Chapter 2

Cone Programming

Cone programming is a natural abstraction of semidefinite programming in which we can conveniently develop some basic theory, most notably semidefinite programming duality. According to Definition 1.4.2, a semidefinite program in equational form is an optimization problem of the form

$$\begin{array}{ll} \text{Maximize} & \mathrm{Tr}(C^T X) \\ \text{subject to} & A(X) = \mathbf{b} \\ & X \succeq 0, \end{array}$$

in the matrix variable $X \in \mathcal{S}_n$, the vector space of real symmetric $n \times n$ matrices. Furthermore, $C \in \mathcal{S}_n$, $\mathbf{b} \in \mathbb{R}^m$, and A is a linear operator from $\mathcal{S}_n$ to $\mathbb{R}^m$.

Equipped with the scalar product $\langle X, Y \rangle = \mathrm{Tr}(X^T Y)$, $\mathcal{S}_n$ is a real *Hilbert space.*[1]

The link to cone programming is established by the fact that the set $\{X \in \mathcal{S}_n : X \succeq 0\}$ of positive semidefinite matrices is a *closed convex cone*. Before we can prove this, we need a formal definition. In what follows, the abstract framework of Hilbert spaces is not the most general one that is possible, but the most convenient one for our purposes.

Throughout this chapter, we fix real Hilbert spaces V and W.

2.1 Closed Convex Cones

2.1.1 Definition. *Let $K \subseteq V$ be a nonempty closed set.*[2] *K is called a* closed convex cone *if the following two properties hold.*

(i) For all $\mathbf{x} \in K$ and all nonnegative real numbers λ, we have $\lambda\mathbf{x} \in K$.

[1] A vector space V with scalar product is called a Hilbert space if it is complete, i.e. if every Cauchy sequence in V converges to a point in V; for $V = \mathcal{S}_n$, this easily follows from completeness of $\mathbb{R}^{n^2}$.

[2] We refer to the standard topology on V, induced by the open balls $B(\mathbf{c}, \rho) = \{\mathbf{x} \in V : \|\mathbf{x} - \mathbf{c}\| < \rho\}$, where $\|\mathbf{x}\| = \sqrt{\langle X, X \rangle}$.

(ii) For all $\mathbf{x}, \mathbf{x}' \in K$*, we have* $\mathbf{x} + \mathbf{x}' \in K$*.*

Property (i) ensures that K is a cone, while property (ii) guarantees convexity of K. Indeed, if $\mathbf{x}, \mathbf{x}' \in K$ and $\lambda \in [0,1]$, then $(1-\lambda)\mathbf{x}$ and $\lambda\mathbf{x}'$ are both in K by (i), and then (ii) shows that $(1-\lambda)\mathbf{x} + \lambda\mathbf{x}' \in K$, as required by convexity.

2.1.2 Lemma. *The set* $\{X \in \mathcal{S}_n : X \succeq 0\}$ *of positive semidefinite matrices is a closed convex cone.*

Proof. Using the charaterization of positive semidefinite matrices provided by Fact 1.4.1 (ii), this is easy. If $\mathbf{x}^T M\mathbf{x} \geq 0$ and $\mathbf{x}^T M'\mathbf{x} \geq 0$, then also $\mathbf{x}^T \lambda M\mathbf{x} = \lambda\mathbf{x}^T M\mathbf{x} \geq 0$ for $\lambda \geq 0$ and $\mathbf{x}^T(M + M')\mathbf{x} = \mathbf{x}^T M\mathbf{x} + \mathbf{x}^T M'\mathbf{x} \geq 0$. To show closedness, we need to show that the complement is open. Indeed, if we have a symmetric matrix M that is not positive semidefinite, there exists $\tilde{\mathbf{x}} \in \mathbb{R}^n$ such that $\tilde{\mathbf{x}}^T M\tilde{\mathbf{x}} < 0$, and this inequality still holds for all matrices M' in a sufficiently small neighborhood of M. $\square$

Let's look at some other examples of closed convex cones. It is obvious that the nonnegative orthant $\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}$ is a closed convex cone; even more trivial examples of closed convex cones in $\mathbb{R}^n$ are $K = \{\mathbf{0}\}$ and $K = \mathbb{R}^n$. We can also get new cones as direct sums of cones (the proof of the following fact is left to the reader).

2.1.3 Fact. *Let* $K \subseteq V, L \subseteq W$ *be closed convex cones. Then*

$$K \oplus L := \{(\mathbf{x}, \mathbf{y}) \in V \oplus W : \mathbf{x} \in K, \mathbf{y} \in L\}$$

is again a closed convex cone, the direct sum *of* K *and* L*.*

Let us remind the reader that $V \oplus W$ is the set $V \times W$, turned into a Hilbert space via

$$\begin{aligned}
(\mathbf{x}, \mathbf{y}) + (\mathbf{x}', \mathbf{y}') &:= (\mathbf{x} + \mathbf{x}', \mathbf{y} + \mathbf{y}'),\\
\lambda(\mathbf{x}, \mathbf{y}) &:= (\lambda\mathbf{x}, \lambda\mathbf{y}),\\
\langle(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}')\rangle &:= \langle\mathbf{x}, \mathbf{x}'\rangle + \langle\mathbf{y}, \mathbf{y}'\rangle.
\end{aligned}$$

Now we get to some more concrete cones.

2.1.1 The Ice Cream Cone in $\mathbb{R}^n$

This cone is defined as

$$\bigtriangledown_n = \{(\mathbf{x}, r) \in \mathbb{R}^{n-1} \times \mathbb{R} : \|\mathbf{x}\| \leq r\},$$

see Figure 2.1 for an illustration in $\mathbb{R}^3$ that (hopefully) explains the name. It is closed because of the "$\leq$" (this argument is similar to the one in the proof of Lemma 2.1.2), and its convexity follows from the triangle inequality $\|\mathbf{x} + \mathbf{x}'\| \leq \|\mathbf{x}\| + \|\mathbf{x}'\|$.

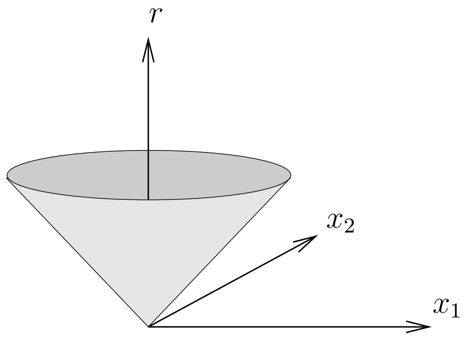

Figure 2.1: The (lower part of the boundary of the) ice cream cone $\bigtriangledown_3$

2.1.2 The Toppled Ice Cream Cone in $\mathbb{R}^3$

Here is another closed convex cone, but its convexity is less obvious. It is defined as

$$\lhd\!) = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, xy \geq z^2\}. \tag{2.1}$$

Exercise 2.6.1 formally explains why we call this the toppled ice cream cone: it is obtained by "dropping" and "vertically squeezing" the ice cream cone in $\mathbb{R}^3$, see Figure 2.2.

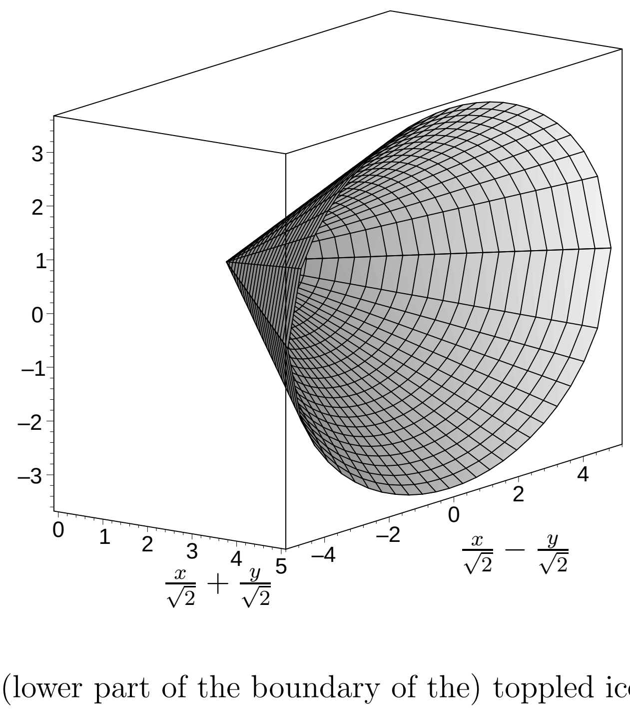

Figure 2.2: The (lower part of the boundary of the) toppled ice cream cone $\lhd\!)$

We remark that $\lhd\!)$ can alternatively be defined as the set of all (x, y, z) such that the symmetric matrix

$$\begin{pmatrix} x & z \\ z & y \end{pmatrix}$$

is positive semidefinite.

2.1.4 Lemma. $\lhd\!()$ *is a closed convex cone.*

Proof. Again, the closedness of $\lhd\!()$ comes from the "$\geq$'s", and the cone Property (i) in Definition 2.1.1 is immediate. Now let $(x, y, z), (x', y', z') \in \lhd\!()$. To show Property (ii) (the sum is also in $\lhd\!()$), we compute

$$\begin{aligned}
(x + x')(y + y') &= xy + x'y' + xy' + x'y \\
&\geq z^2 + z'^2 + 2\frac{xy' + x'y}{2} \\
&\geq z^2 + z'^2 + 2\sqrt{xy'x'y} \\
&\geq z^2 + z'^2 + 2|z||z'| \\
&\geq z^2 + z'^2 + 2zz' = (z + z')^2.
\end{aligned}$$

In the second inequality of this derivation, we use the *AGM inequality*. For nonnegative numbers $a_1, \ldots, a_k$, this inequality says that

$$\frac{1}{k}(a_1 + a_2 + \cdots + a_k) \geq \sqrt[k]{a_1 a_2 \cdots a_k},$$

meaning that the *arithmetic mean* of the numbers (left-hand side expression) is at least the *geometric mean* (right-hand side expression). □

2.2 Dual Cones

2.2.1 Definition. *Let $K \subseteq V$ be a closed convex cone. The set*

$$K^* := \{\mathbf{y} \in V : \langle \mathbf{y}, \mathbf{x} \rangle \geq 0 \ \forall \mathbf{x} \in K\}$$

is called the dual cone *of K.*

In fact, K^* is again a closed convex cone (we omit the simple proof; it uses bilinearity of the scalar product to show the cone properties, and the Cauchy-Schwarz inequality (that holds in Hilbert spaces) for closedness).

Let us illustrate this notion on the examples that we have seen earlier. What is the dual of the nonnegative orthant $\mathbb{R}^n_+$? This is the set of all $\mathbf{y}$ such that

$$\mathbf{y}^T \mathbf{x} \geq 0 \ \forall \mathbf{x} \geq \mathbf{0}.$$

This set certainly contains the nonnegative orthant $\{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} \geq \mathbf{0}\}$ itself, but not more: Given $\mathbf{y} \in \mathbb{R}^n$ with $y_i < 0$, we have $\mathbf{y}^T\mathbf{e}_i < 0$, where $\mathbf{e}_i$ is the i-th unit vector (a member of $\mathbb{R}^n_+$), and this proves that $\mathbf{y}$ is not a member of the dual cone $(\mathbb{R}^n_+)^*$. It follows that the dual of $\mathbb{R}^n_+$ is $\mathbb{R}^n_+$: the nonnegative orthant is *self-dual*.

For the "even more trivial" cones, the situation is as follows.

K	K^*
$\{\mathbf{0}\}$	$\mathbb{R}^n$
$\mathbb{R}^n$	$\{\mathbf{0}\}$

We leave the computation of the dual of the ice cream cone $\triangledown_n$ to the reader (see Exercise 2.6.2) and proceed with the toppled ice cream cone (2.1).

2.2.2 Lemma. *The dual of the toppled ice cream cone* $\triangleleft\!\!)$ *is*

$$\triangleleft\!\!)^* = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, xy \geq \frac{z^2}{4}\} \subseteq \mathbb{R}^3,$$

a "vertically stretched" version of $\triangleleft\!\!)$.

Proof. We first want to show the inclusion "$\supseteq$" in the statement of the lemma, and this again uses the AGM inequality. Let us fix $\tilde{\mathbf{y}} = (\tilde{x}, \tilde{y}, \tilde{z})$ such that $\tilde{x} \geq 0, \tilde{y} \geq 0, \tilde{x}\tilde{y} \geq \tilde{z}^2/4$. For $\mathbf{x} = (x, y, z) \in K$ chosen arbitrarily, we get

$$\begin{aligned}
\tilde{\mathbf{y}}^T\mathbf{x} &= \tilde{x}x + \tilde{y}y + \tilde{z}z \\
&= 2\frac{\tilde{x}x + \tilde{y}y}{2} + \tilde{z}z \\
&\geq 2\sqrt{\tilde{x}x\tilde{y}y} + \tilde{z}z \\
&\geq 2\frac{|\tilde{z}|}{2}|z| + \tilde{z}z \geq 0.
\end{aligned}$$

This means that $\mathbf{y} \in \triangleleft\!\!)^*$.

For the other inclusion, let us fix $\tilde{\mathbf{y}} = (\tilde{x}, \tilde{y}, \tilde{z})$ such that $\tilde{x} < 0$ or $\tilde{y} < 0$ or $\tilde{x}\tilde{y} < \tilde{z}^2/4$; we now need to find a proof for $\tilde{\mathbf{y}} \notin \triangleleft\!\!)^*$. If $\tilde{x} < 0$, we choose $\mathbf{x} = (1, 0, 0) \in K$ and get the desired proof $\tilde{\mathbf{y}}^T\mathbf{x} < 0$. If $\tilde{y} < 0$, $\mathbf{x} = (0, 1, 0)$ will do. In the case of $\tilde{x}, \tilde{y} \geq 0$ but $\tilde{x}\tilde{y} < \tilde{z}^2/4$, let us first assume that $\tilde{z} \geq 0$ and set

$$\mathbf{x} = (\tilde{y}, \tilde{x}, -\sqrt{\tilde{x}\tilde{y}}) \in K.$$

We then compute

$$\tilde{\mathbf{y}}^T\mathbf{x} = 2\tilde{x}\tilde{y} - \tilde{z}\sqrt{\tilde{x}\tilde{y}} < 2\tilde{x}\tilde{y} - 2\tilde{x}\tilde{y} = 0.$$

For $\tilde{z} < 0$, we pick $\mathbf{x} = (\tilde{y}, \tilde{x}, \sqrt{\tilde{x}\tilde{y}}) \in K$. $\square$

We conclude this section with the following intuitive fact: the dual of a direct sum of cones is the direct sum of the dual cones. This fact is easy but not entirely trivial. It actually requires a small proof, see Exercise 2.6.3.

2.2.3 Lemma. *Let $K \subseteq V, L \subseteq W$ be closed convex cones. Then*

$$(K \oplus L)^* = K^* \oplus L^*.$$

2.3 The Farkas Lemma, Cone Version

Under any meaningful notion of duality, you expect the dual of the dual to be the primal (original) object. For cone duality (Definition 2.2.1), this indeed works.

2.3.1 Lemma. *Let $K \subseteq V$ be a closed convex cone. Then $(K^*)^* = K$.*

Maybe surprisingly, the proof of this innocent-looking fact already requires the machinery of separation theorems that will also be essential for cone programming duality below. Separation theorems generally assert that disjoint convex sets can be separated by a hyperplane.

2.3.1 A Separation Theorem for Closed Convex Cones

The following is arguably the simplest nontrivial separation theorem. On top of elementary calculations, the proof only requires one standard result from Hilbert space theory.

2.3.2 Theorem. *Let $K \subseteq V$ be a closed convex cone, $\mathbf{b} \in V \setminus K$. Then there exists a vector $\mathbf{y} \in V$ such that*

$$\langle \mathbf{y}, \mathbf{x} \rangle \geq 0 \ \forall \mathbf{x} \in K, \ \langle \mathbf{y}, \mathbf{b} \rangle < 0.$$

The statement is illustrated in Figure 2.3 (left) for $V = \mathbb{R}^2$. The hyperplane $h = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{y}^T\mathbf{x} = 0\}$ (that passes through the origin) strictly separates $\mathbf{b}$ from K.[3] We also say that $\mathbf{y}$ is a *witness* for $\mathbf{b} \notin K$.

Proof. The plan of the proof is straightforward: We let $\mathbf{z}$ be the point of K nearest to $\mathbf{b}$ (in the distance $\|\mathbf{z} - \mathbf{b}\| = \sqrt{\langle \mathbf{z} - \mathbf{b}, \mathbf{z} - \mathbf{b} \rangle}$ induced by the scalar product), and we check that the vector $\mathbf{y} = \mathbf{z} - \mathbf{b}$ is as required; see Figure 2.3 (right). The existence of $\mathbf{z}$ follows from V being a Hilbert space. In general, every nonempty closed and convex subset of a Hilbert space has a unique point nearest to any given point.[2, Section 2.2b].

With $\mathbf{z}$ being a nearest point of K to $\mathbf{b}$, we set $\mathbf{y} = \mathbf{z} - \mathbf{b}$. First we check that $\langle \mathbf{y}, \mathbf{z} \rangle = 0$. This is clear for $\mathbf{z} = \mathbf{0}$. For $\mathbf{z} \neq \mathbf{0}$, if $\mathbf{z}$ were not perpendicular to $\mathbf{y}$, we could move $\mathbf{z}$ slightly along the ray $\{t\mathbf{z} : t \geq 0\} \subseteq K$ and get a point closer to $\mathbf{b}$ (here we use the fact that K is a cone).

More formally, let us first assume that $\langle \mathbf{y}, \mathbf{z} \rangle > 0$, and let us set $\mathbf{z}' = (1 - \alpha)\mathbf{z}$ for a small $\alpha > 0$. We calculate $\|\mathbf{z}' - \mathbf{b}\|^2 = \langle (\mathbf{y} - \alpha\mathbf{z}), (\mathbf{y} - \alpha\mathbf{z}) \rangle = \|\mathbf{y}\|^2 -$

[3]By this we mean that $\mathbf{b}$ is strictly on one side of the hyperplane.

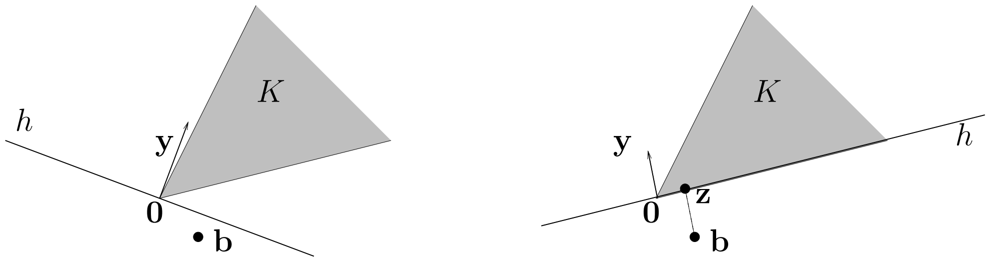

Figure 2.3: A point $\mathbf{b}$ not contained in a closed convex cone $K \subseteq \mathbb{R}^2$ can be separated from K by a hyperplane $h = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{y}^T\mathbf{x} = 0\}$ through the origin (left). The separating hyperplane resulting from the proof of Theorem 2.3.2 (right).

$2\alpha\langle \mathbf{y}, \mathbf{z}\rangle + \alpha^2\|\mathbf{z}\|^2$. We have $2\alpha\langle \mathbf{y}, \mathbf{z}\rangle > \alpha^2\|\mathbf{z}\|^2$ for all sufficiently small $\alpha > 0$, and thus $\|\mathbf{z}' - \mathbf{b}\|^2 < \|\mathbf{y}\|^2 = \|\mathbf{z} - \mathbf{b}\|^2$. This contradicts $\mathbf{z}$ being a nearest point. The case $\langle \mathbf{y}, \mathbf{z}\rangle < 0$ is handled through $\mathbf{z}' = (1+\alpha)\mathbf{z}$.

To verify $\langle \mathbf{y}, \mathbf{b}\rangle < 0$, we recall that $\mathbf{y} \neq \mathbf{0}$, and we compute $0 < \langle \mathbf{y}, \mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{z}\rangle - \langle \mathbf{y}, \mathbf{b}\rangle = -\langle \mathbf{y}, \mathbf{b}\rangle$.

Next, let $\mathbf{x} \in K$, $\mathbf{x} \neq \mathbf{z}$. The angle $\angle\mathbf{bzx}$ has to be at least 90 degrees, for otherwise, points on the segment $\mathbf{zx}$ sufficiently close to $\mathbf{z}$ would lie closer to $\mathbf{b}$ than $\mathbf{z}$ (here we use convexity of K); equivalently, $\langle(\mathbf{b} - \mathbf{z}), (\mathbf{x} - \mathbf{z})\rangle \leq 0$ (this is similar to the above argument for $\langle \mathbf{y}, \mathbf{z}\rangle = 0$ and we leave a formal verification to the reader). Thus $0 \geq \langle(\mathbf{b} - \mathbf{z}), (\mathbf{x} - \mathbf{z})\rangle = -\langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle = -\langle \mathbf{y}, \mathbf{x}\rangle$. □

Using this result, we can now show that $(K^*)^* = K$ for any closed convex cone.

Proof of Lemma 2.3.1. For the direction $K \subseteq (K^*)^*$, we just need to apply the definition of duality: Let us choose $\mathbf{b} \in K$. By definition of K^*, $\langle \mathbf{y}, \mathbf{b}\rangle = \langle \mathbf{b}, \mathbf{y}\rangle \geq 0$ for all $\mathbf{y} \in K^*$; this in turn shows that $\mathbf{b} \in (K^*)^*$.

For the other direction, we let $\mathbf{b} \notin K$. According to Theorem 2.3.2, we find a vector $\mathbf{y}$ such $\langle \mathbf{y}, \mathbf{x}\rangle \geq 0$ for all $\mathbf{x} \in K$ and $\langle \mathbf{y}, \mathbf{b}\rangle = \langle \mathbf{b}, \mathbf{y}\rangle < 0$. The former inequality shows that $\mathbf{y} \in K^*$, but then the latter inequality witnesses that $\mathbf{b} \notin (K^*)^*$. □

2.3.2 Adjoint Operators

We will now bring a linear operator $A : V \to W$ into the game. If V and W are finite-dimensional (as in our semidefinite programming application), then A can be represented as a matrix, w.r.t. a fixed choice of orthogonal bases for V and W. But even in the general case, A "behaves" like a matrix for our purposes, and we still write $A\mathbf{x}$ instead of $A(\mathbf{x})$ for $\mathbf{x} \in V$. Here is the generalization of the transpose of a matrix.

2.3.3 Definition. *Let $A : V \rightarrow W$ be a linear operator. A linear operator $A^T : W \rightarrow V$ is called an* adjoint *of A if*

$$\langle \mathbf{y}, A\mathbf{x}\rangle = \langle A^T\mathbf{y}, \mathbf{x}\rangle \ \forall \mathbf{x} \in V, \mathbf{y} \in W.$$

If V and W are finite-dimensional (more generally, if A is continuous), there is an adjoint A^T of A. In general, *if* there is an adjoint, then it is easy to see that it is unique which justifies the notation A^T. But in exotic situations, it may happen that there is no adjoint. In any case, we only consider operators A that have an adjoint.

In order to stay as close as possible to the familiar matrix terminology, we will also introduce the following notation. If $V_1, V_2, \ldots, V_n$ and $W_1, W_2, \ldots, W_m$ are Hilbert spaces with linear operators $A_{ij} : V_j \rightarrow W_i$, then we write the "matrix"

$$\left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1n} \\ \hline A_{21} & A_{22} & \cdots & A_{2n} \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{m1} & A_{m2} & \cdots & A_{mn} \end{array}\right) \tag{2.2}$$

for the linear operator $\bar{A} : V_1 \oplus V_2 \oplus \cdots \oplus V_n \rightarrow W_1 \oplus W_2 \oplus \cdots \oplus W_m$ defined by

$$\bar{A}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = (\sum_{j=1}^{n} A_{1j}\mathbf{x}_j, \sum_{j=1}^{n} A_{2j}\mathbf{x}_j, \ldots, \sum_{j=1}^{n} A_{mj}\mathbf{x}_j).$$

A simple calculation then shows that

$$\bar{A}^T = \left(\begin{array}{c|c|c|c} A_{11}^T & A_{21}^T & \cdots & A_{m1}^T \\ \hline A_{12}^T & A_{22}^T & \cdots & A_{m2}^T \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{1n}^T & A_{2n}^T & \cdots & A_{mn}^T \end{array}\right), \tag{2.3}$$

just as with matrices.

For the remainder of this chapter, we fix a linear operator A from V to W.

2.3.3 The Farkas Lemma, Bogus Version

We will next try to use Theorem 2.3.2 to derive the "cone version" of the Farkas Lemma. Our goal is to find a witness for the unsolvability of a system

$$A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \in K, \tag{2.4}$$

where $K \subseteq V$ is some closed convex cone, and $\mathbf{b} \in W$. If $A = \text{id}$, Theorem 2.3.2 already tells us how to do this. Here is the proof idea for the general case. System (2.4) can equivalently be written as

$$b \in C = \{A\mathbf{x} : \mathbf{x} \in K\},$$

where C is again a convex cone (a fact that follows from linearity of A). Now we apply Theorem 2.3.2 and find a witness for $\mathbf{b} \notin C$: a vector $\mathbf{y} \in W$ such that $\langle \mathbf{y}, A\mathbf{x} \rangle = \langle A^T\mathbf{y}, \mathbf{x} \rangle \geq 0 \ \forall \mathbf{x} \in K$ and $\langle \mathbf{y}, \mathbf{b} \rangle < 0$. By definition of K^*, this can equivalently be written as

$$A^T\mathbf{y} \in K^*, \ \langle \mathbf{b}, \mathbf{y} \rangle < 0. \tag{2.5}$$

Indeed, for the case $K = K^* = \mathbb{R}^n_+$, we recover the "linear" Farkas lemma: Either the system $A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$ has a solution $\mathbf{x}$, or the system $A^T\mathbf{y} \geq \mathbf{0}, \mathbf{b}^T\mathbf{y} < 0$ has a solution $\mathbf{y}$, but not both.

The trouble is that the above proof does not work in general. Our separation Theorem 2.3.2 requires $C = \{A\mathbf{x} : \mathbf{x} \in K\}$ to be a *closed* convex cone, but we forgot to check the closedness of C. For $K = \mathbb{R}^n_+$, we are lucky: C turns out to be closed, and the above derivation of the Farkas lemma is therefore correct in the linear case (see also Exercise 2.6.6). But there are closed convex cones $K \subseteq \mathbb{R}^n$ and matrices A for which the cone C is not closed. Here is an example. Let $K = \lhd\!\!\!\bigcirc$, the toppled ice cream cone (2.1) in $\mathbb{R}^3$, and let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$C = \{(x, z) \in \mathbb{R}^2 : (x, y, z) \in \lhd\!\!\!\bigcirc\} = \mathbf{0} \cup (\{x \in \mathbb{R} : x > 0\} \times \mathbb{R}),$$

a set that is obviously not closed.

In such cases, Theorem 2.3.2 is not applicable, and the Farkas lemma as envisioned above may fail. Indeed, such a failure can already be constructed from our previous example. If we in addition set $\mathbf{b} = (0, 1) \in \mathbb{R}^2$, *both* (2.4) and (2.5) are unsolvable, and we have no theorem of alternatives (we encourage the reader to go through this example in detail).

To save the situation, we need to work with the closure of C.

2.3.4 The Farkas Lemma, Corrected Version

2.3.4 Lemma. *Let $K \subseteq V$ be a closed convex cone, and $C = \{A\mathbf{x} : \mathbf{x} \in K\}$. Then $\bar{C}$, the closure of C, is a closed convex cone.*

Proof. By definition, the closure of C is the set of all limit points of C. Formally, $\mathbf{b} \in \bar{C}$ if and only if there exists a sequence $(\mathbf{y}_k)_{k\in\mathbb{N}}$ such that $\mathbf{y}_k \in C$ for all k and $\lim_{k\to\infty} \mathbf{y}_k = \mathbf{b}$. This yields that $\bar{C}$ is a convex cone, using that C is a convex cone. In addition, $\bar{C}$ is closed. $\square$

The fact "$\mathbf{b} \in \bar{C}$" can be formulated without reference to the cone C, and this will be more convenient in what follows.

2.3.5 Definition. *Let $K \subseteq V$ be a closed convex cone. The system*

$$A\mathbf{x} = \mathbf{b},\ \mathbf{x} \in K$$

is called subfeasible *if there exists a sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ such that $\mathbf{x}_k \in K$ for all $k \in \mathbb{N}$ and*

$$\lim_{k\to\infty} A\mathbf{x}_k = \mathbf{b}.$$

It is clear that subfeasibility of $A\mathbf{x} = \mathbf{b},\ \mathbf{x} \in K$ implies $\mathbf{b} \in \bar{C}$, but the other direction also holds. If $(\mathbf{y}_k)_{k\in\mathbb{N}}$ is a sequence in C converging to $\mathbf{b}$, then any sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ such that $\mathbf{y}_k = A\mathbf{x}_k$ for all k proves subfeasibility of our system. Here is the correct Farkas lemma for equation systems over closed convex cones.

2.3.6 Lemma. *Let $K \subseteq V$ be a closed convex cone, and $\mathbf{b} \in W$. The system $A\mathbf{x} = \mathbf{b}, \mathbf{x} \in K$ is subfeasible if and only if every $\mathbf{y} \in W$ with $A^T\mathbf{y} \in K^*$ also satisfies $\langle \mathbf{y}, \mathbf{b}\rangle \geq 0$.*

Proof. If $A\mathbf{x} = \mathbf{b},\ \mathbf{x} \in K$ is subfeasible, we choose any sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ that witnesses its subfeasibility. For $\mathbf{y} \in W$, we compute

$$\langle \mathbf{y}, \mathbf{b}\rangle = \langle \mathbf{y}, \lim_{k\to\infty} A\mathbf{x}_k\rangle = \lim_{k\to\infty} \langle \mathbf{y}, A\mathbf{x}_k\rangle = \lim_{k\to\infty} \langle A^T\mathbf{y}, \mathbf{x}_k\rangle.$$

If $A^T\mathbf{y} \in K^*$, then $\mathbf{x}_k \in K$ yields $\langle A^T\mathbf{y}, \mathbf{x}_k\rangle \geq 0$ for all $k \in \mathbb{N}$, and $\langle \mathbf{y}, \mathbf{b}\rangle \geq 0$ follows.

If $A\mathbf{x} = \mathbf{b},\ \mathbf{x} \in K$ is not subfeasible, this can equivalently be expressed as $\mathbf{b} \notin \bar{C}$, where $C = \{A\mathbf{x} : \mathbf{x} \in K\}$. Since $\bar{C}$ is a closed convex cone, we can apply Theorem 2.3.2 and obtain a hyperplane that strictly separates $\mathbf{b}$ from $\bar{C}$ (and in particular from C). This means that we find $\mathbf{y} \in W$ such that

$$\langle \mathbf{y}, A\mathbf{x}\rangle = \langle A^T\mathbf{y}, \mathbf{x}\rangle \geq 0\ \forall \mathbf{x} \in K,\ \langle \mathbf{y}, \mathbf{b}\rangle < 0.$$

It remains to observe that the statement $\langle A^T\mathbf{y}, \mathbf{x}\rangle \geq 0\ \forall \mathbf{x} \in K$ is equivalent to $A^T\mathbf{y} \in K^*$. □

2.4 Cone Programs

Here is the definition of a cone program, in a somewhat more general format than we would need for semidefinite programming. This will introduce symmetry between the primal and the dual program.

2.4.1 Definition. *Let $K \subseteq V, L \subseteq W$ be closed convex cones, $\mathbf{b} \in W$, $\mathbf{c} \in V$, $A : \mathcal{V} \to W$ a linear operator. A* cone program *is a constrained optimization problem of the form*

$$\begin{array}{ll} \textit{Maximize} & \langle \mathbf{c}, \mathbf{x} \rangle \\ \textit{subject to} & \mathbf{b} - A\mathbf{x} \in L \\ & \mathbf{x} \in K. \end{array} \tag{2.6}$$

For $L = \{\mathbf{0}\}$, we get cone programs in equational form.

Following the linear programming case, we call the cone program *feasible* if there is some *feasible solution*, a vector $\tilde{\mathbf{x}}$ with $\mathbf{b} - A\tilde{\mathbf{x}} \in L, \tilde{\mathbf{x}} \in K$. The *value* of a feasible cone program is defined as

$$\sup\{\langle \mathbf{c}, \mathbf{x} \rangle : \mathbf{b} - A\mathbf{x} \in L, \mathbf{x} \in K\}, \tag{2.7}$$

which includes the possibility that the value is ∞.

An *optimal solution* is a feasible solution $\mathbf{x}^*$ such that $\langle \mathbf{c}, \mathbf{x}^* \rangle \geq \langle \mathbf{c}, \mathbf{x} \rangle$ for all feasible solutions $\mathbf{x}$. Consequently, if there is an optimal solution, the value of the cone program is finite, and that value is attained, meaning that the supremum in (2.7) is a maximum.

For people accustomed to the behavior of linear programs, cone programs have some surpises in store. In fact, they seem to misbehave in various ways that we discuss next. On a large scale, however, these are small blunders only: Cone programs will turn out to be almost as civilized as linear programs.

2.4.1 Reachability of the Value

There are cone programs with finite value but no optimal solution. Here is an example involving the toppled ice cream cone, see page 21 (we leave it to the reader to put it into the form (2.6) via suitable $A, \mathbf{b}, \mathbf{c}, K, L$):

$$\begin{array}{ll} \text{Maximize} & -x \\ \text{subject to} & z = 1 \\ & (x, y, z) \in \lhd\!\!\!\!\bigcirc. \end{array} \tag{2.8}$$

After substituting $z = 1$ into the definition of the toppled ice cream cone, we obtain the following equivalent constrained optimization problem in two variables.

$$\begin{array}{ll} \text{Minimize} & x \\ \text{subject to} & x \geq 0 \\ & xy \geq 1. \end{array}$$

It is clear that the value of x is bounded from below by 0, and that values arbitrarily close to 0 can be attained. Due to the constraint $xy \geq 1$, however, the value 0 itself cannot be attained. This means that the cone program (2.8) has value 0 but no optimal solution.

2.4.2 The Subvalue

Another aspect that we don't see in linear programming is that of subfeasibilty, a notion that we have already introduced for equation systems over cones, see Definition 2.3.5. If a linear program is infeasible, then it will remain infeasible under any sufficiently small perturbation of the right-hand side $\mathbf{b}$. In contrast, there are infeasible cone programs that become feasible under an *arbitrarily* small perturbation of $\mathbf{b}$.

The following is merely a repetition of Definition 2.3.5 for the linear operator $(A \mid \mathrm{id}) : V \oplus W \to W$ (recall the matrix notation introduced on page 26) and the cone $K \oplus L$.

2.4.2 Definition. *The cone program (2.6) is called subfeasible if there exist sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}$ and $(\mathbf{x}'_k)_{k\in\mathbb{N}}$ such that $\mathbf{x}_k \in K$ and $\mathbf{x}'_k \in L$ for all $k \in N$, and*

$$\lim_{k\to\infty} (A\mathbf{x}_k + \mathbf{x}'_k) = \mathbf{b}.$$

Such sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}$ and $(\mathbf{x}'_k)_{k\in\mathbb{N}}$ are called feasible sequences *of (2.6).*

Every feasible program is subfeasible, but the converse is not true in general. Proper subfeasibility occurs for some $\mathbf{b}$ if and only if the cone $C = \{A\mathbf{x} + \mathbf{x}' : \mathbf{x} \in K, \mathbf{x}' \in L\}$ is not closed. For $L = \{\mathbf{0}\}$, we have seen such an example on page 27.

We can assign a value even to a subfeasible program, and we consequently call it the subvalue.

2.4.3 Definition. *Given a feasible sequence $(\mathbf{x}_k)_{k\in\mathbb{N}}$ of a subfeasible cone program (2.6), we define its* value *as*

$$\langle \mathbf{c}, (\mathbf{x}_k)_{k\in\mathbb{N}} \rangle := \limsup_{k\to\infty} \langle \mathbf{c}, \mathbf{x}_k \rangle.$$

The subvalue *of (2.6) is then defined as*

$$\sup\{\langle \mathbf{c}, (\mathbf{x}_k)_{k\in\mathbb{N}} \rangle : (\mathbf{x}_k)_{k\in\mathbb{N}} \textit{ is a feasible sequence of (2.6)}\}.$$

2.4.3 Value vs. Subvalue

By definition, the value of a feasible cone program is always upper-bounded by its subvalue, and it is tempting to think that the two are equal. But this is not true in general. There are feasible programs with finite value but infinite subvalue. Here is one such program in two variables x and z:

$$\begin{array}{ll} \text{Maximize} & z \\ \text{subject to} & (x, 0, z) \in \lhd\!\!\!\bigcirc. \end{array}$$

The program is feasible (choose $x \geq 0, z = 0$). In fact, every feasible solution must satisfy $z = 0$, so the value of the program is 0. For the following computation of the subvalue, let us explicitly write down the parameters of this program in the form of (2.6):

$$\mathbf{c} = (0,1),\ A = \left(\begin{array}{cc} -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{array}\right),\ \mathbf{b} = (0,0,0),\ L = \lhd\!\!\!),\ K = \mathbb{R}^2.$$

Now let us define

$$\mathbf{x}_k = (k^3, k) \in \mathbb{R}^2,\ \mathbf{x}'_k = (k^3, 1/k, k) \in \lhd\!\!\!), \quad k \in \mathbb{N}.$$

Then we get

$$\lim_{k\to\infty} A\mathbf{x}_k + \mathbf{x}'_k = \lim_{k\to\infty} \big((-k^3, 0, -k) + (k^3, 1/k, k)\big) = \lim_{k\to\infty} (0, 1/k, 0) = \mathbf{0} = \mathbf{b}.$$

This means that $(\mathbf{x}_k)_{k\in\mathbb{N}}$ and $(\mathbf{x}'_k)_{k\in\mathbb{N}}$ are feasible sequences. The program's subvalue is therefore at least

$$\limsup_{k\to\infty} \mathbf{c}^T\mathbf{x}_k = \limsup_{k\to\infty} k = \infty.$$

We may also have a gap if both value and subvalue are finite. Adding the constraint $z \leq 1$ to the previous program (which can be done by changing $A, \mathbf{b}, L$ accordingly) results in subvalue 1, while the value itself stays at 0.

Interestingly, such pathologies disappear if the program has an *interior point.* In general, requiring additional conditions, with the goal of avoiding exceptional situations is called *constraint qualification.* Requiring an interior point is known as *Slater's constraint qualification.*

2.4.4 Definition. *An interior point (or* Slater *point) of the cone program (2.6) is a point* $\mathbf{x} \in K$ *with the property that*

$$\mathbf{b} - A\mathbf{x} \in \mathrm{int}(L).$$

Let us remind the reader that $\mathrm{int}(L)$ is the set of all points of L that have a small ball around it that is completely contained in L.

Now we can prove the following.

2.4.5 Theorem. *If the cone program (2.6) has an interior point (which in particular means that it is feasible), the value equals the subvalue.*

Proof. We choose feasible sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}, (\mathbf{x}'_k)_{k\in\mathbb{N}}$ that witness the program's subvalue γ' (possibly ∞), and we take an arbitrary interior point $\tilde{\mathbf{x}}$ of (2.6).

For fixed $\delta \in (0,1]$, we next consider the sequence $(\mathbf{w}_k)_{k\in\mathbb{N}}$ defined by

$$\mathbf{w}_k := (1-\delta)\mathbf{x}_k + \delta\tilde{\mathbf{x}}, \ k \in \mathbb{N}.$$

In other words, $\mathbf{w}_k$ is obtained by moving $\mathbf{x}_k$ into the direction of the interior point $\tilde{\mathbf{x}}$. Now we prove (and this is the main technical step, see below) that for k sufficiently large, $\mathbf{w}_k$ is a feasible solution of (2.6). We further have

$$\limsup_{k\to\infty}\langle \mathbf{c}, \mathbf{w}_k\rangle = (1-\delta)\limsup_{k\to\infty}\langle \mathbf{c}, \mathbf{x}_k\rangle + \delta\langle \mathbf{c}, \tilde{\mathbf{x}}\rangle = (1-\delta)\gamma' + \delta\langle \mathbf{c}, \tilde{\mathbf{x}}\rangle,$$

and since δ can be made arbitrarily small, we conclude that there are feasible solutions with objective function value arbitrarily close to γ'. On the other hand, the value γ is always upper-bounded by γ', so $\gamma = \gamma'$ follows.

It remains to show that $\mathbf{w}_k$ is feasible for k sufficiently large. It is clear that $\mathbf{w}_k \in K$. To establish $\mathbf{b} - A\mathbf{w}_k \in L$, we use the definition of feasible sequences and compute

$$\mathbf{b} - A\mathbf{w}_k = (1-\delta)(\underbrace{\mathbf{b} - A\mathbf{x}_k - \mathbf{x}'_k}_{\to 0} + \underbrace{\mathbf{x}'_k}_{\in L}) + \delta\,\underbrace{(\mathbf{b} - A\tilde{\mathbf{x}})}_{\in \mathrm{int}(L)}.$$

For $\delta > 0$, $\delta(\mathbf{b} - A\tilde{\mathbf{x}})$ is still an interior point of L, with a small ball around it (of radius ρ, say) contained in L. For any k, adding $(1-\delta)\mathbf{x}'_k \in L$ is easily seen to take us to another interior point, again with a ball of radius ρ around it contained in L. Calling this point $\mathbf{p}_k$, we thus have

$$\mathbf{b} - A\mathbf{w}_k = \underbrace{(1-\delta)(\mathbf{b} - A\mathbf{x}_k - \mathbf{x}'_k)}_{\to 0} + \mathbf{p}_k.$$

For k large enough, $\|(1-\delta)(\mathbf{b} - A\mathbf{x}_k - \mathbf{x}'_k)\| < \rho$, meaning that $\mathbf{b} - A\mathbf{w}_k \in L$. $\square$

2.5 Cone Programming Duality

For this section, let us call the cone program (2.6) the primal program and name it (P):

$$\text{(P)} \quad \begin{array}{ll} \text{Maximize} & \langle \mathbf{c}, \mathbf{x}\rangle \\ \text{subject to} & \mathbf{b} - A\mathbf{x} \in L \\ & \mathbf{x} \in K. \end{array}$$

Then we define its *dual* as the cone program

$$\text{(D)} \quad \begin{array}{ll} \text{Minimize} & \langle \mathbf{b}, \mathbf{y}\rangle \\ \text{subject to} & A^T\mathbf{y} - \mathbf{c} \in K^* \\ & \mathbf{y} \in L^*. \end{array}$$

Formally, this does not have the cone program format (2.6), but we could easily achieve this if necessary by rewriting (D) as follows.

$$\begin{array}{lll} \text{(D')} & \text{Maximize} & -\langle \mathbf{b}, \mathbf{y} \rangle \\ & \text{subject to} & -\mathbf{c} + A^T \mathbf{y} \in K^* \\ & & \mathbf{y} \in L^*. \end{array}$$

Having done this, we can also compute the dual of (D') which takes us (not surprisingly) back to (P).

For the dual program (D) which is now a minimization problem, value and subvalue are defined through inf's and lim inf's in the canonical way.

Now we are approaching the Duality Theorem of cone programming. As for linear programming, we assume that one of the two programs, (D) say, is feasible. First, we prove *weak duality*: If the primal program (P) is subfeasible, the subvalue of (P) is upper-bounded by the value of (D). The proof of this is a no-brainer and follows more or less directly from the definitions of the primal and the dual. Still, weak duality has the important consequence that (D)—a minimization problem—has finite value if (P) is subfeasible.

Then we prove *regular duality*: there is in fact no gap between the subvalue of (P) and the value of (D). For linear programming where there is no difference between value and subvalue (see Exercise 2.6.6), we would be done at this point and would have proved the strong duality theorem. Here we call it regular duality since the following scenario is possible: both (P) and (D) are feasible, but there is a gap between their values γ and β, see Figure 2.4. We can indeed derive an example for this scenario from the cone program with value 0 and subvalue 1 that we have constructed in Section 2.4.3.

Figure 2.4: Gap between the values of the primal and the dual cone program.

In order to get *strong duality*, we apply Slater's constraint qualification: If one of the programs has an interior point (Definition 2.4.4), then the other program is feasible as well, and there is no gap between their values. This result is a trivial consequence of regular duality together with Theorem 2.4.5.

The presentation essentially follows Duffin's original article [1] in the classic book *Linear Inequalities and Related Systems* from 1956. This book, edited by Harold W. Kuhn and Albert W. Tucker, contains articles by many of the "grand old men", including Dantzig, Ford, Fulkerson, Gale, and Kruskal.

2.5.1 Weak Duality

Let us start with the *weak Duality Theorem.*

2.5.1 Theorem. *If the dual program (D) is feasible, and if the primal program (P) is subfeasible, then the subvalue of (P) is upper-bounded by the value of (D).*

If (P) is feasible as well, this implies that the value of (P) is upper-bounded by the value of (D), and that both values are finite. This is weak duality as we know it from linear programming.

Proof. We pick any feasible solution $\mathbf{y}$ of (D) and any feasible sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}, (\mathbf{x}'_k)_{k\in\mathbb{N}}$ of (P). Then we know that

$$0 \leq \langle \underbrace{A^T\mathbf{y} - \mathbf{c}}_{\in K^*}, \underbrace{\mathbf{x}_k}_{\in K} \rangle + \langle \underbrace{\mathbf{y}}_{\in L^*}, \underbrace{\mathbf{x}'_k}_{\in L} \rangle = \langle \mathbf{y}, A\mathbf{x}_k + \mathbf{x}'_k \rangle - \langle \mathbf{c}, \mathbf{x}_k \rangle, \quad k \in \mathbb{N}.$$

Hence,

$$\limsup_{k\to\infty} \langle \mathbf{c}, \mathbf{x}_k \rangle \leq \limsup_{k\to\infty} \langle \mathbf{y}, A\mathbf{x}_k + \mathbf{x}'_k \rangle = \lim_{k\to\infty} \langle \mathbf{y}, A\mathbf{x}_k + \mathbf{x}'_k \rangle = \langle \mathbf{y}, \mathbf{b} \rangle.$$

Since the feasible sequences were arbitrary, this means that the subvalue of (P) is upper-bounded by $\langle \mathbf{y}, \mathbf{b} \rangle$, and since $\mathbf{y}$ was an arbitrary feasible solution of (D), the lemma follows. □

You may have expected the following stronger version of weak duality: *if both (P) and (D) are subfeasible, then the subvalue of (P) is upper-bounded by the subvalue of (D).* Is this true, and we have just been too lazy to prove it, or is it false? Exercise 2.6.4 asks you to settle this question.

2.5.2 Regular Duality

Here is the *regular Duality Theorem.* The proof essentially consists of applications of the Farkas lemma to carefully crafted systems.

2.5.2 Theorem. *The dual program (D) is feasible and has finite value β if and only if the primal program (P) is subfeasible and has finite subvalue γ. Moreover, $\beta = \gamma$.*

Proof. If (D) is feasible and has value β, we know that

$$A^T\mathbf{y} - \mathbf{c} \in K^*,\ \mathbf{y} \in L^* \quad \Rightarrow \quad \langle \mathbf{b}, \mathbf{y} \rangle \geq \beta. \tag{2.9}$$

We also know that

$$A^T\mathbf{y} \in K^*,\ \mathbf{y} \in L^* \quad \Rightarrow \quad \langle \mathbf{b}, \mathbf{y} \rangle \geq 0. \tag{2.10}$$

Indeed, if we had some $\mathbf{y}$ that fails to satisfy the latter implication, we could add a large positive multiple of it to any feasible solution of (D) and in this way obtain a feasible solution of value smaller than β.

We can now merge (2.9) and (2.10) into the single implication

$$A^T\mathbf{y} - z\mathbf{c} \in K^*,\ \mathbf{y} \in L^*,\ z \geq 0 \quad \Rightarrow \quad \langle \mathbf{b}, \mathbf{y}\rangle \geq z\beta. \tag{2.11}$$

For $z > 0$, we obtain this from (2.9) by multiplication of all terms with z and then renaming $z\mathbf{y} \in L^*$ back to $\mathbf{y}$. For $z = 0$, it is simply (2.10). In matrix form as introduced on page 26, we can rewrite (2.11) as follows.

$$\left(\begin{array}{c|c} A^T & -\mathbf{c} \\ \hline \mathrm{id} & \mathbf{0} \\ \hline 0 & 1 \end{array}\right)(\mathbf{y}, z) \in K^* \oplus L^* \oplus \mathbb{R}_+ \quad \Rightarrow \quad (\mathbf{b}^T| - \beta)(\mathbf{y}, z) \geq 0. \tag{2.12}$$

Here and in the following, we use a column vector $\mathbf{c} \in V$ as the linear operator $z \mapsto z\mathbf{c}$ from $\mathbb{R}$ to V and the row vector $\mathbf{c}^T$ as the (adjoint) linear operator $\mathbf{x} \mapsto \langle \mathbf{c}, \mathbf{x}\rangle$ from V to $\mathbb{R}$.

The form (2.12) now allows us to apply the Farkas lemma. According to Lemma 2.3.6, and using (2.3), the implication (2.12) holds if and only if the system

$$\left(\begin{array}{c|c|c} A & \mathrm{id} & 0 \\ \hline -\mathbf{c}^T & \mathbf{0}^T & 1 \end{array}\right)(\mathbf{x}, \mathbf{x}', z) = (\mathbf{b}, -\beta),\ (\mathbf{x}, \mathbf{x}', z) \in (K^* \oplus L^* \oplus \mathbb{R}_+)^* = K \oplus L \oplus \mathbb{R}_+ \tag{2.13}$$

is subfeasible. The latter equality uses Lemma 2.2.3 together with Lemma 2.3.1.

System (2.13) is subfeasible if and only if there are sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}, (\mathbf{x}'_k)_{k\in\mathbb{N}}, (z_k)_{k\in\mathbb{N}}$ with $\mathbf{x}_k \in K, \mathbf{x}'_k \in L, z_k \geq 0$ for all k, such that

$$\lim_{k\to\infty} A\mathbf{x}_k + \mathbf{x}'_k = \mathbf{b}, \tag{2.14}$$

and

$$\lim_{k\to\infty} \langle \mathbf{c}, \mathbf{x}_k\rangle - z_k = \beta. \tag{2.15}$$

Now (2.14) shows that (P) is subfeasible, and (2.15) shows that the subvalue of (P) is at least β. Weak duality (Theorem 2.5.1) shows that it is at most β, concluding the "only if" direction.

For the "if" direction, let (P) be subfeasible with finite subvalue γ and assume for the purpose of obtaining a contradiction that (D) is infeasible. This yields the implication

$$A^T\mathbf{y} - z\mathbf{c} \in K^*,\ \mathbf{y} \in L^*, \quad \Rightarrow \quad z \leq 0, \tag{2.16}$$

since for any pair $(\mathbf{y}, z)$ that violates it, $\frac{1}{z}\mathbf{y}$ would be a feasible solution of (D). We now play the same game as before and write this in Farkas-lemma-compatible matrix form:

$$\left(\begin{array}{c|c} A^T & -\mathbf{c} \\ \hline \mathrm{id} & 0 \end{array}\right)(\mathbf{y}, z) \in K^* \oplus L^* \quad \Rightarrow \quad (\mathbf{0}^T \mid -1)(\mathbf{y}, z) \geq 0. \tag{2.17}$$

According to Lemma 2.3.6, this means that the system

$$\left(\begin{array}{c|c} A & \mathrm{id} \\ \hline -\mathbf{c}^T & 0 \end{array}\right)(\mathbf{x}, \mathbf{x}') = (\mathbf{0}, -1), \ (\mathbf{x}, \mathbf{x}') \in K \oplus L \tag{2.18}$$

is subfeasible, which in turn means that there are sequences $(\mathbf{x}_k)_{k\in\mathbb{N}}, (\mathbf{x}'_k)_{k\in\mathbb{N}}$ with $\mathbf{x}_k \in K, \mathbf{x}'_k \in L$ for all k, such that

$$\lim_{k\to\infty} A\mathbf{x}_k + \mathbf{x}'_k = \mathbf{0} \tag{2.19}$$

and

$$\lim_{k\to\infty} \langle \mathbf{c}, \mathbf{x}_k \rangle = 1. \tag{2.20}$$

But this is a contradiction: Elementwise addition of $(\mathbf{x}_k)_{k\in\mathbb{N}}, (\mathbf{x}'_k)_{k\in\mathbb{N}}$ to any feasible sequences of (P) that witness subvalue γ would result in feasible sequences that yield subvalue at least $\gamma + 1$.

Consequently, the dual program (D) must have been feasible. Weak duality (Theorem 2.5.1) yields that (D) has finite value $\beta \geq \gamma$. But then $\beta = \gamma$ follows from the previous "only if" direction. □

2.5.3 Strong Duality

Here is the **strong Duality Theorem of Cone Programming**, under Slater's constraint qualification.

2.5.3 Theorem. *If the primal program (P) is feasible, has finite value* γ and *has an interior point* $\tilde{\mathbf{x}}$, *then the dual program (D) is feasible and has finite value* $\beta = \gamma$.

Proof. (P) is in particular subfeasible, and since the program has an interior point, Theorem 2.4.5 shows that the subvalue of (P) is also γ. Using the regular Duality Theorem 2.5.2 ("if" direction), the statement follows. □

The strong Duality Theorem 2.5.3 is not applicable if the primal cone program (P) is in equational form ($L = \{\mathbf{0}\}$), since the cone $L = \{\mathbf{0}\}$ has no interior points. But there is a different variant of Slater's constraint qualification that we can use in this case, given that V, W are finite-dimensional. Exercise 2.6.5 asks you to prove the following.

2.5.4 Theorem. *Assume that the Hilbert spaces* V *and* W *are finite-dimensional. If the primal program*

$$\begin{array}{lll} \textit{(P)} & \textit{Maximize} & \langle \mathbf{c}, \mathbf{x} \rangle \\ & \textit{subject to} & A\mathbf{x} = \mathbf{b} \\ & & \mathbf{x} \in K \end{array}$$

is feasible, has finite value γ and *has a point* $\tilde{\mathbf{x}} \in \text{int}(K)$ *such that* $A\tilde{\mathbf{x}} = \mathbf{b}$, *then the dual program*

$$\begin{array}{lll} (D) & \textit{Minimize} & \langle \mathbf{b}, \mathbf{y} \rangle \\ & \textit{subject to} & A^T\mathbf{y} - \mathbf{c} \in K^* \end{array}$$

is feasible and has finite value $\beta = \gamma$.

We remark that for general L, just requiring a point $\tilde{\mathbf{x}} \in \text{int}(K)$ with $\mathbf{b} - A\tilde{\mathbf{x}} \in L$ is not enough to achieve strong duality. To see this, we recall the example program on page 30 that has $K = \mathbb{R}^2$. It is feasible, has finite value 0 and a feasible solution in $\text{int}(K) = K$. But the fact that the subvalue is ∞ implies by weak duality that the dual program must be infeasible.

2.6 Exercises

2.6.1 Exercise. *Let*

$$\triangleleft\!\!\bigcirc' = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, 2xy \geq z^2\}$$

be a vertically stretched version of the toppled ice cream cone $\triangleleft\!\!\bigcirc$. *Prove that there is an orthogonal matrix* M *(i.e.* $M^{-1} = M^T$*) such that*

$$\triangleleft\!\!\bigcirc' = \{M \begin{pmatrix} \mathbf{x} \\ \hline r \end{pmatrix} : (\mathbf{x}, r) \in \bigtriangledown_3\}.$$

This means, the toppled ice cream cone is an isometric image of the ice cream, plus some additional vertical squeezing (which naturally happens when soft objects topple over). To understand why the isometric image indeed corresponds to "toppling", analyze what M *does to the "axis"* $\{(0, 0, r) \in \mathbb{R}^3 : r \in \mathbb{R}_+\}$ *of the ice cream cone.*

2.6.2 Exercise. *What is the dual of the ice cream cone* $\bigtriangledown_n$ *(see Page 20)?*

2.6.3 Exercise. *Prove Lemma 2.2.3: The dual of a direct sum of cones* K *and* L *is the direct sum of the dual cones* K^* *and* L^*.

2.6.4 Exercise. *Here is a statement that (if true) would strengthen the weak Duality Theorem 2.5.1: If both (P) and (D) are subfeasible, then the subvalue of (P) is upper-bounded by the subvalue of (D). Prove or disprove the statement.*

2.6.5 Exercise. *Prove Theorem 2.5.4!*

Hint: *First reprove Theorem 2.4.5 under the given constraint qualification.*

2.6.6 Exercise. *Prove the following statement that introduces yet another kind of constraint qualification.*

If the primal program (P) is feasible, has finite value γ and *the cone*

$$C = \{ \left(\begin{array}{c|c} A & \mathrm{id} \\ \hline \mathbf{c}^T & \mathbf{0}^T \end{array} \right) (\mathbf{x}, \mathbf{x}') : \, (\mathbf{x}, \mathbf{x}') \in K \times L \}$$

is closed, then the dual program (D) is feasible and has finite value $\beta = \gamma$*. (Here, we refer again to the matrix notation of operators introduced on Page 26.)*

We remark that if K *and* L *are "linear programming cones" (meaning that they are iterated direct sums of one-dimensional cones* $\{0\}, \mathbb{R}, \mathbb{R}^+$*), then the above cone* C *is indeed closed, see e.g.[5, Chapter 6.5]. It follows that strong linear programming duality requires no constraint qualification.*

Hint*: First reprove Theorem 2.4.5 under the given constraint qualification.*