CS 766/QIC 820 Theory of Quantum Information (Fall 2011)

Lecture 7: Semidefinite programming

This lecture is on semidefinite programming, which is a powerful technique from both an analytic and computational point of view. It is not a technique that is specific to quantum information, and in fact there will be almost nothing in this lecture that directly concerns quantum information, but we will later see that it has several very interesting applications to quantum information theory.

7.1 Definition of semidefinite programs and related terminology

We begin with a formal definition of the notion of a semidefinite program. Various terms connected with semidefinite programs are also defined.

There are two points regarding the definition of semidefinite programs that you should be aware of. The first point is that it represents just one of several formalizations of the semidefinite programming concept, and for this reason definitions found in other sources may differ from the one found here. (The differing formalizations do, however, lead to a common theory.) The second point is that semidefinite programs to be found in applications of the concept are typically not phrased in the precise form presented by the definition: some amount of massaging may be required to fit the semidefinite program to the definition.

These two points are, of course, related in that they concern variations in the forms of semidefinite programs. This issue will be discussed in greater detail later in the lecture, where conversions of semidefinite programs from one form to another are discussed. For now, however, let us consider that semidefinite programs are as given by the definition below.

Before proceeding to the definition, we need to define one term: a mapping $\Phi \in \mathrm{T}(\mathcal{X}, Y)$ is *Hermiticity preserving* if it holds that $\Phi(X) \in \mathrm{Herm}(\mathcal{Y})$ for all choices of $X \in \mathrm{Herm}(\mathcal{X})$. (This condition happens to be equivalent to the three conditions that appear in the third question of Assignment 1.)

Definition 7.1. A semidefinite program is a triple (Φ, A, B), where

1. $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is a Hermiticity-preserving linear map, and
2. $A \in \mathrm{Herm}(\mathcal{X})$ and $B \in \mathrm{Herm}(\mathcal{Y})$ are Hermitian operators,

for some choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$.

We associate with the triple (Φ, A, B) two optimization problems, called the *primal* and *dual* problems, as follows:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}). \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Phi^*(Y) \geq A, \\ & Y \in \mathrm{Herm}(\mathcal{Y}). \end{aligned}$$

The primal and dual problems have a special relationship to one another that will be discussed shortly.

An operator $X \in \text{Pos}(\mathcal{X})$ satisfying $\Phi(X) = B$ is said to be *primal feasible*, and we let $\mathcal{A}$ denote the set of all such operators:

$$\mathcal{A} = \{X \in \text{Pos}(\mathcal{X}) \,:\, \Phi(X) = B\}.$$

Following a similar terminology for the dual problem, an operator $Y \in \text{Herm}(\mathcal{Y})$ satisfying $\Phi^*(Y) \geq A$ is said to be *dual feasible*, and we let $\mathcal{B}$ denote the set of all dual feasible operators:

$$\mathcal{B} = \{Y \in \text{Herm}(\mathcal{Y}) \,:\, \Phi^*(Y) \geq A\}.$$

The linear functions $X \mapsto \langle A, X\rangle$ and $Y \mapsto \langle B, Y\rangle$ are referred to as the primal and dual *objective functions*. The *primal optimum* or *optimal primal value* of a semidefinite program is defined as

$$\alpha = \sup_{X \in \mathcal{A}} \langle A, X\rangle$$

and the *dual optimum* or *optimal dual value* is defined as

$$\beta = \inf_{Y \in \mathcal{B}} \langle B, Y\rangle.$$

The values α and β may be finite or infinite, and by convention we define $\alpha = -\infty$ if $\mathcal{A} = \varnothing$ and $\beta = \infty$ if $B = \varnothing$. If an operator $X \in \mathcal{A}$ satisfies $\langle A, X\rangle = \alpha$ we say that X is an *optimal primal solution*, or that X *achieves* the optimal primal value. Likewise, if $Y \in \mathcal{B}$ satisfies $\langle B, Y\rangle = \beta$ then we say that Y is an *optimal dual solution*, or that Y achieves the optimal dual value.

Example 7.2. A simple example of a semidefinite program may be obtained, for an arbitrary choice of a complex Euclidean space $\mathcal{X}$ and a Hermitian operator $A \in \text{Herm}(\mathcal{X})$, by taking $\mathcal{Y} = \mathbb{C}$, $B = 1$, and $\Phi(X) = \text{Tr}(X)$ for all $X \in \text{L}(\mathcal{X})$. The primal and dual problems associated with this semidefinite program are as follows:

<u>Primal problem</u>

$$\begin{aligned} \text{maximize:} \quad & \langle A, X\rangle \\ \text{subject to:} \quad & \text{Tr}(X) = 1, \\ & X \in \text{Pos}(\mathcal{X}). \end{aligned}$$

<u>Dual problem</u>

$$\begin{aligned} \text{minimize:} \quad & y \\ \text{subject to:} \quad & y\mathbb{1} \geq A, \\ & y \in \mathbb{R}. \end{aligned}$$

To see that the dual problem is as stated, we note that $\text{Herm}(\mathcal{Y}) = \mathbb{R}$ and that the adjoint mapping to the trace is given by $\text{Tr}^*(y) = y\mathbb{1}$ for all $y \in \mathbb{C}$. The optimal primal and dual values α and β happen to be equal (which is not unexpected, as we will soon see), coinciding with the largest eigenvalue $\lambda_1(A)$ of A.

There can obviously be no optimal primal solution to a semidefinite program when α is infinite, and no optimal dual solution when β is infinite. Even in cases where α and β are finite, however, there may not exist optimal primal and/or optimal dual solutions, as the following example illustrates.

Example 7.3. Let $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^2$, and define $A \in \text{Herm}(\mathcal{X})$, $B \in \text{Herm}(\mathcal{Y})$, and $\Phi \in \text{T}(\mathcal{X}, \mathcal{Y})$ as

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \Phi(X) = \begin{pmatrix} 0 & X(1,2) \\ X(2,1) & 0 \end{pmatrix}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It holds that $\alpha = 0$, but there does not exist an optimal primal solution to (Φ, A, B).

To establish this fact, suppose first that $X \in \mathrm{Pos}(\mathcal{X})$ is primal-feasible. The condition $\Phi(X) = B$ implies that X takes the form

$$X = \begin{pmatrix} X(1,1) & 1 \\ 1 & X(2,2) \end{pmatrix} \tag{7.1}$$

Given that X is positive semidefinite, it must hold that $X(1,1) \geq 0$ and $X(2,2) \geq 0$, because the diagonal entries of positive semidefinite operators are always nonnegative. Moreover, $\mathrm{Det}(X) = X(1,1)X(2,2) - 1$, and given that the determinant of every positive semidefinite operator is nonnegative it follows that $X(1,1)X(2,2) \geq 1$. It must therefore be the case that $X(1,1) > 0$, so that $\langle A, X \rangle < 0$. On the other hand, one may consider the operator

$$X_n = \begin{pmatrix} \frac{1}{n} & 1 \\ 1 & n \end{pmatrix}$$

for each positive integer n. It holds that $X_n \in \mathcal{A}$ and $\langle A, X_n \rangle = -1/n$, and therefore $\alpha \geq -1/n$, for every positive integer n. Consequently one has that $\alpha = 0$, while no primal-feasible X achieves this supremum value.

7.2 Duality

We will now discuss the special relationship between the primal and dual problems associated with a semidefinite program, known as *duality*. A study of this relationship begins with *weak duality*, which simply states that $\alpha \leq \beta$ for every semidefinite program.

Proposition 7.4 (Weak duality for semidefinite programs)**.** *For every semidefinite program (Φ, A, B) it holds that $\alpha \leq \beta$.*

Proof. The proposition is trivial in case $\mathcal{A} = \varnothing$ (which implies $\alpha = -\infty$) or $\mathcal{B} = \varnothing$ (which implies $\beta = \infty$), so we will restrict our attention to the case that both $\mathcal{A}$ and $\mathcal{B}$ are nonempty. For every primal feasible $X \in \mathcal{A}$ and dual feasible $Y \in \mathcal{B}$ it holds that

$$\langle A, X \rangle \leq \langle \Phi^*(Y), X \rangle = \langle Y, \Phi(X) \rangle = \langle Y, B \rangle = \langle B, Y \rangle .$$

Taking the supremum over all $X \in \mathcal{A}$ and the infimum over all $Y \in \mathcal{B}$ establishes that $\alpha \leq \beta$ as required. □

One implication of weak duality is that every dual-feasible operator $Y \in \mathcal{B}$ establishes an upper bound of $\langle B, Y \rangle$ on the optimal primal value α, and therefore an upper bound on $\langle A, X \rangle$ for every primal-feasible operator $X \in \mathcal{A}$. Likewise, every primal-feasible operator $X \in \mathcal{A}$ establishes a lower-bound of $\langle A, X \rangle$ on the optimal dual value β. In other words, it holds that

$$\langle A, X \rangle \leq \alpha \leq \beta \leq \langle B, Y \rangle ,$$

for every $X \in \mathcal{A}$ and $Y \in \mathcal{B}$. If one finds a primal-feasible operator $X \in \mathcal{A}$ and a dual-feasible operator $Y \in \mathcal{B}$ for which $\langle A, X \rangle = \langle B, Y \rangle$, it therefore follows that $\alpha = \beta$ and both X and Y must be optimal: $\alpha = \langle A, X \rangle$ and $\beta = \langle B, Y \rangle$.

The condition that $\alpha = \beta$ is known as *strong duality*. Unlike weak duality, strong duality does not hold for every semidefinite program, as the following example shows.

Example 7.5. Let $\mathcal{X} = \mathbb{C}^3$ and $\mathcal{Y} = \mathbb{C}^2$, and define

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad \Phi(X) = \begin{pmatrix} X(1,1) + X(2,3) + X(3,2) & 0 \\ 0 & X(2,2) \end{pmatrix}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The mapping Φ is Hermiticity-preserving and A and B are Hermitian, so (Φ, A, B) is a semidefinite program.

The primal problem associated with the semidefinite program (Φ, A, B) represents a maximization of $-X(1,1)$ subject to the constraints $X(1,1) + X(2,3) + X(3,2) = 1$, $X(2,2) = 0$, and $X \geq 0$. The constraints $X(2,2) = 0$ and $X \geq 0$ force the equality $X(2,3) = X(3,2) = 0$. It must therefore hold that $X(1,1) = 1$, so $\alpha \leq -1$. The fact that $\alpha = -1$, as opposed to $\alpha = -\infty$, is established by considering the primal feasible operator $X = E_{1,1}$.

To analyze the dual problem, we begin by noting that

$$\Phi^*(Y) = \begin{pmatrix} Y(1,1) & 0 & 0 \\ 0 & Y(2,2) & Y(1,1) \\ 0 & Y(1,1) & 0 \end{pmatrix}.$$

The constraint $\Phi^*(Y) \geq A$ implies that

$$\begin{pmatrix} Y(2,2) & Y(1,1) \\ Y(1,1) & 0 \end{pmatrix} \geq 0,$$

so that $Y(1,1) = 0$, and therefore $\beta \geq 0$. The fact that $\beta = 0$ is established by choosing the dual feasible operator $Y = 0$.

Thus, strong duality fails for this semidefinite program: it holds that $\alpha = -1$ while $\beta = 0$.

While strong duality does not hold for every semidefinite program, it does typically hold for semidefinite programs that arise in applications of the concept. Informally speaking, if one does not *try* to make strong duality fail, it will probably hold. There are various conditions on semidefinite programs that allow for an easy verification that strong duality holds (when it does), with one of the most useful conditions being given by the following theorem.

Theorem 7.6 (Slater's theorem for semidefinite programs)**.** *The following implications hold for every semidefinite program* (Φ, A, B).

1. *If $\mathcal{A} \neq \varnothing$ and there exists a Hermitian operator Y for which $\Phi^*(Y) > A$, then $\alpha = \beta$ and there exists a primal feasible operator $X \in \mathcal{A}$ for which $\langle A, X \rangle = \alpha$.*
2. *If $\mathcal{B} \neq \varnothing$ and there exists a positive semidefinite operator X for which $\Phi(X) = B$ and $X > 0$, then $\alpha = \beta$ and there exists a dual feasible operator $Y \in \mathcal{B}$ for which $\langle B, Y \rangle = \beta$.*

(The condition that $X \in \mathrm{Pd}(\mathcal{X})$ satisfies $\Phi(X) = B$ is called *strict primal feasibility*, while the condition that $Y \in \mathrm{Herm}(\mathcal{Y})$ satisfies $\Phi^*(Y) > A$ is called *strict dual feasibility*. In both cases, the "strictness" concerns the positive semidefinite ordering.)

There are two main ideas behind the proof of this theorem, and the proof (of each of the two implications) splits into two parts based on the two ideas. The ideas are as follows.

1. By the hyperplane separation theorem stated in the notes for Lecture 2, every closed convex set can be separated from any point not in that set by a hyperplane.

2. Linear mappings do not always map closed convex sets to closed sets, but under some conditions they do.

We will not prove the hyperplane separation theorem: you can find a proof in one of the references given in Lecture 2. That theorem is stated for a real Euclidean space of the form $\mathbb{R}^{\Sigma}$, but we will apply it to the space of Hermitian operators $\text{Herm}\left(\mathbb{C}^{\Gamma}\right)$ for some finite and nonempty set Γ. As we have already noted more than once, we may view $\text{Herm}\left(\mathbb{C}^{\Gamma}\right)$ as being isomorphic to $\mathbb{R}^{\Gamma\times\Gamma}$ as a real Euclidean space.

The second idea is quite vague as it has been stated above, so let us now state it more precisely as a lemma.

Lemma 7.7. *Let Σ and Γ be finite, nonempty sets, let $\Psi : \mathbb{R}^{\Sigma} \rightarrow \mathbb{R}^{\Gamma}$ be a linear mapping, and let $\mathcal{P} \subseteq \mathbb{R}^{\Sigma}$ be a closed convex cone possessing the property that $\ker(\Psi) \cap \mathcal{P}$ is a linear subspace of $\mathbb{R}^{\Sigma}$. It holds that $\Psi(\mathcal{P})$ is closed.*

To prove this lemma, we will make use of the following simple proposition, which we will take as given. It is straightforward to prove using basic concepts from analysis.

Proposition 7.8. *Let Σ be a finite and nonempty set and let $\mathcal{S} \subset \mathbb{R}^{\Sigma}$ be a compact subset of $\mathbb{R}^{\Sigma}$ such that $0 \notin \mathcal{S}$. It holds that $\text{cone}(\mathcal{S}) \triangleq \{\lambda v : v \in \mathcal{S},\, \lambda \geq 0\}$ is closed.*

Proof of Lemma 7.7. First we consider the special case that $\ker(\Psi) \cap \mathcal{P} = \{0\}$. Let

$$\mathcal{R} = \{v \in \mathcal{P} : \|v\| = 1\}.$$

It holds that $\mathcal{P} = \text{cone}(\mathcal{R})$ and therefore $\Psi(\mathcal{P}) = \text{cone}(\Psi(\mathcal{R}))$. The set $\mathcal{R}$ is compact, and therefore $\Psi(\mathcal{R})$ is compact as well. Moreover, it holds that $0 \notin \Psi(\mathcal{R})$, for otherwise there would exist a unit norm (and therefore nonzero) vector $v \in \mathcal{P}$ for which $\Psi(v) = 0$, contradicting the assumption that $\ker(\Psi) \cap \mathcal{P} = \{0\}$. It follows from Proposition 7.8 that $\Psi(\mathcal{P})$ is closed.

For the general case, let us denote $\mathcal{V} = \ker(\Psi) \cap \mathcal{P}$, and define $\mathcal{Q} = \mathcal{P} \cap \mathcal{V}^{\perp}$. It holds that $\mathcal{Q}$ is a closed convex cone, and $\ker(\Psi) \cap \mathcal{Q} = \mathcal{V} \cap \mathcal{V}^{\perp} = \{0\}$. We therefore have that $\Psi(\mathcal{Q})$ is closed by the analysis of the special case above. It remains to prove that $\Psi(\mathcal{Q}) = \Psi(\mathcal{P})$. To this end, choose $u \in \mathcal{P}$, and write $u = v + w$ for $v \in \mathcal{V}$ and $w \in \mathcal{V}^{\perp}$. Given that $\mathcal{V} \subseteq \mathcal{P}$ and $\mathcal{P}$ is a convex cone, it follows that $\mathcal{P} + \mathcal{V} = \mathcal{P}$, so $w = u - v \in \mathcal{P} + \mathcal{V} = \mathcal{P}$. Consequently $w \in \mathcal{Q}$. As $\Psi(w) = \Psi(u) - \Psi(v) = \Psi(u)$, we conclude that $\Psi(\mathcal{P}) \subseteq \Psi(\mathcal{Q})$. The reverse containment $\Psi(\mathcal{Q}) \subseteq \Psi(\mathcal{P})$ is trivial, and so we have proved $\Psi(\mathcal{P}) = \Psi(\mathcal{Q})$ as required. □

Now we are ready to prove Theorem 7.6. The two implications are proved in the same basic way, although there are technical differences in the proofs. Each implication is split into two lemmas, along the lines suggested above, which combine in a straightforward way to prove the theorem.

Lemma 7.9. *Let (Φ, A, B) be a semidefinite program. If $\mathcal{A} \neq \varnothing$ and the set*

$$\mathcal{K} = \left\{ \begin{pmatrix} \Phi(X) & 0 \\ 0 & \langle A, X\rangle \end{pmatrix} : X \in \text{Pos}\left(\mathcal{X}\right) \right\} \subseteq \text{Herm}\left(\mathcal{Y} \oplus \mathbb{C}\right)$$

is closed, then $\alpha = \beta$ and there exists a primal feasible operator $X \in \mathcal{A}$ such that $\langle A, X\rangle = \alpha$.

Proof. Let $\varepsilon > 0$ be chosen arbitrarily. Observe that the operator

$$\begin{pmatrix} B & 0 \\ 0 & \alpha + \varepsilon \end{pmatrix} \tag{7.2}$$

is not contained in $\mathcal{K}$, for there would otherwise exist an operator $X \in \text{Pos}(\mathcal{X})$ with $\Phi(X) = B$ and $\langle A, X \rangle > \alpha$, contradicting the optimality of α. The set $\mathcal{K}$ is convex and (by assumption) closed, and therefore there must exist a hyperplane that separates the operator (7.2) from $\mathcal{K}$ in the sense prescribed by Theorem 2.8 of Lecture 2. It follows that there must exist an operator $Y \in \text{Herm}(\mathcal{Y})$ and a real number λ such that

$$\langle Y, \Phi(X) \rangle + \lambda \langle A, X \rangle > \langle Y, B \rangle + \lambda(\alpha + \varepsilon) \tag{7.3}$$

for all $X \in \text{Pos}(\mathcal{X})$.

The set $\mathcal{A}$ has been assumed to be nonempty, so one may select an operator $X_0 \in \mathcal{A}$. As $\Phi(X_0) = B$, we conclude from (7.3) that

$$\lambda \langle A, X_0 \rangle > \lambda(\alpha + \varepsilon),$$

and therefore $\lambda < 0$. By dividing both sides of (7.3) by $|\lambda|$ and renaming variables, we see that there is no loss of generality in assuming that $\lambda = -1$. Substituting $\lambda = -1$ into (7.3) yields

$$\langle \Phi^*(Y) - A, X \rangle > \langle Y, B \rangle - (\alpha + \varepsilon) \tag{7.4}$$

for every $X \in \text{Pos}(\mathcal{X})$. The quantity on the right-hand-side of this inequality is a real number independent of X, which implies that $\Phi^*(Y) - A$ must be positive semidefinite; for if it were not, one could choose X appropriately to make the quantity on the left-hand-side smaller than $\langle Y, B \rangle - (\alpha + \varepsilon)$ (or any other fixed real number independent of X). It therefore holds that

$$\Phi^*(Y) \geq A,$$

so that Y is a dual-feasible operator. Setting $X = 0$ in (7.4) yields $\langle B, Y \rangle < \alpha + \varepsilon$. It has therefore been shown that for every $\varepsilon > 0$ there exists a dual feasible operator $Y \in \mathcal{B}$ such that $\langle B, Y \rangle < \alpha + \varepsilon$. This implies that $\alpha \leq \beta < \alpha + \varepsilon$ for every $\varepsilon > 0$, from which it follows that $\alpha = \beta$ as claimed.

To prove the second part of the lemma, a similar methodology to the first part of the proof is used, except that ε is set to 0. More specifically, we consider whether the operator

$$\begin{pmatrix} B & 0 \\ 0 & \alpha \end{pmatrix} \tag{7.5}$$

is contained in $\mathcal{K}$. If this operator is in $\mathcal{K}$, then there exists an operator $X \in \text{Pos}(\mathcal{X})$ such that $\Phi(X) = B$ and $\langle A, X \rangle = \alpha$, which is the statement claimed by the lemma.

It therefore suffices to derive a contradiction from the assumption that the operator (7.5) is not contained in $\mathcal{K}$. Under this assumption, there must exist a Hermitian operator $Y \in \text{Herm}(\mathcal{Y})$ and a real number λ such that

$$\langle Y, \Phi(X) \rangle + \lambda \langle A, X \rangle > \langle Y, B \rangle + \lambda\alpha \tag{7.6}$$

for all $X \in \text{Pos}(\mathcal{X})$. As before, one concludes from the existence of a primal feasible operator X_0 that $\lambda < 0$, so there is again no loss of generality in assuming that λ and Y are re-scaled so that $\lambda = -1$. After this re-scaling, one finds that

$$\langle \Phi^*(Y) - A, X \rangle > \langle Y, B \rangle - \alpha \tag{7.7}$$

for every $X \in \text{Pos}(\mathcal{X})$, and therefore $\Phi^*(Y) \geq A$ (i.e., Y is dual-feasible). Setting $X = 0$ in (7.7) implies $\langle Y, B \rangle < \alpha$. This, however, implies $\beta < \alpha$, which is in contradiction with weak duality. It follows that the operator (7.5) is contained in $\mathcal{K}$ as required. □

Lemma 7.10. *Let (Φ, A, B) be a semidefinite program. If there exists an operator $Y \in \text{Herm}(\mathcal{Y})$ for which $\Phi^*(Y) > A$, then the set*

$$\mathcal{K} = \left\{ \begin{pmatrix} \Phi(X) & 0 \\ 0 & \langle A, X \rangle \end{pmatrix} : X \in \text{Pos}(\mathcal{X}) \right\} \subseteq \text{Herm}(\mathcal{Y} \oplus \mathbb{C})$$

is closed.

Proof. The set $\text{Pos}(\mathcal{X})$ is a closed convex cone, and $\mathcal{K}$ is the image of this set under the linear mapping

$$\Psi(X) = \begin{pmatrix} \Phi(X) & 0 \\ 0 & \langle A, X \rangle \end{pmatrix}.$$

If $X \in \ker(\Psi)$, then $\Phi(X) = 0$ and $\langle A, X \rangle = 0$, and therefore

$$\langle \Phi^*(Y) - A, X \rangle = \langle Y, \Phi(X) \rangle - \langle A, X \rangle = 0.$$

If, in addition, it holds that $X \geq 0$, then $X = 0$ given that $\Phi^*(Y) - A > 0$. Thus, $\ker(\Psi) \cap \text{Pos}(\mathcal{X}) = \{0\}$, so $\mathcal{K}$ is closed by Lemma 7.7. □

The two lemmas above together imply that the first implication of Theorem 7.6 holds. The second implication is proved by combining the following two lemmas, which are closely related to the two lemmas just proved.

Lemma 7.11. *Let (Φ, A, B) be a semidefinite program. If $\mathcal{B} \neq \varnothing$ and the set*

$$\mathcal{L} = \left\{ \begin{pmatrix} \Phi^*(Y) - Z & 0 \\ 0 & \langle B, Y \rangle \end{pmatrix} : Y \in \text{Herm}(\mathcal{Y}),\, Z \in \text{Pos}(\mathcal{X}) \right\} \subseteq \text{Herm}(\mathcal{X} \oplus \mathbb{C})$$

is closed, then $\alpha = \beta$ and there exists a dual feasible operator $Y \in \mathcal{B}$ such that $\langle B, Y \rangle = \beta$.

Proof. Let $\varepsilon > 0$ be chosen arbitrarily. Along similar lines to the proof of Lemma 7.9, one observes that the operator

$$\begin{pmatrix} A & 0 \\ 0 & \beta - \varepsilon \end{pmatrix} \tag{7.8}$$

is not contained in $\mathcal{L}$; for if it were, there would exist an operator $Y \in \text{Herm}(\mathcal{Y})$ with $\Phi^*(Y) \geq A$ and $\langle B, Y \rangle < \beta$, contradicting the optimality of β. As the set $\mathcal{L}$ is closed (by assumption) and is convex, there must exist a hyperplane that separates the operator (7.8) from $\mathcal{L}$; that is, there must exist a real number λ and an operator $X \in \text{Herm}(\mathcal{X})$ such that

$$\langle X, \Phi^*(Y) - Z \rangle + \lambda \langle B, Y \rangle < \langle X, A \rangle + \lambda(\beta - \varepsilon) \tag{7.9}$$

for all $Y \in \text{Herm}(\mathcal{Y})$ and $Z \in \text{Pos}(\mathcal{X})$.

The set $\mathcal{B}$ has been assumed to be nonempty, so one may select a operator $Y_0 \in \mathcal{B}$. It holds that $\Phi^*(Y_0) \geq A$, and setting $Z = \Phi^*(Y_0) - A$ in (7.9) yields

$$\lambda \langle B, Y_0 \rangle < \lambda(\beta - \varepsilon),$$

implying $\lambda < 0$. There is therefore no loss of generality in re-scaling λ and X in (7.9) so that $\lambda = -1$, which yields

$$\langle \Phi(X) - B, Y \rangle < \langle X, A + Z \rangle - (\beta - \varepsilon) \tag{7.10}$$

for every $Y \in \text{Herm}(\mathcal{Y})$ and $Z \in \text{Pos}(\mathcal{X})$. The quantity on the right-hand-side of this inequality is a real number independent of Y, which implies that $\Phi(X) = B$; for if this were not so, one could choose a Hermitian operator $Y \in \text{Herm}(\mathcal{Y})$ appropriately to make the quantity on the left-hand-side larger than $\langle X, A + Z \rangle - (\beta - \varepsilon)$ (or any other real number independent of Y). It therefore holds that X is a primal-feasible operator. Setting $Y = 0$ and $Z = 0$ in (7.10) yields $\langle A, X \rangle > \beta - \varepsilon$. It has therefore been shown that for every $\varepsilon > 0$ there exists a primal feasible operator $X \in \mathcal{A}$ such that $\langle A, X \rangle > \beta - \varepsilon$. This implies $\beta \geq \alpha > \beta - \varepsilon$ for every $\varepsilon > 0$, and therefore $\alpha = \beta$ as claimed.

To prove the second part of the lemma, we may again use essentially the same methodology as for the first part, but setting $\varepsilon = 0$. That is, we consider whether the operator

$$\begin{pmatrix} A & 0 \\ 0 & \beta \end{pmatrix} \tag{7.11}$$

is in $\mathcal{L}$. If so, there exists an operator $Y \in \text{Herm}(\mathcal{Y})$ for which $\Phi^*(Y) - Z = A$ for some $Z \in \text{Pos}(\mathcal{X})$ (i.e., for which $\Phi^*(Y) \geq A$) and for which $\langle B, Y \rangle = \beta$, which is the statement we aim to prove.

It therefore suffices to derive a contradiction from the assumption the operator (7.11) is not in $\mathcal{L}$. Under this assumption, there must exist a real number λ and a Hermitian operator $X \in \text{Herm}(\mathcal{X})$ such that

$$\langle X, \Phi^*(Y) - Z \rangle + \lambda \langle B, Y \rangle < \langle X, A \rangle + \lambda\beta \tag{7.12}$$

for all $Y \in \text{Herm}(\mathcal{Y})$ and $Z \in \text{Pos}(\mathcal{X})$. We conclude, as before, that it must hold that $\lambda < 0$, and so there is no loss of generality in assuming $\lambda = -1$. Moreover, we have

$$\langle \Phi(X) - B, Y \rangle < \langle X, A + Z \rangle - \beta$$

for every $Y \in \text{Herm}(\mathcal{Y})$ and $Z \in \text{Pos}(\mathcal{X})$, and therefore $\Phi(X) = B$. Finally, taking $Y = 0$ and $Z = 0$ implies $\langle A, X \rangle > \beta$. This, however, implies $\alpha > \beta$, which is in contradiction with weak duality. It follows that the operator (7.11) is contained in $\mathcal{L}$ as required. □

Lemma 7.12. *Let (Φ, A, B) be a semidefinite program. If there exists a operator $X \in \text{Pos}(\mathcal{X})$ for which $\Phi(X) = B$ and $X > 0$, then the set*

$$\mathcal{L} = \left\{ \begin{pmatrix} \Phi^*(Y) - Z & 0 \\ 0 & \langle B, Y \rangle \end{pmatrix} : Y \in \text{Herm}(\mathcal{Y}), Z \in \text{Pos}(\mathcal{X}) \right\} \subseteq \text{Herm}(\mathcal{X} \oplus \mathbb{C})$$

is closed.

Proof. The set

$$\mathcal{P} = \left\{ \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} : Y \in \text{Herm}(\mathcal{Y}), Z \in \text{Pos}(\mathcal{X}) \right\}$$

is a closed, convex cone. For the linear map

$$\Psi \begin{pmatrix} Y & \cdot \\ \cdot & Z \end{pmatrix} = \begin{pmatrix} \Phi^*(Y) - Z & 0 \\ 0 & \langle B, Y \rangle \end{pmatrix}$$

it holds that $\mathcal{L} = \Psi(\mathcal{P})$.

Suppose that

$$\begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix} \in \ker(\Psi) \cap \mathcal{P}.$$

It must then hold that

$$0 = \langle B, Y \rangle - \langle X, \Phi^*(Y) - Z \rangle = \langle B - \Phi(X), Y \rangle + \langle X, Z \rangle = \langle X, Z \rangle,$$

for X being the positive definite operator assumed by the statement of the lemma, implying that $Z = 0$. It follows that

$$\ker(\Psi) \cap \mathcal{P} = \left\{ \begin{pmatrix} Y & 0 \\ 0 & 0 \end{pmatrix} : Y \in \{B\}^\perp \cap \ker(\Phi^*) \right\},$$

which is a linear subspace of $\text{Herm}(\mathcal{Y} \oplus \mathcal{X})$. It follows that $\mathcal{L}$ is closed by Lemma 7.7. □

7.3 Alternate forms of semidefinite programs

As was mentioned earlier in the lecture, semidefinite programs can be specified in ways that differ from the formal definition we have been considering thus far in the lecture. A few such ways will now be discussed.

7.3.1 Semidefinite programs with inequality constraints

Suppose $\Phi \in \text{T}(\mathcal{X}, \mathcal{Y})$ is a Hermiticity-preserving map and $A \in \text{Herm}(\mathcal{X})$ and $B \in \text{Herm}(\mathcal{Y})$ are Hermitian operators, and consider this optimization problem:

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \Phi(X) \leq B, \\ & X \in \text{Pos}(\mathcal{X}). \end{aligned}$$

It is different from the primal problem associated with the triple (Φ, A, B) earlier in the lecture because the constraint $\Phi(X) = B$ has been replaced with $\Phi(X) \leq B$. There is now more freedom in the valid choices of $X \in \text{Pos}(\mathcal{X})$ in this problem, so its optimum value may potentially be larger.

It is not difficult, however, to phrase the problem above using an equality constraint by using a so-called *slack variable*. That is, the inequality constraint

$$\Phi(X) \leq B$$

is equivalent to the equality constraint

$$\Phi(X) + Z = B \qquad (\text{for some } Z \in \mathrm{Pos}(\mathcal{Y})).$$

With this in mind, let us define a new semidefinite program, by which we mean something that conforms to the precise definition given at the beginning of the lecture, as follows. First, let $\Psi \in \mathrm{T}(\mathcal{X} \oplus \mathcal{Y}, \mathcal{Y})$ be defined as

$$\Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = \Phi(X) + Z$$

for all $X \in \mathrm{L}(\mathcal{X})$ and $Z \in \mathrm{L}(\mathcal{Y})$. (The dots indicate elements of $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $\mathrm{L}(\mathcal{Y}, \mathcal{X})$ that we don't care about, because they don't influence the output of the mapping Ψ.) Also define $C \in \mathrm{Herm}(\mathcal{X} \oplus \mathcal{Y})$ as

$$C = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$

The primal and dual problems associated with the semidefinite program (Ψ, C, B) are as follows.

<u>Primal problem</u>

$$\begin{aligned} \text{maximize:} \quad & \left\langle \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} \right\rangle \\ \text{subject to:} \quad & \Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = B, \\ & \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}). \end{aligned}$$

<u>Dual problem</u>

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Psi^*(Y) \geq \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \\ & Y \in \mathrm{Herm}(\mathcal{Y}). \end{aligned}$$

Once again, we are using dots to represent operators we don't care about. The fact that we don't care about these operators in this case is a combination of the fact that Ψ is independent of them and the fact that the objective function is independent of them as well. (Had we made a different choice of C, this might not be so.)

The primal problem simplifies to the problem originally posed above. To simplify the dual problem, we must first calculate Ψ^*. It is given by

$$\Psi^*(Y) = \begin{pmatrix} \Phi^*(Y) & 0 \\ 0 & Y \end{pmatrix}.$$

To verify that this is so, we simply check the required condition:

$$\left\langle \begin{pmatrix} X & W \\ V & Z \end{pmatrix}, \Psi^*(Y) \right\rangle = \langle X, \Phi^*(Y) \rangle + \langle Z, Y \rangle = \langle \Phi(X) + Z, Y \rangle = \left\langle \Psi \begin{pmatrix} X & W \\ V & Z \end{pmatrix}, Y \right\rangle,$$

for all $X \in \mathrm{L}(\mathcal{X})$, $Y, Z \in \mathrm{L}(\mathcal{Y})$, $V \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, and $W \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$. This condition uniquely determines Ψ^*, so we know we have it right. The inequality

$$\begin{pmatrix} \Phi^*(Y) & 0 \\ 0 & Y \end{pmatrix} \geq \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix},$$

for $Y \in \mathrm{Herm}(\mathcal{Y})$, is equivalent to $\Phi^*(Y) \geq A$ and $Y \geq 0$ (i.e., $Y \in \mathrm{Pos}(\mathcal{Y})$). So, we may simplify the problems above so that they look like this:

$$\begin{array}{ll} \underline{\text{Primal problem}} & \\ \text{maximize:} & \langle A, X\rangle \\ \text{subject to:} & \Phi(X) \leq B, \\ & X \in \operatorname{Pos}(\mathcal{X}). \end{array} \qquad \begin{array}{ll} \underline{\text{Dual problem}} & \\ \text{minimize:} & \langle B, Y\rangle \\ \text{subject to:} & \Phi^*(Y) \geq A, \\ & Y \in \operatorname{Pos}(\mathcal{Y}). \end{array}$$

This is an attractive form, because the primal and dual problems have a nice symmetry between them.

Note that we could equally well convert a primal problem having an equality constraint into one with an inequality constraint, by using the simple fact that $\Phi(X) = B$ if and only if

$$\begin{pmatrix} \Phi(X) & 0 \\ 0 & -\Phi(X) \end{pmatrix} \leq \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}.$$

So, it would have been alright had we initially defined the primal and dual problems associated with (Φ, A, B) to be the ones with inequality constraints as just discussed: one can convert back and forth between the two forms. (Note, however, that we are better off with Slater's theorem for semidefinite programs with equality constraints than we would be for a similar theorem for inequality constraints, which is why we have elected to start with equality constraints.)

7.3.2 Equality and inequality constraints

It is sometimes convenient to consider semidefinite programming problems that include both equality and inequality constraints, as opposed to just one type. To be more precise, let $\mathcal{X}$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$ be complex Euclidean spaces, let $\Phi_1 : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y}_1)$ and $\Phi_2 : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y}_2)$ be Hermiticity-preserving maps, let $A \in \operatorname{Herm}(\mathcal{X})$, $B_1 \in \operatorname{Herm}(\mathcal{Y}_1)$, and $B_2 \in \operatorname{Herm}(\mathcal{Y}_2)$ be Hermitian operators, and consider these optimization problems:

$$\begin{array}{ll} \underline{\text{Primal problem}} & \\ \text{maximize:} & \langle A, X\rangle \\ \text{subject to:} & \Phi_1(X) = B_1, \\ & \Phi_2(X) \leq B_2, \\ & X \in \operatorname{Pos}(\mathcal{X}). \end{array} \qquad \begin{array}{ll} \underline{\text{Dual problem}} & \\ \text{minimize:} & \langle B_1, Y_1\rangle + \langle B_2, Y_2\rangle \\ \text{subject to:} & \Phi_1^*(Y_1) + \Phi_2^*(Y_2) \geq A, \\ & Y_1 \in \operatorname{Herm}(\mathcal{Y}_1), \\ & Y_2 \in \operatorname{Pos}(\mathcal{Y}_2). \end{array}$$

The fact that these problems really are dual may be verified in a similar way to the discussion of inequality constraints in the previous subsection. Specifically, one may define a linear mapping

$$\Psi : \operatorname{Herm}(\mathcal{X} \oplus \mathcal{Y}_2) \to \operatorname{Herm}(\mathcal{Y}_1 \oplus \mathcal{Y}_2)$$

as

$$\Psi \begin{pmatrix} X & \cdot \\ \cdot & Z \end{pmatrix} = \begin{pmatrix} \Phi_1(X) & 0 \\ 0 & \Phi_2(X) + Z \end{pmatrix}$$

for all $X \in \mathrm{L}(\mathcal{X})$ and $Z \in \mathrm{L}(\mathcal{Y}_2)$, and define Hermitian operators $C \in \operatorname{Herm}(\mathcal{X} \oplus \mathcal{Y}_2)$ and $D \in \operatorname{Herm}(\mathcal{Y}_1 \oplus \mathcal{Y}_2)$ as

$$C = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}.$$

The primal problem above is equivalent to the primal problem associated with the semidefinite program (Ψ, C, D). The dual problem above coincides with the dual problem of (Ψ, C, D), by virtue of the fact that

$$\Psi^* \begin{pmatrix} Y_1 & \cdot \\ \cdot & Y_2 \end{pmatrix} = \begin{pmatrix} \Phi_1^*(Y_1) + \Phi_2^*(Y_2) & 0 \\ 0 & Y_2 \end{pmatrix}$$

for every $Y_1 \in \text{Herm}(\mathcal{Y}_1)$ and $Y_2 \in \text{Herm}(\mathcal{Y}_2)$.

Note that strict primal feasibility of (Ψ, C, D) is equivalent to the condition that $\Phi_1(X) = B_1$ and $\Phi_2(X) < B_2$ for some choice of $X \in \text{Pd}(\mathcal{X})$, while strict dual feasibility of (Ψ, C, D) is equivalent to the condition that $\Phi_1^*(Y_1) + \Phi_2^*(Y_2) > A$ for some choice of $Y \in \text{Herm}(\mathcal{Y}_1)$ and $Y_2 \in \text{Pd}(\mathcal{Y}_2)$. In other words, the "strictness" once again refers to the positive semidefinite ordering—every time it appears.

7.3.3 The standard form

The so-called *standard form* for semidefinite programs is given by the following pair of optimization problems:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \langle B_1, X \rangle = \gamma_1 \\ & \vdots \\ & \langle B_m, X \rangle = \gamma_m \\ & X \in \text{Pos}(\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \textstyle\sum_{j=1}^m \gamma_j y_j \\ \text{subject to:} \quad & \textstyle\sum_{j=1}^m y_j B_j \geq A \\ & y_1, \ldots, y_m \in \mathbb{R} \end{aligned}$$

Here, $B_1, \ldots, B_m \in \text{Herm}(\mathcal{X})$ take the place of Φ and $\gamma_1, \ldots, \gamma_m \in \mathbb{R}$ take the place of B in semidefinite programs as we have defined them.

It is not difficult to show that this form is equivalent to our form. First, to convert a semidefinite program in the standard form to our form, we define $\mathcal{Y} = \mathbb{C}^m$,

$$\Phi(X) = \sum_{j=1}^m \langle B_j, X \rangle E_{j,j},$$

and

$$B = \sum_{j=1}^m \gamma_j E_{j,j}.$$

The primal problem above is then equivalent to a maximization of $\langle A, X \rangle$ over all $X \in \text{Pos}(\mathcal{X})$ satisfying $\Phi(X) = B$. The adjoint of Φ is given by

$$\Phi^*(Y) = \sum_{j=1}^m Y(j,j) B_j,$$

and we have that

$$\langle B, Y \rangle = \sum_{j=1}^m \gamma_j Y(j,j).$$

The off-diagonal entries of Y are irrelevant for the sake of this problem, and we find that a minimization of $\langle B, Y \rangle$ subject to $\Phi^*(Y) \geq A$ is equivalent to the dual problem given above.

Working in the other direction, the equality constraint $\Phi(X) = B$ may be represented as

$$\langle H_{a,b}, \Phi(X) \rangle = \langle H_{a,b}, B \rangle,$$

ranging over the Hermitian operator basis $\{H_{a,b} : a, b \in \Gamma\}$ of $\mathrm{L}(\mathcal{Y})$ defined in Lecture 1, where we have assumed that $\mathcal{Y} = \mathbb{C}^{\Gamma}$. Taking $B_{a,b} = \Phi^*(H_{a,b})$ and $\gamma_{a,b} = \langle H_{a,b}, B \rangle$ allows us to write the primal problem associated with (Φ, A, B) as a semidefinite program in standard form. The standard-form dual problem above simplifies to the dual problem associated with (Φ, A, B).

The standard form has some positive aspects, but for the semidefinite programs to be encountered in this course we will find that it is less convenient to use than the forms we discussed previously.