Strong Duality - Why do We Need Slater's Condition?

Conic Programming Scenario

An Euclidean space $\mathbb{E}$ is a finite-dimensional vector space over $\mathbb{R}$ endowed with an inner product $\langle \cdot, \cdot \rangle$. If $\mathbb{E}$ and $\mathbb{Y}$ are Euclidean spaces, the *direct sum* of $\mathbb{E}$ and $\mathbb{Y}$ is the Euclidean space

$$\mathbb{E} \oplus \mathbb{Y} := \{x \oplus y : x \in \mathbb{E} \text{ and } y \in \mathbb{Y}\},$$

where we consider for each $x_1 \oplus y_1$, $x_2 \oplus y_2 \in \mathbb{E} \oplus \mathbb{Y}$ and $\alpha \in \mathbb{R}$:

(i) $(x_1 \oplus y_1) + (x_2 \oplus y_2)_{\mathbb{E}\oplus\mathbb{Y}} = (x_1 +_{\mathbb{E}} x_2) \oplus (y_1 +_{\mathbb{Y}} y_2)$;

(ii) $\alpha(x_1 \oplus y_1) = \alpha x_1 \oplus \alpha y_1$;

(iii) $\langle x_1 \oplus y_1, x_2 \oplus y_2 \rangle_{\mathbb{E}\oplus\mathbb{Y}} = \langle x_1, x_2 \rangle_{\mathbb{E}} +_{\mathbb{R}} \langle y_1, y_2 \rangle_{\mathbb{Y}}$.

Definition. Let $\mathbb{E}$ be an Euclidean space. A *cone* is a set $K \subseteq \mathbb{E}$ such that $\alpha x \in K$ for each $x \in K$ and $\alpha \in \mathbb{R}_{++}$. A *hyperplane* is a set of the form $\{x \in \mathbb{E} : \langle x, a \rangle = \beta\}$ for some $0 \neq a \in \mathbb{E}$ and $\beta \in \mathbb{R}$. Similarly, a (closed) *halfspace* is a set of the form $\{x \in \mathbb{E} : \langle a, b \rangle \leq \beta\}$ for some $0 \neq a \in \mathbb{E}$ and $\beta \in \mathbb{R}$. A *polyhedron* is the intersection of finitely many halfspaces.

Definition. Let $\mathbb{E}$ be an Euclidean space. A cone $K \subseteq \mathbb{E}$ is *pointed* if $K \cap -K = \{0\}$. We also say that K is proper if K is *convex*, *closed*, pointed, and $\operatorname{int}(K) \neq \varnothing$. Moreover, the cone K is *polyhedral* if it is a polyhedron.

Definition. Let $\mathbb{E}$ and let $K \subseteq \mathbb{E}$ be a cone. The *dual cone* of K is the set

$$K^* := \{x \in \mathbb{E} : \langle x, k \rangle \geq 0 \text{ for each } k \in K\}.$$

Example. $\mathbb{R}^n$ and $\mathbb{S}^n$ are the classic examples of Euclidean spaces. Some basic examples of proper cones are the n-dimensional semidefinite cone $\mathbb{S}^n_+$, the n-dimensional p-norm cones

$$\mathbb{L}^p_n := \{x \oplus \lambda \in \mathbb{R}^n \oplus \mathbb{R}_+ : \|x\|_p \leq \lambda\},$$

and the exponential cone

$$\mathbb{G}_n := \left\{ (x \oplus \theta \oplus \beta) \in \mathbb{R}^n \oplus \mathbb{R}_+ \oplus \mathbb{R}_+ : \theta \sum_{i \in [n]} \exp\left(\frac{-x_i}{\theta}\right) \leq \beta \right\},$$

where we consider $0 \exp(\frac{\alpha}{0}) = 0$ for each $\alpha \in \mathbb{R}$. In fact, under certain (not that specific) conditions, the epigraph of a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a proper cone. The set $\mathbb{R}^n_+$ is an important polyhedral cone.

Cone Partial Order

Definition. Let S be a set. A *partial order* on S is a binary relation $\leq$ such that, for each $a, b, c \in S$:

(i) $a \leq a$;

(ii) if $a \leq b$ and $b \leq a$, then $a = b$;

(iii) if $a \leq b$ and $b \leq c$, then $a \leq c$.

Example. Let S be any set. If we consider for each $A, B \in \mathcal{P}(S)$ that $A \leq B$ if $A \subseteq B$, then we have a partial order on $\mathcal{P}(S)$.

Moreover, if for each $x, y \in \mathbb{R}^n$ we consider $x \leq y$ if $x_i \leq y_i$ for each $i \in [n]$, then we have a partial order on $\mathbb{R}^n$.

Definition. Let $\mathbb{E}$ be an Euclidean space, $K \subseteq \mathbb{E}$ be a proper cone and $x, y \in \mathbb{E}$. Then, the cone K induces an order in $\mathbb{E}$ as follows:

$$x \succeq_K y \text{ if } x - y \in K.$$

Moreover,

$$x \succ_K y \text{ if } x - y \in \operatorname{int}(K).$$

The expression $x \succeq_K y$ may be read as x is greater or equal to y in K.

The reader should notice that $x \in K$ if, and only if $x \succeq_K 0$. Thus,

$$K = \{x \in \mathbb{E} \colon x \succeq_K 0\}.$$

Also note that the order does not depend on $\langle \cdot, \cdot \rangle$. Therefore, it could be said that the cone K induces an order in the vector space V. Here, this will not be done for simplicity. Moreover, the order considered at the very beginning of this text for $\mathbb{R}^n$ can be seen as the special case of this definition where $\mathbb{E} = \mathbb{R}^n$ equipped with any inner product and $K = \mathbb{R}^n_+$.

Proposition 1. Let $\mathbb{E}$ be an Euclidean space and let $K \subseteq \mathbb{E}$ be a proper cone. Then, $\succeq_K$ is a partial order on $\mathbb{E}$.

Proof. For reflexivity, let $x \in \mathbb{E}$. Since K is closed, we have that $0 \in K$ and hence $x - x = 0 \in K$. Thus, $x \succeq_K x$. Antisymmetry follows from the fact that, whenever $x, y \in \mathbb{E}$, $x \succeq_K y$ and $y \succeq_K x$, we have $x - y \in K$, and $y - x = -(x - y) \in K$. Since K is pointed, this implies that $x - y = 0$. For transitivity, let $x, y, z \in K$ such that $x \succeq_K y$ and $y \succeq_K z$. We have :

$$x - y \in K \text{ and } y - z \in K$$

Since K is convex $(x - y) + (y - z) = x - z \in K$. Hence, $x \succeq_K z$. $\square$

Hopefully, Proposition 1 makes it clear why a proper cone K being required to be convex, pointed and closed contributes for the definition of this partial order. Requiring $\mathrm{int}(K)$ to be nonempty allows us to consider strict inequalities. Furthermore, we can similarly consider $x \preceq_K y$ if, and only if $-x \succeq_K -y$. Also, we remark that because K^* is a proper cone as well, we can also define a partial order using K^*. Next, we present an example illustrating why '$\succeq_K$' is not a total order on $\mathbb{E}$:

Example. Let $\mathbb{E}$ be $\mathbb{R}^m$ and K be $\mathbb{R}^m_+$. Then, take $x = e_1$ and $y = e_2$ and see that $x - y \notin K$ and $y - x \notin K$ as both have a negative coordinate.

Overview of Duality Theory

Definition. Let $\mathbb{E}_1, \mathbb{E}_2, \mathbb{E}_3$ and $\mathbb{Y}_1, \mathbb{Y}_2, \mathbb{Y}_3$ be Euclidean spaces. Let $K \subseteq \mathbb{E}_1$ and $L \subseteq \mathbb{Y}_1$ both be proper cones. Let $K_p \subseteq \mathbb{E}_2$ and $L_p \subseteq \mathbb{Y}_2$ both be polyhedral cones. Consider $c_1 \oplus c_2 \oplus c_3 \in \mathbb{E}_1 \oplus \mathbb{E}_2 \oplus \mathbb{E}_3$ and $b_1 \oplus b_2 \oplus b_3 \in \mathbb{Y}_1 \oplus \mathbb{Y}_2 \oplus \mathbb{Y}_3$. Let $A \colon \mathbb{E}_1 \oplus \mathbb{E}_2 \oplus \mathbb{E}_3 \to \mathbb{Y}_1 \oplus \mathbb{Y}_2 \oplus \mathbb{Y}_3$ be a linear function.

A *conic optimization problem* is an optimization problem of the form:

$$\begin{aligned} \text{minimize} \quad & \langle x_1 \oplus x_2 \oplus x_3, c_1 \oplus c_2 \oplus c_3 \rangle \\ \text{subject to} \quad & A(x_1 \oplus x_2 \oplus x_3) \succeq_{L^* \oplus L_p^* \oplus \{0\}} b_1 \oplus b_2 \oplus b_3, \\ & x_1 \oplus x_2 \oplus x_3 \in K \oplus K_p \oplus \mathbb{E}_3. \end{aligned} \tag{1}$$

The set $G \coloneqq \{x_1 \oplus x_2 \oplus x_3 \in K \oplus K_p \oplus \mathbb{E}_3 \colon A(x_1 \oplus x_2 \oplus x_3) \succeq_{L^* \oplus L_p^* \oplus \{0\}} b_1 \oplus b_2 \oplus b_3\}$ is the feasible set of (1). According to the notation presented in the Preliminaries, a conic optimization problem can be represented simply as $(G, \langle \cdot, c_1 \oplus c_2 \oplus c_3 \rangle)$.

We now define the dual problem of (1).

Definition. Consider the conic optimization problem (1). The *dual problem* of (1) is the conic optimization problem

$$\begin{aligned} \text{maximize} \quad & \langle b_1 \oplus b_2 \oplus b_3, y_1 \oplus y_2 \oplus y_3 \rangle \\ \text{subject to} \quad & A^*(y_1 \oplus y_2 \oplus y_3) \preceq_{K^* \oplus K_p^* \oplus \{0\}} c_1 \oplus c_2 \oplus c_3, \\ & y_1 \oplus y_2 \oplus y_3 \in L \oplus L_p \oplus \mathbb{Y}_3. \end{aligned} \tag{2}$$

To simplify the notation on the following propositions, we will denote $x \coloneqq x_1 \oplus x_2 \oplus x_3$, $y \coloneqq y_1 \oplus y_2 \oplus y_3$, $c \coloneqq c_1 \oplus c_2 \oplus c_3$, and $b \coloneqq b_1 \oplus b_2 \oplus b_3$ on (1) and (2) until the end of this section.

Example. Semidefinite programming, linear programming, second order cone programming and many other fancy types of optimization problems fit into the scope of the previous definition

Example. Many problems in graph theory are tackled using conic programming. For example, the graph coloring problem can be studied via semidefinite programming. Moreover, the classic regression problems in statistical learning can also be approached via conic optimization.

Theorem 2 (Weak duality)**.** Let α be the optimal value of (1) and let β be the optimal value of (2). If x is feasible in (1) and y is feasible in (2), then $\langle b, y\rangle \leq \langle c, x\rangle$. In particular, $\alpha \geq \beta$. Moreover, if $\langle x, c\rangle = \langle b, y\rangle$ then x and y are optimal solutions for their respective problems and $\alpha = \beta$.

Proof. Since y is a feasible point in (2), we have that $A^*(y) \preceq_{K^* \oplus K_p^* \oplus \{0\}} c$. By definition, this is equivalent to $c - A^*(y) \in K^* \oplus K_p^* \oplus \{0\}$, implying that

$$\langle c - A^*(y), x\rangle = \langle c, x\rangle - \langle A^*(y), x\rangle \geq 0.$$

Similarly, since x is feasible in (1) we have that $A(x) - b \in L^* \oplus L_p^* \oplus \{0\}$, which yields

$$\langle A(x) - b, y\rangle = \langle A(x), y\rangle - \langle b, y\rangle \geq 0.$$

Thus, we conclude that $\langle b, y\rangle \leq \langle c, x\rangle$ by the definition of an adjoint operator. Obviously, this implies that $\alpha \geq \beta$. Finally, assume that $\langle x, c\rangle = \langle y, b\rangle$ and note that in this case we have $\langle x, c\rangle \leq \langle \bar{x}, c\rangle$ for each $\bar{x}$ feasible in (1). That is, x is optimal. Symmetrically, we obtain that that $\langle y, b\rangle \geq \langle \bar{y}, b\rangle$ for each $\bar{y}$ feasible in (2), concluding that y is optimal as well. Clearly, this implies that $\alpha = \beta$. $\square$

Corollary 3 (Complementary Slackness)**.** Let α be the optimal value of (1) and let β be the optimal value of (2). Let x be a feasible solution in (1) and y be a feasible solution in (2). Then x and y are optimal in their respective problems and $\alpha = \beta$ if, and only if,

$$\langle x, c - A^*(y)\rangle = \langle A(x) - b, y\rangle = 0.$$

Proof. From the proof of Theorem 2 we obtain that, whenever x is feasible in (1) and y is feasible in (2):

$$\langle x, c\rangle \geq \langle A^*(y), x\rangle = \langle y, A(x)\rangle \geq \langle b, y\rangle.$$

In particular, if x and y are optimal and $\alpha = \beta$, then $\langle c, x\rangle = \langle b, y\rangle$, which forces $\langle x, c\rangle = \langle x, A^*(y)\rangle$ and hence $\langle x, c - A^*(y)\rangle = 0$. Symmetrically, $\langle A(x), y\rangle = \langle b, y\rangle$ implies that $\langle A(x) - b, y\rangle = 0$.

Conversely, assume that

$$\langle x, c - A^*(y)\rangle = 0 \text{ and } \langle A(x) - b, y\rangle = 0.$$

Then, $\langle x, c\rangle = \langle x, A^*(y)\rangle$ and $\langle A(x), y\rangle = \langle b, y\rangle$. Applying the definition of an adjoint operator and Theorem 2 produces the desired result. $\square$

Proposition 4. Let $\mathbb{E}$ be an Euclidean space, let $K \subseteq \mathbb{E}$ be a proper cone, let $K_p \subseteq \mathbb{E}$ be a polyhedral cone and $S \subseteq \mathbb{E}$ a linear subspace. If $\operatorname{int}(K) \cap K_p \cap S \neq \varnothing$, then

$$(K \cap K_p \cap S)^* = (K^* + K_p^* + S^\perp).$$

Proof. The inclusion $(K^* + K_p^* + S^\perp) \subseteq (K \cap K_p \cap S)^*$ is easy to prove. Let $a + b + c \in (K^* + K_p^* + S^\perp)$. Then, since $a \in K^*, b \in K_p^*$, and $c \in S^\perp$, we have for each $x \in K \cap K_p \cap S$

$$\langle a + b + c, x\rangle = \langle a, x\rangle + \langle b, x\rangle + \langle c, x\rangle \geq 0,$$

which implies that $x \in (K \cap K_p \cap S)^*$.

For the reverse inclusion, we show that $(K \cap K_p \cap S) \supseteq (K^* + K_p^* + S^\perp)^*$ then, the desired result will follow because:

(i) If $\mathbb{E}$ is an Euclidean space and $K_1 \subseteq K_2 \subseteq \mathbb{E}$ are cones. Then $K_2^* \subseteq K_1^*$.

(ii) Let $\mathbb{E}$ be an Euclidean space, Let $\{K_i\}_{i \in I} \subseteq \mathbb{E}$ be a finite family of convex cones. Assume that there exists $I_0 \subseteq I$ such that K_i is polyhedral for each $i \in I_0$. If $\bigcap_{i \in I_0} K_i \cap \bigcap_{i \in I \setminus I_0} \operatorname{ri}(K_i) \neq \varnothing$, then $\sum_{i \in I} K_i^*$ is closed.

(iii) If $K \subseteq \mathbb{E}$ is a closed cone then $K^{**} = K$.

Let $x \in (K^* + K_p^* + S^\perp)^*$. By definition,

$$\langle a + b + c, x\rangle = \langle a, x\rangle + \langle b, x\rangle + \langle c, x\rangle \geq 0 \text{ for each } a + b + c \in (K^* + K_p^* + S^\perp).$$

We want to conclude $x \in (K \cap K_p \cap S)$. First, note that $0 \in (K^* \cap K_p^* \cap S^\perp)$ and hence either a, b or c can be zero. If $b = c = 0$ we obtain that $\langle a, x\rangle \geq 0$ for each $x \in K^*$ and thus $x \in (K^*)^* = K$. Similarly, setting $a = c = 0$ we conclude that $x \in K_p$. Finally, setting $a = b = 0$, we obtain that $\langle x, c\rangle$ should be nonnegative. However, since $S^\perp$ is a linear subspace, it is true that $-c \in S^\perp$. This forces $\langle x, c\rangle$ to be zero for each $c \in S^\perp$ and hence we conclude that $x \in (S^\perp)^\perp = S$. Therefore, $x \in (K \cap K_p \cap S)$. $\square$

Theorem 5 (Strong duality)**.** Consider the optimization problem (1). If (1) is bounded below and has a restricted Slater point, then the optimal values of (1) and its dual (2) are equal and (2) has an optimal solution.

Proof. Let $\alpha \in \mathbb{R}$ be the optimal value of (1) and consider the following objects:

(i) $\bar{\mathbb{E}} \coloneqq \mathbb{E}_1 \oplus \mathbb{E}_2 \oplus \mathbb{E}_3$

(ii) $\mathbb{K} \coloneqq K \oplus \mathbb{E}_2 \oplus \mathbb{E}_3 \oplus \mathbb{R}_+$;

(iii) $\mathbb{K}_p \coloneqq \mathbb{E}_1 \oplus K_p \oplus \mathbb{E}_3 \oplus \mathbb{R}_+$;

(iv) $\mathbb{L} \coloneqq L \oplus L_p \oplus \mathbb{Y}_3$

(v) $S \coloneqq \{x \oplus t \in \bar{\mathbb{E}} \oplus \mathbb{R} : A(x) \succeq_{\mathbb{L}^*} tb\}$.

First, we check that S is indeed a linear subspace. Note that $0 \oplus 0 \in S$ since $0 \in \mathbb{L}^*$. Also, if $x \oplus t \in S$ and $\lambda \in \mathbb{R}$, then it is clear that $A(\lambda x) = \lambda A(x) \succeq_{\mathbb{L}^*} \lambda tb$ and hence $\lambda(x \oplus t) \in S$. Finally, if $x_1 \oplus t_1$ and $x_2 \oplus t_2$ belong to S, then we have that $A(x_1) - t_1 b$ and $A(x_2) - t_2 b$ belong to $\mathbb{L}^*$. Since this set is a convex cone, it follows that $A(x_1 + x_2) - (t_1 + t_2)b \in \mathbb{L}^*$ and then we conclude that $(x_1 + x_2) \oplus (t_1 + t_2) \in S$. Therefore S is a subspace. Together with the definition of S, this fact implies that $x \oplus t \in S$ if, and only if

$$\langle A(x) - tb, y \rangle = \langle x, A^*(y) \rangle - t\langle b, y \rangle = 0 \text{ for each } y \in \mathbb{L}.$$

Thus, we have that

$$S^\perp = \{A^*(y) \oplus (-\langle b, y \rangle) : y \in \mathbb{L}\}$$

.

Moreover, note that $\mathbb{K}^* = K^* \oplus \{0\} \oplus \{0\} \oplus \mathbb{R}_+$ and $\mathbb{K}_p^* = \{0\} \oplus K_p^* \oplus \{0\} \oplus \mathbb{R}_+$ and also that $c \in (\mathbb{K} \cap \mathbb{K}_p \cap S)^*$ because α is the optimal value of Problem (1). Since there is a restricted Slater point by hypothesis, we can apply Proposition 4 to conclude that there exists $z \in \mathbb{K}^*$, $w \in \mathbb{K}_p^*$ and $v \in S^\perp$ such that $c = z + v + w$. The last coordinate of this equation gives us:

$$-\alpha = (\beta + \gamma) - \langle b, y \rangle \text{ for some } \beta, \gamma \in \mathbb{R}_+ \text{ and } y \in \mathbb{L}$$

This equality implies that $\alpha \leq \langle b, y \rangle$. Since $\alpha \geq \langle b, y \rangle$ by Theorem 2, the result follows. Of course, y in an optimal solution for Problem (2). □

Race for Closedness

Our proof of strong duality relies heavily on Proposition 4, which in turn needs the set $K^* + K_p^* + S^\perp$ (in that context) to be closed. As we shall discuss, this fact is closely related with the commutativity between closures and linear images of convex sets and their duals.

Let $\mathbb{E}$ be an Euclidean space and let $\varnothing \neq C_1, C_2 \subseteq \mathbb{E}$ both be convex sets. Then $C_1 + C_2$ can be seen as $A(C_1 \oplus C_2)$ where $A \colon \mathbb{E} \oplus \mathbb{E} \to \mathbb{E}$ is given by $A(x_1 \oplus x_2) = x_1 + x_2$.

If we imagine an ideal scenario where we had $\overline{A(C_1 \oplus C_2)} = A(\overline{C_1 \oplus C_2})$, we could apply A to $K^* + K_p^* + S^\perp$ and nothing we will see next would be required.

Remark. In the case C_1 and C_2 are polyhedra, the ideal scenario is actually real. That's why there is no need for constraint qualifications in linear programming.

Unfortunately, the ideal scenario does not hold in general. In fact, we have that

Proposition 6. Let $\mathbb{E}$ and $\mathbb{Y}$ be Euclidean spaces, let $C \subseteq \mathbb{E}$ be a convex set, and let $A \colon \mathbb{E} \to \mathbb{Y}$ be a linear function. Then:

(i) $A(\overline{C}) \subseteq \overline{A(C)}$;

(ii) $\operatorname{ri}(A(C)) = A(\operatorname{ri}(C))$.

Proof.

(i) Let $x \in A(\overline{C})$ and consider $y \in \overline{C}$ such that $A(y) = x$. Let $\varepsilon \in \mathbb{R}_{++}$. Then, there exists $\delta \in \mathbb{R}_{++}$ such that $f(z) \in x + \varepsilon\mathbb{B}$ for each $z \in y + \delta\mathbb{B}$. Since $y \in \overline{C}$, we know that $y + \delta\mathbb{B} \cap C \neq \varnothing$. Thus, we can assume that $z \in C$, obtaining that $f(z) \in A(C)$. Therefore, $(x + \varepsilon\mathbb{B}) \cap A(C) \neq \varnothing$ for each $\varepsilon \in \mathbb{R}_{++}$. That is, $x \in \overline{A(C)}$.

(ii) Let $x \in \operatorname{ri}(A(C)) \subseteq A(C)$ and assume that $x \notin A(\operatorname{ri}(C))$. Thus, for each $y \in C$ such that $A(y) = x$ we have that $y \in C \setminus \operatorname{ri}(C)$. Then, for each $\varepsilon \in \mathbb{R}_{++}$, there exists $z \in y + \varepsilon\mathbb{B}$ such that $z \in \mathbb{E} \setminus C$. We know that for each $\gamma \in \mathbb{R}_{++}$ there exists $w \in x + \gamma\mathbb{B}$ such that $w \in \mathbb{Y} \setminus A(C)$. Therefore, $x \notin \operatorname{ri}(A(C))$.

We now prove that $A(\operatorname{ri}(C)) \subseteq \operatorname{ri}(A(C))$. Let $x_1 \in A(\operatorname{ri}(C))$ and let $y_1 \in A(C)$. Consider $x_2 \in \operatorname{ri}(C)$ such that $A(x_2) = x_1$ and $y_2 \in C$ such that $A(y_2) = y_1$. Then, there exists $\lambda > 1$ such that $(1-\lambda)y_2 + \lambda x_2 \in C$. Thus,

$$(A((1-\lambda)y_2 + \lambda x_2) = (1-\lambda)A(y_2) + \lambda A(x_2) = (1-\lambda)y_1 + \lambda x_1 \in C.$$

Therefore, $x_1 \in \operatorname{ri}(A(C))$. $\square$

The first question one asks is: When does item (ii) of Proposition 6 hold with equality? In order to have an satisfactory answer, we need one additional tool.

Definition. Let $\mathbb{E}$ be an Euclidean space and let $\varnothing \neq C \subseteq \mathbb{E}$ be a convex set. The *recession cone* of C is the set

$$0^+C \coloneqq \{y \in \mathbb{E} : x + \alpha y \in C \text{ for each } x \in C \text{ and } \alpha \in \mathbb{R}_{++}\}.$$

The *lineality space* of C is the set $\operatorname{lin}(C) \coloneqq 0^+C \cap (-0^+C)$.

Proposition 7. Let $\mathbb{E}$ and $\mathbb{Y}$ be Euclidean spaces, let $A\colon \mathbb{E} \to \mathbb{Y}$ be a linear function, and let $\varnothing \neq C \subseteq \mathbb{E}$ be a closed convex set. If $0^+C \cap \operatorname{Null}(A) \subseteq \operatorname{lin}(C)$, then $A(C)$ is closed.

Proof. Let $x \in \overline{A(C)}$. By definition, for each $\varepsilon \in \mathbb{R}_{++}$ we have $(x + \varepsilon\mathbb{B}) \cap A(C) \neq \varnothing$. For each $k \in \mathbb{Z}_{++}$, consider $\varepsilon_k \coloneqq \frac{1}{k}$ and $x_k \in x + \varepsilon_k\mathbb{B}$. Observe that the limit of the sequence $\{x_k\}_{k\in\mathbb{Z}_{++}}$ is x and that $x_k \in x + \bigcap_{i\leq k} \varepsilon_i\mathbb{B}$. Consider, for each $k \in \mathbb{Z}_{++}$:

$$C_k \coloneqq \{y \in C : A(y) \in x + \varepsilon_k\mathbb{B}\} = C \cap A^{-1}(x + \varepsilon_k\mathbb{B}).$$

Note that $x_k \in C_k$ for each $k \in \mathbb{Z}_{++}$ and thus C_k is always nonempty. Moreover,

$$\bigcap_{k\in\mathbb{Z}_{++}} C_k = \{y \in C : A(y) \in x + \varepsilon_k\mathbb{B}, \text{ for each } k \in \mathbb{Z}_{++}\} = \{y \in C : A(y) = x\}.$$

Thus, it suffices to show that $\bigcap_{k\in\mathbb{Z}_{++}} C_k$ is nonempty. Then we have that, $0^+C_k = 0^+C \cap \operatorname{Null}(A)$, that $\operatorname{lin}(C_k) = \operatorname{lin}(C) \cap \operatorname{Null}(A)$, and that C_k is closed and convex for each $k \in \mathbb{Z}_{++}$. Also, by hypothesis we know that $0^+C \cap \operatorname{Null}(A) \subseteq \operatorname{lin}(C)$, which implies that

$$0^+C \cap \operatorname{Null}(A) \subseteq \operatorname{lin}(C) \cap \operatorname{Null}(A).$$

Since the converse is always true, we conclude that these sets are actually equal. Thus, we obtain that $\bigcap_{k\in\mathbb{Z}_{++}} C_k \neq \varnothing$ and thus $x \in A(C)$. Therefore $\overline{A(C)} \subseteq A(C)$. That is, $A(C)$ is closed. $\square$

Corollary 8. Let $\mathbb{E}$ be an Euclidean space, let $\{C_i\}_{i\in I} \subseteq \mathbb{E}$ be a finite family of nonempty convex sets. If $x_i \in C_i$ for each $i \in I$ and $\sum_{i\in I} x_i = 0$ implies that $x_i \in \operatorname{lin}(C_i)$ for each $i \in I$, then $\overline{\sum_{i\in I} C_i} = \sum_{i\in I} \overline{C_i}$.

Proof. Consider the Euclidean space $\mathbb{E}^I$ and the linear transformation $A\colon \mathbb{E}^I \to \mathbb{E}$ where $A(x) = \sum_{i\in I} x_i$. Then $\operatorname{Null}(A) = \{x \in \mathbb{E}^I : \sum_{i\in I} x_i = 0\}\}$. Moreover, if $C \coloneqq \bigoplus_{i\in I} C_i$, then trivially $0^+C = \bigoplus_{i\in I} 0^+C_i$ and, needless to say, $\operatorname{lin}(C) = \bigoplus_{i\in I} \operatorname{lin}(C_i)$. Thus, our hypothesis implies that $0^+C \cap \operatorname{Null}(A) \subseteq \operatorname{lin}(C)$. Therefore, Proposition 7 yields the desired result. $\square$

Finally, we need to relate our previous achievements with duality theory and the next result is the key to it.

Proposition 9. Let $\mathbb{E}$ be an Euclidean space, let $S \subseteq \mathbb{E}$ be a subspace of $\mathbb{E}$, and let $K \subseteq \mathbb{E}$ be a convex cone. Then exactly one of the following two statements is true:

(i) There is no hyperplane separating S and K properly;

(ii) There exists x such that $x \in S^\perp$, $x \in -K^*$, and $x \notin K^*$.

Proof. We know that there exists a hyperplane separating S and K properly if, and only if there exists $x \in \mathbb{E} \setminus \{0\}$ such that

$$\inf_{s\in S}\{\langle x, s\rangle\} \geq \sup_{k\in K}\{\langle x, k\rangle\}$$

and

$$\sup_{s\in S}\{\langle x, s\rangle\} > \inf_{k\in K}\{\langle x, k\rangle\}.$$

These inequalities are equivalent to

$$-\delta(-x \,|\, S^{\perp}) \geq \delta(x \,|\, -K^*) \tag{3}$$

and

$$\delta(x \,|\, S^{\perp}) > -\delta(-x \,|\, -K^*). \tag{4}$$

Analyzing each possible case, we easily conclude that (3) and (4) hold if, and only if $x \in S^{\perp}$, $x \in -K^*$, and $x \notin K^*$. □

Now things start to connect:

Corollary 10. Let $\mathbb{E}$ be an Euclidean space and let $\{K_i\}_{i\in I} \subseteq \mathbb{E}$ be a finite family of convex cones. Then exactly one of the following two statements is true:

(i) There exists $y \in \bigcap_{i\in I} \operatorname{ri}(K_i)$;

(ii) There exists a family $\{x_i\}_{i\in I}$ such that $\sum_{i\in I} x_i = 0$, $x_i \in -K_i^*$ for each $i \in I$, and $x_i \notin K_i^*$ for some $i \in I$.

Proof. Consider the Euclidean space $\mathbb{E}^I$ and the cone $\bigoplus_{i\in I} K_i \subseteq \mathbb{E}^I$. By Corollary 6, we have that $\operatorname{ri}(K) = \bigoplus_{i\in I} \operatorname{ri}(K_i)$. Set $S \coloneqq \{x \in \mathbb{E}^I : x_i = x_j \text{ for each } i, j \in I\}$. Then , for each $x \in S$ and $y \in \mathbb{E}^I$:

$$\langle x, y\rangle = \sum_{i\in I}\langle x_i, y_i\rangle = |I|\langle x_i, \sum_{i\in I} y_i\rangle.$$

The latter expression is zero for each $x \in S$ if and only if $\sum_{i\in I} y_i = 0$. Thus,

$$S^{\perp} = \{y \in \mathbb{E}^I : \sum_{i\in I} y_i = 0\}.$$

Applying Proposition 9 for S and K, we conclude that exactly one of the following statements is true:

(i) There exists $y \in S \cap \operatorname{ri}(K)$;

(ii) There exists x such that $x \in S^{\perp}$, $x \in -K_i^*$, and $x \notin K^*$.

Note that $K^* = \bigoplus_{i\in I} K_i^*$. Also, we have that $S \cap \operatorname{ri}(K) \neq \varnothing$ if, and only if $\bigcap_{i\in I} \operatorname{ri}(K_i) \neq \varnothing$. Thus, the former alternatives are equivalent to

(i) There exists $y \in \bigcap_{i\in I} \operatorname{ri}(K_i)$;

(ii) There exists a family $\{x_i\}_{i\in I}$ such that $\sum_{i\in I} x_i = 0$, $x_i \in -K^*$ for each $i \in I$, and $x_i \notin K_i^*$ for some $i \in I$. □

Corollary 11. Let $\mathbb{E}$ be an Euclidean space and let $\{K_i\}_{i\in I} \subseteq \mathbb{E}$ be a finite family of convex cones. If $\bigcap_{i\in I} \operatorname{ri}(K_i) \neq \varnothing$ then $\sum_{i\in I} K_i^*$ is closed.

Proof. Since $\bigcap_{i\in I} \operatorname{ri}(K_i) \neq \varnothing$, we know that item (ii) of Corollary 10 is false. Thus, we can apply Corollary 8 to the family $\{K_i^*\}_{i\in I}$, obtaining the desired result. □

Proposition 9 and Corollaries 10 and 11 can all trivially be adapted to consider a finite family of polyhedral cones. In this case, similar statements can be proved if we replace proper separation for strong separation and weaken the relative interior requirements. The following theorem is a refinement of these results.

Theorem 12. Let $\mathbb{E}$ be an Euclidean space, Let $\{K_i\}_{i\in I} \subseteq \mathbb{E}$ be a finite family of convex cones. Assume that there exists $I_0 \subseteq I$ such that K_i is polyhedral for each $i \in I_0$. If $\bigcap_{i\in I_0} K_i \cap \bigcap_{i\in I\setminus I_0} \operatorname{ri}(K_i) \neq \varnothing$, then $\sum_{i\in I} K_i^*$ is closed.

Proof. We already know that the result is valid when $I_0 = I$ and $I_0 = \varnothing$. Then, we conclude that the result is true for the families $\{K_i\}_{i \in I_0}$ and $\{K_i\}_{i \in I \setminus I_0}$. Hence, it suffices to show the result for cones $K, K_p \subseteq \mathbb{E}$, where K_p is polyhedral and $\operatorname{ri}(K) \cap K_p \neq \varnothing$.

In this context, let $S = \{x \in E^2 : x_1 = x_2\}$. We know that there exists a hyperplane properly separating S and $K \oplus K_p$ if, and only if there exists a hyperplane properly separating K and K_p. This happens if, and only if $\operatorname{ri}(K) \cap K_p = \varnothing$. Since $\operatorname{ri}(K) \cap K_p \neq \varnothing$ by hypothesis, we conclude that there is no hyperplane separating S and $K \oplus K_p$ properly. Applying Proposition 9 for these sets, we conclude that item (ii) is false.

Just as in Corollaries 10 and 11, we obtain that the statement "There exists $x \in K$ and $p \in K_p$ such that $x + p = 0$, $x \in -K^*, p \in -K_p^*$, and ($x \notin K_i^*$ or $p \notin K_p^*$)" is false. Then, Applying Corollary 8 to K and K_p yields the desired result. □

Homomorphisms - What Does Equivalence Mean in Optimization?

Definition. An *equivalence relation* on S is a binary relation $\sim$ such that, for each $a, b, c \in S$:

(i) $a \sim a$;

(ii) if $a \sim b$, then $b \sim a$;

(iii) if $a \sim b$ and $b \sim c$, then $a \sim c$.

Remark. An equivalence relation induces a partition of it's ground-set. Each element of the partition is called *equivalence class.*

Example. Consider $S = \mathbb{R}$ and the usual equality. Note that we have an equivalence relation where the equivalence classes are singletons.

Example. Let $G = (V, E)$ be a graph and consider the equivalence relation $v \sim w$ if v reaches w. The equivalence classes under this relation are the connected components of G.

Example. Let V be a vector space and let W be a subspace of V. Consider $u \sim v$ if $u - v \in W$. We have just defined an equivalence relation on V. Note that the equivalence classes under this relation are affine sets parallel to W. This relation is used to define the quotient space U/W.

Definition. An *optimization problem* is an ordered pair $P = (X, f)$, where X is a set and $f\colon X \to \overline{\mathbb{R}}$ is an extended real-valued function. The problem P is more commonly denoted as

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X. \end{array}$$

The set X is the *feasible region* of P and the function f is the *objective function* of P. The elements of X are the *feasible points* or *feasible solutions* of P; everything else is *infeasible.* The optimization problem P is *feasible* if $X \neq \varnothing$. Otherwise it is *infeasible.* The *objective value* of $x \in X$ is $f(x)$. The *optimal value* of P is $\inf_{x \in X} f(x) \in \overline{\mathbb{R}}$. A feasible solution $\bar{x}$ is *optimal* if $f(\bar{x})$ is the optimal value of the problem. If the optimal value of P is $-\infty$, the problem is *unbounded.*

When we write

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & x \in X \end{array}$$

we are referring to the optimization problem $(X, -f)$ and, besides the definition of optimal value which becomes $-\inf_{x \in X} -f(x)$, we use the same terminology as above.

Definition. Let $P = (X, f)$ be an optimization problem. A feasible point $\bar{x}$ is *locally optimal* if there exists $V \subseteq X$ such that $\bar{x}$ is the optimal solution of the problem (V, f).

Remark. We remark that an optimization problem P always has exactly one of the following outcomes:

1. The problem P is infeasible;
2. The problem P is unbounded;
3. The problem P is bounded and has no optimal solution;
4. The problem P is bounded and has optimal solution.

Homomorphisms and Equivalence

Definition. Let $P = (X, f)$ and $Q = (Y, g)$ be optimization problems. A *homomorphism* from P to Q is a function $\varphi\colon X \to Y$ such that $g(\varphi(x)) \leq f(x)$ for each $x \in X$.

Proposition 13. Let $P = (X, f)$, $Q = (Y, g)$, and $R = (Z, h)$ be optimization problems. If $\varphi\colon X \to Y$ be a homomorphism from P to Q and $\psi\colon Y \to Z$ be a homomorphism from Q to R, then $\psi \circ \varphi\colon X \to Z$ is a homomorphism from P to R.

Proof. Since ψ is a homomorphism from Q to S, we have that $h(\psi(y)) \leq g(y)$ for each $y \in Y$. Consider the set $\varphi(X) =: Y' \subseteq Y$. Then, $h(\psi(y')) \leq g(y')$ for each $y' \in Y'$. The latter implies that $h(\psi(\varphi(x))) \leq g(\varphi(x))$ for every $x \in X$. Since φ is a homomorphism from P to Q, we have that $g(\varphi(x)) \leq f(x)$ for each $x \in X$. Hence, $h(\psi(\varphi(x))) \leq f(x)$ for each $x \in X$. Therefore, $\psi \circ \varphi$ is a homomorphism from P to S. ☐

Proposition 14. Let $P = (X, f)$ and $Q = (Y, g)$ be optimization problems. If $\varphi\colon X \to Y$ is a bijective function such that $g(\varphi(x)) = f(x)$ for each $x \in X$, then $\varphi^{-1}\colon Y \to X$ is a homomorphism from Q to P.

Proof. Let $y \in Y$ and consider $x := \varphi^{-1}(y)$. Then,

$$g(y) = f(x) = f(\varphi^{-1}(y)) \geq f(\varphi^{-1}(y)).$$

Therefore, φ^{-1} a homomorphism from Q to P. ☐

Corollary 15. Let $P = (X, f)$ and $Q = (Y, g)$ be optimization problems. If there exists a bijective function $\varphi\colon X \to Y$ such that $g(\varphi(x)) = f(x)$ for each $x \in X$ then P and Q are equivalent.

Proof. From Proposition 14, we easily conclude that φ is an homomorphism from P to Q and that φ^{-1} is an homomorphism from Q to P. Hence, P and Q are equivalent. ☐

Next, we define an equivalence relation between optimization problems

Proposition 16. Consider, for every optimization problems A and B,

$$A \sim B \text{ if, and only if there exists a homomorphism from} A \text{ to } B \text{ and vice-versa.}$$

Then $\sim$ is an equivalence relation in $\mathcal{C}$.

Proof. Let $P = (X, f)$, $Q = (Y, g)$, and $S = (Z, h)$ be optimization problems.

(i) For reflexivity, consider the function $\varphi\colon X \to X$ given by $\varphi(x) = x$ for each $x \in X$. Then, we have that $f(\varphi(x)) = f(x)$ for each $x \in X$. Thus, $P \sim P$ by Corollary 15.

(ii) Symmetry follows immediately from the definition.

(iii) For transitivity, assume that $P \sim Q$ and $Q \sim S$. Then, let φ_1, φ_2 be homomorphisms from P to Q and from Q to P, respectively. Similarly, let ψ_1, ψ_2 be a homomorphism from Q to S and from S to Q, respectively. By Proposition 13 it follows that $\varphi_1 \circ \psi_1$ is a homomorphism from P to S and $\psi_2 \circ \varphi_2$ is a homomorphism from S to P. Therefore, $P \sim S$. ☐

First Results

Proposition 17. Let $P = (X, f)$ and $Q = (Y, g)$ be equivalent optimization problems. If either P or Q has finite optimal value $\alpha \in \mathbb{R}$, then α is the optimal value of both problems.

Proof. Consider the homomorphisms $\varphi\colon X \to Y$ and $\psi\colon Y \to X$, which exist by hypothesis. With no loss of generality, assume that P has optimal value α. Assume that the optimal value β of Q is different from α. If $\beta > \alpha$, for each $\bar{x} \in X$ such that $f(\bar{x}) < \alpha + \frac{\beta-\alpha}{2}$ we have that $g(\varphi(\bar{x})) \leq \alpha + \frac{\beta-\alpha}{2}$, and then β is greater then the optimal value of Q. If $\beta < \alpha$, for each $\bar{y} \in Y$ such that $g(\bar{y}) \leq \beta + \frac{\alpha-\beta}{2}$ we have that $f(\psi(\bar{y})) \leq \beta + \frac{\alpha-\beta}{2}$ and thus α is not the optimal value of P. Therefore, $\alpha = \beta$. ☐

Proposition 18. Let $P = (X, f)$ and $Q = (Y, g)$ be equivalent optimization problems. Then P has an optimal solution if, and only if Q has an optimal solution.

Proof. Consider the homomorphism $\varphi\colon X \to Y$, which exists by hypothesis. Let $\alpha \in \mathbb{R}$ be the optimal value of P and assume that there exists $x^* \in X$ such that $f(x^*) = \alpha$. By definition, $g(\varphi(x^*)) \leq \alpha$. By Proposition 17, we know that $g(y) \geq \alpha$ for each $y \in Y$. Thus, $g(\varphi(x^*)) = \alpha$. That is, $\varphi(x^*)$ is an optimal solution in Q. Since the converse follows by the exact same argument, the proof is completed. ☐

Proposition 19. Let $P = (X, f)$ and $Q = (Y, g)$ be equivalent optimization problems. Then P and Q have the same outcome. That is:

(i) P is infeasible if and only if Q is infeasible.

(ii) P is unbounded if and only if Q is unbounded.

(iii) P has finite optimal value and does not have optimal solution if and only if Q has finite optimal value and does not have optimal solution.

(iv) P has finite optimal value and an optimal solution if and only if Q has finite optimal value and an optimal solution.

Proof. Let $\varphi\colon X \to Y$ be a homomorphism from P to Q and let $\psi\colon Y \to X$ be a homomorphism from Q to P. We will show each of the items in our statement.

For the first item, assume that P is infeasible and Q is feasible. Thus, there exists $y \in Y$ and we have by definition that $\psi(y) \in X$. This contradicts the hypothesis that $X = \varnothing$. Hence, Q is infeasible. Clearly, the converse is proven using the exact same reasoning.

For the second item, assume that P is unbounded. Then, the set $L_n \coloneqq \{x \in X : f(x) \leq -n\}$ is nonempty for each $n \in \mathbb{N}$. Consider sequence $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n \in L_n$ for each $n \in \mathbb{N}$ and note that

$$\lim_{n\to\infty} f(v_n) = -\infty.$$

Set $w_n \coloneqq \varphi(v_n)$ for each $n \in \mathbb{N}$ so that w_n is always feasible in Q. Observe that $\lim_{n\to\infty} g(w_n) = -\infty$ and thus Q is unbounded. Again, the converse is shown by the exact same argument.

The two remaining items follow immediately from Propositions 17 and 18. $\square$

Further Developments