QIC 820 / CO 781 / CO 486 / CS 867

Supplementary notes on SDP for Oct 17 lecture.

This heavily expands the proof of Slater's Theorem by Jamie Sikora

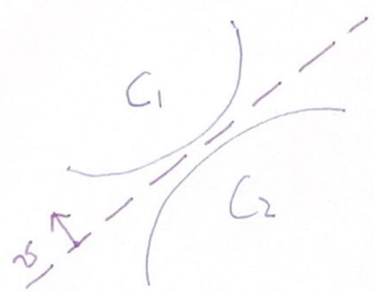
Alternative proof is available in LN 2011.

Back to 2nd or 3rd lecture, end of Chapter 2 in LN 2011,
there are 3 theorems on convexity. Here is the last one (without proof):

Hyperplane separation theorem:

Let C_1, C_2 be two disjoint, non-empty, convex subsets of $\mathbb{R}^n$.

Then $\exists\, v \in \mathbb{R}^n$, $v \neq 0$, $\exists\, c \in \mathbb{R}$
s.t. $\forall\, x \in C_1$, $y \in C_2$, $\langle x, v\rangle \geq c$ and $\langle y, v\rangle \leq c$.



NB We do not need this here but, in addition:

If C_1, C_2 are closed, and at least one is compact
then $\langle y, v\rangle < c' < c < \langle x, v\rangle$.

Slater's Theorem:

If ① $\exists$ strictly feasible primal sol'n (ie $\tilde{X}$ s.t. $\Phi(\tilde{X}) = B$, $\tilde{X} > 0$, not nec opt)

② α finite

then ③ $\alpha = \beta$

④ $\exists$ optimal dual feasible Y (ie $\Phi^*(Y) \geq A$, $\beta = \langle B, Y \rangle$).

Note the asymmetry – primal SDP still may not have optimal solution
dual SDP may not be strictly feasible.

Also all conditions in the hypothesis pertain to the primal SDP,
" " " " conclusion " " dual ".

Nomenclature: If ③④ hold, then we say ~~strong duality~~ holds for the SDP.

Pf: Relies on 2 non-empty convex sets (living in the same space):

$$M := \{ (S, V, t) : \exists X \in \text{Herm}(\mathcal{X}),\ S \leq X,\ V = B - \Phi(X),\ t \leq \langle A, X \rangle \}$$

S: Herm($\mathcal{X}$), V: Image(Φ), t: in $\mathbb{R}$; big S

$$N := \{ (0, 0, s) : s > \alpha \}$$

small s

ⓐ N is non-empty $\because$ α is finite !! N is clearly convex.

ⓑ M is non-empty $\because$ $(0, 0, \langle A, \tilde{X} \rangle) \in M$

(check, use $\tilde{X}$, $S = 0 \leq \tilde{X}$, $V = 0 = B - \Phi(\tilde{X})$, $t = \langle A, \tilde{X} \rangle$)

To check M is convex, take $p_1 \geq 0$, $p_2 \geq 0$, $p_1 + p_2 = 1$

For $i = 1, 2$, let $(S_i, V_i, t_i) \in M$ with X_i s.t. $S_i \leq X_i$, $V_i \leq B - \Phi(X_i)$, $t_i \leq \langle A, X_i \rangle$

X_i: Herm($\mathcal{X}$)

Then $(\sum_i p_i S_i, \sum_i p_i V_i, \sum_i p_i t_i) \in M$

with $X = \sum_i p_i X_i \in \text{Herm}(\mathcal{X})$, $\sum_i p_i S_i \leq \sum_i p_i X_i = X$

$\sum_i p_i V_i \leq B - \sum_i p_i \Phi(X_i) = B - \Phi(\sum_i p_i X_i)$

$\sum_i p_i t_i \leq \langle A, \sum_i p_i X_i \rangle$.

ⓒ $M \cap N = \phi$

- because if not $\exists\ (0,0,s) \in N \cap M$
- $(0,0,s) \in M \Rightarrow \quad \exists X$ s.t. $X \geq 0$, $\Phi(X) = B$, $s \leq \langle A, X \rangle$

  $\Rightarrow \quad X$ feasible so $s \leq \alpha$
- but $(0,0,s) \in N \Rightarrow s > \alpha$ $\quad \therefore$ contradiction.

ⓓ By ⓐⓑⓒ, can apply hyperplane thm to M, N.

$\therefore \exists$ <u>non-zero</u> $(\Lambda, Y, \lambda) \in (\text{Herm}(X), \text{Image}(\Phi), \mathbb{R})$

$\exists\ c \in \mathbb{R}$

s.t. $(S, V, t) \in M \Rightarrow \langle \Lambda, S \rangle + \langle Y, V \rangle + \lambda t \leq c$ (H1)

$(0, 0, s) \in N \Rightarrow \qquad \lambda s \geq c$ (H2)

Idea: derive a lot of constraints on Λ, Y, λ using (H1), (H2)

ⓔ From (H2) $\forall s > \alpha$, $\lambda s \geq c$. $\quad \therefore \underbrace{\lambda \geq 0}_{\substack{\text{If not } \lambda < 0 \\ s \leq c/\lambda \\ \text{contradicting} \\ \text{satisfiability } \forall s > \alpha.}}$ and $\underbrace{c \leq \lambda \alpha}_{\text{Choose } s = \alpha.}$.

fixed

4 Consider an interesting subset of M:

$$M_2 = \{ (S, 0, \langle A, X \rangle) : X \text{ feasible}, S \leq 0 \}$$

Check $M_2 \subseteq M$: If X fasible, $S \leq 0 \leq X$ ✓

and $V = 0 = B - \Phi(X)$ ✓

and $t = \langle A, X \rangle$ ✓

Apply (H1) to M_2: $\underbrace{\underset{\uparrow\ \text{any } S \leq 0}{\langle \Lambda, S \rangle}}{} + \underbrace{\underset{\geq 0}{\lambda} \underbrace{\langle A, X \rangle}_{\leq \alpha}}_{\text{finite}} \leq C$ (HM2)

Claim: $\Lambda \geq 0$

Pf by contradiction:

If not, let $\Lambda = -\gamma |e\rangle\langle e| + \Lambda'$ $\left(\begin{array}{l} \text{spec decomp,} \\ \langle e | \Lambda' | e \rangle = 0 \end{array} \right)$

$\gamma > 0$

Then, let $S = r |e\rangle\langle e| \leq 0$ if $r < 0$.

Then (HM2) $\Rightarrow$ $\underbrace{-\underset{\text{arbitrary negative}}{r}\ \underset{+ve}{\gamma}}_{\text{unbound positive}} + \text{finite} \leq \underset{\text{fixed}}{C}$

$\Rightarrow$ contradiction

$\therefore \Lambda \geq 0$.

⑨ Consider 2 more subsets of M:

- $M_3 = \{ (X, 0, \langle A, X \rangle) : X \text{ feasible} \}$

Here $S = X \therefore S \leq X$ ✓

X feasible $\Leftrightarrow V = 0$ ✓

$t \leq \langle A, X \rangle$ ✓

$\therefore M_3 \subseteq M$.

- $M_4 = \{ (S, B - \Phi(S), \langle A, S \rangle) : S \in \mathrm{Herm}(\mathcal{X}) \}$

To see that $M_4 \subseteq M$,

$\forall S \in \mathrm{Herm}(\mathcal{X})$, take $X = S$ so $S \leq X$. ✓

now $V = B - \Phi(X) = B - \Phi(S)$ ✓

and $t = \langle A, S \rangle = \langle A, X \rangle$ ✓

(NB: in M_2, M_3, X feasible. In M_4, X need not be feasible.)

Idea: use $\lambda \geq 0$, $\Lambda \geq 0$, M_3, M_4 to show $\lambda > 0$ by contradiction.

(**) So, suppose $\lambda = 0$.

Apply (H1) to M_3: $\langle \Lambda, X \rangle \leq c \leq \lambda s = 0$

only 1 term on LHS of (H1) $\because \lambda = 0, V = 0$

Above hold $\forall X$ feasible. Use $\tilde{X} \in \mathrm{Pd}(\mathcal{X})$ full rank.

We already know $\Lambda \geq 0$, but now $\Lambda = 0$ else $\langle \Lambda, \tilde{X} \rangle > 0$.

full rank with component overlapping with +ve eigenspace of Λ if $\Lambda \neq 0$

Now $\lambda = 0$, $\Lambda = 0$.

Apply H1 to M_4: $\langle Y, B - \Phi(S) \rangle \le c \le \lambda S = 0$

But $B = \Phi(\tilde{X})$ $\therefore$ $\langle Y, \Phi(\tilde{X} - S) \rangle \le 0$ $\forall S \in \mathrm{Herm}(X)$.

Now, the set $\{\tilde{X} - S : S \in \mathrm{Herm}(X)\} = \mathrm{Herm}(X)$.

$\therefore$ the set $\{\Phi(\tilde{X} - S) : S \in \mathrm{Herm}(X)\} = \mathrm{Image}(\Phi)$

$\therefore$ $Y = 0$

Now $(\Lambda, Y, \lambda) = (0, 0, 0)$ contradict Hyperplane separation thm!

$\therefore$ (**) must be wrong $\therefore$ $\lambda \neq 0$ $\therefore$ $\lambda > 0$

Note: (e) $\Rightarrow \lambda \geq 0$
(f) $\Rightarrow \Lambda \geq 0$
(g) $\Rightarrow$ if $\lambda = 0$ then $\Lambda = 0$ and $Y = 0$ contradiction
$\therefore \lambda > 0$
and now we only know $\Lambda \geq 0$, nothing about Y.

h $\because \lambda > 0$, rescale H1:

$$\underbrace{\left\langle \frac{\Lambda}{\lambda}, S \right\rangle}_{= \Lambda'} + \underbrace{\left\langle \frac{Y}{\lambda}, V \right\rangle}_{= Y'} + t \overset{\text{H1}}{\leq} \frac{c}{\lambda} \overset{\text{e}}{\leq} \alpha$$

Apply to M4:

$$\langle \Lambda', S \rangle + \langle Y', B - \Phi(S) \rangle + \langle A, S \rangle \leq \alpha$$

Pulling out the dependence on S:

$$\underbrace{\langle Y', B \rangle}_{\text{fixed}} + \langle S, \underbrace{\Lambda' - \Phi^*(Y') + A}_{\text{fixed}} \rangle \leq \underset{\text{fixed}}{\alpha} \quad \leftarrow \text{HM4}$$

ranges over Herm(X), unbounded and containing it's own negatin

$\therefore$ $\Lambda' - \Phi^*(Y') + A = 0$ else inner product unbounded leading to contradiction.

But $\Lambda' \geq 0$ $\therefore$ Y' is dual feasible

Also HM4 $\Rightarrow$ $\langle Y', B \rangle \leq \alpha$

$\langle Y', B \rangle \geq \beta$ (Y' dual feasible) $\alpha \geq \beta$ (weak duality ($\because \tilde{X}, Y'$ feasible))

$\Rightarrow$ Y' optimal, $\alpha = \beta$.

Slater's Thm (dual)

If ① $\exists$ strictly feasible dual solution ($\tilde{Y}$ s.t. $\Phi^*(\tilde{Y}) > A$)
② β is finite

then ③ $\alpha = \beta$
④ α attained by some optimal feasible X.

NB If ③④ holds above, we say strong duality holds for the dual.
NB Dual of the dual SDP is the primal SDP. Pf: Ex.

Corollary: If an SDP and its dual are both strictly feasible then $\alpha = \beta$ and both can be attained.
NB no need to assume α finite $\because$ dual feasible $\Rightarrow$ α finite. Same for primal feasible $\Rightarrow$ β finite.

Complementary Slackness:

① If $\alpha = \beta$, X, Y, S optimal primal & dual solutions, then $XS = SX = 0$.
ie $X\Phi^*(Y) = XA$ and $\Phi^*(Y)X = AX$.

Pf: In the proof for weak duality, $\langle X, S\rangle = \langle B, Y\rangle - \langle A, X\rangle = \beta - \alpha = 0$

But $X \succeq 0$, $S \succeq 0$

$\therefore \langle X, S\rangle = 0 \iff XS = SX = 0$

② If X, Y, S feasible, $XS = SX = 0$

then X, Y, S are optimal primal & dual solutions, and $\alpha = \beta$.

Application:

Let H be a Hamiltonian.

Ground state energy = min eigenvalue of $H \simeq \gamma$.

If $\gamma \neq -\infty$, we can consider $H' = H + r\mathbb{1}_X$ for sufficiently large r so that $H' > 0$

$\therefore$ WLOG, $H \geq 0$, $\gamma > 0$.

Consider: $\alpha = \sup t$ s.t. $t\mathbb{1} \leq H$, $t \geq 0$.

$$\begin{aligned} \alpha = \ & \sup t \\ \text{s.t. } & t\mathbb{1} \leq H \\ & t \geq 0. \end{aligned}$$

Dual:

$$\begin{aligned} \beta = \ & \inf \langle \rho, H \rangle \\ \text{s.t. } & \operatorname{tr}(\rho) = 1 \\ & \rho \geq 0 \end{aligned}$$

To see that these are primal-dual SDPs, rewrite the constraints in the primal SDP:

$$t\mathbb{1} + K = H$$

$$t \geq 0 \quad (t \in \mathbb{R})$$

$$K \geq 0 \quad (K \in \operatorname{Pos}(X))$$

Combine the variables into $X = \begin{pmatrix} t & \\ & K \end{pmatrix}$

rewrite the constraints: $\Phi(X) = t\mathbb{1} + K = H = B$

rewrite the obj fcn: $A = \begin{pmatrix} 1 & 0 \cdots \\ 0 & \\ \vdots & 0 \end{pmatrix}$

$$\langle A, X \rangle = t$$

Note that each feasible solution $t \geq 0$, $K \geq 0$ for origin SDP is included in $X = \begin{bmatrix} t & 0\,0\cdots 0 \\ \begin{matrix}0\\0\\\vdots\\0\end{matrix} & K \end{bmatrix} \geq 0$.

Conversely for each $X \geq 0$, $X_{11} \geq 0$ and lower right block ≥ 0

$\therefore$ Switch variable to $X \geq 0$ preserves the original feasible set.

$\therefore$ Primal SDP can be rewritten as:

$$\alpha = \sup \langle A, X \rangle$$
$$\text{s.t. } \Phi(X) = H$$
$$X \geq 0 \qquad (\text{note } X \in \mathrm{Pos}(\mathcal{X} \oplus \mathbb{C}))$$

$\forall Y \in \mathrm{Herm}(\mathcal{X})$, $\forall X \in \mathrm{Pos}(\mathcal{X} \oplus \mathbb{C})$,

$$\langle \underbrace{\Phi^*(Y)}_{\text{in Herm}(\mathcal{X}\oplus\mathbb{C})}, X \rangle = \langle \underset{\text{in Herm}(\mathcal{X})}{Y}, \underset{\text{in Herm}(\mathcal{X})}{\Phi(X)} \rangle$$

$$= \langle Y, t\mathbb{1} + K \rangle$$

$$= t \cdot \mathrm{tr}\, Y + \langle Y, K \rangle = \left\langle \underset{=\Phi^*(Y)}{\begin{bmatrix} \mathrm{tr}Y & 0\,0\,0 \\ \begin{matrix}0\\0\\0\end{matrix} & Y \end{bmatrix}}, \underset{=X}{\begin{bmatrix} t & //// \\ // & K \end{bmatrix}} \right\rangle$$

$$\therefore \Phi^*(Y) = \begin{bmatrix} \mathrm{tr}Y & 0\,0\cdots 0 \\ \begin{matrix}0\\\vdots\\0\end{matrix} & Y \end{bmatrix}$$

Dual $= \inf \langle Y, H \rangle$ s.t. $\Phi^*(Y) \geq A$

i.e. $\inf \langle Y, H \rangle$ s.t. $\mathrm{tr}Y \geq 1$, $Y \geq 0$ $\iff$ $\inf \langle Y, H \rangle$, $\mathrm{tr}Y = 1$, $Y \geq 0$

(relabel Y as ρ)

$t = \frac{\gamma}{4}$, $K = \frac{\gamma}{4} \mathbb{1}_X$ strictly primal feasible

$\rho = \frac{\mathbb{1}_X}{\dim(X)}$ strictly dual feasible.

$\therefore$ strong duality holds, $\alpha = \beta$, both attained.

The primal SDP characterizes min eigenvalue of H.

Strong duality says there are ground states (optimal ρ's).

Complementary slackness says that, for optimal X, Y:

$$X\, \Phi^*(Y) = XA \quad \text{and} \quad \Phi^*(Y)\, X = A\, X$$

$$\begin{pmatrix} t & \\ & K \end{pmatrix} \cdot \begin{pmatrix} \mathrm{Tr}\,Y & 0\,0 \cdots 0 \\ 0 & \\ 0 & Y \\ \vdots & \end{pmatrix} \qquad \begin{pmatrix} t & \\ & K \end{pmatrix} \begin{pmatrix} 1 & 0\,0\,0 \\ 0 & \\ 0 & 0 \\ \vdots & \end{pmatrix}$$

$\mathrm{Tr}\,Y = 1$, $KY = 0$

Recall $K = H - t\mathbb{1}$

$\therefore$ gs $\rho \in \ker(H - \gamma \mathbb{1}_X)$

Reading Ex: Sec 7.3 in LN 2011 on alternate forms of SDP.

Here, brief discussion on multiple variables or multiple constraints.

eg. $\sup \langle A, X\rangle$

s.t. $\Phi_1(X) = B_1 \in L(\mathcal{Y}_1)$

$\Phi_2(X) = B_2 \in L(\mathcal{Y}_2)$

$X \geq 0$

Define $\Phi \in L(\mathcal{X}, \mathcal{Y}_1 \oplus \mathcal{Y}_2)$

$$\Phi(X) = \begin{bmatrix} \Phi_1(X) & 0 \\ 0 & \Phi_2(X) \end{bmatrix}$$

if $\mathcal{Y}_1 \sim \mathcal{Y}_2 \sim \mathcal{Y}$, we can write this as $\sum_i |i\rangle\langle i| \otimes \Phi_i(X)$.

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

Then above SDP equivalent to

$\sup \langle A, X\rangle$

s.t. $\Phi(X) = B$

$X \geq 0$.

With dual:

$\inf \langle B, Y\rangle$

s.t. $\Phi^*(Y) \geq A$

What is $\Phi^*(Y)$?. Let $Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \in L(\mathcal{Y}_1 \oplus \mathcal{Y}_2)$, $Y_{11} \in L(\mathcal{Y}_1)$

$Y_{12} \in L(\mathcal{Y}_2, \mathcal{Y}_1)$

$Y_{21} \in L(\mathcal{Y}_1, \mathcal{Y}_2)$

$Y_{22} \in L(\mathcal{Y}_2, \mathcal{Y}_2)$

$\forall X, Y$

$$\langle \Phi^*(Y), X\rangle = \langle Y, \Phi(X)\rangle = \langle Y_{11}, \Phi_1(X)\rangle + \langle Y_{22}, \Phi_2(X)\rangle$$

$$= \langle \Phi_1^*(Y_{11}), X\rangle + \langle \Phi_2^*(Y_{22}), X\rangle$$

$\therefore \Phi^*(Y) = \Phi_1^*(Y_{11}) + \Phi_2^*(Y_{22})$. If $\mathcal{Y}_1 \sim \mathcal{Y}_2 \sim \mathcal{Y}$,

$$\Phi^*(Y) = \sum_i \Phi_i^*\Big(\overbrace{(\langle i| \otimes \mathbb{1}_{\mathcal{Y}})\, Y\, (|i\rangle \otimes \mathbb{1}_{\mathcal{Y}})}^{Y_{ii}}\Big)$$

$\langle B, Y\rangle = \sum_i \langle B_i, Y_{ii}\rangle$

$\therefore$ Dual SDP: $\inf \sum_i \langle B_i, Y_{ii}\rangle$

s.t. $\sum_i \Phi_i^*(Y_{ii}) \geq A$

eg. $\sup \langle A_1, X_1\rangle + \langle A_2, X_2\rangle$

s.t. $\Phi_1(X_1) = B_1$

$\Phi_2(X_2) = B_2$

$X_1 \geq 0$

$X_2 \geq 0$

Define: $X = \begin{bmatrix} X_1 & * \\ * & X_2 \end{bmatrix}$, $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$, $B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$

$$\Phi(X) = \begin{bmatrix} \Phi_1(X_1) & 0 \\ 0 & \Phi_2(X_2) \end{bmatrix}$$

Then above SDP is equiv to:

$\sup \langle A, X\rangle$

s.t. $\Phi(X) = B$

$X \geq 0$

Dual:

$\inf \langle Y_{11}, B_1\rangle + \langle Y_{22}, B_2\rangle$

s.t. $\Phi^*(Y)$

$$\begin{bmatrix} \Phi_1^*(Y_1) & 0 \\ 0 & \Phi_2^*(Y_2) \end{bmatrix} \geq \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

ie $\Phi_1^*(Y_1) \geq A_1$

$\Phi_2^*(Y_2) \geq A_2$.

Note difference from previous case.