := \{ \mathsf{Z} \in \mathcal{S}^n \mid \langle \mathsf{X}, \mathsf{Z} \rangle \geq 0 \text{ for all } \mathsf{X} \in \mathcal{K} \}.$$

(P) and (D) are dual problems to each other.
Slater's condition $\Rightarrow$ Strong duality.

Table of Contents

Positive semidefinite cone

Positive **S**emi**d**efinite cone,

$$\mathcal{PSD}^n := \{\mathsf{X} \in \mathcal{S}^n \mid \mathbf{v}^\mathsf{T}\mathsf{X}\mathbf{v} \geq 0 \; \forall \mathbf{v} \in \mathbb{R}^n\}$$
$$= \operatorname{conic}\{\mathbf{b}\mathbf{b}^\mathsf{T} \mid \mathbf{b} \in \mathbb{R}^n\}.$$

Ex. 9.2 Prove that $(\mathcal{PSD}^n)^* = \mathcal{PSD}^n$.

Ex. 9.3 Prove that $\mathcal{PSD}^n$ is a proper cone.

Denote $\mathsf{A} \succeq \mathsf{B}$ iff $\mathsf{A} - \mathsf{B} \in \mathcal{PSD}^n$.

Solvers

Commercial: MOSEK. Free: SDPT3, SEDUMI

Correlation matrices

Definition 9.4

For two random variables X_1, X_2, define their correlation to be
$\operatorname{corr}(X_1, X_2) := \mathbb{E}\left[\left(\frac{X_1-\mu_1}{\sigma_1}\right)\left(\frac{X_2-\mu_2}{\sigma_2}\right)\right]$.
For vector of random variables $\boldsymbol{X} = \begin{pmatrix} X_1 & \cdots X_n \end{pmatrix}^{\mathsf{T}}$ define correlation matrix $\operatorname{corr}(\boldsymbol{X}) \in \mathcal{S}^n$ s.t. $[\operatorname{corr}(\boldsymbol{X})]_{ij} = \operatorname{corr}(X_i, X_j)$.

Theorem 9.5

For $\mathrm{A} \in \mathbb{R}^{n\times n}$, there exists a distribution such that $\mathrm{A} = \operatorname{corr}(\boldsymbol{X})$ iff $\mathrm{A} \in \mathcal{PSD}^n$ and $a_{ii} = 1$ for all i.

Example

What is the maximum possible $\operatorname{corr}(X_1, X_2)$ given $\operatorname{corr}(X_1, X_3) = 0.6$ and $\operatorname{corr}(X_2, X_3) = 0$?

$$\begin{aligned} \max \quad & y_1 \\ \text{s. t.} \quad & \begin{pmatrix} 1 & y_1 & 0.6 \\ y_1 & 1 & 0 \\ 0.6 & 0 & 1 \end{pmatrix} \succeq 0 \end{aligned}$$

Ex. 9.4 What is the dual problem to the example above?

Ex. 9.5 Formulate as a semidefinite problem, the problem of finding the minimum possible $\operatorname{corr}(X_1, X_2)$ given $0.5 \le \operatorname{corr}(X_1, X_3) \le 0.6$ and $-0.1 \le \operatorname{corr}(X_2, X_3) \le 0$.

Eigenvalue problems

Lemma 9.6

For $\mathsf{A} \in \mathcal{S}^n$ and $s, t \in \mathbb{R} \cup \{\pm\infty\}$ with $s \leq t$, all the eigenvalues of A are between s and t if and only if $s\mathsf{I} \preceq \mathsf{A} \preceq t\mathsf{I}$.
(Where $\mathsf{I} \in \mathcal{S}^n$ is the identity matrix.)

Example

Find $\mathbf{x} \in \mathbb{R}^m$ such that the absolute values of the eigenvalues of $C - \sum_{i=1}^n \mathsf{A}_i x_i$ are as small as possible:

$$\min_{\mathbf{x},t} \quad t$$
$$\text{s.t.} \quad -t\mathsf{I} \preceq \mathsf{C} - \sum_{i=1}^n \mathsf{A}_i x_i \preceq t\mathsf{I}.$$

Approximating quadratic problems

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \mathbf{x}^\mathsf{T} \mathsf{Q}\mathbf{x} + \mathbf{q}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}^\mathsf{T} \mathsf{A}_i\mathbf{x} + \mathbf{a}_i^\mathsf{T}\mathbf{x} = \alpha_i \text{ for all } i = 1,\dots,m, \end{aligned} \tag{1}$$

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \left\langle \mathsf{Q}, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle + \mathbf{q}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \left\langle \mathsf{A}_i, \mathbf{x}\mathbf{x}^\mathsf{T} \right\rangle + \mathbf{a}_i^\mathsf{T}\mathbf{x} = \alpha_i \text{ for all } i = 1,\dots,m \\ & \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathbf{x}\mathbf{x}^\mathsf{T} \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}^\mathsf{T} \in \mathcal{PSD}^{n+1}, \end{aligned} \tag{2}$$

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n,\, \mathsf{X}\in\mathcal{S}^n} \quad & \langle \mathsf{Q}, \mathsf{X} \rangle + \mathbf{q}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \langle \mathsf{A}_i, \mathsf{X} \rangle + \mathbf{a}_i^\mathsf{T}\mathbf{x} = \alpha_i \text{ for all } i = 1,\dots,m \\ & \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathsf{X} \end{pmatrix} \in \mathcal{PSD}^{n+1} \end{aligned} \tag{3}$$

$val(1) = val(2) \geq val(3).$

Examples

Example

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 - 4x_1x_2 + 2x_2 \\ \text{s. t.} \quad & x_1^2 + x_2^2 - x_1x_2 - 4x_1 = 1, \quad \mathbf{x} \in \mathbb{R}^2, \end{aligned}$$

$$\begin{aligned} \min \quad & \left\langle \begin{pmatrix} 1 & -2 \\ -2 & 2 \end{pmatrix}, \mathrm{X} \right\rangle + 2x_2 \\ \text{s. t.} \quad & \left\langle \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}, \mathrm{X} \right\rangle - 4x_1 = 1, \quad \begin{pmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & \mathrm{X} \end{pmatrix} \in \mathcal{PSD}^3. \end{aligned}$$

Ex. 9.6 For the following problem, give a finite upper bound and formulate a PSD problem which would give a lower bound.

$$\begin{aligned} \min \quad & 2x_1^2 + 2x_2^2 - 6x_1x_2 + 2x_2 - 4x_3 \\ \text{s. t.} \quad & 2x_1^2 + x_2^2 - 2x_1x_2 - 4x_1 = 1 \\ & 2x_2 - x_3^2 = 0, \quad \mathbf{x} \in \mathbb{R}^3. \end{aligned}$$