Semidefinite programming

Péter Hajnal

2024. Fall

Semidefinite Programming, SDP

General Formulation of Semidefinite Programming

Minimize	$c^{\mathsf{T}}x$-t
subject to	$\sum_{i=1}^{n} x_i A_i \preceq B$
	$Dx = e$,

where $c, x \in \mathbb{R}^n$, $A_i, B \in \mathcal{S}^k = \{M \in \mathbb{R}^{k\times k} : M^{\mathsf{T}} = M\} \subset \mathbb{R}^{k\times k}$, $D \in \mathbb{R}^{\ell\times n}$, and $e \in \mathbb{R}^{\ell}$.

$\mathcal{S}^n$ denotes the set of real symmetric $n \times n$ matrices, i.e., $M \in \mathbb{R}^{n\times n}$ belongs to $\mathcal{S}^n$ if and only if $M^{\mathsf{T}} = M$. Specifically, $\mathcal{S}^n \subset \mathbb{R}^{n\times n}$.

Notation

$A \preceq B$ if and only if $A, B \in \mathcal{S}^n$ and $0 \preceq B - A$, i.e., $B - A$ is positive semidefinite, denoted as $B - A \in \mathcal{S}^n_+$.

$\sum_{i=1}^{n} x_i A_i$

$$\begin{pmatrix} A_{1,1}^{(1)}x_1 + \ldots + A_{1,1}^{(n)}x_n & A_{1,2}^{(1)}x_1 + \ldots + A_{1,2}^{(n)}x_n & \ldots & A_{1,k}^{(1)}x_1 + \ldots + A_{1,k}^{(n)}x_n \\ A_{2,1}^{(1)}x_1 + \ldots + A_{2,1}^{(n)}x_n & A_{2,2}^{(1)}x_1 + \ldots + A_{2,2}^{(n)}x_n & \ldots & A_{2,k}^{(1)}x_1 + \ldots + A_{2,k}^{(n)}x_n \\ \vdots & \vdots & & \vdots \\ A_{k,1}^{(1)}x_1 + \ldots + A_{k,1}^{(n)}x_n & A_{k,2}^{(1)}x_1 + \ldots + A_{k,2}^{(n)}x_n & \ldots & A_{k,k}^{(1)}x_1 + \ldots + A_{k,k}^{(n)}x_n \end{pmatrix}$$

Example

$$\begin{pmatrix} 2x_1 - 3x_2 + x_3 & 5x_1 + 2x_2 - x_3 & x_1 - x_2 + 8x_3 & 6x_1 + 5x_2 + x_3 \\ 5x_1 + 2x_2 - x_3 & -x_1 + 7x_2 - 2x_3 & 9x_1 - 3x_2 + x_3 & -2x_1 - x_2 + 4x_3 \\ x_1 - x_2 + 8x_3 & 9x_1 - 3x_2 + x_3 & 10x_1 - 2x_2 + 2x_3 & 8x_1 + x_2 + x_3 \\ 6x_1 + 5x_2 + x_3 & -2x_1 - x_2 + 4x_3 & 8x_1 + x_2 + x_3 & -11x_1 + 2x_2 - 3x_3 \end{pmatrix}$$

Linear Inequalities as SDP Constraints

The non-negativity constraints can be expressed as the positivity of suitable matrices:

$$x_1, \ldots x_n \geq 0 \iff \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix} \succeq 0 \iff \begin{pmatrix} -x_1 & & 0 \\ & \ddots & \\ 0 & & -x_n \end{pmatrix} \preceq 0$$

Observation

The LP problem is a special case of an SDP problem.

$\mathcal{S}^n$

$\mathcal{S}^n$ is a linear subspace of $\mathbb{R}^{n\times n}/\mathbb{R}^{[n]\times[n]}/\mathbb{R}^{H\times H}$. Its dimension is $\left(\!\binom{n}{2}\!\right)$.

Notation

$\langle .,.\rangle : \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}$ is the following inner product: For $M, N \in \mathcal{S}^n$,

$$\langle M, N\rangle = \mathrm{Tr}(M^{\mathrm{T}}N) = \sum_j (M^{\mathrm{T}}N)_{jj} = \sum_{i,j} M^{\mathrm{T}}_{ji}N_{ij} = \sum_{i,j} M_{ij}N_{ij}.$$

An alternative: Let $vec : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ be the vectorization of a matrix (i.e., stacking the columns of a table into a vector). Then $\langle M, N\rangle = vec(M)^{\mathsf{T}} vec(N)$.

Below we summarize the properties of the introduced inner product. Interested students can verify these properties themselves.

Lemma

Let $M, N, P \in \mathcal{S}^n$

(i) $\langle M, N\rangle = \langle N, M\rangle$

(ii) $\langle MN, P\rangle = \langle M, PN\rangle$

(iii) $\langle M, M\rangle = \mathrm{Tr} M^2 = \sum_{i=1}^{n} \lambda_i^2 \geq 0$

Notation

$\|M\|_F = \sqrt{\langle M, M\rangle}$, the Frobenius norm of a symmetric matrix M.

$\mathcal{S}^n_+$, Positive Semidefinite Matrices

Lemma

Let $M \in \mathcal{S}^n$ be arbitrary. The following are equivalent:

(i) M is positive semidefinite, i.e., $x^\mathsf{T} M x \geq 0$ for all $x \in \mathbb{R}^n$,

(ii) The eigenvalues of M are nonnegative,

(iii) M is a Gram matrix, i.e., there exists a matrix $V \in \mathbb{R}^{k \times n}$ such that $M = V^\mathsf{T} V$,

(iv) Every symmetric submatrix of M (obtained by deleting s rows and the corresponding s columns) has a nonnegative determinant.

Another Lemma providing an equivalent description

Lemma

If $M \succeq 0$, then there exists $M^{\frac{1}{2}} \succeq 0$ such that $M = M^{\frac{1}{2}} M^{\frac{1}{2}}$.

- Let the eigenvalues of M be $\lambda_1, \ldots, \lambda_n \geq 0$.
- Define

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \qquad \Lambda^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}.$$

- Since M is symmetric, there exists an orthogonal matrix Q such that $Q^{\mathrm{T}} M Q = \Lambda$.
- Thus,

$$M = Q \Lambda Q^{\mathsf{T}} = Q \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} = Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} = (Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}})(Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}}).$$

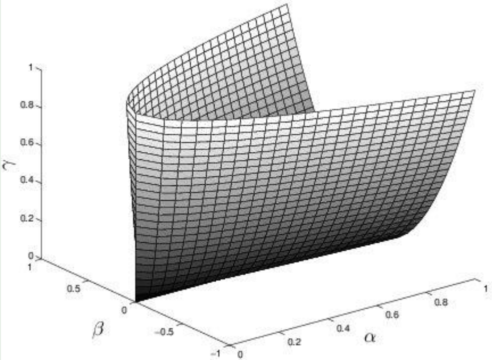
- Choosing $M^{\frac{1}{2}} = Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}}$ proves the Lemma.

$\mathcal{S}^n_+$ Geometrically

Observation

$\mathcal{S}^n_+$ is a cone in $\mathcal{S}^n \subset \mathbb{R}^{n\times n}$.

The Set of Positive Semidefinite Matrices $\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$



Normal Forms of SDP

SDP: I. Normal Form

Minimize	$\langle C, X\rangle$-t
subject to	$\langle A_i, X\rangle = b_i, \quad i = 1, \ldots k$
	$X \succeq 0,$

where $C, X, A_i \in \mathcal{S}^n$ and $b \in \mathbb{R}^k$.

SDP: II. Normal Form

Minimize	$c^{\mathsf{T}}x$-t
subject to	$\sum_{i=1}^n x_i A_i \preceq B,$

where $c, x \in \mathbb{R}^n$ and $A_i, B \in \mathcal{S}^k$.

Break

Duality (Lagrange Method)

We examine the primal problem in the first normal form:

Minimize	$\langle C, X\rangle$-t
subject to	$\langle A_i, X\rangle = b_i, \quad i = 1, \dots k,$
	$X \succeq 0.$

We proceed with the usual idea. We "embed" the constraints into the Lagrangian function.

There is no issue with the finitely many linear constraints.

However, we cannot incorporate the positive semidefiniteness constraint into L, so we "constrain" the domain of L with this condition.

$$L(X, \mu) = \langle C, X\rangle + \sum \mu_i(\langle A_i, X\rangle - b_i), \quad \mathrm{dom}\, L = \{X :\, X \succeq 0\}.$$

Duality (Lagrange Method)

Minimizing this function leads to the objective function of the dual problem.

$$\widehat{c}(\mu) = \inf_{X:\ X \succeq 0} L(X, y) = -\sum b_i \mu_i + \inf \left\langle C + \sum \mu_i A_i, X \right\rangle .$$

We need to solve the optimization problem

$$\inf_{X \succeq 0} \langle M, X \rangle = ?.$$

Duality (Lagrange Method)

Lemma

M is positive semidefinite if and only if $\langle M, X \rangle \geq 0$ for all $X \succeq 0$.

Before proving this lemma, we highlight a consequence.

Corollary

If $M, N \succeq 0$, then $\langle M, N \rangle \geq 0$.

We need a simple equality:

Auxiliary Lemma

Let $A \in \mathbb{R}^{k \times \ell}$ and $B \in \mathbb{R}^{\ell \times k}$. Then

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA).$$

Proof of the Lemma

- Let $M, X \succeq 0$. Then

$$\begin{aligned}
\langle M, X\rangle &= \operatorname{Tr}(MX) = \operatorname{Tr}(M^{1/2}M^{1/2}X^{1/2}X^{1/2}) \\
&= \operatorname{Tr}((M^{1/2}M^{1/2}X^{1/2})X^{1/2}) = \operatorname{Tr}(X^{1/2}(M^{1/2}M^{1/2}X^{1/2})) \\
&= \operatorname{Tr}((X^{1/2}M^{1/2})(M^{1/2}X^{1/2})) \\
&= \operatorname{Tr}((M^{1/2}X^{1/2})^{\mathsf{T}}(M^{1/2}X^{1/2})) \geq 0,
\end{aligned}$$

since $M^{1/2}$ and $X^{1/2}$ are symmetric.

- Conversely, assume that $\langle M, X\rangle \geq 0$ for all $X \succeq 0$ matrices. Let $x \in \mathbb{R}^n$ be arbitrary. Apply our assumption to the positive semidefinite matrix $X = xx^{\mathsf{T}}$:

$$\begin{aligned}
0 &\leq \langle M, X\rangle = \operatorname{Tr}(M(xx^{\mathsf{T}})) = \operatorname{Tr}((Mx)x^{\mathsf{T}}) = \operatorname{Tr}(x^{\mathsf{T}}(Mx)) \\
&= \operatorname{Tr}(x^{\mathsf{T}}Mx) = x^{\mathsf{T}}Mx.
\end{aligned}$$

Thus, M is positive semidefinite.

Application of the Lemma

It follows from the lemma that

$$\inf\langle M, X\rangle = \begin{cases} 0, & M \succeq 0 \\ -\infty, & \text{otherwise} \end{cases}.$$

• Indeed, if $M \succeq 0$, then by the lemma $\langle M, X\rangle$ cannot be negative, but $M = 0$ shows it can be 0.

• Similarly, by the lemma, if $M \not\succeq 0$, then $\langle M, X\rangle$ can be negative. Multiplying X by a positive scalar, we can make it arbitrarily large in absolute value and negative.

Geometric Interpretation of the Lemma

The following definition provides an opportunity to rephrase our lemma in a useful manner:

Definition

The dual of a cone $K \subseteq \mathbb{R}^N$ is:

$$K^* = \{x \in \mathbb{R}^N : \; x^{\mathrm{T}} k \geq 0 \; \forall k \in K\}.$$

Then our proved lemma can be stated as follows:

Lemma

(i) $\mathcal{S}_+^n$ is a cone.

(ii) $\mathcal{S}_+^n$ is self-dual, $(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$.

Dual Problem

As a result of the detour, we obtain that

$$\inf_{X \succeq 0} L(X,\mu) = \begin{cases} -\sum b_i \mu_i, & C + \sum \mu_i A_i \succeq 0 \\ -\infty, & \text{otherwise} \end{cases}.$$

Thus, we find the dual of the starting SDP (primal) problem.

Maximize	$-\sum b_i \mu_i$-t
subject to	$-\sum \mu_i A_i \preceq C$

Equivalently,

Minimize	$\sum b_i \mu_i$-t
subject to	$C + \sum \mu_i A_i = S$,
	$S \succeq 0$.

Weak Duality Theorem

Weak Duality

If p^* is the optimum of the primal SDP and d^* is the optimum of the dual SDP, then $d^* \leq p^*$.

Assume $p^* < \infty$, $d^* > -\infty$.

Let $X \in \mathcal{L}_P$ be a feasible solution of the primal problem and $(\mu, S) \in \mathcal{L}_D$ be a feasible solution of the dual.

Claim

$\langle C, X \rangle \geq -b^\mathsf{T}\mu$.

The claim states that any primal objective value is at least as large as any dual objective value. From this, the theorem trivially follows.

Proof of Weak Duality

We have seen that the product of two positive semidefinite matrices is nonnegative. Hence,

$$\begin{aligned} 0 \leq& \langle X, S\rangle = \langle X, C + \sum \mu_i A_i\rangle = \langle X, C\rangle + \sum \mu_i \langle A_i, X\rangle \\ =& \langle C, X\rangle + \sum \mu_i b_i = \langle C, X\rangle + b_i^{\mathsf{T}} \mu_i. \end{aligned}$$

From this, the claim follows by rearrangement.

Example

Minimize	x_{12}-t
subject to	$\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1+x_{12} \end{pmatrix} \succeq 0$

Example: Converting to Standard Form, Introduction

To dualize the problem, we first bring the problem to the standard form of semidefinite programming I:

We introduce the following variable matrix:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix}$$

The matrix in the original form is $\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1 + x_{12} \end{pmatrix}$.

This is equivalent to setting $x_{11} = 0$, $x_{13} = 0$, $x_{23} = 0$, $x_{33} = 1 + x_{12}$.

Example: Rewriting the Constraints

Rewriting these constraints in the standard form:

- Setting $x_{11} = 0$: $\left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 0$ (this equality constraint corresponds to the dual variable μ_1).
- Setting $x_{13} = 0$: $\left\langle \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0$ (this equality constraint corresponds to the dual variable μ_2).
- Setting $x_{23} = 0$: $\left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, X \right\rangle = 0$ (this equality constraint corresponds to the dual variable μ_3).
- Setting $x_{33} = 1 + x_{12}$: $\left\langle \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1$ (this equality constraint corresponds to the dual variable μ_4).

The Example in Normal Form

The new form of our problem:

Minimize
$$\left\langle \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle\text{-t}$$

subject to
$$\left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 0, \left\langle \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0,$$
$$\left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0, \left\langle \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1.$$
$$X \succeq 0$$

Example: Dualization

Maximize	$-\mu_4$-t
subject to	$\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} +$
	$\mu_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \mu_4 \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \succeq 0.$

So,

Maximize	$-\mu_4$-t
subject to	$\begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} & \mu_2 \\ \frac{1-\mu_4}{2} & 0 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix} \succeq 0.$

p^* with Elementary Methods

Minimize	x_{12}-t
subject to	$\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1+x_{12} \end{pmatrix} \succeq 0$

If there exists an $X \in \mathcal{L}$, then the upper-left 2×2 principal minor of the matrix in the example's condition must be positive semidefinite. Thus, the product of its eigenvalues (determinant) must be greater than zero.

In our case, $-x_{12}^2 \geqslant 0$, so $x_{12} = 0$. Hence $p^* = 0$.

d^* with Elementary Methods

$$\begin{array}{ll} \text{Maximize} & -\mu_4\text{-t} \\ \text{subject to} & \begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} & \mu_2 \\ \frac{1-\mu_4}{2} & 0 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix} \succeq 0. \end{array}$$

d^*: Based on the previous considerations, it must hold that
$\det \begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} \\ \frac{1-\mu_4}{2} & 0 \end{pmatrix} \geq 0.$

This only holds for $\mu_4 = 1$. So, $d^* = -1$.

Thus, indeed the weak duality theorem holds, so $p^* \geqslant d^*$ $(0 \geqslant -1)$.

However, there is no strong duality.

Strong Duality

Notation

$X \in \mathcal{L}_P^+$ if and only if $\langle A_i, X \rangle = b_i, i = 1, \ldots, \ell$ and $X \succ 0$ (i.e., X is positive definite).

Similarly defined is $\mathcal{L}_D^+$.

Strong Duality

If $\mathcal{L}_P^+, \mathcal{L}_D^+ \neq \emptyset$, then $p^* = d^*$.

Moreover, the set of optimal points is nonempty and compact.

Strong duality holds even with slightly weakened conditions.

Strong Duality

If $\mathcal{L}_P^+, \mathcal{L}_D \neq \emptyset$, then $p^* = d^*$.

We won't prove the theorems here.

This is the End!

Thank you for your attention!