SDP in Quantum Information

Spring 2011, U. Waterloo.
Ashwin Nayak
May 3, 2011

Consider the following classic problem in QI

<u>State distinguishability</u> (All Hilbert spaces – finite dim.)

<u>Given</u> a single copy of quantum state $\rho_i \in L(\mathcal{H})$, with probability p_i, $i = 1, \dots, n$, what measurement gives us the maximum probability of correctly identifying the given state?

<u>Goal</u>: Find a POVM $\{E_i\}_{i=1}^{n}$ such that $\sum_{i=1}^{n} p_i \, \mathrm{Tr}(E_i \rho_i)$

is maximized

Mathematically we may state this as an optimization problem as follows:

$$\sup \sum_{i=1}^{n} p_i \, \mathrm{Tr}(E_i \rho_i)$$

$$\text{subject to:} \quad \sum_{i=1}^{n} E_i = \mathbb{1}$$

$$E_i \geq 0, \quad i = 1, \dots, n.$$

The variables here are the E_i, all PSD matrices that sum to $\mathbb{1}$ on $\mathcal{H}$ — linear constraints & linear obj. fn.
Notation: $A \geq B$ if $A - B$ is PSD; $A > B$ if $A - B$ is positive definite.

This is an **instance** of a semi-definite program (SDP).
In general, an SDP looks like so:

$$\boxed{\begin{array}{lll} \inf & \mathrm{Tr}(C^\dagger X) & \\ \text{subject to} & \Phi(X) \geq B & (P) \quad \underline{\text{primal}} \\ & X \geq 0 & \end{array}}$$

where X is the variable over which we optimize, and $X \in L(\mathcal{H})$ ($\dim(\mathcal{H}) = N$, say), C is a given Hermitian matrix in $L(\mathcal{H})$, (finite dim) $\Phi : L(\mathcal{H}) \to L(\mathcal{K})$, ($\dim(\mathcal{K}) = M$ say), maps Hermitian matrices to Hermitian matrices and is linear, $B \in L(\mathcal{K})$ is Hermitian.

Note:
- (0) Since $X \geq 0$, & C is Hermitian, obj fn is real.
- (i) $\mathrm{Tr}(A^\dagger B) = \langle A, B \rangle$ is the standard inner product on $L(\mathcal{H})$.
- (ii) This allows for complex SD matrices and is in a form ~~most~~ suitable for our applications, but is equivalent to the "standard form" in the literature (which deals with real SDP). (HW 1).
- (iii) LPs are a special case of SDPs, where all matrices are diagonal

$\boxed{\text{Example}}$: <u>State distinguishability</u>: we may bring it in the form above

$$\sum_{i=1}^n E_i \geq \mathbb{1} \quad \text{and} \; -\sum_{i=1}^n E_i \geq -\mathbb{1} \quad \Leftrightarrow \quad \sum_{i=1}^n E_i = \mathbb{1}.$$

$$X = \begin{bmatrix} E_1 & & ? \\ & E_2 & \\ & ? & \ddots \\ & & & E_n \end{bmatrix}$$

? → may assume $\emptyset$, as these do not occur in the constraints or obj-fn value; variable over $\oplus^n \mathcal{H}$

$$\Phi(X) = \begin{bmatrix} \sum_{i=1}^n E_i & \emptyset \\ \emptyset & -\sum_{i=1}^n E_i \end{bmatrix} \qquad B = \begin{bmatrix} \mathbb{1} & \emptyset \\ \emptyset & -\mathbb{1} \end{bmatrix}$$

maps $\bigoplus_{i=1}^n \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}$

$$C = -\begin{bmatrix} p_1 \rho_1 & & \emptyset \\ & \ddots & \\ \emptyset & & p_n \rho_n \end{bmatrix}$$

(negative so that sup $\Leftrightarrow$ inf)

(as in any SDP)
In our example, the set of feasible solutions is convex, and the objective function is linear. Further, the feasible region is closed and bounded. Therefore, the optimum exists, and is achieved. This need not always be the case.

<u>Example</u> (i) $\inf -x$

such that $x \geq 0$

is clearly unbdd. (<u>Like</u> LPs, SDPs may be infeasible, or when feasible, the opt may be unbdd.)

(ii) $\inf x_1$

subject to $\begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} \geq \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

& $\begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix} \geq 0$

has optimum $x_1 = 0$, however, this is not achieved.
(The constraints are $x_1 \geq 0$, $x_2 \geq 0$ & $x_1 \cdot x_2 \geq 1$).

<u>Unlike</u> LPs, when the opt is finite, it is not necessarily achieved.
What does the SDP tell us about state distinguishability?
We can establish certain linear algebraic properties of the optimum measurement by considering the <u>dual</u> SDP. For the general SDP (P) above, the dual is given by

$$\begin{aligned} &\sup \langle B, Y \rangle \\ &\text{subject to } \Phi^*(Y) \leq C \qquad (D) \ \underline{\text{dual}} \\ &\qquad Y \geq 0 \end{aligned}$$

where the variable $Y \in L(K)$, $\Phi^*: L(K) \to L(H)$ is the adjoint of the linear operator Φ, and B and C are as above.

Note: This is also an SDP equivalent to one of the form given by the primal. We can convert the sup to an inf by taking ~~the~~ negatives. Same with the direction of the inequality. Need to verify that Φ^* maps Hermitian matrices to Hermitian matrices (HW).

Example State distinguishability

The dual program is given by a variable Y over $L(\mathcal{H}\oplus\mathcal{H})$

$$\Phi^* : L(\mathcal{H}\oplus\mathcal{H}) \longrightarrow L(\oplus^n \mathcal{H})$$

By definition,
$$\langle \Phi^*(Y), X\rangle = \langle Y, \Phi(X)\rangle$$

$$= \mathrm{Tr}\begin{bmatrix} Y_1 & ? \\ ? & Y_2 \end{bmatrix}^{+} \Phi\begin{bmatrix} X_1 & & ? \\ & \ddots & \\ ? & & X_n \end{bmatrix}$$

$$= \mathrm{Tr}\begin{bmatrix} Y_1 & ? \\ ? & Y_2 \end{bmatrix}^{+} \begin{bmatrix} \sum_{i=1}^{n} X_i & 0 \\ 0 & -\sum_i X_i \end{bmatrix}$$

$$= \mathrm{Tr}\left[Y_1^{+} \sum_i X_i\right] - \mathrm{Tr}\left(Y_2^{+} \sum_i X_i\right]$$

$$= \mathrm{Tr}\left((Y_1 - Y_2)^{+} \sum_i X_i\right)$$

$$= \mathrm{Tr}\begin{bmatrix} Y_1 - Y_2 & & 0 \\ & Y_1 - Y_2 & \\ 0 & & \ddots \\ & & & Y_1 - Y_2 \end{bmatrix} \begin{bmatrix} X_1 & & ? \\ & \ddots & \\ ? & & X_n \end{bmatrix}$$

So
$$\Phi^*\begin{bmatrix} Y_1 & ? \\ ? & Y_2 \end{bmatrix} = \begin{bmatrix} Y_1 - Y_2 & & 0 \\ & Y_1 - Y_2 & \\ 0 & & \ddots \\ & & & Y_1 - Y_2 \end{bmatrix}$$

So the dual is (equivalent to) :

$$\begin{aligned} &\sup \quad \mathrm{Tr}(Y_1 - Y_2) \\ &\text{subject to} \quad Y_1 - Y_2 \leq -p_i \rho_i \quad \forall\, i = 1, \dots, n \\ &\qquad\qquad\quad Y_1 \geq 0,\ Y_2 \geq 0 \end{aligned}$$

(Can be simplified further...)

What does the dual tell us about the original problem?
Under certain conditions, the optima are equal. Moreover, the optimum feasible solutions are related to each other by complementary slackness. This helps us characterize optimal measurement in some cases.

Theorem

(i) Weak duality : If X & Y are feasible for P & D respectively, then $\langle C, X\rangle \geq \langle B, Y\rangle$. (Primal & Dual optima are finite & primal opt $\geq$ dual opt.)

(Slater)
(ii) Strong duality : If, further, the dual has a strictly feasible solution Y (i.e. $Y > 0$ s.t. $\Phi^*(Y) < C$), then the primal opt is achieved, and equals the dual opt.

(Since the dual of the dual is the primal, a symmetric condition in terms of a primal strictly feasible solution may be stated.)

Note: The strictly feasible solution is called a Slater point.

Let's see weak duality; proof of (ii) in refs.

Complementary slackness:

(iii) If X, Y are primal and dual opt. solutions, then
$$Y(\Phi(X) - C) = 0 = (\Phi(X) - C) \cdot Y$$
$$\& \quad X(B - \Phi^*(Y)) = 0 = (B - \Phi^*(Y)) \cdot X$$
iff primal opt = dual opt.

Proof :

Proposition
(c) If $A, B \geq 0$, $\mathrm{Tr}(AB) = 0$ iff $AB = 0 = BA$.
(b) If $A \geq B$, and $C \geq 0$, then $\langle A, C\rangle \geq \langle B, C\rangle$.
(a) $\langle A,B\rangle = \langle B,A\rangle$, if A,B are Hermitian.

$$\begin{aligned}
\langle C, X\rangle &\geq \langle \Phi^*(Y), X\rangle, && \text{since } X \geq 0 \ \& \ \Phi^*(Y) \leq C. \\
&= \langle Y, \Phi(X)\rangle, && \text{by defn of } \Phi^* \\
&= \langle \Phi(X), Y\rangle && \text{(by (ii) above)} \\
&\geq \langle B, Y\rangle && \text{since } Y \geq 0 \ \& \ \Phi(X) \geq B.
\end{aligned}$$

(i) ∎

(iii) follows from the proof of (i) & (c) above. ∎

Example State distinguishability. Primal is feasible: take $E_1 = \mathbb{1}$, $E_i = 0$, $\forall\, i > 1$. Dual is feasible: take $Y_2 = \mathbb{1}$, $Y_1 = 0$.
Dual has a Slater point: take $\begin{cases} Y_2 = +2\mathbb{1} \\ Y_1 = \frac{1}{2}\mathbb{1} \end{cases}$, so strong duality holds.

Complementary slackness conditions: If (E_i) and (Y_1, Y_2) are optimal feasible solutions, then: (let $Y_0 = Y_2 - Y_1$).

$$Y_j\Big(\sum_i E_i - \mathbb{1}\Big) = 0 = \Big(\sum_i E_i - \mathbb{1}\Big)Y_j, \quad j = 1, 2.$$

(trivial, as $\sum_i E_i = \mathbb{1}$ for any feasible soln)

and $E_i(p_i\rho_i - Y_0) = 0 = (p_i\rho_i - Y_0)E_i \quad \forall\, i = 1, 2, \dots, n.$

(★) $\begin{cases} E_i(\rho_i - Y_0) = 0 \quad (\rho_i - Y_0)E_i \\ \& \ \mathrm{Tr}(Y_0) = \sum_{i=1}^{n} p_i\,\mathrm{Tr}(E_i\rho_i) \end{cases}$

(2nd equality follows from Hermiticity as well) Complementary Slackn

(Strict duality)

(★) $\Rightarrow E_1 Y_0 = p_1 E_1\rho_1$ $(E_2Y_0 = p_2E_2\rho_2$ etc$)$, $Y_0E_i = p_i\rho_iE_i$

$\Rightarrow Y_0 = \sum_{i=1}^{n} p_i E_i\rho_i = \sum_i p_i\rho_i E_i$. — (**)

(#) Since (**) implies dual obj = primal obj, Y_0 is opt & feasible.

So, (E_i) are optimal iff $Y_0 = \sum_{i=1}^{n} p_i E_i \rho_i$ is feasible & (**)
(i.e. $Y_0 \geq p_i\rho_i \ \forall\, i = 1, \dots, n$ & $Y_0^\dagger = Y_0$)

<u>Proposition</u> Let $A, B \geq 0$.

Then $A \cdot B \geq 0$ iff $(AB)^{\dagger} = BA = AB$ (AB is Hermitian)

(($\Rightarrow$) $A \cdot B \geq 0 \Rightarrow (AB)^{\dagger} = AB$

$$(AB)^{\dagger} = B^{\dagger}A^{\dagger} = BA.$$

($\Leftarrow$) If $AB = BA$, they are simultaneously (unitarily) diagonalizable.

$\therefore$ if $A = \sum_i \lambda_i |a_i\rangle\langle a_i|$, $B = \sum_i \mu_i |a_i\rangle\langle a_i|$

$$AB = \sum_i \lambda_i \mu_i |a_i\rangle\langle a_i| \quad)$$

(Don't need the above for the <u>Holevo & Yuen, Kennedy, Lax conditions</u>)

Holevo; Yuen, Kennedy, Lax conditions

<u>Theorem</u> (E_i) is an optimal POVM for measuring (p_i, ρ_i) iff

(i) $\sum_{i=1}^{n} p_i \rho_i E_i = \sum_{i=1}^{n} p_i E_i \rho_i$, and

(ii) $\forall j = 1, 2, \ldots, n$, $\sum_{i=1}^{n} p_i E_i \rho_i \geq p_j \rho_j$.

The complementary slackness & above conditions have been used to show that

(a) The Fourier basis is the optimal one for distinguishing hidden subgroup states (by Ip 2003) for abelian groups

(b) The "pretty good measurement" is optimal for Dihedral groups (Bacon, Childs, van Dam '05)

(c) Have been studied further ~~-~~ Sarah might talk about this in her module.

Example State Distinguishability

$n = 2$

E_1 = projection onto the +ve eigenspace of $p_1\rho_1 - p_2\rho_2$

E_2 = —"— ≤ 0 . —"—

Consider $p_1\rho_1E_1 + p_2\rho_2E_2$

$$= p_1\rho_1E_1 - p_2\rho_2E_1 + p_2\rho_2E_1 + p_2\rho_2E_2$$
$$= (p_1\rho_1 - p_2\rho_2)E_1 + p_2\rho_2(E_1+E_2)$$
$$= E_1(p_1\rho_1 - p_2\rho_2) + p_2\rho_2$$
$$= p_1E_1\rho_1 + p_2E_2\rho_2 \quad \text{by similar steps in reverse.}$$

Moreover, $p_1\rho_1E_1 + p_2\rho_2E_2 \geq p_1\rho_1$

$$\Leftrightarrow \quad p_1\rho_1(E_1 - 1) + p_2\rho_2E_2 \geq 0$$
$$\Leftrightarrow \quad -p_1\rho_1E_2 + p_2\rho_2E_2 \geq 0$$
$$\Leftrightarrow \quad -(p_1\rho_1 - p_2\rho_2)E_2 \geq 0 \qquad \text{True!}$$

So this is the **optimal** measurement.

Example Strong Duality doesn't always hold.

Primal

$\inf(x_3 - x_1)$

subject to
$$\begin{pmatrix} 0 & x_3 - x_1 & 0 \\ x_3 - x_1 & x_2 & 0 \\ 0 & 0 & x_3 - x_1 + 1 \end{pmatrix} \succeq 0 \quad (x_1, x_2, x_3 \geq 0)$$

Feasible region : $x_1 = x_3$, $x_2 \geq 0$.

Opt $= 0$

Dual (after simplification)

$\sup \; -y_{33}$

subject to

$Y \succeq 0$ (3x3 matrix)

$$\begin{pmatrix} -y_{21} - y_{12} - y_{33} & 0 & 0 \\ 0 & y_{22} & 0 \\ 0 & 0 & y_{21} + y_{12} + y_{33} \end{pmatrix} \preceq \begin{pmatrix} -1 & 0 & \\ 0 & 0 & \\ 0 & 0 & \end{pmatrix}$$

i.e. $y_{22} \leq 0$, $y_{12} + y_{21} + y_{33} = 1$

$y_{22} \geq 0$, $y_{33} \geq 0$

Opt $= -1$