Lecture 1

Course overview and semidefinite programming basics

This is a seminar-style course centered around the use of semidefinite programming in quantum information and computation.

The notes for this first lecture will be shorter than for typical lectures: the lecture started with an overview of the course, a summary of the expected background knowledge for students taking the course, and a review of linear algebra notation and terminology to be used to discuss the basics of semidefinite programming, all of which are described elsewhere. In particular, the logistics of the course are described in handouts posted on the course web page, and for a summary of linear algebra notation and terminology I will refer you to Chapter 1 my book Theory of Quantum Information, which is freely available at this URL:

https://cs.uwaterloo.ca/~watrous/TQI/

(For this lecture and the next, the material in Sections 1.1.1 and 1.1.2 is sufficient. Some of the material in these first two lectures is also discussed in Section 1.2.2.) The particular notation and terminology used in this book is common in mathematics, but it has some differences from what is most common in physics and quantum information and computation. (For instance, I generally avoid the Dirac notation and prefer to represent adjoints with a star rather than a dagger.)

Regarding background knowledge, students taking this course are assumed to have already taken an introductory course in quantum information and computation, or to have comparable background knowledge through self-study. We will frequently make use of the mathematical formalism through which quantum states are represented by density operators, measurements are represented by collections of positive semidefinite operators summing to the identity, and quantum channels are represented by completely positive and trace-preserving maps, so a familiarity with this formalism is required.
1.1 Semidefinite programming definitions

A *semidefinite program* is specified by a triple \((\Phi, A, B)\), where \(\Phi \in T(\mathcal{X}, \mathcal{Y})\) is a Hermitian-preserving map and \(A \in \text{Herm}(\mathcal{X})\) and \(B \in \text{Herm}(\mathcal{Y})\) are Hermitian operators, for some choice of complex Euclidean spaces \(\mathcal{X}\) and \(\mathcal{Y}\) (which you may take as \(\mathcal{X} = \mathbb{C}^n\) and \(\mathcal{Y} = \mathbb{C}^m\) if you like). With this triple we associate the following pair of optimization problems:

<table>
<thead>
<tr>
<th>Primal problem</th>
<th>Dual problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>maximize: (\langle A, X \rangle)</td>
<td>minimize: (\langle B, Y \rangle)</td>
</tr>
<tr>
<td>subject to: (\Phi(X) = B,) (X \in \text{Pos}(\mathcal{X}).)</td>
<td>subject to: (\Phi^*(Y) \geq A,) (Y \in \text{Herm}(\mathcal{Y}).)</td>
</tr>
</tbody>
</table>

These problems should be viewed as computational tasks, where the aim is to maximize the value \(\langle A, X \rangle\) or minimize the value \(\langle B, Y \rangle\), both of which are necessarily real numbers, subject to the indicated constraints. The following terminology is used when discussing these problems:

1. *Feasible sets*. One defines the *primal feasible* set \(\mathcal{A}\) as

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

and the *dual feasible* set \(\mathcal{B}\) as

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

Operators \(X \in \mathcal{A}\) and \(Y \in \mathcal{B}\) are also said to be *primal feasible* and *dual feasible*, respectively.

2. *Objective functions*. The function \(X \mapsto \langle A, X \rangle\), from \(\text{Herm}(\mathcal{X})\) to \(\mathbb{R}\), is the *primal objective function*, while the function \(Y \mapsto \langle B, Y \rangle\), from \(\text{Herm}(\mathcal{Y})\) to \(\mathbb{R}\), is the *dual objective function* of \((\Phi, A, B)\).

3. *Optimal values*. The *optimal values* associated with the primal and dual problems are defined as

\[
\alpha = \sup \{ \langle A, X \rangle : X \in \mathcal{A} \} \quad \text{and} \quad \beta = \inf \{ \langle B, Y \rangle : Y \in \mathcal{B} \},
\]

respectively. If it is the case that \(\mathcal{A} = \emptyset\) or \(\mathcal{B} = \emptyset\), then one defines \(\alpha = -\infty\) and \(\beta = \infty\), respectively. It is also possible that \(\alpha = \infty\) or \(\beta = -\infty\).
1.2 A simple example

Suppose $H \in \text{Herm}(\mathbb{C}^n)$ is an arbitrarily chosen Hermitian operator, and consider the semidefinite program $(\Phi, A, B)$ obtained by setting $X = \mathbb{C}^n$, $y = \mathbb{C}$, and

$$\Phi = \text{Tr}, \quad A = H, \quad B = 1. \quad (1.4)$$

(One is viewing the trace defined over $X$ as a mapping of the form $\text{Tr} \in T(X, \mathbb{C})$, which is reasonable because it is linear and we can identify $L(\mathbb{C})$ with $\mathbb{C}$.)

The primal and dual problems associated with this semidefinite program, after being simplified slightly, are as follows:

**Primal problem**

maximize: $\langle H, X \rangle$

subject to: $\text{Tr}(X) = 1$, $X \in \text{Pos}(X)$.

**Dual problem**

minimize: $y$

subject to: $y\mathbb{1} \geq H$, $y \in \mathbb{R}$.

To see that the dual problem is as stated, we note that $\text{Herm}(\mathbb{C}) = \mathbb{R}$ and that the adjoint mapping to the trace (with respect to $X$) is given by $\text{Tr}^*(y) = y\mathbb{1}_X$ for all $y \in \mathbb{C}$. This may be verified through the equality

$$\langle y\mathbb{1}_X, X \rangle = \langle y, \text{Tr}(X) \rangle, \quad (1.5)$$

which evidently holds for all $X \in L(X)$ and $y \in L(\mathbb{C}) = \mathbb{C}$.

The optimal primal and dual values $\alpha$ and $\beta$ happen to be equal (which is not unexpected, as we will soon see), coinciding with the largest eigenvalue of $H$. This is not too hard to verify, in both cases, by considering a spectral decomposition

$$H = \sum_{k=1}^{n} \lambda_k x_k x_k^*. \quad (1.6)$$