and $\mathrm{s2vec} : \mathbb{S}^n \to \mathbb{R}^{\frac{n(n-1)}{2}}$ by

$$\mathrm{s2vec}(X) := \begin{bmatrix} X_{1,1} \\ \sqrt{2}X_{2,1} \\ \vdots \\ \sqrt{2}X_{n,1} \\ X_{2,2} \\ \sqrt{2}X_{3,2} \\ \vdots \\ \sqrt{2}X_{n,2} \\ \vdots \\ X_{n,n} \end{bmatrix}.$$

# 2 Duality Theory

THEOREM 81. [THEOREM 2.8: A SEPARATION THEOREM] *Let* $n \in \mathbb{N}$ *and let* $G \subseteq \mathbb{R}^n$. *If* $G$ *is a non-empty, closed convex set with* $0 \notin G$, *then there exists an* $a \in \mathbb{R}^n \setminus \{0\}$ *and an* $\alpha \in \mathbb{R}_{++}$ *such that* $a^\mathsf{T} x \geq \alpha$ *for all* $x \in G$.

*Proof.* Since $G$ is non-empty, let $\bar{x} \in G$. Set

$$G_{\bar{x}} := \{x \in G;\ \|x\|_2 \leq \|\bar{x}\|_2\}.$$

Note that $G_{\bar{x}} = G \cap \{x \in \mathbb{R}^n;\ \|x\|_2 \leq \|\bar{x}\|_2\}$ in the intersection of compact convex sets and is hence compact and convex. Moreover, it is non-empty because $\bar{x} \in G_{\bar{x}}$, and $0 \notin G_{\bar{x}}$ since $0 \notin G$. As such, there exists a unique point $a$ in $G_{\bar{x}}$ which is closest to the origin. (Indeed, consider minimizing the continuous, strictly convex function $\|\cdot\|_2^2$ on the non-empty compact convex set $G_{\bar{x}}$.)

Since $a \in G_{\bar{x}}$ and $0 \notin G_{\bar{x}}$, $a \neq\in \mathbb{R}^n \setminus \{0\}$. Set $\alpha := \|a\|_2^2 > 0$. Let $x \in G$. Since $G$ is convex, for every $\lambda \in (0,1)$ we have $\lambda x + (1-\lambda)a \in G$. By choice of $a$, we have $\|a\|_2 \leq \|z\|_2^2$ for all $z \in G$. Therefore

$$\|\lambda x + (1-\lambda)a\|_2^2 \geq \|a\|_2^2.$$

But expanding the left side yields

$$\|\lambda(x-a) + a\|_2^2 = \lambda^2 \|x-a\|_2^2 + 2\lambda \langle x-a, a\rangle + \|a\|_2^2.$$

Hence the first inequality becomes

$$\lambda^2 \|x-a\|_2^2 + \lambda \langle x-a, a\rangle \geq 0,$$

which shows, since $\lambda > 0$,

$$(x-a)^\mathsf{T} a \geq -\tfrac{\lambda}{2} \|x-a\|_2^2.$$

As this holds for all $\lambda \in (0,1)$, we get

$$x^T a - a^T a \geq 0,$$

or equivalently

$$x^T a \geq \alpha.$$

∎

Corollary 82. [Corollary 2.9] *Let $n \in \mathbb{N}$ and let $G_1, G_2 \subseteq \mathbb{R}^n$. If $G_1$ and $G_2$ are non-empty, disjoint, closed convex sets, and at least one of $G_1$ and $G_2$ is bounded, then there exists an $a \in \mathbb{R}^n \setminus \{0\}$ such that*

$$\inf\{a^T u; u \in G_1\} > \sup\{a^T v; v \in G_2\}.$$

*Proof.* Note that $G_1 - G_2$ is clearly a non-empty convex set, which does not contain the origin because $G_1$ and $G_2$ are disjoint. Now we claim that it is closed. We will assume $G_2$ si bounded (if $G_2$ is unbounded then $G_1$ must be bounded, and the proof is similar). Suppose $(g^{(k)})_{k\in\mathbb{N}}$ is a sequence in $G_1 - G_2$ which converges to some $g \in \mathbb{R}^n$. Then there exists sequences $(u^{(k)})_{k\in\mathbb{N}}$ and $(v^{(k)})_{k\in\mathbb{N}}$ in $G_1$ and $G_2$ respectively such that $g^{(k)} = u^{(k)} - v^{(k)}$ for each $k \in \mathbb{N}$. Since $G_2$ is compact (it is closed and bounded), there is a subsequence $(v^{(k_m)})_{m\in\mathbb{N}}$ of $(v^{(k)})_{k\in\mathbb{N}}$ which converges to some $\bar{g}_2 \in G_2$. Now note that the subsequence $(u^{(k_m)})_{m\in\mathbb{N}} = (g^{(k_m)})_{m\in\mathbb{N}} + (v^{(k_m)})_{m\in\mathbb{N}}$ converges to $g + \bar{g}_2$. Since $G_1$ is closed, $g + \bar{g}_2 \in G_1$. Thus $g = (g + \bar{g}_2) - \bar{g}_2 \in G_1 - G_2$ and so $G_1 - G_2$ is closed, as claimed.

By the previous theorem, there exists an $a \in \mathbb{R}^n \setminus \{0\}$ and an $\alpha \in \mathbb{R}_{++}$ such that $a^T g \geq \alpha$ for all $g \in G_1 - G_2$. So let $u \in G_1$ and $v \in G_2$. Then $u - v \in G_1 - G_2$, so

$$a^T u - a^T v = a^T (u - v) \geq \alpha,$$

or equivalently

$$a^T u \geq \alpha + a^T v.$$

Taking the infimum over all such $u \in G_1$ yields

$$\inf\{a^T u; u \in G_1\} \geq \alpha + a^T v$$

for all $v \in G_2$. Taking the supremum over all such $v \in G_2$ yields

$$\inf\{a^T u; u \in G_1\} \geq \alpha + \sup\{a^T v; v \in G_2\} > \sup\{a^T v; v \in G_2\}.$$

∎

Theorem 83. [Theorem 2.11] *Let $n \in \mathbb{N}$ and let $G \subseteq \mathbb{R}^n$. If $G$ is a non-empty convex set with $0 \notin G$ then there exists an $a \in \mathbb{R}^n \setminus \{0\}$ such that $a^T x \geq 0$ for all $x \in G$.*

*Proof.* For $x \in G$, define

$$HS(x) := \{s \in \mathbb{R}^n; s^\mathsf{T} x \geq 0, \|s\|_2 = 1\}.$$

We need to show that

$$\bigcap_{x \in G} HS(x) \neq \varnothing.$$

Since each $HS(x)$, for $x \in G$, is compact, it suffices to prove that the intersection is non-empty for any finite subset of G. So suppose $x^{(1)}, \ldots, x^{(k)} \in G$. Note that $\operatorname{conv}\{x^{(1)}, \ldots, x^{(k)}\}$ is a non-empty, closed, convex set. Moreover, it is a subset of G, so since $0 \notin G$, we get $0 \notin \operatorname{conv}\{x^{(1)}, \ldots, x^{(k)}\}$. Thus by theorem 2.8, there exists an $a \in \mathbb{R}^n \setminus \{0\}$ such that $a^\mathsf{T} x \geq 0$ for all $x \in \operatorname{conv}\{x^{(1)}, \ldots, x^{(k)}\}$. So we see

$$\tfrac{1}{\|a\|_2} a \in \bigcap_{x \in \{x^{(1)}, \ldots, x^{(k)}\}} HS(x) \neq \varnothing,$$

as required. ∎

Corollary 84. [Corollary 2.12] *Let $n \in \mathbb{N}$ and let $G_1, G_2 \subseteq \mathbb{R}^n$. If $G_1$ and $G_2$ are non-empty, disjoint convex sets, then there exists an $a \in \mathbb{R}^n \setminus \{0\}$ such that*

$$\inf\{a^\mathsf{T} u; u \in G_1\} \geq \sup\{a^\mathsf{T} v; v \in G_2\}.$$

*Proof.* Note that $G_1 - G_2$ is a non-empty convex set, which does not contain the origin because $G_1$ and $G_2$ are disjoint. By the previous theorem, there exists an $a \in \mathbb{R}^n \setminus \{0\}$ such that $a^\mathsf{T} x \geq 0$ for all $x \in G_1 - G_2$. So let $u \in G_1$ and $v \in G_2$. Then $u - v \in G_1 - G_2$, so $a^\mathsf{T}(u - v) \geq 0$, which shows

$$a^\mathsf{T} u \geq a^\mathsf{T} v.$$

Taking the infimum over $u \in G_1$ yields

$$\inf\{a^\mathsf{T} u; u \in G_1\} \geq a^\mathsf{T} v$$

for all $v \in G_2$. Taking the supremum over all $v \in G_2$ yields

$$\inf\{a^\mathsf{T} u; u \in G_1\} \geq \sup\{a^\mathsf{T} v; v \in G_2\}.$$

∎

Theorem 85. [Theorem 2.14: A Strong Duality Theorem] *Let $n, m \in \mathbb{N}$, let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear, let $b \in \mathbb{R}^m$, and let $C \in \mathbb{S}^n$. If (D) has a Slater point and its objective value is bounded from above (on its feasible region) then (P) has an optimal solution and the optimal objective values of (P) and (D) are the same.*

*Proof.* First suppose that $b = 0$. Then $\bar{X} := 0$ is a feasible solution for (P), with objective value 0. Also, the given Slater point is a feasible solution for (D), with objective value 0 also. Hence by weak duality, $\bar{S}$ is an optimal solution and the objective values agree.

So we may assume $b \neq 0$. Since the objective value of (D) is bounded above, it has an optimal value, say $z^*$. Define $G_2 := \mathbb{S}^n_{++}$ and

$$G_1 := \{C - \mathcal{A}^*(y);\, y \in \mathbb{R}^m,\, b^\mathsf{T} y \geq z^*\} \subseteq \mathbb{S}^n.$$

It is easy to see that $G_1$ and $G_2$ are both convex, and clearly $G_2$ is non-empty.

We claim that $G_1$ is also non-empty. Since $b \neq 0$, $\frac{z^*}{\|b\|_2^2} b \in \mathbb{R}^m$ and

$$b^\mathsf{T} \left( \frac{z^*}{\|b\|_2^2} b \right) = \frac{z^*}{b^\mathsf{T} b} \left( b^\mathsf{T} b \right) = z^* \geq z^*.$$

Hence $\frac{z^*}{\|b\|_2^2} b \in G_1$.

Next we claim that $G_1$ and $G_2$ are disjoint. Assume, for a contradiction, that they are not. This means that there exists a $\tilde{y} \in \mathbb{R}^m$ such that $b^\mathsf{T}\tilde{y} \geq z^*$ and $C - \mathcal{A}^*(\tilde{y}) \succ 0$. But then $\tilde{y}$ is feasible for (D), so since $z^*$ is the optimal objective value, we must have $b^\mathsf{T}\tilde{y} \leq z^*$, and hence $b^\mathsf{T}\tilde{y} = z^*$. Consider $y(\epsilon) := \tilde{y} + \epsilon b$, for $\epsilon > 0$. Since $\mathcal{A}^*$ is linear and hence continuous, and since $C - \mathcal{A}^*(\tilde{y}) \in \mathbb{S}^n_{++} = \int(\mathbb{S}^n_+)$, there is some $\epsilon > 0$ such that $C - \mathcal{A}^*(y(\epsilon)) \in \mathbb{S}^n_{++}$. In particular, $y(\epsilon)$ is feasible for (D). But it has objective value

$$b^\mathsf{T} y(\epsilon) = b^\mathsf{T} (\tilde{y} + \epsilon b) = b^\mathsf{T}\tilde{y} + \epsilon b^\mathsf{T} b > b^\mathsf{T}\tilde{y} = z^*,$$

contradiction that $z^*$ is the optimal objective value of (D). Thus the claim holds.

Now we can apply the previous corollary to get that there exists an $\tilde{X} \in \mathbb{S}^n \setminus \{0\}$ such that

$$\sup\{\langle \tilde{X}, S\rangle;\, S \in G_1\} \leq \inf\{\langle \tilde{X}, S\rangle;\, S \in G_2\}.$$

Assume, for a contradiction, that $\inf\{\langle \tilde{X}, S\rangle;\, S \in G_2\} < 0$. This means that there is an $S \in G_2 = \mathbb{S}^n_{++}$ such that $\langle \tilde{X}, S\rangle < 0$. But now for any $\alpha \in \mathbb{R}_{++}$, $\alpha S \in \mathbb{S}^n_{++} = G_2$ and $\langle \tilde{X}, \alpha S\rangle = \alpha \langle \tilde{X}, S\rangle \to -\infty$ as $\alpha \to \infty$. This contradicts that $\{\langle \tilde{X}, S\rangle;\, S \in G_2\}$ is bounded below. Thus $\inf\{\langle \tilde{X}, S\rangle;\, S \in G_2\} \geq 0$. On the other hand, $\alpha I \in \mathbb{S}^n_{++}$ for all $\alpha \in \mathbb{R}_{++} = G_2$ and $\langle \tilde{X}, \alpha I\rangle = \alpha \langle \tilde{X}, I\rangle \to 0$ as $\alpha \to 0^+$. This shows $\inf\{\langle \tilde{X}, S\rangle;\, S \in G_2\} \leq 0$ and thus $\inf\{\langle \tilde{X}, S\rangle;\, S \in G_2\} = 0$.

From this we get $\sup\{\langle \tilde{X}, S\rangle;\, S \in G_1\} \leq 0$. This means that $\langle \tilde{X}, S\rangle \leq 0$ for all $S \in G_1$, or equivalently, $\langle \tilde{X}, C - \mathcal{A}^*(y)\rangle \leq 0$ for all $y \in \mathbb{R}^m$ such that $b^\mathsf{T} y \geq z^*$. Rearranging and using the definition of the adjoint yields

$$\mathcal{A}(\tilde{X})^\mathsf{T} y \geq \langle \tilde{X}, C\rangle \tag{4}$$

for all $y \in \mathbb{R}^m$ such that $b^\mathsf{T} y \geq z^*$. This shows that the LP

$$\begin{aligned} &\min\ \mathcal{A}(\tilde{X}) y \\ &\text{s.t.:}\ b^\mathsf{T} y \geq z^* \end{aligned}$$

is bounded below, and hence its dual

$$\begin{aligned} \max\ & z^*\alpha \\ \text{s.t.:}\ & \alpha b = \mathcal{A}(\tilde{X}) \\ & \alpha \geq 0 \end{aligned}$$

is feasible. That is to say, there exists an $\alpha \in \mathbb{R}_+$ such that $\alpha b = \mathcal{A}(\tilde{X})$.

Assume, for a contradiction, that $\alpha = 0$. Then $\mathcal{A}(\tilde{X}) = \alpha b = 0$. We are given that (D) has a Slater point, say $(\bar{S}, \bar{y})$. And since we saw $\inf\{\langle \tilde{X}, S\rangle ; S \in G_2\} = 0$, we have $\langle \tilde{X}, S\rangle \geq 0$ for all $S \in \mathbb{S}^n_{++}$. It follows that $\langle \tilde{X}, S\rangle \geq 0$ for all $S \in \mathbb{S}^n_+$, and thus $\tilde{X} \in \mathbb{S}^n_+$. Now we see, since $\tilde{X} \in \mathbb{S}^n_+ \setminus \{0\}$ and $\bar{S} \in \mathbb{S}^n_{++}$,

$$0 < \langle \bar{S}, \tilde{X}\rangle = \langle \bar{S}, \tilde{X}\rangle + \langle \bar{y}, \mathcal{A}(\tilde{X})\rangle = \langle \bar{S}, \tilde{X}\rangle + \langle \mathcal{A}^*(\bar{y}), \tilde{X}\rangle = \langle \bar{S} - \mathcal{A}^*(\bar{y}), \tilde{X}\rangle = \langle C, \tilde{X}\rangle \leq 0,$$

where the last inequality follows because

$$\langle C, \tilde{X}\rangle = \langle C - \mathcal{A}^*(\bar{y}), \tilde{X}\rangle + \langle \mathcal{A}^*(\bar{y}), \tilde{X}\rangle \leq \langle \bar{y}, \mathcal{A}(\tilde{X})\rangle = \langle \bar{y}, 0\rangle = 0.$$

But this is absurd.

Thus $\alpha > 0$. Set $\hat{X} := \frac{1}{\alpha}\tilde{X}$. By linearity of $\mathcal{A}$, $\mathcal{A}(\hat{X}) = \frac{1}{\alpha}\mathcal{A}(\tilde{X}) = \frac{1}{\alpha}\alpha b = b$. And $\hat{X} \in \mathbb{S}^n_+$ since $\tilde{X} \in \mathbb{S}^n_+$. Therefore $\hat{X}$ is feasible for (P). Moreover, the objective value of $\hat{X}$ is

$$\langle C, \hat{X}\rangle = \frac{1}{\alpha}\langle C, \tilde{X}\rangle \leq \frac{1}{\alpha}\mathcal{A}(\tilde{X})^T \frac{z^*}{\|b\|_2^2} b = \frac{1}{\alpha}\alpha \frac{z^*}{\|b\|_2^2} b^T b = z^*.$$

By weak duality, since $z^*$ was the optimal value of (D), we must have $\langle C, \hat{X}\rangle \geq z^*$. Therefore $\langle C, \hat{X}\rangle = z^*$ as desired. ∎

Remark 86. The statement and the proof of this strong duality theorem generalize to convex optimization problems in conic form.

Even though the optimal objective values of (P) and (D) are the same, the optimal objective value of (D) may not be attained by any feasible solution.

The statement requires us to know that the objective function of (D) is bounded above on the feasible region. This is typically done by demonstrating a primal feasible solution.

Finally, note that if (D) has a Slater point, then the set of optimal solutions of (P) is compact.

Corollary 87. *Let $n, m \in \mathbb{N}$, let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear, let $b \in \mathbb{R}^m$, and let $C \in \mathbb{S}^n$. If* (P) *has a Slater point and its objective value is bounded from above (on its feasible region) then* (D) *has an optimal solution and the optimal objective values of* (P) *and* (D) *are the same.*

*Proof.* This follows immediately because the dual of (D) is equivalent to (P). ∎

Corollary 88. *Let $n, m \in \mathbb{N}$, let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear, let $b \in \mathbb{R}^m$, and let $C \in \mathbb{S}^n$. If both* (P) *and* (D) *have Slater points then* (P) *and* (D) *have optimal solutions and the optimal objective values of* (P) *and* (D) *are the same.*

*Proof.* This is immediate by applying both the above strong duality theorem and its corollary (and using weak duality). ∎

Example 89. There were two examples at this point which are omitted from this document, and a third referenced in the textbook.

There was also an infeasible SDP with no LP-like infeasibility certificate, motivating the following definition.

Definition 90. Let $n, m \in \mathbb{N}$, let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear, let $b \in \mathbb{R}^m$, and let $C \in \mathbb{S}^n$. We say $\mathcal{A}^*(y) \preccurlyeq C$ is *almost feasible* if for every $\epsilon > 0$, there is a $C' \in \mathbb{S}^n$ such that $\|C - C'\| < \epsilon$ and $\mathcal{A}^*(y) \preccurlyeq C'$ is feasible.

Theorem 91. [Theorem 2.21] *Let $n, m \in \mathbb{N}$, let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear, and let $C \in \mathbb{S}^n$.*

*(1) If there is a $D \in \mathbb{S}^n_+$ such that $\mathcal{A}(D) = 0$ and $\langle C, D \rangle < 0$, then there is no $y \in \mathbb{R}^m$ such that $\mathcal{A}^*(y) \preccurlyeq C$.*

*(2) If there is no $D \in \mathbb{S}^n_+$ such that $\mathcal{A}(d) = 0$ and $\langle C, D \rangle < 0$, then $\mathcal{A}^*(y) \preccurlyeq C$ is almost feasible.*

*Proof.* First we verify (1). Suppose $D \in \mathbb{S}^n_+$ is such that $\mathcal{A}(D) = 0$ and $\langle C, D \rangle < 0$. Assume, for a contradiction, that there is a $y \in \mathbb{R}^m$ such that $\mathcal{A}^*(y) \preccurlyeq C$. Then

$$0 > \langle C, D \rangle = \langle C - \mathcal{A}^*(y), D \rangle + \langle \mathcal{A}^*(y), D \rangle \geq \langle \mathcal{A}^*(y), D \rangle = \langle y, \mathcal{A}(D) \rangle = \langle y, 0 \rangle = 0,$$

which is absurd. Thus no such $y$ exists.

Next we verify (2). Suppose there is no $D \in \mathbb{S}^n_+$ such that $\mathcal{A}(D) = 0$ and $\langle C, D \rangle < 0$. Consider the SDP (D)

$$\begin{aligned} \sup \eta& \\ \mathcal{A}^*(y) + \eta I &\preccurlyeq C \\ \eta &\leq 0 \end{aligned}$$

and its dual (P)

$$\begin{aligned} \inf \langle C, X \rangle& \\ \mathcal{A}(x) &= 0 \\ \langle I, X \rangle &\leq 1 \\ X &\succcurlyeq 0. \end{aligned}$$

Note that $(0, -\|C\|_2 - 1)$ is a Slater point for (D). Moreover, the objective value is clearly bounded above on the feasible region because one of the constraints is $\eta \leq 0$. Therefore our strong duality theorem applies, and we conclude that (P) has an optimal solution and its value agrees with the optimal objective value of (D). Since $X = 0$ is a feasible solution for (P), the optimal objective value of (P) is at least 0. Suppose the optimal objective value of (P) were less than zero. Then there would be an $X \geq 0$ with $\langle X, I \rangle \leq 1$, $\mathcal{A}(X) = 0$, and $\langle C, X \rangle < 0$. But this contradicts our initial assumption. Thus the optimal objective value of (D) is 0.

So the optimal objective value of (P) is 0. This means that there exists a sequence $((y^{(k)}, \eta^{(k)}))_{k=1}^{\infty}$ such that for all $k \in \mathbb{N}$ we have $\mathcal{A}^*(y^{(k)}) + \eta^{(k)} I \preccurlyeq C$ and $\eta^{(k)} \leq 0$, and $\eta^{(k)} \to 0^-$. But now, for any $\epsilon > 0$, we can pick $k \in \mathbb{N}$ such that $\|\eta^{(k)} I\|_2 < \epsilon$ to get $\mathcal{A}^*(y^{(k)}) \preccurlyeq C - \eta^{(k)} I$ with $\|C - (C - \eta^{(k)} I)\|_2 = \|\eta^{(k)} I\|_2 < \epsilon$. Thus $\mathcal{A}^*(y) \preccurlyeq C$ is almost feasible. ∎

**Theorem 92.** [Theorem 2.22] *Let $n \in \mathbb{N}$, let $C \in \mathbb{S}^n$, and let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be linear. Then there exists a $D \in \mathbb{S}^n$ such that $D \succcurlyeq 0$, $\mathcal{A}(D) = 0$, and $\langle C, D \rangle < 0$, if and only if $\mathcal{A}^*(y) \preccurlyeq C$ is not almost feasible.*

Note that (D) has a Slater point $\bar{y}=0$, $\eta=-\|C\|_2-1$ and its objective value is bounded above. Therefore Strong Duality Theorem applies.

By Strong Duality Theorem, the optimal objective values of this primal-dual pair are the same. Moreover, $\bar{X}:=0$ is a feasible solution with objective value 0, and by assumption (about $\nexists D$) it is optimal. Then either $\exists(\bar{y},\bar{\eta})$ feasible with $\bar{\eta}=0$ ($A^*(y)\preceq C$ is feasible) or at least there is a sequence $\{y^{(k)}, \eta_k\}$ such that $A^*(y^{(k)})+\eta_k I\preceq C$, $\eta_k\to 0^-$. Therefore, in either case, $A^*(y)\preceq C$ is almost feasible. ∎ 2014 05 29

Theorem 2.22: There exists $D\succeq 0$ such that $A(D)=0$, $\langle C,D\rangle<0$ if and only if the system $A^*(y)\preceq C$ is not almost feasible.

Proof: ($\Leftarrow$) was proved in thm 2.21 (b).

($\Rightarrow$): Suppose there exists such a (D). WLOG, pick such $D$ with $\langle C,D\rangle=-1$. Then for all $C'$ such that $\|C-C'\|<\|D\|^{-1}$, our choice for $D$ proves $A^*(y)\preceq C'$ has no
☆ solutions. ∎

What can we do if Slater condition fails? In many applications, we formulate say a primal SDP and that is the object of main interest.

$$
(P)\quad \begin{aligned} \inf\ & \langle C,X\rangle \\ & A(X)=b \\ & X\succeq 0 \end{aligned}
$$

Suppose neither (P) nor (D) has a Slater point. Borwein-Wolkowicz:
If $\{X\succeq 0;\ A(X)=b\}\neq\emptyset$ then it has a relative interior; restrict your space to the affine hull of

$$\operatorname{relint}\{X\succeq 0;\ A(X)=b\}.$$

In this smaller dimensional space, we do have Slater points. Then our strong duality theorem applies.
To apply this idea to (SDP), we need to look at the geometry of $S^n_+$.

(2014 05 29)

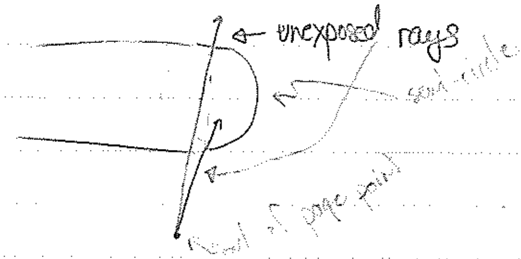
$K \subset \mathbb{R}^d$ is a closed convex cone, a closed convex cone $G \subseteq K$ such that $\forall u, v \in K$ with $u+v \in G$ we have $u, v \in G$, is called a <u>face of $K$</u>.
A face $G$ of $K$ is <u>exposed</u> if $\exists a \in \mathbb{R}^d \setminus \{0\}$ st

$$G = \{x \in K ; \langle a, x \rangle = 0\}$$

and

$$K \subseteq \{x \in \mathbb{R}^d ; \langle a, x \rangle \geq 0\}$$

$G$ is a proper face of $K$
if $G$ is a face of $K$
and $G \neq \emptyset$ and $G \neq K$.




Theorem 2.25.
(a) Every face $G$ of $\mathbb{S}^n_+$ is identified by a linear subspace $L \subseteq \mathbb{R}^n$:

$$G = \{X \in \mathbb{S}^n_+ ; \ \mathrm{Null}(X) \supseteq L\},$$
$$\mathrm{relint}(G) = \{X \in \mathbb{S}^n_+ ; \ \mathrm{Null}(X) = L\};$$

(b) Every proper face of $\mathbb{S}^n_+$ is exposed;
(c) Every face $G$ of $\mathbb{S}^n_+$ is <u>projectionally exposed</u> (ie:

$$G = (I-Q)\,\mathbb{S}^n_+\,(I-Q)$$

where $Q$ is the orthogonal projection onto linear subspace $L$ defining $G$).

In fact, every face $G$ of $\mathbb{S}^n_+$ is linearly isomorphic to $\mathbb{S}^k_+$ for some $k \leq n$.
$G$ is the image of

$$\left\{\begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{S}^n_+ ; \ X \in \mathbb{S}^k_+\right\} \text{ for some } k \leq n.$$

Recall the extreme rays (one dimensional faces) of $\mathbb{S}^n_+$

$$hh^T \text{ for } h \in \mathbb{R}^n.$$

Every extreme ray is linearly isomorphic under an automorphism of $\mathbb{S}^n_+$ to

$$\left\{\begin{bmatrix} \alpha & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{S}^n ; \ \alpha \in \mathbb{R}_+\right\} \qquad (h = e_1)$$

Thus, once we identify the minimal face (wrt inclusion) of $\mathbb{S}^n_+$ which contains our feasible region $\{X \in \mathbb{S}^n_+ ; \mathcal{A}(X) = b\}$, we can restrict our problem to $\mathbb{S}^k_+$ and for this smaller dimensional cone, we do ~~not~~ have a Slater point.

Note that in non-linear optimization adding redundant constraints can change the behaviour of the dual.

ex
$$C := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_1 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_2 := \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} \quad b := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(P) $\inf X_{11}$
$$X = \begin{bmatrix} 1 & 0 & X_{31} \\ 0 & 0 & 0 \\ X_{31} & 0 & X_{33} \end{bmatrix} \succeq 0$$
The optimal value is clearly 1.

(D) $\sup y_2$
$$\begin{bmatrix} 1-y_2 & & \\ & -y_1 & -y_2 \\ & -y_2 & 0 \end{bmatrix} \succeq 0$$
The optimal value is clearly 0.

Consider adding the redundant constraint $\langle A_3, X \rangle = 0$ for
$$A_3 := \begin{bmatrix} & & \\ & & 1 \\ & 1 & \end{bmatrix}$$
to (P). Note (P) remains unchanged. New dual:
$$\sup y_2$$
$$\begin{bmatrix} 1-y_2 & & \\ & -y_1 & -y_2-y_3 \\ & -y_2-y_3 & 0 \end{bmatrix} \succeq 0$$
has optimal value 1 as we no longer require $-y_2 = 0$.

2014 06 03

$$(P) \quad \begin{aligned} \inf\ & \langle C, X \rangle \\ & \mathcal{A}(X) = b \\ & X \succeq 0 \end{aligned}$$

(*) Is a hypothesis to the lemma

(*) Suppose (P) has feasible solution(s).
Restrict the feasible region to the minimal face of $\mathbb{S}^n_+$.

A key lemma to help us arrive at the minimal face:

compare with proof?

see (*)

(2014 06 03)

Lemma 2.27: Let $\mathcal{A}: \mathbb{S}^n \to \mathbb{R}^m$ linear, $b \in \mathbb{R}^m$, such that (P) is feasible. Then exactly one of the following holds:

(I) $\mathcal{A}(X) = b$, $X \in \mathbb{S}^n_{++}$;

(II) $\mathcal{A}^*(y) \in \mathbb{S}^n_+ \setminus \{0\}$ and $b^T y = 0$

If (I) does not hold, then $\exists \bar{y} \in \mathbb{R}^m$ such that $\mathcal{A}^*(\bar{y}) =: \bar{S} \in \mathbb{S}^n_+ \setminus \{0\}$ and $0 = b^T \bar{y}$. So

$$0 = b^T \bar{y} \underset{\substack{\uparrow \\ \text{since (P) is} \\ \text{feasible}}}{=} \mathcal{A}(\hat{X})^T \bar{y} = \langle \hat{X}, \mathcal{A}^*(\bar{y}) \rangle = \langle \bar{S}, \hat{X} \rangle.$$

We proved

$$\langle \bar{S}, X \rangle = 0 \quad \forall \text{ feasible } X.$$

Ramana [1997] gave an explicit construction. We will work with the dual form as our primal.

$$(D) \quad \begin{array}{l} \sup b^T y \\ \mathcal{A}^*(y) \preceq C \end{array}$$

$$(ELSD) \quad \begin{array}{ll} \inf & \langle C, U + W \rangle \\ \text{s.t.:} & \mathcal{A}(U+W) = b \\ & W \in \mathcal{W}_n \\ & U \succeq 0 \end{array}$$

(Extended Lagrange Slater Dual)

where $\mathcal{W}_n$ is a linear subspace of $\mathbb{S}^n$ explicitly represented as the feasible region of an SDP with $n$ $2n \times 2n$ matrix variables. (See textbook.)

Theorem 2.28 ~~If (D) is feasible~~ Suppose (D) has a finite optimal value. Then (ELSD) has the same optimal value and $\exists (\bar{U}, \bar{W})$ feasible in (ELSD) which attains that optimal value.

Recall theorem 2.22. Using theorem 2.8, we obtain

Theorem 2.29: For every $\mathcal{A}: \mathbb{S}^n \to \mathbb{R}^m$ linear, and $C \in \mathbb{S}^n$, exactly one of the following holds.

(I) $\mathcal{A}^*(y) \preceq C$ is feasible;

(II) $\exists U \succeq 0$, $W \in \mathcal{W}_n$ such that $\mathcal{A}(U+W) = 0$, $\langle C, U+W \rangle = -1$.

The previous theorem implies

Theorem 2.30: SDP feasibility is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ in the real number machine model.

Theorem 2.31: In the Turing Machine Model, if SDP-feasibility is in $\mathcal{NP}$ then it is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$.

It turns out in almost all "applications" we can find SDP formulations which satisfy the Slater condition.

Suppose $c \in \mathbb{R}^n$ is given and we want to solve

$$\inf_{x \in F} c^T x$$

where $F$ is a compact subset of $\mathbb{R}^n$. Suppose $\exists \mathcal{A}: \mathbb{S}^{n+1} \to \mathbb{R}^m$ linear such that

$$F = \left\{ x \in \mathbb{R}^n ; \mathcal{A}\left( \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} \right) = 0 \right\} \leftarrow \text{Homogeneous Equality form}$$

Let $Q \in \mathbb{S}^n$, $q \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ be given. Consider the quadratic equation

$$x^T Q x + 2q^T x + \gamma = 0.$$

Note any quadratic in $n$-variables can be written this way. But this is just

$$\left\langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} \right\rangle = 0.$$

Consider an arbitrary quadratic inequality

$$x^T Q x + 2q^T x + \gamma \leq 0$$

$$\iff x^T Q x + 2q^T x + \gamma + \tilde{s}^2 = 0$$

$$\iff \left\langle \begin{bmatrix} \gamma & q^T & 0 \\ q & Q & 0 \\ 0 & 0^T & 1 \end{bmatrix}, \begin{bmatrix} 1 & x^T & \tilde{s} \\ x & xx^T & \tilde{s}x \\ \tilde{s} & \tilde{s}x^T & \tilde{s}^2 \end{bmatrix} \right\rangle = 0.$$

Note that every polynomial inequality can be beaten down to a system of quadratic inequalities.

$$10x_1^4 x_2^3 + x_2^2 + 7x_3^5 - 3 \leq 0$$

$$\iff x_4 = x_1^2,\ x_5 = x_4^2,\ x_6 = x_2^2,\ x_7 = x_2 x_6,\ x_8 = x_3^2,\ x_9 = x_8^2,\ x_{10} = x_9 x_3$$

$$10x_5 x_7 + x_2^2 + 7x_{10} - 3 \leq 0$$

2014 06 10

Proposition 2.32: Solution set of every finite system of polynomial equations and inequalities can be expressed in Homogeneous Equality Form.
Note that due to $x_j^2 - x_j = 0 \; \forall j \in \{1, \dots, n\}$ being a system of quadratic equations, 0,1 IP and 0,1 MIP are both special cases.

$$\mathcal{F} := \operatorname{conv}\left\{ \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} : x \in F \right\},$$

$$c^T x \longleftrightarrow \left\langle \begin{bmatrix} 0 & \frac{1}{2}c^T \\ \frac{1}{2}c & 0 \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \right\rangle$$

anything symmetric

An SDP relaxation of $\mathcal{F}$ is

check

$$\hat{\mathcal{P}} := \left\{ \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}^{n+1}_+ : \mathcal{A}\left( \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \right) = 0 \right\} \supseteq \mathcal{F}$$

Theorem 2.33: Let $F \subset \mathbb{R}^n$ be a set which admits a Homogeneous Equality Form representation and $\operatorname{conv}(F)$ is full dimensional. Then $\hat{\mathcal{P}}$ has a Slater point.
Proof: $\exists \{v^{(1)}, \dots, v^{(n)}, v^{(n+1)}\} \subseteq F$ such that

$$\left\{ \begin{pmatrix} 1 \\ v^{(1)} \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ v^{(n+1)} \end{pmatrix} \right\}$$

is linearly independent in $\mathbb{R}^{n+1}$. Therefore

$$\begin{pmatrix} 1 \\ v^{(i)} \end{pmatrix} \begin{pmatrix} 1 & v^{(i)T} \end{pmatrix} \in \hat{\mathcal{P}} \qquad \forall i \in \{1, \dots, n+1\}.$$

Since $\hat{\mathcal{P}}$ is convex, every convex combination of these matrices lie in $\hat{\mathcal{P}}$. So

$$V := \frac{1}{n+1} \sum_{i=1}^{n+1} \begin{pmatrix} 1 \\ v^{(i)} \end{pmatrix} \begin{pmatrix} 1 & v^{(i)T} \end{pmatrix} \in \hat{\mathcal{P}} \subseteq \mathbb{S}^{n+1}_+.$$

By proposition 1.11, $V \in \mathbb{S}^{n+1}_{++}$ and is a Slater point for $\hat{\mathcal{P}}$.

What if $\dim(\operatorname{conv}(F)) = d < n$?
Once we know what $d$ is, in a constructive way, then we can apply the above construction of Slater points for a suitable SDP relaxation.
Suppose we know the affine hull of $F$, $\exists \, l \in \mathbb{R}^n$ and $L \in \mathbb{R}^{d \times n}$ such that

$$x \in F \Rightarrow x = l + L^T y \text{ for some } y \in \mathbb{R}^d.$$

read chap 2,3

Define

$$\mathcal{L}: \mathcal{S}^{n+1} \to \mathcal{S}^{d+1}$$

$$\mathcal{L}(Z) := \begin{bmatrix} 1 & \ell^T \\ 0 & L \end{bmatrix} Z \begin{bmatrix} 1 & 0 \\ \ell & L^T \end{bmatrix}$$

$$\mathcal{L}^*: \mathcal{S}^{d+1} \to \mathcal{S}^{n+1}$$

$$\mathcal{L}^*(W) = \begin{bmatrix} 1 & 0 \\ \ell & L^T \end{bmatrix} W \begin{bmatrix} 1 & \ell^T \\ 0 & L \end{bmatrix}$$

$$\bar{\mathcal{A}}: \mathcal{S}^{d+1} \to \mathbb{R}^m$$

$$\bar{\mathcal{A}}(W) := \mathcal{A}(\mathcal{L}^*(W))$$

We have an equivalent formulation of F in Homogeneous Equality Form with linear transformation $\bar{\mathcal{A}}$.
Note that the new formulation corresponds to a full-dimensional set (the d-dimensional convex hull of F). The corresponding SDP relaxation has the feasible region

$$\hat{P}_{\mathcal{L}} := \left\{ \begin{bmatrix} 1 & y^T \\ y & Y \end{bmatrix} \in \mathcal{S}_+^{d+1} ; \bar{\mathcal{A}}\left( \begin{bmatrix} 1 & y^T \\ y & Y \end{bmatrix} \right) = 0 \right\}.$$

Theorem 2.34: Slater condition holds for $\hat{P}_{\mathcal{L}}$.