7 Equivalence Between Optimization and Separation

The Ellipsoid Method provides a way to do (linear) optimization using a separation oracle, i.e., we can do optimization through separation. Interestingly, there are also approaches to solve the separation problem given an oracle for the optimization problem. In this section we highlight some main ideas that underlie this relation. As in the Ellipsoid Method, there are some technical requirements that need to be fulfilled for this approach to work. Still, this very strong relation between optimization and separation is often simply referred to as the "equivalence between optimization and separation".

Assume we want to solve the separation problem over a polytope $P \subseteq \mathbb{R}^n$, and the only way how we can access P is via an optimization oracle, which, for any $c \in \mathbb{R}^n$, returns an optimal solution to $\max\{c^\top x \colon x \in P\}$. To do separation through optimization, we define the so-called *polar* P° of P, which is a polyhedron over which we can separate if we can optimize over P. More formally, for any set $X \subseteq \mathbb{R}^n$, we define its polar to be

$$X^\circ = \{y \in \mathbb{R}^n \colon x^\top y \leq 1 \ \forall x \in X\} \ .$$

Figure 7.1 shows the polar of a polytope $P \subseteq \mathbb{R}^2$.

Given an optimization oracle for P, the following is a separation oracle for its polar P°. Let $y \in \mathbb{R}^n$ be the input of the separation problem over P°. We obtain a maximizer $z \in \operatorname{argmax}\{y^\top x \colon x \in P\}$ using the optimization oracle for P. If $y^\top z \leq 1$, then $y \in P^\circ$, because for any $x \in P$ we have $x^\top y \leq z^\top y \leq 1$. Otherwise, $\{x \in \mathbb{R}^n \colon z^\top x \leq 1\}$ is a half-space separating y from P°.

This simple observation together with some additional properties of the polar can be used to separate also over P, in many natural settings. For this we first state some properties of the polar.

Lemma 7.1

Let $X \subseteq \mathbb{R}^n$ be a compact (i.e., closed and bounded) convex set, containing the origin in its interior. Then

(a) X° is a compact convex set with the origin in its interior.
(b) $(X^\circ)^\circ = X$.

Proof. (a) We start by showing that X° is convex and compact. X° is clearly convex and closed since it is the intersection of half-spaces. To show boundedness of X°, notice that since the origin is in the interior of X, there exists $\delta > 0$ such that $B(0, \delta) \subseteq X$, where $B(0, \delta)$ is the closed ball of radius δ that is centered at the origin. We show boundedness

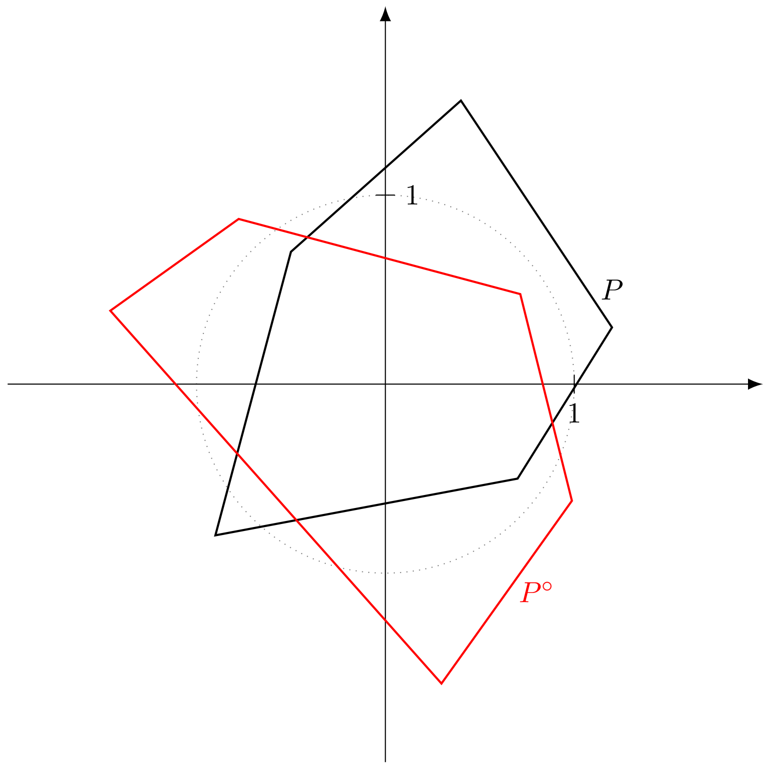

Figure 7.1: A polytope P and its polar P°.

of X° by showing that $X^\circ \subseteq B(0,1/\delta)$. Indeed, for any $y \in \mathbb{R}^n$ with $\|y\|_2 > \frac{1}{\delta}$, we obtain $x^\top y > 1$ where $x = \delta \cdot \frac{y}{\|y\|_2} \in B(0,\delta) \subseteq X$; hence, $y \notin X^\circ$.
It remains to show that the origin is in the interior of X°. Since X is bounded, there exists some $M > 0$ such that $X \subseteq B(0,M)$. We complete this part of the proof by showing $B(0,\frac{1}{M}) \subseteq X^\circ$. Let $y \in B(0,\frac{1}{M})$. We will show that y does not violate any constraint in the description of X°. Indeed, for any $x \in X \subseteq B(0,M)$, we obtain

$$y^\top x \leq \underbrace{\|y\|_2}_{\leq 1/M} \underbrace{\|x\|_2}_{\leq M} \leq 1 \quad ,$$

and hence $y \in X^\circ$.

(b) We clearly have $X \subseteq (X^\circ)^\circ$, because for any $x \in X$ and $y \in X^\circ$ we have $x^\top y \leq 1$. Thus, x fulfills all the constrains of $(X^\circ)^\circ$.
To show $(X^\circ)^\circ \subseteq X$, assume by contradiction that $\exists z \in (X^\circ)^\circ \setminus X$. Because X is convex and closed, there is a hyperplane separating z from X, i.e., there is $c \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$ such that $c^\top x \leq b \; \forall x \in X$ and $c^\top z > b$. Notice that $b > 0$ since the origin is in the interior of X. Let $y = \frac{1}{b}c$. We have $y \in X^\circ$, because for any $x \in X$,

$$x^\top y = \frac{1}{b} x^\top c \leq 1 \quad .$$

However, $y \in X^\circ$ implies $z \notin (X^\circ)^\circ$ because

$$y^\top z = \frac{1}{b} c^\top z > 1 \ ,$$

which contradicts the assumption $z \in (X^\circ)^\circ$.

□

Lemma 7.2

Let $P \subseteq \mathbb{R}^n$ be a polytope containing the origin in its interior. Then P° is a polytope. Moreover, for any $x \in \mathbb{R}^n$, we have

$$x \text{ is a vertex of } P \quad \Leftrightarrow \quad \{y \in \mathbb{R}^n \colon x^\top y \leq 1\} \text{ is facet-defining for } P^\circ.$$

Proof. Let

$$Q = \{y \in \mathbb{R}^n \colon x^\top y \leq 1 \ \forall x \in \text{vertices}(P)\} \ . \tag{7.1}$$

We first show $P^\circ = Q$. Notice that this implies that P° is a polytope (boundedness follows from Lemma 7.1).

We clearly have $P^\circ \subseteq Q$. To show $Q \subseteq P^\circ$, consider a point $y \in Q$. We will show $y \in P^\circ$ by showing that $x^\top y \leq 1$ for all $x \in P$. Hence, let $x \in P$. Since P is a polytope, x can be written as a convex combination of its vertices, i.e., $x = \sum_{i=1}^k \lambda_i x_i$, where x_i is a vertex of P for $i \in [k]$, $\sum_{i=1}^k \lambda_i = 1$, and $\lambda_i \geq 0 \ \forall i \in [k]$. Thus,

$$x^\top y = \sum_{i=1}^k \lambda_i x_i^\top y \leq \sum_{i=1}^k \lambda_i = 1 \ ,$$

as desired, where the inequality follows from $x_i^\top y \leq 1 \ \forall i \in [k]$ since $y \in Q$.

$\Leftarrow$) Notice that $P^\circ = Q$ also implies the '$\Leftarrow$' implication of the statement, because the description (7.1) of $Q = P^\circ$ contains facet-defining inequalities for each facet of P°. Hence, if for some $x \in \mathbb{R}^n$ we have that $\{y \in \mathbb{R}^n \colon x^\top y \leq 1\}$ is facet-defining, then it has to be a possibly scaled version of one of the constraints in (7.1). However, because all right-hand sides of the considered constraints are one, this implies that the constraint $x^\top y \leq 1$ must be one of the constraints in the description (7.1), thus implying $x \in \text{vertices}(P)$.

$\Rightarrow$) Let $x_1, \dots, x_k$ be the vertices of P. We will show that $\{y \in \mathbb{R}^n \colon x_k^\top y \leq 1\}$ is facet-defining for P°. As shown in the first part, we have $P^\circ = Q := \{y \in \mathbb{R}^n \colon x_i^\top y \leq 1 \ \forall i \in [k]\}$. Hence, we have to show that $\{y \in \mathbb{R}^n \colon x_k^\top y \leq 1\}$ is facet-defining for Q.

Assume for the sake of deriving a contradiction that $\{y \in \mathbb{R}^n \colon x_k^\top y \leq 1\}$ is not facet-defining for Q. Notice that $Q = P^\circ$ is full-dimensional because it contains the origin in its interior due to Lemma 7.1. Together with Lemma 1.20, this implies that the non-facet-defining constraint $x_k^\top y \leq 1$ in (7.1) is redundant. Hence, the constraint $x_k^\top y \leq 1$ is implied by the constraints $x_i^\top y \leq 1$ for $i \in [k-1]$, i.e., a conic combination of the constraints $x_i^\top y \leq 1$ for $i \in [k-1]$ leads to the constraint $x_k^\top y \leq \alpha$ for some $\alpha \leq 1$. Formally, this means that there are multipliers

$\lambda_i \geq 0$ for $i \in [k-1]$ such that

$$\sum_{i=1}^{k-1} \lambda_i x_i = x_k \ , \text{ and}$$
$$\sum_{i=1}^{k-1} \lambda_i \leq 1 \ .$$

However, this implies that x_k is a convex combination of $x_1, \ldots, x_{k-1}$ and the origin. Moreover, notice that x_k is different from $x_1, \ldots, x_{k-1}$, and x_k is not the origin, because x_k is a vertex of P and the origin is in the interior of P and thus cannot be a vertex of P. In other words, x_k is a non-trivial convex combination of other points in P, namely $x_1, \ldots, x_{k-1}$ and the origin. This contradicts the fact of x_k being a vertex of P. □

Based on the above results, there is a natural strategy to construct a separation oracle for a polytope $P \subseteq \mathbb{R}^n$ that contains the origin in its interior. Notice that if P is full-dimensional, there is always a way to translate P such that it contains the origin in its interior. As discussed, an optimization oracle for P implies a separation oracle for P°. Now, if the technical requirements are fulfilled to apply the Ellipsoid Method to P°, we can use this separation oracle to optimize over P°. Finally, an optimization oracle for P° implies a separation oracle for $(P^\circ)^\circ = P$. Furthermore, Lemma 7.2 implies that if the optimization oracle always returns a vertex solution, then the separation oracle constructed with the above approach does not just return any separating hyperplane, but even one that is facet-defining.

Figure 7.2 summarizes this interplay between optimization and separation through the polar.

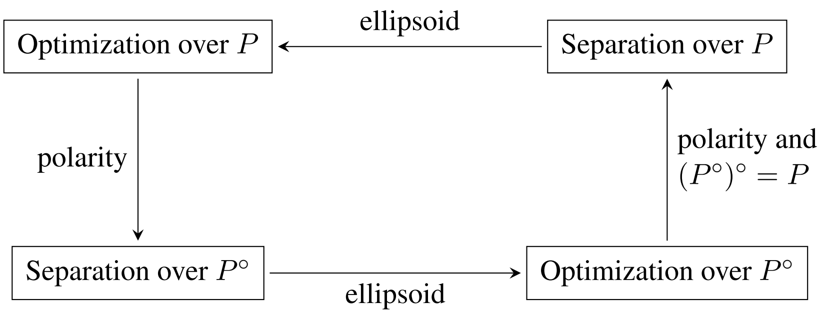

Figure 7.2: A graphical representation of the relation between optimization and separation via the Ellipsoid Method and polar polytopes.