On duality theory of conic linear problems

Alexander Shapiro*

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA
e-mail: ashapiro@isye.gatech.edu

Abstract

In this paper we discuss duality theory of optimization problems with a linear objective function and subject to linear constraints with cone inclusions, referred to as conic linear problems. We formulate the Lagrangian dual of a conic linear problem and survey some results based on the conjugate duality approach where the questions of "no duality gap" and existence of optimal solutions are related to properties of the corresponding optimal value function. We discuss in detail applications of the abstract duality theory to the problem of moments, linear semi-infinite and continuous linear programming problems.

Key words: Conic linear programs, Lagrangian and conjugate duality, optimal value function, problem of moments, semi-infinite programming, continuous linear programming.

AMS subject classification: 49N15, 90C34, 90C46

*Work performed at Argonne National Laboratory while a Faculty Research Participant. Program administered by the Argonne Division of Educational Programs with funding provided by the U.S. Department of Energy

1 Introduction

In this paper we discuss duality theory of conic linear optimization problems of the form

$$\operatorname*{Min}_{x \in C} \langle c, x \rangle \ \text{ subject to } \ Ax + b \in K, \tag{1.1}$$

where X and Y are linear spaces (over real numbers), $C \subset X$ and $K \subset Y$ are convex cones, $b \in Y$ and $A : X \to Y$ is a linear mapping. We assume that spaces X and Y are paired with some linear spaces X' and Y', respectively, in the sense that bilinear forms $\langle \cdot, \cdot \rangle : X' \times X \to I\!\!R$ and $\langle \cdot, \cdot \rangle : Y' \times Y \to I\!\!R$ are defined. In other words, for any $x^* \in X'$ and $x \in X$, we have that $\langle x^*, \cdot \rangle$ and $\langle \cdot, x \rangle$ are linear functionals on spaces X and X', respectively, and similarly for the pair Y and Y'.

Many interesting examples of optimization problems can be formulated in the above framework. We discuss, in particular, applications of the general theory to the problem of moments, semi-infinite programming and continuous linear programming problems.

There are numerous studies devoted to duality theory of optimization problems. We adopt here an approach which is based on conjugate duality pioneered by Rockafellar [20],[22]. The standard way how the conjugate duality is developed in an infinite dimensional setting is based on pairing of locally convex topological vector spaces, [4],[5],[22]. We outline some basic results of that theory in the next section. In particular, we indicate where exactly topological assumptions about the involved spaces are essential.

This paper is organized as follows. In section 2 we discuss duality theory of abstract conic linear problems. Sections 3, 4 and 5 are devoted to applications of the general theory. In section 3 we discuss the classical problem of moments. In the setting of the problem of moments the space X is an infinite dimensional space of measures while the space Y is finite dimensional. The corresponding dual can be written as a semi-infinite programming problem which we discuss in section 4. Finally, in section 5 we study the so-called continuous linear programming problems where both spaces X and Y are infinite dimensional.

Recall that a vector space Y is said to be a locally convex topological vector space if it is equipped with a Hausdorff (i.e., satisfying the separation axiom) topology which is compatible with the algebraic operations of Y and such that any neighborhood of $0 \in Y$ includes an open, convex, balanced and absorbing subset of Y. It is said that two locally convex topological vector spaces Y and Y' are paired locally convex topological vector spaces if their topologies are compatible with the corresponding bilinear form $\langle \cdot, \cdot \rangle$, i.e., the set of linear continuous functionals on Y coincides with the set $\{\langle y^*, \cdot \rangle : y^* \in Y'\}$ and the set of linear continuous functionals on Y' coincides with the set $\{\langle \cdot, y \rangle : y \in Y\}$. If Y is a Banach space, then we can equip Y either with its strong (i.e., norm) or its weak topology and pair Y with its standard dual space Y^* (of continuous linear functionals) equipped with the weak star (weak*) topology. The interested reader can look in almost any standard text book on functional analysis (e.g., [11]) for a theory of locally convex topological vector spaces.

We use "Min" and "Max" to denote the respective minimization and maximization operators. Their appearance does not automatically imply that the corresponding minimum or maximum is attained.

2 Conic linear problems

In this section we discuss a duality theory of the conic linear problem (1.1). We associate with the cone C its polar (positive dual) cone

$$C^* := \{x^* \in X' : \langle x^*, x\rangle \geq 0,\ \forall x \in C\},$$

and similarly with the cone K we associate its polar cone $K^* \subset Y'$. Furthermore, the bipolar of the cone K is defined as

$$K^{**} := \{y \in Y : \langle y^*, y\rangle \geq 0,\ \forall y^* \in K^*\}.$$

Note that at this point we do not introduce any particular topology in the considered spaces. We only require that the dual space X' be large enough such that the adjoint mapping of A does exist. That is, we make the following assumption throughout the paper.

(A1) For any $y^* \in Y'$ there exists unique $x^* \in X'$ such that $\langle y^*, Ax\rangle = \langle x^*, x\rangle$ for all $x \in X$.

Then we can define the adjoint mapping $A^* : Y' \to X'$ by the equation

$$\langle y^*, Ax\rangle = \langle A^* y^*, x\rangle, \quad \forall\, x \in X.$$

Existence of $x^* = A^* y^*$ in the assumption (A1) means that the space X' is sufficiently large to include all linear functionals of the form $x \mapsto \langle y^*, Ax\rangle$. Uniqueness of x^* means that any two points of X' can be separated by a linear functional $\langle \cdot, x\rangle$, $x \in X$. That is, for any $x^* \in X'$ there exists $x \in X$ such that $\langle x^*, x\rangle \neq 0$.

Consider the Lagrangian function

$$L(x, y^*) := \langle c, x\rangle + \langle y^*, Ax + b\rangle$$

of the problem (1.1), and the following optimization problem

$$\operatorname*{Min}_{x \in C} \left\{ \psi(x) := \sup_{y^* \in -K^*} L(x, y^*) \right\}. \tag{2.1}$$

Let us observe that the above problem (2.1) is equivalent to the problem

$$\operatorname*{Min}_{x \in C} \langle c, x\rangle \ \text{ subject to } \ Ax + b \in K^{**}. \tag{2.2}$$

Indeed, if $Ax + b \in K^{**}$, then $\langle y^*, Ax + b\rangle \leq 0$ for any $y^* \in -K^*$, and hence the maximum of $\langle y^*, Ax + b\rangle$, over $y^* \in -K^*$, is zero. Therefore, in that case $\psi(x) = \langle c, x\rangle$. If $Ax + b \notin K^{**}$, then $\langle y^*, Ax + b\rangle > 0$ for some $y^* \in -K^*$ and hence $\psi(x) = +\infty$.

Note that the inclusion $K \subset K^{**}$ always holds. Therefore, the optimal value of the problem (2.2) is less than or equal to the optimal value of the problem (1.1).

By changing the Min and Max operators in (2.1) we can calculate the dual of problems (1.1) and (2.2). That is, consider the problem

$$\operatorname*{Max}_{y^* \in -K^*} \left\{ \phi(y^*) := \inf_{x \in C} L(x, y^*) \right\}. \tag{2.3}$$

Since $L(x, y^*) = \langle c + A^* y^*, x \rangle + \langle y^*, b \rangle$, we have that $\phi(y^*) = \langle y^*, b \rangle$ if $c + A^* y^* \in C^*$, and $\phi(y^*) = -\infty$ otherwise. Therefore, problem (2.3) is equivalent to the following problem

$$\operatorname*{Max}_{y^* \in -K^*} \langle y^*, b \rangle \ \text{ subject to } \ A^* y^* + c \in C^*. \tag{2.4}$$

We denote problems (1.1) and (2.2) by (P) and (P'), respectively, and refer to them as *primal* problems. Problem (2.4) is denoted by (D) and referred to as the *dual* problem. By $\text{val}(P)$, $\text{val}(P')$ and $\text{val}(D)$ we denote their respective optimal values, and by $\text{Sol}(P)$, $\text{Sol}(P')$, $\text{Sol}(D)$ their sets of optimal solutions.

Let us remark that one can construct a dual problem by introducing a Lagrange multiplier for the constraint $x \in C$ as well. That is, the dual problem is constructed by minimizing $L(x, x^*, y^*)$ over $x \in X$, where

$$L(x, x^*, y^*) := \langle c, x \rangle + \langle x^*, x \rangle + \langle y^*, Ax + b \rangle.$$

This leads to the dual problem

$$\operatorname*{Max}_{\substack{x^* \in -C^* \\ y^* \in -K^*}} \langle y^*, b \rangle \ \text{ subject to } \ A^* y^* + c + x^* = 0. \tag{2.5}$$

Clearly the above dual problem (2.5) is equivalent to the dual problem (2.4).

The dual problem (2.4) is also a conic linear problem. Its dual can be written in the form (1.1) but with the cones C and K replaced by C^{**} and K^{**}, respectively, and with A replaced by A^{**}, provided A^{**} does exist. Therefore, there is a symmetry between problems (P) and (D) (i.e., the dual of (D) coincides with (P)), if $C = C^{**}$, $K = K^{**}$ and $A = A^{**}$.

The following weak duality relations hold

$$\text{val}(P) \geq \text{val}(P') \geq \text{val}(D). \tag{2.6}$$

Indeed, we already mentioned that the inequality $\text{val}(P) \geq \text{val}(P')$ follows from the inclusion $K \subset K^{**}$. The other inequality $\text{val}(P') \geq \text{val}(D)$ follows by the standard min-max duality. Recall that the min-max problem (2.1) is equivalent to the problem (P'), and not to the original problem (P). In the subsequent analysis we deal with problem (P), while problem (P') is introduced only in order to demonstrate this point.

The weak duality (2.6) is also not difficult to show directly. Let x be a feasible point of the problem (P) (of the problem (P')) and y^* be a feasible points of the problem (D). We have then

$$\langle c, x \rangle \geq \langle c, x \rangle + \langle y^*, Ax + b \rangle = \langle c + A^* y^*, x \rangle + \langle y^*, b \rangle \geq \langle y^*, b \rangle. \tag{2.7}$$

Since the inequality $\langle c, x\rangle \geq \langle y^*, b\rangle$ holds for any feasible pair x and y^*, it follows that $\mathrm{val}(P) \geq \mathrm{val}(D)$. It also follows from (2.7) that the equality $\langle c, x\rangle = \langle y^*, b\rangle$ holds iff the following complementarity conditions are satisfied

$$\langle y^*, Ax + b\rangle = 0 \ \text{ and } \ \langle c + A^*y^*, x\rangle = 0. \tag{2.8}$$

The above complementarity conditions simply mean that a feasible pair (x, y^*) is a saddle point of the Lagrangian. We obtain the following result.

Proposition 2.1 *There is no duality gap between the primal and dual problems, i.e.,* $\mathrm{val}(P) = \mathrm{val}(D)$*, and both problems have optimal solutions iff there exists a feasible pair* (x, y^*) *satisfying the complementarity conditions* (2.8).

Now let us associate with problem (1.1) the optimal value function

$$v(y) := \inf\{\langle c, x\rangle : x \in C,\ Ax + y \in K\}. \tag{2.9}$$

By the definition $v(y) = +\infty$ if the corresponding feasible set is empty. Clearly $v(b) = \mathrm{val}(P)$. Since the primal problem (P) is convex we have that the (extended real valued) function $v(y)$ is convex ([22]). It is also not difficult to see that $v(y)$ is positively homogeneous, i.e., $v(ty) = tv(y)$ for any $t > 0$ and $y \in Y$.

The *conjugate* of $v(y)$ is defined as

$$v^*(y^*) := \sup_{y\in Y}\{\langle y^*, y\rangle - v(y)\}. \tag{2.10}$$

Let us calculate the conjugate function $v^*(y^*)$. We have

$$\begin{aligned} v^*(y^*) &= \sup\{\langle y^*, y\rangle - \langle c, x\rangle : (x, y^*) \in X \times Y^*,\ x \in C,\ Ax + y \in K\} \\ &= \sup_{x\in C} \sup_{Ax+y\in K}\{\langle y^*, y\rangle - \langle c, x\rangle\} \\ &= \sup_{x\in C} \sup_{y\in K}\{\langle y^*, y - Ax\rangle - \langle c, x\rangle\} \\ &= \sup_{x\in C} \sup_{y\in K}\{\langle y^*, y\rangle - \langle A^*y^* + c, x\rangle\}. \end{aligned}$$

We obtain that $v^*(y^*) = 0$ if $y^* \in -K^*$ and $A^*y^* + c \in C^*$, and $v^*(y^*) = +\infty$ otherwise. That is, $v^*(y^*)$ is the indicator function of the feasible set of the dual problem. Therefore, we can write the dual problem (D) as

$$\operatorname*{Max}_{y^*\in Y^*}\{\langle y^*, b\rangle - v^*(y^*)\}. \tag{2.11}$$

It follows that $\mathrm{val}(D) = v^{**}(b)$, where

$$v^{**}(y) := \sup_{y^*\in Y^*}\{\langle y^*, y\rangle - v^*(y^*)\} \tag{2.12}$$

is the biconjugate of $v(y)$.

So far we did not use any topology in the considered spaces. We make the following assumption in the consequent analysis.

(A2) The spaces Y and Y' are paired locally convex topological vector spaces.

All consequent topological statements are made with respect to the considered topologies of Y and Y'.

Recall that the domain of an extended real valued function $f : Y \to \overline{I\!R}$ is defined as

$$\operatorname{dom} f := \{y \in Y : f(y) < +\infty\}$$

and its subdifferential is defined at a point y, where $f(y)$ is finite valued, as

$$\partial f(y) := \{y^* \in Y' : f(z) - f(y) \geq \langle y^*, z - y\rangle, \ \ \forall\, z \in Y\}.$$

It immediately follows from the definitions that

$$y^* \in \partial f(y) \ \text{ iff } \ f^*(y^*) = \langle y^*, y\rangle - f(y).$$

By applying that to the function f^{**}, instead of f, we obtain that $y^* \in \partial f^{**}(y)$ iff $f^{***}(y^*) + f^{**}(y) = \langle y^*, y\rangle$. Now by the Fenchel-Moreau theorem we have that $f^{***} = f^*$. Consequently we obtain (cf., [5, Lemma 2.4, p.52], [22, p.35])

$$\partial f^{**}(y) = \arg\max_{y^* \in Y^*} \{\langle y^*, y\rangle - f^*(y^*)\}. \tag{2.13}$$

This leads to the following results, [5],[22].

Proposition 2.2 *The following holds:* (i) $\mathrm{val}(D) = v^{**}(b)$*,* (ii) *if* $\mathrm{val}(D)$ *is finite, then* $\mathrm{Sol}(D) = \partial v^{**}(b)$*.*

Proof. We already showed that $\mathrm{val}(D) = v^{**}(b)$. In order to prove assertion (ii) we have to show that $\bar{y}^*$ is a maximizer of the right hand side of (2.12), for $y = b$, iff $\bar{y}^* \in \partial v^{**}(b)$. This follows immediately from (2.13). ∎

Note that assertion (i) in the above proposition does not depend on the assumption (A2), while assertion (ii) is based on (2.13) which in turn involves topological properties of the paired spaces Y and Y'.

Consider the feasible set Φ^* of the dual problem (2.4), i.e.,

$$\Phi^* = \{y^* \in Y^* : y^* \in -K^*, \ A^* y^* + c \in C^*\}.$$

We can right then that $\mathrm{val}(D) = \sup_{y^* \in \Phi^*} \langle y^*, b\rangle$. Since $\mathrm{val}(D) = v^{**}(b)$ it follows that

$$v^{**}(y) = \sup_{y^* \in \Phi^*} \langle y^*, y\rangle,$$

i.e., $v^{**}(y)$ is the support function of the set Φ^*.

By $\operatorname{lsc} v$ we denote the lower semicontinuous hull of the function v. That is, $\operatorname{lsc} v$ is the supremum of all lower semicontinuous functions majorized by v, i.e.,

$$\operatorname{lsc} v(y) = \min\left\{v(y), \liminf_{z \to y} v(z)\right\}.$$

The problem (P) is said to be *subconsistent* if $\operatorname{lsc} v(b) < +\infty$. Of course, if (P) is consistent (i.e., its feasible set is nonempty), then $v(b) < +\infty$, and hence $\operatorname{lsc} v(b) < +\infty$. Therefore, if (P) is consistent, then it is subconsistent. By $\operatorname{cl} v$ we denote the closure of the function v,

$$\operatorname{cl} v(\cdot) := \begin{cases} \operatorname{lsc} v(\cdot), & \text{if } \operatorname{lsc} v(y) > -\infty \text{ for all } y \in Y, \\ -\infty, & \text{if } \operatorname{lsc} v(y) = -\infty \text{ for at least one } y \in Y. \end{cases}$$

By the Fenchel-Moreau theorem we know that $v^{**} = \operatorname{cl} v$, and we have that $\operatorname{cl} v(b) = \operatorname{lsc} v(b)$ if $\operatorname{lsc} v(b) < +\infty$, [22, Theorem 5]. Therefore, we have the following results.

Proposition 2.3 *The following holds:* (i) $\operatorname{val}(D) = \operatorname{cl} v(b)$, (ii) *if* (P) *is subconsistent, then* $\operatorname{val}(D) = \operatorname{lsc} v(b)$.

The above proposition shows that if (P) is subconsistent, then there is no duality gap between the primal and dual problems iff the optimal value function $v(y)$ is lower semicontinuous at $y = b$. Note that $v(y)$ is lower semicontinuous at every $y \in Y$ iff its epigraph $\operatorname{epi} v := \{(y, \alpha) : \alpha \geq v(y)\}$ is a closed subset of $Y \times I\!R$.

It may be not easy to verify lower semicontinuity of $v(y)$ directly. Therefore, we discuss now conditions ensuring the "no duality gap" property and existence of optimal solutions.

Proposition 2.4 *Suppose that the space X is a topological vector space, the functional $\langle c, \cdot\rangle$ is lower semicontinuous, the linear mapping $A : X \to Y$ is continuous, the cones C and K are closed in the respective topologies, and the following, so-called* inf-compactness, *condition holds: there exist $\alpha \in I\!R$ and a compact set $S \subset X$ such that for every y in a neighborhood of b the level set*

$$\{x \in X : \langle c, x\rangle \leq \alpha,\ x \in C,\ Ax + y \in K\} \tag{2.14}$$

is contained in S, and for $y = b$ this level set is nonempty. Then $\operatorname{Sol}(P)$ *is nonempty and compact and the optimal value function $v(y)$ is lower semicontinuous at $y = b$.*

Proof. The inf-compactness condition implies that there is a level set of (P) which is nonempty and compact. Since the objective function $\langle c, \cdot\rangle$ is lower semicontinuous, it follows that $\operatorname{Sol}(P)$ is nonempty and compact, and hence $\operatorname{val}(P)$ is finite. We can take $\alpha = \operatorname{val}(P)$. Since $\langle c, \cdot\rangle$ is lower semicontinuous and $\operatorname{Sol}(P)$ is compact, it follows that for any $\varepsilon > 0$ there exists a neighborhood N of $\operatorname{Sol}(P)$ such that $\langle c, x\rangle \geq \operatorname{val}(P) - \varepsilon$ for all $x \in N$. Consider the multifunction $\mathcal{M}(y)$ which maps $y \in Y$ into the level set (2.14). The multifunction $\mathcal{M}(y)$ is closed, $\mathcal{M}(b) = \operatorname{Sol}(P)$ and for all y in a neighborhood of b we have that $\mathcal{M}(y)$ is a subset of a compact set. It follows that $\mathcal{M}(y)$ is upper semicontinuous at $y = b$, and hence $\mathcal{M}(y) \subset N$ for all y in a neighborhood of b. We obtain that for all y in that neighborhood of b the inequality $v(y) \geq \operatorname{val}(P) - \varepsilon$ holds. Since $\varepsilon > 0$ was arbitrary it follows that $v(y)$ is lower semicontinuous at $y = b$. ∎

The inf-compactness condition implies that the problem (P) is consistent, and hence it follows from the lower semicontinuity of $v(y)$ that $\operatorname{val}(P) = \operatorname{val}(D)$. Note that the above proposition depends on the chosen topology of the space X.

It is said that the function $v(y)$ is *subdifferentiable* at $y = b$ if $v(b)$ is finite and the subdifferential $\partial v(b)$ is nonempty. By convex analysis we have that if $v(y)$ is subdifferentiable at $y = b$, then $v^{**}(b) = v(b)$, and conversely if $v^{**}(b) = v(b)$ and is finite, then $\partial v(b) = \partial v^{**}(b)$, [22]. This leads to the following results.

Proposition 2.5 *If the optimal value function $v(y)$ is subdifferentiable at $y = b$, then* $\text{val}(P) = \text{val}(D)$ *and the set of optimal solutions of (D) coincides with $\partial v(b)$. Conversely, if* $\text{val}(P) = \text{val}(D)$ *and is finite, then* $\text{Sol}(D) = \partial v(b)$.

The above proposition shows that there is no duality gap between the primal and dual problems and the dual problem has an optimal solution iff $v(y)$ is subdifferentiable at $y = b$. Yet it may be not easy to verify subdifferentiability of the optimal value function directly.

Let us consider the following set

$$M := \{(y, \alpha) \in Y \times I\!R : y = k - Ax,\ \alpha \geq \langle c, x \rangle,\ x \in C,\ k \in K\}. \tag{2.15}$$

It is not difficult to see that M is a convex cone in the space $Y \times I\!R$. Moreover, we have that $\alpha \in I\!R$ is greater than or equal to $\langle c, x \rangle$ for some feasible point x of the primal problem (1.1) iff $(b, \alpha) \in M$. Therefore, the optimal value of (1.1) is equal to the optimal value of the following problem

$$\text{Min } \alpha \quad \text{subject to} \quad (b, \alpha) \in M. \tag{2.16}$$

It follows that M is the epigraph of the optimal value function $v(\cdot)$ iff the minimum in the above problem (2.16) is attained for every b such that $v(b)$ is finite. This happens, in particular, if the cone M is closed (in the product topology of $Y \times I\!R$). Sometimes the following cone

$$M' := \{(y, \alpha) \in Y \times I\!R : y = k - Ax,\ \alpha = \langle c, x \rangle,\ x \in C,\ k \in K\} \tag{2.17}$$

is used rather than the cone M (see (3.6) and proposition 3.1). Note that the cone M is closed iff the cone M' is closed

Proposition 2.6 *Suppose that* $\text{val}(P)$ *is finite and the cone M is closed in the product topology of $Y \times I\!R$. Then* $\text{val}(P) = \text{val}(D)$ *and the primal problem (P) has an optimal solution.*

Proof. Since M is closed we have that the set $\{\alpha : (b, \alpha) \in M\}$ is closed, and since $\text{val}(P)$ is finite this set is nonempty. Moreover, for any α in that set we have that $\alpha \geq \text{val}(P)$. Therefore, this set has a minimal element $\bar{\alpha}$. Let $\bar{x}$ be a corresponding feasible point of the primal problem (1.1). Then by the above construction we have that $\bar{\alpha} = \text{val}(P)$ and $\bar{x} \in \text{Sol}(P)$. This proves that the primal problem (P) has an optimal solution.

Since M is closed we have that $\text{epi}\, v = M$, and hence $\text{epi}\, v$ is closed. It follows that $v(y)$ is lower semicontinuous, and hence $\text{val}(P) = \text{val}(D)$ by proposition 2.3. ∎

By convex analysis we know that if $v(y)$ is finite valued and continuous at $y = b$, then $v(y)$ is subdifferentiable at $y = b$ and $\partial v(b)$ is compact in the paired topology of Y' (e.g., [11, p. 84]). In particular, if Y is a Banach space and Y^* is its standard dual, then $\partial v(b)$ is closed and bounded in the dual norm topology of Y^*. This leads to the following results.

Proposition 2.7 *Suppose that the optimal value function $v(y)$ is continuous at $y = b$. Then* $\mathrm{val}(P) = \mathrm{val}(D)$ *and the set of optimal solutions of (D) is nonempty. Moreover, if Y is a Banach space paired with its standard dual Y^*, then the set of optimal solutions of (D) is bounded in the dual norm topology of Y^*.*

Suppose now that $v(y)$ is bounded from above on an open subset of Y, i.e., the following condition holds.

(A3) There exist an open set $N \subset Y$ and $c \in I\!R$ such that $v(y) \le c$ for all $y \in N$.

Note that the above condition holds iff the cone M, defined in (2.15), has a nonempty interior in the product topology of $Y \times I\!R$. In particular, condition (A3) holds if the cone K has a nonempty interior (in the considered topology of Y). Indeed, let N be an open subset of Y included in K. Then $v(y) \le \langle c, 0 \rangle = 0$ for any $y \in N$.

Suppose that either assumption (A3) is satisfied or the space Y is finite dimensional. Then the optimal value function $v(y)$ is subdifferentiable at $y = b$ iff

$$\liminf_{t \downarrow 0} \frac{v(b + td) - v(b)}{t} > -\infty, \quad \forall\, d \in Y \tag{2.18}$$

(e.g., [4, Proposition 2.134]).

Suppose that either one of the following conditions holds: (i) the space Y is finite dimensional, (ii) assumption (A3) is satisfied, (iii) X and Y are Banach spaces (equipped with strong topologies), the cones C and K are closed and $\langle c, \cdot \rangle$ and $A : X \to Y$ are continuous. Then the function $v(y)$ is continuous at $y = b$ iff

$$b \in \mathrm{int}(\mathrm{dom}\, v). \tag{2.19}$$

It is well known that a convex function over a finite dimensional space is continuous at every interior point of its domain. In a locally convex topological vector space this holds under assumption (A3) (e.g., [11]). The result that under the above assumption (iii) condition (2.19) is necessary and sufficient for continuity of the optimal value function $v(y)$ at $y = b$ is due to Robinson [19]. Clearly we have that $\mathrm{dom}\, v = -A(C) + K$. Therefore condition (2.19) can be written in the following equivalent form

$$-b \in \mathrm{int}[A(C) - K]. \tag{2.20}$$

Suppose that either the space Y is finite dimensional or Y is a Banach space and assumption (A3) holds. Then $v(y)$ is continuous at $y = b$ iff $\partial v(b)$ is nonempty and bounded (e.g., [4, Proposition 2.131]).

The above discussion leads to the following results. We assume that if the space Y (the space X) is a Banach space, then it is equipped with its strong (norm) topology and is paired with its standard dual Y^* equipped with the weak* topology.

Proposition 2.8 *Suppose that* val(P) *is finite and either the space* Y *is finite dimensional or assumption* (A3) *is satisfied. Then the following holds.*
(i) val$(P) =$ val(D) *and* Sol(D) *is nonempty iff condition* (2.18) *holds.*
(ii) *If condition* (2.20) *holds, then* val$(P) =$ val(D) *and* Sol(D) *is nonempty, and, moreover, if* Y *is a Banach space, then* Sol(D) *is bounded.*
(iii) *If the space* Y *is a Banach space and* Sol(D) *is nonempty and bounded, then condition* (2.20) *holds and* val$(P) =$ val(D).

Proposition 2.9 *Suppose that* X *and* Y *are Banach spaces, the cones* C *and* K *are closed and* $\langle c, \cdot \rangle$ *and* $A : X \to Y$ *are continuous and condition* (2.20) *holds. Then* val$(P) =$ val(D) *and* Sol(D) *is nonempty and bounded.*

It is said that the (generalized) Slater condition for the problem (1.1) is satisfied if

$$\exists\, \bar{x} \in C \text{ such that } A\bar{x} + b \in \operatorname{int}(K). \tag{2.21}$$

Clearly (2.21) implies (2.20). The converse of that is also true if the cone K has a nonempty interior. That is, if K has a nonempty interior, then the (generalized) Slater condition (2.21) is equivalent to the condition (2.20) (e.g, [4, Proposition 2.106]).

In some cases there are equality as well as inequality type constraints involved in the definition of the considered problem. That is, the feasible set of the primal problem is defined by the constraints

$$A_1 x + b_1 = 0, \quad A_2 x + b_2 \in K_2, \quad x \in C, \tag{2.22}$$

where $A_i : X \to Y_i$, $i = 1, 2$, are linear mappings and $C \subset X$ and $K_2 \subset Y_2$ are convex cones. Clearly such constraints can be written in the form of the problem (1.1) by defining the cone $K := \{0\} \times K_2$ and the mapping $Ax := (A_1 x, A_2 x)$ from X into $Y := Y_1 \times Y_2$. In the case of such constraints and if the cones C and K_2 have nonempty interiors, the regularity condition (2.19) can be written in the following equivalent form (cf., [4, section 2.3.4]):

$$\begin{array}{l} A_1 \text{ is onto, i.e., } A_1(X) = Y_1, \\ \exists\, \bar{x} \in \operatorname{int}(C) \text{ such that } A_1\bar{x} + b_1 = 0, \quad A_2\bar{x} + b_2 \in \operatorname{int}(K_2). \end{array} \tag{2.23}$$

We have by proposition 2.9 that if X and Y are Banach spaces, the cones C and K_2 are closed and $\langle c, \cdot \rangle$ and $A_i : X \to Y_i$, $i = 1, 2$, are continuous and conditions (2.23) hold, then val$(P) =$ val(D) and Sol(D) is nonempty and bounded.

3 Problem of moments

In this section we discuss duality theory of the following conic linear problems. Let Ω be a nonempty set, $\mathcal{F}$ be a sigma algebra of subsets of Ω, and $\phi(\omega), \psi_1(\omega), ..., \psi_p(\omega)$, be real valued measurable functions on $(\Omega, \mathcal{F})$. Consider the set $\mathcal{M}^+$ of all probability measures on the measurable space $(\Omega, \mathcal{F})$ such that each function $\phi, \psi_1, ..., \psi_p$ is μ-integrable for all $\mu \in \mathcal{M}^+$. Let $\mathcal{X}$ be the linear space of signed measures generated by $\mathcal{M}^+$, and $\mathcal{X}'$ be the linear space of functions $f : \Omega \to I\!\!R$ generated by the functions $\phi, \psi_1, ..., \psi_p$ (i.e., elements of $\mathcal{X}'$ are formed by linear combinations of the functions $\phi, \psi_1, ..., \psi_p$). The spaces $\mathcal{X}$ and $\mathcal{X}'$ are paired by the bilinear form (scalar product)

$$\langle f, \mu \rangle := \int_\Omega f(\omega) d\mu(\omega). \tag{3.1}$$

Since in the sequel we deal with signed measures we say that a measure $\mu \in \mathcal{X}$ is nonnegative, and write $\mu \succeq 0$, if $\mu(\Xi) \geq 0$ for any $\Xi \in \mathcal{F}$. If $\mu \succeq 0$ and $\mu(\Omega) = 1$, then it is said that μ is a probability measure. In that notation

$$\mathcal{M}^+ = \{\mu \in \mathcal{X} : \mu \succeq 0,\ \mu(\Omega) = 1\}.$$

Consider the problem

$$\operatorname*{Max}_{\mu \in \mathcal{C}} \langle \phi, \mu \rangle \ \text{ subject to } \ A\mu - b \in K, \tag{3.2}$$

where $\mathcal{C}$ is a convex cone in $\mathcal{X}$, K is a closed convex cone in $I\!\!R^p$, $b = (b_1, ..., b_p) \in I\!\!R^p$ and $A : \mu \mapsto (\langle \psi_1, \mu \rangle, ..., \langle \psi_p, \mu \rangle)$ is the linear mapping from $\mathcal{X}$ into $I\!\!R^p$. In particular, suppose that the cone $\mathcal{C}$ is generated by a convex set $\mathcal{A} \subset \mathcal{M}^+$ of probability measures on $(\Omega, \mathcal{F})$, and set $\psi_1(\cdot) \equiv 1$, $b_1 = 1$ and $K = \{0\}$. Then problem (3.2) becomes

$$\begin{array}{ll} \text{Max}_{\mu \in \mathcal{A}} & I\!\!E_\mu[\phi(\omega)] \\ \text{subject to} & I\!\!E_\mu[\psi_i(\omega)] = b_i,\ i = 2, ..., p, \end{array} \tag{3.3}$$

where $I\!\!E_\mu$ denotes the expected value with respect to the probability measure μ. In case $\mathcal{A} = \mathcal{M}^+$ the above problem (3.3) is the classical problem of moments. For a discussion of the historical background of the problem of moments the interested reader is referred to the monograph [16] and references therein.

Problem (3.2) is a conic linear problem of the form (1.1) with the "min" operator replaced by the "max" operator and the space $Y := I\!\!R^p$ being finite dimensional. The space Y is paired with itself with respect to the standard scalar product in $I\!\!R^p$ denoted by " $\cdot$ ". The Lagrangian of the problem (3.2) can be written in the form

$$L(\mu, x) := \langle \phi, \mu \rangle - x \cdot (A\mu - b) = \langle \phi - A^* x, \mu \rangle + x \cdot b,$$

where $A^* x = \sum_{i=1}^p x_i \psi_i$ is the corresponding adjoint mapping from $I\!\!R^p$ into $\mathcal{X}'$.

Therefore the dual of (3.2) is the conic linear problem

$$\operatorname*{Min}_{x \in -K^*} b \cdot x \ \text{ subject to } \ \sum_{i=1}^p x_i \psi_i - \phi \in \mathcal{C}^*. \tag{3.4}$$

Recall that

$$\mathcal{C}^* := \{f \in \mathcal{X}' : \langle f, \mu \rangle \geq 0,\ \forall\, \mu \in \mathcal{C}\}.$$

Suppose that the cone $\mathcal{C}$ is generated by a convex set $\mathcal{A} \subset \mathcal{X}$ (written $\mathcal{C} = \text{cone}(\mathcal{A})$), e.g., by a convex set of probability measures. Then, since $\langle f, \mu \rangle$ is linear in μ, it follows that the condition "$\forall\, \mu \in \mathcal{C}$" in the above definition of $\mathcal{C}^*$ can be replaced by "$\forall\, \mu \in \mathcal{A}$". Consequently, the dual problem (3.4) can be written as follows

$$\operatorname*{Min}_{x \in -K^*} b \cdot x \ \text{ subject to } \sum_{i=1}^{p} x_i \langle \psi_i, \mu \rangle \geq \langle \phi, \mu \rangle, \ \ \forall\, \mu \in \mathcal{A}. \tag{3.5}$$

We refer in this section to (3.2) as the primal (P) and to (3.4) (or (3.5)) as the dual (D) problems, respectively. The above problem (3.5) is a linear semi-infinite programming problem. This is because the optimization space $\mathbb{R}^p$ is finite dimensional while the number of constraints is infinite.

Consider the cone

$$M_{p+1} := \Big\{ (x, \alpha) \in \mathbb{R}^{p+1} : (x, \alpha) = (A\mu - k, \langle \phi, \mu \rangle), \ \mu \in \mathcal{C}, \ k \in K \Big\} \tag{3.6}$$

associated with the primal problem (3.2). The above cone M_{p+1} is a specification of the cone defined in (2.17) to the present problem. The next result follows by proposition 2.6.

Proposition 3.1 *Suppose that* $\text{val}(P)$ *is finite and the cone* M_{p+1} *is closed in the standard topology of* $\mathbb{R}^{p+1}$*. Then* $\text{val}(P) = \text{val}(D)$ *and the primal problem* (3.2) *has an optimal solution.*

Of course, it is inconvenient to parameterize the inequality constraints in (3.5) by measures. Suppose that every finite subset of Ω is $\mathcal{F}$-measurable, which means that for every $\omega \in \Omega$ the corresponding Dirac measure $\delta(\omega)$ (of mass one at the point ω) belongs to $\mathcal{M}^+$. That assumption certainly holds if Ω is a Hausdorff topological (e.g., metric) space and $\mathcal{F}$ is the Borel sigma algebra of Ω. Suppose further that $\mathcal{C} = \text{cone}(\mathcal{A})$, where $\mathcal{A} \subset \mathcal{M}^+$ is a convex set such that $\delta(\omega) \in \mathcal{A}$ for every $\omega \in \Omega$. Then $\mathcal{C}^*$ is formed by the nonnegative valued functions, that is

$$\mathcal{C}^* = \{ f \in \mathcal{X}' : f(\omega) \geq 0, \ \forall\, \omega \in \Omega \}. \tag{3.7}$$

Indeed, since $\mathcal{A} \subset \mathcal{M}^+$ we have that the right hand side of (3.7) is included in $\mathcal{C}^*$. Since $\langle f, \delta(\omega) \rangle = f(\omega)$ for any Dirac measure $\delta(\omega)$, $\omega \in \Omega$, we obtain that the other inclusion necessarily holds, and hence (3.7) follows. Therefore, in that case we obtain that the dual problem (3.5) can be written in the form:

$$\begin{array}{ll} \text{Min}_{x \in -K^*} & b \cdot x \\ \text{subject to} & x_1 \psi_1(\omega) + ... + x_p \psi_p(\omega) \geq \phi(\omega), \ \forall\, \omega \in \Omega. \end{array} \tag{3.8}$$

Note that if $\mathcal{A} = \mathcal{M}^+$ and every finite subset of Ω is $\mathcal{F}$-measurable, then it suffices to take only measures with a finite support of at most $p+1$ points in the definition of the cone M_{p+1}, i.e., the obtained cone will coincide with the one defined in (3.6). This follows from lemma 3.1 below. If moreover $K = \{0\}$, then this cone can be written in the form

$$M_{p+1} = \text{cone}\{ (\psi_1(\omega), ..., \psi_p(\omega), \phi(\omega)), \ \omega \in \Omega \}. \tag{3.9}$$

In that form the cone M_{p+1} and proposition 3.1 are well known in the theory of semi-infinite programming, [6, p. 79].

As it was shown in section 2, the weak duality $\text{val}(D) \geq \text{val}(P)$ always holds for the primal problem (P) and its Lagrangian dual (D). (Note that in this section the primal is a maximization problem, therefore the inequality relation between the optimal values of the primal and dual problems is reverse from the one of section 2.) If the set Ω is finite, $K = \{0\}$ and the set $\mathcal{A}$ is given by the set of all probability measures on Ω, then problems (3.3) and (3.8) are linear programming problems and it is well known that in that case $\text{val}(D) = \text{val}(P)$ unless both problems are inconsistent. If Ω is infinite the situation is more subtle of course. We apply now the conjugate duality approach described in section 2.

Let us associate with the primal problem (3.2) the corresponding optimal value function

$$v(y) := \inf\{\langle -\phi, \mu\rangle : \mu \in \mathcal{C},\ A\mu - y \in K\}. \tag{3.10}$$

(We use the infimum rather than the supremum in the definition of the optimal value function in order to deal with a convex rather than concave function.) We have here $\text{val}(P) = -v(b)$ and by the derivations of section 2 the following results follow.

Proposition 3.2 *The following holds.*
(i) *The optimal value function $v(y)$ is convex.*
(ii) $\text{val}(D) = -v^{**}(b)$.
(iii) *If the primal problem (P) is subconsistent, then* $\text{val}(D) = -\text{lsc}\, v(b)$.
(iv) *If* $\text{val}(D)$ *is finite, then the (possibly empty) set of optimal solutions of (D) coincides with $-\partial v^{**}(b)$.*
(v) *If $v(y)$ is subdifferentiable at b, then* $\text{val}(D) = \text{val}(P)$ *and the set of optimal solutions of (D) coincides with $-\partial v(b)$.*
(vi) *If* $\text{val}(D) = \text{val}(P)$ *and is finite, then the (possibly empty) set of optimal solutions of (D) coincides with $-\partial v(b)$.*

It follows from the assertion (iii) of the above proposition that in the subconsistent case there is no duality gap between the primal and dual problems iff the optimal value function $v(y)$ is lower semicontinuous at $y = b$. It follows from (v) and (vi) that $\text{val}(D) = \text{val}(P)$ and the set of optimal solutions of (D) is nonempty iff $\text{val}(D)$ is finite and $v(y)$ is subdifferentiable at $y = b$. Moreover, since the space $Y = I\!R^p$ is finite dimensional we have that $v(y)$ is subdifferentiable at b iff condition (2.18) holds.

Suppose now that the set Ω is a metric space, $\mathcal{F}$ is the Borel sigma algebra of Ω, the functions $\psi_1, ..., \psi_p$ are bounded and continuous and ϕ is upper semicontinuous on Ω. Let us equip the space $\mathcal{X}$ with the weak topology (see, e.g., [3] for a discussion of the weak topology in spaces of probability measures). Under such conditions the mapping $A : \mathcal{X} \to I\!R^p$ is continuous and the functional $\langle -\phi, \cdot\rangle$ is lower semicontinuous. We have then by proposition 2.4 that $\text{val}(P) = \text{val}(D)$ if the inf-compactness condition holds. In particular, consider the moment problem, that is $\mathcal{C} = \text{cone}(\mathcal{A})$ where $\mathcal{A}$ is a convex set of probability measures, $\psi_1(\cdot) \equiv 1$ and $b_1 = 1$. Suppose that the set $\mathcal{A}$ is compact in the weak topology of $\mathcal{X}$. Take a closed interval $[\alpha, \beta] \subset I\!R$ such that $0 < \alpha < 1 < \beta$. Then the set $\mathcal{S} := \bigcup_{t \in [\alpha,\beta]} t\mathcal{A}$ is also compact and $\{\mu \in \mathcal{C} : A\mu - y \in K\} \subset \mathcal{S}$ for all vectors y such that their first component belongs to the interval (α, β). Therefore, if moreover the problem (P) is consistent, then the inf-compactness condition holds. We obtain the following result.

Proposition 3.3 *Consider the moment problem (i.e., $\mathcal{A}$ is a subset of $\mathcal{M}^+$). Suppose that:* (i) *the set Ω is a metric space and $\mathcal{F}$ is its Borel sigma algebra,* (ii) *the functions $\psi_2, ..., \psi_p$ are bounded continuous and ϕ is upper semicontinuous on Ω,* (iii) *the problem (P) is consistent,* (iv) *the set $\mathcal{A}$ is compact in the weak topology of $\mathcal{X}$.*
Then $\text{val}(P) = \text{val}(D)$ *and* $\text{Sol}(P)$ *is nonempty.*

Recall that by Prohorov's theorem a closed (in the weak topology) set $\mathcal{A}$ of probability measures is compact if it is tight, i.e., for any $\varepsilon > 0$ there exists a compact set $\Xi \subset \Omega$ such that $\mu(\Xi) > 1 - \varepsilon$ for any $\mu \in \mathcal{A}$. In particular, if Ω is a compact metric space, then the set of all probability measures on $(\Omega, \mathcal{F})$, is weakly compact. Therefore, we obtain the following corollary, [14].

Corollary 3.1 *Consider the moment problem. Suppose that Ω is a compact metric space, the functions $\psi_2, ..., \psi_p$ are continuous and ϕ is upper semicontinuous, and the primal problem is consistent. Then there is no duality gap between the primal and the corresponding dual problem and* $\text{Sol}(P)$ *is nonempty.*

Consider now condition (2.19). Since the objective function is real valued we have that $y \in \text{dom}\, v$ iff the corresponding feasible set $\{\mu \in \mathcal{C} : A\mu - y \in K\}$ is nonempty. That is,

$$\text{dom}\, v = A(\mathcal{C}) - K. \tag{3.11}$$

Since the space Y is finite dimensional here, we have that the following conditions are equivalent: (i) $v(y)$ is continuous at $y = b$, (ii) $b \in \text{int}(\text{dom}\, v)$, (iii) $\partial v(b)$ is nonempty and bounded. Because of (3.11), the condition $b \in \text{int}(\text{dom}\, v)$ (i.e., condition (2.19)) can be written in the form

$$b \in \text{int}[A(\mathcal{C}) - K]. \tag{3.12}$$

We also have that if $\text{lsc}\, v(b)$ is finite, then $v(y)$ is continuous at b iff $\partial v^{**}(b)$ is nonempty and bounded. Therefore we obtain the following results.

Proposition 3.4 *If condition* (3.12) *holds, then* $\text{val}(D) = \text{val}(P)$*, and, moreover, if the common optimal value of problems (D) and (P) is finite, then the set of optimal solutions of (D) is nonempty and bounded. Conversely, if* $\text{val}(D)$ *is finite and the set of optimal solutions of (D) is nonempty and bounded, then condition* (3.12) *holds.*

Sufficiency of condition (3.12) to ensure the property: "$\text{val}(D) = \text{val}(P)$ and $\text{Sol}(D)$ is nonempty and bounded", is well known ([12],[14]). In fact the above proposition shows that condition (3.12) is equivalent to that property, provided that $\text{val}(P)$ is finite.

A condition weaker than (3.12) is that b belongs to the relative interior of the convex set $A(\mathcal{C}) - K$. Under such condition $v(\cdot)$ is subdifferentiable at b, and hence $\mathrm{val}(D) = \mathrm{val}(P)$ and the set of optimal solutions of (D) is nonempty (although may be unbounded), provided that $\mathrm{val}(D)$ is finite.

By the above discussion we have that if $v(b)$ is finite, then $\partial v(b)$ is nonempty and bounded iff condition (3.12) holds. We also have by convex analysis that the convex function $v(\cdot)$ is differentiable at b iff $\partial v(b)$ is a singleton, [20]. Therefore, we obtain that $v(\cdot)$ is differentiable at b iff the dual problem (D) has a unique optimal solution. Since $v(\cdot)$ is a convex function on a finite dimensional space it follows that there is a set $S \subset \mathbb{R}^p$ of Lebesgue measure zero such that for every $b \in \mathbb{R}^p \setminus S$ either $v(b) = \pm\infty$ or $v(y)$ is differentiable at b, [20]. Therefore, we obtain the following result.

Proposition 3.5 *For almost every $b \in \mathbb{R}^p$ (with respect to the Lebesgue measure) either $\mathrm{val}(P) = \pm\infty$ or $\mathrm{val}(P) = \mathrm{val}(D)$ and $\mathrm{Sol}(D)$ is a singleton.*

The result that for almost every b such that $\mathrm{val}(D)$ is finite, the set $\mathrm{Sol}(D)$ is a singleton was obtained in [14].

Let us remark that the assumption that every finite subset of Ω is measurable was used in the above arguments only in order to derive formula (3.7) for the polar of the cone $\mathcal{C}$, and hence to calculate the dual problem in the form (3.8). The following lemma shows that if $\mathcal{A} = \mathcal{M}^+$ and every finite subset of Ω is $\mathcal{F}$-measurable, then it suffices to solve the problem of moments (3.3) with respect to discrete probability measures with a finite support. This lemma is due to Rogosinsky [23], we quickly outline its proof for the sake of completeness.

Lemma 3.1 *Suppose that every finite subset of Ω is $\mathcal{F}$-measurable. Let $f_1, ..., f_n$ be measurable on $(\Omega, \mathcal{F})$ real valued functions, and let μ be a nonnegative measure on $(\Omega, \mathcal{F})$ such that $f_1, ..., f_n$ are μ-integrable. Then there exists a nonnegative measure μ' on $(\Omega, \mathcal{F})$ with a finite support of at most n points such that $\langle f_i, \mu \rangle = \langle f_i, \mu' \rangle$ for all $i = 1, ..., n$.*

Proof. The proof proceeds by induction on n. It can be easily shown that the assertion holds for $n = 1$. Consider the set $S \subset \mathbb{R}^n$ generated by vectors of the form $(\langle f_1, \mu' \rangle, ..., \langle f_n, \mu' \rangle)$ with μ' being a nonnegative measure on Ω with a finite support. We have to show that vector $a := (\langle f_1, \mu \rangle, ..., \langle f_n, \mu \rangle)$ belongs to S. Note that the set S is a convex cone. Suppose that $a \notin S$. Then, by the separation theorem, there exists $c \in \mathbb{R}^n \setminus \{0\}$ such that $c \cdot a \le c \cdot x$, for all $x \in S$. Since S is a cone, it follows that $c \cdot a \le 0$. This implies that $\langle f, \mu \rangle \le 0$ and $\langle f, \mu \rangle \le \langle f, \mu' \rangle$ for any measure $\mu' \succeq 0$ with a finite support, where $f := \sum_{i=1}^n c_i f_i$. In particular, by taking measures of the form $\mu' = \alpha \delta(\omega)$, $\alpha > 0$, $\omega \in \Omega$, we obtain from the second inequality that $f(\omega) \ge 0$ for all $\omega \in \Omega$. Together with the first inequality this implies that $\langle f, \mu \rangle = 0$.

Consider the set $\Omega' := \{\omega \in \Omega : f(\omega) = 0\}$. Note that the function f is measurable and hence $\Omega' \in \mathcal{F}$. Since $\langle f, \mu \rangle = 0$ and $f(\cdot)$ is nonnegative valued, it follows that Ω' is a support of μ, i.e., $\mu(\Omega') = \mu(\Omega)$. If $\mu(\Omega) = 0$, then the assertion clearly holds. Therefore suppose that $\mu(\Omega) > 0$. Then $\mu(\Omega') > 0$, and hence Ω' is nonempty. Moreover, the functions f_i,

$i = 1, ..., n$, are linearly dependent on Ω'. Consequently, by the induction assumption there exists a measure μ' with a finite support on Ω' such that $\langle f_i, \mu^* \rangle = \langle f_i, \mu' \rangle$ for all $i = 1, ..., n$, where μ^* is the restriction of the measure μ to the set Ω'. Moreover, since μ is supported on Ω' we have that $\langle f_i, \mu^* \rangle = \langle f_i, \mu \rangle$, and hence the proof is complete. ∎

Consider the problem of moments (3.3) with $\mathcal{A} = \mathcal{M}^+$. Suppose that every finite subset of Ω is $\mathcal{F}$-measurable. Then it follows by the above lemma that for any nonnegative measure μ satisfying the feasibility equations $\langle \psi_i, \mu \rangle = b_i$, $i = 1, ..., p$, (with $\psi_1(\cdot) \equiv 1$ and $b_1 = 1$) there exists a nonnegative measure μ' with a finite support of at most $p+1$ points satisfying the feasibility equations and such that $\langle \phi, \mu' \rangle = \langle \phi, \mu \rangle$. Consequently it suffices to solve the problem of moments with respect to discrete probability measures with a finite support of at most $p+1$ points. In fact it suffices to solve it with respect to probability measures with a finite support of at most p points. Indeed, consider the problem of moments restricted to a finite subset $\{\omega_1, ..., \omega_m\}$ of Ω. That is,

$$\begin{array}{ll} \text{Max}_{\alpha \in \mathbb{R}^m} & \sum_{j=1}^m \alpha_j \phi(\omega_j) \\ \text{subject to} & \sum_{j=1}^m \alpha_j \psi_i(\omega_j) = b_i, \; i = 2, ..., p, \\ & \sum_{j=1}^m \alpha_j = 1, \;\; \alpha_j \geq 0, \; j = 1, ..., m. \end{array} \tag{3.13}$$

The feasible set of the above problem (3.13) is bounded, and hence, by standard arguments of linear programming, problem (3.13) has an optimal solution with at most p nonzero components provided that this problem is consistent.

It is also possible to approach the duality problem by considering the semi-infinite programming problem (3.4) as the primal problem. We discuss that in the next section.

4 Semi-infinite programming

In this section we consider the semi-infinite problem (3.8) and view it as the *primal* problem (P). We refer to the recent book by Goberna and López [7] for a thorough discussion of the theory of linear semi-infinite programming problems. In order to formulate (3.8) in the general framework of conic linear problems we make the following assumptions. Let X be the finite dimensional space $\mathbb{R}^p$ paired with itself by the standard scalar product in $\mathbb{R}^p$, $\mathcal{Y}$ be a linear space of functions $f : \Omega \to \mathbb{R}$ and $\mathcal{Y}'$ be a linear space of finite signed measures on $(\Omega, \mathcal{F})$ paired with $\mathcal{Y}$ by the bilinear form (3.1). We assume that $\mathcal{Y}$ and $\mathcal{Y}'$ are chosen in such a way that the functions $\phi, \psi_1, ..., \psi_p$, belong to $\mathcal{Y}$, the bilinear form (i.e., the integral) (3.1) is well defined for every $f \in \mathcal{Y}$ and $\mu \in \mathcal{Y}'$, and that the following condition holds.

(B1) The space $\mathcal{Y}'$ includes all measures with a finite support on Ω.

In particular, the above assumption implies that every finite subset of Ω is $\mathcal{F}$-measurable.

Let us consider the cones

$$C^+(\mathcal{Y}) := \{f \in \mathcal{Y} : f(\omega) \geq 0, \; \forall \, \omega \in \Omega\} \;\; \text{and} \;\; C^+(\mathcal{Y}') := \{\mu \in \mathcal{Y}' : \mu \succeq 0\}.$$

Clearly $\langle f, \mu \rangle \geq 0$ for any $f \in C^+(\mathcal{Y})$ and $\mu \in C^+(\mathcal{Y}')$. Therefore, the polar of the cone $C^+(\mathcal{Y})$ includes the cone $C^+(\mathcal{Y}')$, and the polar of the cone $C^+(\mathcal{Y}')$ includes the cone $C^+(\mathcal{Y})$. Assumption (B1) ensures that the polar of the cone $C^+(\mathcal{Y}')$ is equal to the cone $C^+(\mathcal{Y})$. We also assume that $C^+(\mathcal{Y}')$ is equal to the the polar of the cone $C^+(\mathcal{Y})$.

(B2) The cone $C^+(\mathcal{Y}')$ is the polar of the cone $C^+(\mathcal{Y})$.

This assumption can be ensured by taking the space $\mathcal{Y}$ to be sufficiently "large".

Under condition (B2) the Lagrangian dual of the semi-infinite problem (3.8) can be written as follows

$$\operatorname*{Max}_{\mu \succeq 0} \ \langle \phi, \mu \rangle \ \text{ subject to } \ A\mu - b \in K. \tag{4.1}$$

In this section we refer to (4.1) as the *dual* (D) problem. Assumption (B1) ensures that the Lagrangian dual of (4.1) coincides with the semi-infinite problem (3.8), and furthermore by lemma 3.1, it suffices to perform optimization in (4.1) with respect to measures with a finite support of at most $p+1$ points only.

We have the following result by proposition 3.4 (cf., [7, Theorem 8.1]).

Proposition 4.1 *Suppose that the assumption* (B1) *holds and that the primal problem* (3.8) *has a nonempty and bounded set of optimal solutions, then* $\operatorname{val}(P) = \operatorname{val}(D)$.

Consider the optimal value function

$$w(\xi) := \inf \left\{ b \cdot x : x \in -K^*, \ \sum_{i=1}^{p} x_i \psi_i(\omega) \geq \xi(\omega), \ \forall\, \omega \in \Omega \right\}. \tag{4.2}$$

associated with the problem (3.8). Clearly $w(\phi)$ is equal to the optimal value of (3.8).

Suppose now that Ω is a compact metric space and the functions $\phi, \psi_1, ..., \psi_p$, are continuous. In that case we can define $\mathcal{Y}$ to be the Banach space $C(\Omega)$ of continuous real valued functions on Ω equipped with the sup-norm. The dual $C(\Omega)^*$ of the space $C(\Omega)$ is the space of finite signed Borel measures on Ω. By equipping $\mathcal{Y} = C(\Omega)$ and $\mathcal{Y}^* = C(\Omega)^*$ with the strong (norm) and weak* topologies, respectively, we obtain a pair of locally convex topological vector spaces. In that framework the assumptions (B1) and (B2) always hold.

Note that the cone $C^+(\mathcal{Y})$ has a nonempty interior in the Banach space $\mathcal{Y} = C(\Omega)$. Note also that the norm of $\mu \in C(\Omega)^*$, which is dual to the sup-norm of $C(\Omega)$, is given by the total variation norm

$$\|\mu\| := \sup_{\Omega' \in \mathcal{F}} \mu(\Omega') - \inf_{\Omega'' \in \mathcal{F}} \mu(\Omega'').$$

Therefore, we have by proposition 2.8 the following results (cf, [4, section 5.4.1]).

Proposition 4.2 *Suppose that* Ω *is a compact metric space and that the optimal value of the problem* (3.8) *is finite. Then the following holds.*
(i) *There is no duality gap between* (3.8) *and* (4.1) *and the set of optimal solutions of* (4.1) *is nonempty iff the the following condition holds:*

$$\liminf_{t \downarrow 0} \frac{w(\phi + t\eta) - w(\phi)}{t} > -\infty, \ \ \forall\, \eta \in \mathcal{Y}. \tag{4.3}$$

(ii) *There is no duality gap between* (3.8) *and* (4.1) *and the set of optimal solutions of* (4.1) *is nonempty and bounded (with respect to the total variation norm) iff the following condition holds: there exists* $\bar{x} \in -K^*$ *such that*

$$\bar{x}_1\psi_1(\omega) + ... + \bar{x}_p\psi_p(\omega) > \phi(\omega), \ \ \forall\, \omega \in \Omega. \tag{4.4}$$

Recall that in the problem of moments the function ψ_1 is identically 1. Therefore, if also Ω is compact and the functions are continuous, then by taking $\bar{x}_1$ large enough condition (4.4) can be always satisfied. It follows that in such case there is no duality gap between the problem of moments and its dual. This result was already obtained in corollary 3.1.

We come back now to a discussion of the general case, i.e., we do not assume that Ω is compact and the considered functions are continuous. For a given finite set $\{\omega_1, ..., \omega_n\} \subset \Omega$ consider the following discretization of (3.8)

$$\begin{array}{ll} \text{Min}_{x \in -K^*} & b \cdot x \\ \text{subject to} & x_1\psi_1(\omega_i) + ... + x_p\psi_p(\omega_i) \geq \phi(\omega_i), \ i = 1, ..., n. \end{array} \tag{4.5}$$

We denote the above problem (4.5) by (P_n), and make the following assumption.

(B3) The cone K is polyhedral.

Since the cone K is polyhedral, its dual K^* is also polyhedral, and hence (P_n) is a linear programming problem. The dual of (P_n), denoted (D_n), is obtained by restricting the problem (4.1) to discrete measures (with a finite support) of the form $\mu = \sum_{i=1}^n \lambda_i \delta(\omega_i)$. Since the feasible set of (P_n) includes the feasible set of (P) and the feasible set of (D_n) is included in the feasible set of (D), we have that $\text{val}(P) \geq \text{val}(P_n)$ and $\text{val}(D) \geq \text{val}(D_n)$. Moreover, by the duality theorem of linear programming we have that $\text{val}(P_n) = \text{val}(D_n)$ unless both (P_n) and (D_n) are inconsistent.

Proposition 4.3 *Suppose that the assumptions* (B1) *and* (B3) *hold and that* $\text{val}(P)$ *is finite. Then the following holds.*
(i) $\text{val}(P) = \text{val}(D)$ *iff for any* $\varepsilon > 0$ *there exists a finite discretization* (P_n) *such that* $\text{val}(P_n) \geq \text{val}(P) - \varepsilon$.
(ii) $\text{val}(P) = \text{val}(D)$ *and the dual problem* (4.1) *has an optimal solution iff there exists a finite discretization* (P_n) *such that* $\text{val}(P_n) = \text{val}(P)$.

Proof. Suppose that there exists a finite discretization (P_n) such that $\text{val}(P_n) \geq \text{val}(P) - \varepsilon$. We have that $\text{val}(P_n) = \text{val}(D_n)$, and hence

$$\text{val}(P) \leq \text{val}(P_n) + \varepsilon = \text{val}(D_n) + \varepsilon \leq \text{val}(D) + \varepsilon. \tag{4.6}$$

Since $\text{val}(P) \geq \text{val}(D)$, it follows that $|\text{val}(P) - \text{val}(D)| \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this implies that $\text{val}(P) = \text{val}(D)$.

Conversely suppose that $\text{val}(P) = \text{val}(D)$. Since $\text{val}(D)$ is finite, for any $\varepsilon > 0$ the dual problem has an ε-optimal solution μ. By lemma 3.1 there exists a measure μ' with a finite support (of at most $p+1$ points) such that $A\mu = A\mu'$ and $\langle \phi, \mu \rangle = \langle \phi, \mu' \rangle$. It follows that μ' is also an ε-optimal solution of the dual problem. Let (D_n) and (P_n) be discretizations of the dual and primal problems, respectively, corresponding to the support of μ'. It follows that

$$\mathrm{val}(P_n) = \mathrm{val}(D_n) \geq \mathrm{val}(D) - \varepsilon, \tag{4.7}$$

which together with $\mathrm{val}(P) = \mathrm{val}(D)$ implies that $\mathrm{val}(P_n) \geq \mathrm{val}(P) - \varepsilon$. This completes the proof of assertion (i).

Assertion (ii) can be proved in a similar way by taking $\varepsilon = 0$. ∎

Since $\mathrm{val}(P) \geq \mathrm{val}(P_n)$ it follows that the condition: "for any $\varepsilon > 0$ there exists a finite discretization (P_n) such that $\mathrm{val}(P_n) \geq \mathrm{val}(P) - \varepsilon$", holds iff there exists a sequence of finite subsets of Ω such that the optimal values of the corresponding discretized problems converge to $\mathrm{val}(P)$. Linear semi-infinite programming problems satisfying this property are called *discretizable*, and problems satisfying the property that $\mathrm{val}(P_n) = \mathrm{val}(P)$, for some discretization (P_n), are called *reducible* in [7]. The results of proposition 4.3 were proved in [7, section 8.3] by a different method.

Recall that, by proposition 2.5, $\mathrm{val}(P) = \mathrm{val}(D)$ and $\mathrm{Sol}(D)$ is nonempty iff the optimal value function $w(\xi)$ is subdifferentiable at $\xi = \phi$. Therefore, we obtain that $w(\xi)$ is subdifferentiable at $\xi = \phi$ iff there exists a finite discretization of (P) with the same optimal value, i.e., iff (P) is reducible.

5 Continuous linear programming

In this section we discuss the following class of optimization problems

$$\begin{aligned}
\mathrm{Min}_x \quad & \int_0^1 c(t)^T x(t) dt && (5.1)\\
\text{s.t.} \quad & \int_0^t M(s,t)x(s)ds \geq a(t), \quad t \in [0,1], && (5.2)\\
& H(t)x(t) \geq h(t), \quad \text{a.e. } t \in [0,1], && (5.3)\\
& x(t) \geq 0, \quad \text{a.e. } t \in [0,1]. && (5.4)
\end{aligned}$$

Here $c(t)$, $a(t)$ and $h(t)$ are column vector valued functions, defined on the interval $[0,1]$, of dimensions n_1, n_2 and n_3, respectively, and $M(s,t)$ and $H(t)$ are matrix valued functions of dimensions $n_2 \times n_1$ and $n_3 \times n_1$, respectively. The notation "a.e." means that the corresponding inequality holds for almost every $t \in [0,1]$, i.e., there is a set $S \subset [0,1]$ of Lebesgue measure zero such that the corresponding inequality holds for all $t \in [0,1] \setminus S$.

We refer to (5.1)-(5.4) as a *continuous linear programming* (CLP) problem. Continuous linear programming problems, with the constraints (5.2) and (5.3) *mixed* together, were introduced by Bellman [2]. Duality theory of (mixed) continuous linear programming problems was discussed by several authors, notably by Levinson [15] and Grinold [8]. Duality theory of (CLP) problems in the above form is discussed extensively in Anderson and Nash [1], where such problems are called *separable* continuous linear programming problems.

Of course, in order to formalize the above (CLP) problem we need to specify in what class of functions $x(t) = (x_1(t), ..., x_{n_1}(t))^T$ we are looking for an optimal solution. We work with the following Banach spaces, the space $L_1[0,1]$ of Lebesgue integrable on the interval $[0,1]$ functions and the space $C[0,1]$ of continuous on $[0,1]$ functions. Recall that the dual of $L_1[0,1]$ is the space $L_\infty[0,1]$ of essentially bounded measurable functions, and the dual of $C[0,1]$ is the space of finite signed Borel measures on $[0,1]$, denoted by $C[0,1]^*$. Note that $C[0,1]^*$ can be identified with the space of continuous from the right functions $\psi(\cdot)$ of bounded variation on the interval $[0,1]$, normalized by $\psi(1) = 0$. By $C^+[0,1]$ we denote the cone of nonnegative valued on $[0,1]$ functions in the space $C[0,1]$, and by $L_\infty^+[0,1]$ the cone of almost everywhere nonnegative valued on $[0,1]$ functions in the space $L_\infty[0,1]$.

By $L_\infty[0,1]^*$ we denote the dual space of the Banach space $L_\infty[0,1]$. This dual space can be identified with the linear space generated by finite additive measures μ on $[0,1]$ of bounded variation such that if S is a Borel subset of $[0,1]$ of Lebesgue measure zero, then $\mu(S) = 0$. The corresponding linear functional $y^* \in L_\infty[0,1]^*$ is given by the integral $y^*(y) = \int_0^1 y(t)d\mu(t)$ (see, e.g., [13, Ch. VI]). Moreover, every linear functional $y^* \in L_\infty[0,1]^*$ can be expressed uniquely as the sum of an "absolutely continuous" component y_a^* and a "singular" component y_s^*. The "absolutely continuous" component is representable in the form $y_a^*(y) = \int_0^1 y(t)\eta(t)dt$ for some $\eta \in L_1[0,1]$, and the "singular" linear functional y_s^* has the property that the interval $[0,1]$ can be represented as the union of an increasing sequence of Borel sets S_k (i.e., $S_1 \subset S_2 \subset ... \subset [0,1]$ and $\cup_k S_k = [0,1]$) such that for every k and any function $y \in L_\infty[0,1]$ vanishing almost everywhere outside of S_k it follows that $y_s^*(y) = 0$. We denote by $\mathcal{S}[0,1]$ the subspace of $L_\infty[0,1]^*$ formed by singular functionals. By the above discussion we have that $L_1[0,1]$ can be identified with a closed subspace of $L_\infty[0,1]^*$ and this subspace has a natural complement formed by the linear space $\mathcal{S}[0,1]$ of singular functionals. The decomposition $L_\infty[0,1]^* = L_1[0,1] + \mathcal{S}[0,1]$ was used by several authors in studies of duality properties of optimization problems (see [21] and references therein). We identify a functional $y^* \in \mathcal{S}[0,1]$ with the corresponding "singular" finite additive measure and denote by $\mathcal{S}^+[0,1]$ the cone of nonnegative "singular" finite additive measures.

We make the following assumptions about data of the (CLP) problem. We assume that all components of $h(t)$ and $H(t)$ belong to $L_\infty[0,1]$, i.e., are essentially bounded measurable functions, all components of $c(t)$ belong to $L_1[0,1]$, and all components of $a(t)$ belong to $C[0,1]$. We also assume that each component of $M(s,t)$ is a continuous function such that for every $s \in [0,1]$ it is of bounded variation as a function of $t \in [0,1]$. We need these assumptions about $M(s,t)$ in order to use integration by parts later on.

We can view then the (CLP) problem as a conic linear problem, of the form (1.1), if we define the spaces

$$X := L_\infty[0,1]^{n_1} \;\text{ and }\; Y := C[0,1]^{n_2} \times L_\infty[0,1]^{n_3} \tag{5.5}$$

(i.e., X is given by the Cartesian product of n_1 copies of the space $L_\infty[0,1]$), and the cones $C \subset X$ and $K \subset Y$, and the linear mapping $A : X \to Y$ and $b \in Y$, by

$$C := (L_\infty^+[0,1])^{n_1} \;\text{ and }\; K := (C^+[0,1])^{n_2} \times (L_\infty^+[0,1])^{n_3}, \tag{5.6}$$

$$(Ax)(t) := \left(\int_0^t M(s,t)x(s)ds, H(t)x(t) \right) \text{ and } b(t) := (-a(t), -h(t)). \tag{5.7}$$

Note that the above spaces X and Y, equipped with corresponding product (say max) norms, become Banach spaces and that these Banach spaces are not reflexive. We also consider spaces

$$X' := L_1[0,1]^{n_1} \text{ and } Y' := (C[0,1]^*)^{n_2} \times L_1[0,1]^{n_3}. \tag{5.8}$$

The space X is dual of the Banach space X'. Therefore X' and X can be viewed as paired spaces equipped with the weak (or strong) and weak* topologies, respectively. The dual of the Banach space Y is the space

$$Y^* := (C[0,1]^*)^{n_2} \times (L_\infty[0,1]^*)^{n_3}. \tag{5.9}$$

Therefore Y and Y^* can be viewed as paired spaces equipped with the strong and weak* topologies, respectively. Since $L_1[0,1]$ can be isometrically embedded into its bidual $L_\infty[0,1]^*$, the Banach space Y' can be isometrically embedded into the Banach space Y^*, and hence we can view Y' as a subspace of Y^*.

The spaces $L_1[0,1]$ and $L_\infty[0,1]$ can be paired by using their respective weak and weak* topologies, and hence Y and Y' can be paired by equipping them with respective products of paired topologies. The above discussion suggests that we can construct a dual of the (CLP) problem either in Y^* or Y' spaces. From the point of view of the duality theory developed in section 2 it will be preferable to work in the space Y^* which is the dual of the Banach space Y. However, it is inconvenient to deal with finite additive measures of the space $L_\infty[0,1]^*$. Therefore, we first consider the pairing Y and Y'.

Let us calculate now the dual of the (CLP) problem, with respect to the paired spaces X,X' and Y,Y'. The corresponding Lagrangian can be written as follows

$$\begin{aligned} L(x,\pi,\eta) &= \int_0^1 c(t)^T x(t)dt - \int_0^1 \left[\int_0^t M(s,t)x(s)ds - a(t) \right]^T d\pi(t) \\ &\quad - \int_0^1 \eta(t)^T [H(t)x(t) - h(t)]dt, \end{aligned} \tag{5.10}$$

where $x \in L_\infty[0,1]^{n_1}$, $\eta \in L_1[0,1]^{n_3}$ and $\pi \in (C[0,1]^*)^{n_2}$, i.e., each component of π is a right continuous function of bounded variation on $[0,1]$, normalized by $\pi(1) = 0$, and the corresponding integral is understood to be a Lebesgue-Stieltjes integral.

By interchanging order of integration and using integration by parts we obtain

$$\begin{aligned} \int_0^1 \left[\int_0^t M(s,t)x(s)ds \right]^T d\pi(t) = \int_0^1 x(t)^T \int_t^1 M(t,s)^T d\pi(s)dt = \\ - \int_0^1 x(t)^T \left(M(t,t)^T \pi(t) + \int_t^1 [dM(t,s)]^T \pi(s) \right) dt, \end{aligned} \tag{5.11}$$

and hence

$$L(x,\pi,\eta) = \int_0^1 \eta(t)^T h(t)dt + \int_0^1 a(t)^T d\pi(t) + \int_0^1 g_{\pi,\eta}(t)^T x(t)dt, \tag{5.12}$$

where

$$g_{\pi,\eta}(t) := c(t) + M(t,t)^T\pi(t) + \int_t^1 [dM(t,s)]^T\pi(s) - H(t)^T\eta(t) \tag{5.13}$$

and the components $\sum_{i=1}^{n_2}\int_t^1 \pi_i(s)dM_{ij}(t,s)$, $j = 1,...,n_1$, of the integral term $\int_t^1 [dM(t,s)]^T\pi(s)$ are understood as Lebesgue-Stieltjes integrals with respect to $M(t,\cdot)$ Note that if for every $t \in [0,1]$ the function $M(t,\cdot)$ is constant on $[0,1]$, then $\int_t^1 \pi(s)^T dM(t,s) = 0$ and hence the corresponding integral term can be deleted.

It follows that the (Lagrangian) dual of the (CLP) problem can be written in the form

$$\begin{aligned} \operatorname{Max}_{(\pi,\eta)\in Y'} \quad & \int_0^1 a(t)^T d\pi(t) + \int_0^1 h(t)^T\eta(t)dt & (5.14)\\ \text{subject to} \quad & g_{\pi,\eta}(t) \geq 0, \;\; \text{a.e. } t \in [0,1], & (5.15)\\ & \eta(t) \geq 0, \;\; \text{a.e. } t \in [0,1], \quad \pi \succeq 0, & (5.16) \end{aligned}$$

where function $g_{\pi,\eta}(t)$ is defined in (5.13) and $\pi \succeq 0$ means that all components of $\pi(t)$ are monotonically nondecreasing on $[0,1]$ or, equivalently, that the corresponding measures are nonnegative valued. We denote the above dual problem (5.14)-(5.16) by (CLP'). For (CLP) problems with constant matrices $M(s,t)$ and $H(t)$, the dual problem (CLP') was suggested by Pullan [17],[18].

The optimal value of the (CLP) problem is always greater than or equal to the optimal value of its Lagrangian dual (5.14)-(5.16), i.e., val$(CLP) \geq$ val(CLP'). Moreover, we have by proposition 2.1 that if val$(CLP) =$ val(CLP') and (CLP) and (CLP') have optimal solutions $\bar{x}$ and $(\bar{\pi},\bar{\eta})$, respectively, then the following complementarity conditions hold

$$\int_0^1 g_{\pi,\eta}(t)^T\bar{x}(t)dt = 0, \tag{5.17}$$

$$\int_0^1 \left[\int_0^t M(s,t)\bar{x}(s)ds - a(t)\right]^T d\bar{\pi}(t) = 0, \tag{5.18}$$

$$\int_0^1 \bar{\eta}(t)^T\left[H(t)\bar{x}(t) - h(t)\right]dt = 0. \tag{5.19}$$

Conversely if the above complementarity conditions hold for some feasible $\bar{x}$ and $(\bar{\pi},\bar{\eta})$, then val$(CLP) =$ val(CLP') and $\bar{x}$ and $(\bar{\pi},\bar{\eta})$ are optimal solutions of (CLP) and (CLP'), respectively. These complementarity conditions can be written in the following equivalent form (cf., [17])

$$g_{\pi,\eta}(t)^T\bar{x}(t) = 0, \;\; \text{a.e. } t \in [0,1], \tag{5.20}$$

$$\int_0^t M(s,t)\bar{x}(s)ds - a(t) = 0, \;\; t \in \operatorname{supp}(\pi), \tag{5.21}$$

$$\bar{\eta}(t)^T\left[H(t)\bar{x}(t) - h(t)\right] = 0, \;\; \text{a.e. } t \in [0,1], \tag{5.22}$$

where supp(π) denotes the support of the measure defined by the function π.

Let us consider now duality relations with respect to the pairing of X with $X^* := (L_\infty[0,1]^*)^{n_1}$, and Y with Y^*. The spaces X^* and Y^* are dual of the respective Banach spaces X and Y. Therefore, we equip X and Y with their strong (norm) topologies and X^* and Y^* with their weak* topologies. We denote by (CLP^*) the optimization problem which is the Lagrangian dual of (CLP) with respect to such pairing. An explicit form of (CLP^*) and a relation between the dual problems (CLP^*) and (CLP') will be discussed in a moment. Let us remark that since (CLP^*) is a Lagrangian dual of (CLP) we have that $\text{val}(CLP) \geq \text{val}(CLP^*)$, and since X' and Y' can be identified with closed subspaces of X^* and Y^*, respectively, we have that $\text{val}(CLP^*) \geq \text{val}(CLP')$, and hence

$$\text{val}(CLP) \geq \text{val}(CLP^*) \geq \text{val}(CLP'). \tag{5.23}$$

It follows that there is no duality gap between the problems (CLP) and (CLP') if and only if there is no duality gap between the problems (CLP) and (CLP^*) and $\text{val}(CLP') = \text{val}(CLP^*)$. Therefore, the question of whether $\text{val}(CLP) = \text{val}(CLP')$ can be separated into two questions of whether $\text{val}(CLP) = \text{val}(CLP^*)$ and $\text{val}(CLP') = \text{val}(CLP^*)$.

Let us observe that the constraints (5.4) can be absorbed into the constraints (5.3) by adding to $H(t)$ rows of the $n_1 \times n_1$ identity matrix and adding n_1 zero components to $h(t)$, i.e., by replacing $H(t)$ and $h(t)$ with $\bar{H}(t) := \begin{bmatrix} H(t) \\ I \end{bmatrix}$ and $\bar{h}(t) := \begin{bmatrix} h(t) \\ 0 \end{bmatrix}$, respectively. This will not change the corresponding dual problem (see the discussion of section 2 about the dual problem (2.5)). Consider the Lagrangian $L^*(x, \pi, \eta, \sigma, \eta', \sigma')$ of the obtained (CLP) problem with respect to the pairing of X,X^* and Y,Y^*. Here $\sigma \in \mathcal{S}[0,1]^{n_3}$ represents multipliers associated with singular measures of $(L_\infty[0,1]^*)^{n_3}$, and $\eta' \in L_1[0,1]^{n_1}$ and $\sigma' \in \mathcal{S}[0,1]^{n_1}$ are multipliers corresponding to the constraints (5.4). Compared with the Lagrangian $L(x, \pi, \eta)$, given in (5.10) and (5.12), the above Lagrangian can be written as follows

$$\begin{aligned} L^*(x, \pi, \eta, \sigma, \eta', \sigma') &= L(x, \pi, \eta) - \int_0^1 [H(t)x(t) - h(t)]^T d\sigma(t) \\ &\quad - \int_0^1 x(t)^T \eta'(t) dt - \int_0^1 x(t)^T d\sigma'(t). \end{aligned} \tag{5.24}$$

The dual problem (CLP^*) is obtained by minimizing $L^*(x, \pi, \eta, \sigma, \eta', \sigma')$ with respect to $x \in X$ and then maximizing with respect to $\pi \succeq 0$, $\eta \geq 0$, $\sigma \succeq 0$, $\eta' \geq 0$, $\sigma' \succeq 0$.

Since $g_{\pi,\eta} \in L_1[0,1]^{n_1}$ and the space $L_\infty[0,1]^*$ is the direct sum of the spaces $L_1[0,1]$ and $\mathcal{S}[0,1]$, it follows that the dual problem (CLP^*) is obtained from (CLP') by adding to it the term

$$\underset{\sigma \in \mathcal{S}^+[0,1]^{n_3}}{\text{Max}} \int_0^1 h(t)^T d\sigma(t) \ \text{ subject to } \ -H(t) d\sigma(t) \succeq 0. \tag{5.25}$$

Note that the above maximum is either zero or $+\infty$. This leads to the following result.

Proposition 5.1 *Suppose that the problem* (CLP) *is consistent. Then* $\text{val}(CLP') = \text{val}(CLP^*)$ *and* $\text{Sol}(CLP') = \text{Sol}(CLP^*)$.

Proof. Since (CLP) is consistent we have that $\text{val}(CLP) < +\infty$ and since $\text{val}(CLP^*) \leq \text{val}(CLP)$ it follows that $\text{val}(CLP^*) < +\infty$. Consequently the additional term given by the maximum (5.25) is zero. The result then follows by the above discussion. ∎

The above proposition shows that an investigation of duality relations between the problems (CLP) and (CLP') can be reduced to a study of the problems (CLP) and (CLP^*). Let us discuss now the question of "no duality gap" between the problems (CLP) and (CLP^*).

Let $v(y)$ be the optimal value of the problem (CLP). That is, $y = (y_1, y_2)$, $y_1 \in C[0,1]^{n_2}$, $y_2 \in L_\infty[0,1]^{n_3}$ and $v(y)$ is the optimal value of the problem (5.1)-(5.4) with $a(\cdot)$ and $h(\cdot)$ replaced by $y_1(\cdot)$ and $y_2(\cdot)$, respectively. We have by proposition 2.3 that if the problem (CLP) is subconsistent (in particular, consistent), then $\mathrm{val}(CLP) = \mathrm{val}(CLP^*)$ iff $v(y)$ is lower semicontinuous at $y = (a, h)$.

The cone K has a nonempty interior in the Banach space Y. This implies that the optimal value function $v(y)$ is bounded from above on an open subset of Y, i.e., that condition (A3) of section 2 holds, and that the constraint qualification (2.20) is equivalent to the (generalized) Slater condition (2.21). The interior of the cone $C^+[0,1]$ is formed by functions $\phi \in C[0,1]$ such that $\phi(t) > 0$ for all $t \in [0,1]$, and the interior of the cone $L^+_\infty[0,1]$ is formed by functions $\psi \in L_\infty[0,1]$ such that $\psi(t) \geq \varepsilon$ for a.e. $t \in [0,1]$ and some $\varepsilon > 0$. Therefore the (generalized) Slater condition for the (CLP) problem can be formulated as follows.

(Slater condition) There exists $\bar{x} \in (L^+_\infty[0,1])^{n_1}$ and $\varepsilon > 0$ such that:

$$\int_0^t M(s,t)\bar{x}(s)ds > a(t), \ \forall\, t \in [0,1], \tag{5.26}$$

$$H(t)\bar{x}(t) \geq h(t) + \varepsilon, \ \ \text{a.e. } t \in [0,1]. \tag{5.27}$$

Clearly the above Slater condition implies that the problem (CLP) is consistent. The above discussion together with propositions 2.8 and 5.1 imply the following results.

Proposition 5.2 *Suppose that the (generalized) Slater condition* (5.26) - (5.27) *holds. Then* $\mathrm{val}(CLP) = \mathrm{val}(CLP')$, *and moreover if* $\mathrm{val}(CLP)$ *is finite, then the set of optimal solutions of the problem* (CLP') *is nonempty and bounded (in the norm topology of* Y^**). Conversely if the set of optimal solutions of the problem* (CLP') *is nonempty and bounded, then the (generalized) Slater condition holds.*

In some cases (CLP) problems involve equality type constraints. For example, some (all) constraints in (5.2) and/or (5.3) can be equality constraints. Of course, the equality constraints such as $H(t)x(t) = h(t)$ can be split into the inequality constraints $H(t)x(t) \geq h(t)$ and $H(t)x(t) \leq h(t)$. Note, however, that for such split inequality constraints the (generalized) Slater condition can never be satisfied. Nevertheless, the assertion of the above proposition 5.2 still holds if the (generalized) Slater condition is replaced by the regularity condition (2.19), or its equivalent (2.20). Recall that (2.19) and (2.20) are equivalent to conditions (2.23) which can be written in a form similar to (5.26)-(5.27).

In the present case, since the cone K has a nonempty interior in the Banach space Y, Slater condition (5.26)-(5.27) is equivalent to continuity (in the norm topology of Y) of the optimal value function $v(y)$ at $y = (a, h)$. We show now that a variant of the inf-compactness condition is sufficient for lower semicontinuity of the optimal value function. Recall that the Banach space $X = L_\infty[0,1]^{n_1}$ is dual of the Banach space $L_1[0,1]^{n_1}$ and therefore X can be equipped with the corresponding weak* topology. Moreover, if the feasible set of the (CLP) problem is nonempty and bounded in the norm topology of X, then it is compact in the weak* topology of X, and hence the problem (CLP) possesses an optimal solution. Note, however, that in the present case we cannot apply the result of proposition 2.4 in a direct way since the linear mapping $A : X \to Y$, defined in (5.7), is not continuous with respect to the weak* topology of X and norm topology of Y.

Let us consider the following condition. For $y \in Y$, with $y = (y_1, y_2)$, $y_1 \in C[0,1]^{n_2}$, $y_2 \in L_\infty[0,1]^{n_3}$, denote by $\Phi(y)$ the set of all $x \in X$ satisfying the feasibility constraints (5.2)-(5.4) with functions $a(\cdot)$ and $h(\cdot)$ replaced by $y_1(\cdot)$ and $y_2(\cdot)$, respectively. In that notation the set $\Phi(a, h)$ denotes the feasible set of the (CLP) problem.

(C1) The feasible set of the (CLP) problem is nonempty and there exists $\varepsilon > 0$ such that the sets $\Phi(y)$ are uniformly bounded for all $y \in Y$ satisfying $\|y - (a, h)\| < \varepsilon$.

Proposition 5.3 *Suppose that the assumption* (C1) *holds. Then the* (CLP) *problem has an optimal solution, the optimal value function* $v(y)$ *is lower semicontinuous at* $y = (a, h)$ *and* $\text{val}(CLP) = \text{val}(CLP')$.

Proof. Since assumption (C1) implies that the feasible set of the (CLP) problem is nonempty and bounded, it follows that (CLP) has an optimal solution.

By assumption (C1) we have that the problem (CLP) is consistent, and hence it follows by proposition 5.1 that $\text{val}(CLP') = \text{val}(CLP^*)$. Therefore in order to show that $\text{val}(CLP) = \text{val}(CLP')$ it suffices to verify that $v(y)$ is lower semicontinuous, in the norm topology of Y, at the point $y_0 := (a, h)$. If $\liminf_{y \to y_0} v(y) = +\infty$, then clearly $v(y)$ is lower semicontinuous at y_0. Therefore we can assume that $\liminf_{y \to y_0} v(y) < +\infty$. Consider a sequence $y_n \in Y$ along which the lower limit of $v(y)$ is attained as y tends to y_0. By condition (C1) we can assume that for every $n \in \mathbb{N}$ the corresponding feasible set $\Phi(y_n)$ is bounded and, since $\liminf_{y \to y_0} v(y) < +\infty$, is nonempty. It follows that the associated optimization problem attains its optimal value at a point $x_n \in X$, i.e., $x_n \in \Phi(y_n)$ and $v(y_n) = \langle c, x_n \rangle$. Again by assumption (C1) we have that the sequence $\{x_n\}$ is bounded in X. Therefore, by passing to a subsequence if necessary, we can assume that $\{x_n\}$ converges in the weak* topology of X to a point x_0. It follows that x_0 is a feasible point of the (CLP) problem. Consequently we obtain

$$\liminf_{y \to y_0} v(y) = \lim_{n \to \infty} v(y_n) = \langle c, x_0 \rangle \geq \text{val}(CLP) = v(y_0),$$

which proves lower semicontinuity of $v(y)$ at y_0. ∎

The above condition (C1) assumes that the feasible sets of the corresponding continuous linear programming problems are uniformly bounded for small (with respect to the sup-norm) perturbations of the right hand side functions. Condition (C1) is slightly stronger than the condition that the feasible set of the (CLP) problem is nonempty and bounded. The later condition was used in Pullan [17],[18] for deriving the property val(CLP) = val(CLP') under the additional conditions that the matrices $M(s,t)$ and $H(t)$ are constant (i.e., independent of s and t) and the functions $c(t)$,$a(t)$ and $h(t)$ are piecewise analytic.

Let us consider the following condition.

(C2) The feasible set of the (CLP) problem is nonempty, the matrix $H = H(t)$ is constant and the set

$$\{x \in X : Hx(t) \geq h(t) \ \text{ and } \ x(t) \geq 0, \ \text{ a.e. } t \in [0,1]\} \tag{5.28}$$

is bounded in X.

Note that the above set (5.28) is the set defined by constraints (5.3)-(5.4).

Proposition 5.4 *Assumption* (C2) *implies assumption* (C1).

Proof. For $b \in I\!R^{n_3}$ consider the set

$$\Psi(b) := \{x \in I\!R^{n_1} : Hx \geq b, \ x \geq 0\}.$$

Suppose that for some $b_0 \in I\!R^{n_3}$ the set $\Psi(b_0)$ is nonempty. Then by Hoffman's lemma [10] there exists constant $\kappa > 0$, depending on H, such that

$$\Psi(b) \subset \Psi(b_0) + \kappa \|b - b_0\| B, \ \ \forall\, b \in I\!R^{n_3}, \tag{5.29}$$

where $B := \{x \in I\!R^{n_1} : \|x\| \leq 1\}$. Assumption (C2) implies that for almost all $t \in [0,1]$ the sets $\Psi(h(t))$ are nonempty and uniformly bounded. It follows then by (5.29) that for any $\varepsilon > 0$ the subsets of X defined by constraints (5.3)-(5.4) with $h(t)$ replaced by $y_2(t)$ satisfying $\|y_2 - h\| \leq \varepsilon$, are uniformly bounded. This implies that the feasible sets $\Phi(y)$ are uniformly bounded for all $y \in Y$ such that $\|y_2 - h\|$ is bounded. This shows that assumption (C2) implies assumption (C1). ∎

By propositions 5.3 and 5.4 we obtain that if assumption (C2) holds, then val(CLP) = val(CLP') and Sol(CLP) is nonempty.

References

[1] Anderson, E. J. and Nash, P., *Linear Programming in Infinite-Dimensional Spaces. Theory and Applications*, Wiley, Chichester, 1987.

[2] Bellman, R.E., Bottleneck problems and dynamic programming, *Proc. Nat. Acad. Sci. USA*, 39, 947-951, 1953.

[3] Billingsley, P., *Convergence of Probability Measures*, Wiley, New York, 1999 (Second Edition).

[4] Bonnans, J.F. and Shapiro, A., *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[5] Ekeland, I. and Temam, R., *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam, 1976.

[6] Glashoff, K. and Gustafson, S.A., *Linear Optimization and Approximation*, Springer, New York, 1983.

[7] Goberna, M. A. and López, M. A., *Linear Semi-Infinite Optimization*, Wiley, Chichester, 1998.

[8] Grinold, R.C., Symmetric duality for continuous linear programs, *SIAM J. Appl. Math.,* 18, 84-97, 1970.

[9] Hettich, R. and Kortanek, K.O., Semi-infinite programming: theory, methods and applications, *SIAM Review*, 35, 380-429, 1993.

[10] Hoffman, A., On approximate solutions of systems of linear inequalities, *Journal of Research of the National Bureau of Standards, Section B, Mathematical Sciences*, 49, 263-265, 1952.

[11] Holmes, R.B., *Geometric Functional Analysis and its Applications*, Springer, New york, 1975.

[12] Isii, K., On sharpness of Tchebycheff type inequalities, *Ann. Inst. Stat. Math.*, 14, 185-197, 1963.

[13] Kantorovich, I.V. and Akilov, G.P., *Functional Analysis*, Nauka, Moscow, 1984.

[14] Kemperman, J.H.B, The general moment problem, a geometric approach, *Annals Math. Statist.*, 39, 93-122, 1968.

[15] Levinson, N., A class of continuous linear programming problems, *J. Math. Anal. Appl.,* 16, 78-83, 1966.

[16] *Moments in mathematics*, H.J. Landau (ed.), Proc. Sympos. Appl. Math., 37, Amer. Math. Soc., Providence, RI, 1987.

[17] Pullan, M. C., A duality theory for separated continuous linear programs, *SIAM J. Control Optim.*, 34, 931-965, 1996.

[18] Pullan, M. C., Existence and duality theory for separated continuous linear programs, *Math. Model. Systems*, 3, 219-245, 1997.

[19] Robinson, S.M., Regularity and stability for convex multivalued functions, *Mathematics of Operations Research*, 1, 130-143, 1976.

[20] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[21] Rockafellar, R. T. Integrals which are convex functionals. II, *Pacific J. Math.,* 39, 439-469, 1971.

[22] Rockafellar, R.T., *Conjugate Duality and Optimization*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1974.

[23] Rogosinsky, W.W., Moments of non-negative mass, *Proc. Roy. Soc. London Ser. A*, 245, 1-27, 1958.

[24] Smith, J.E., Generalized Chebychev inequalities: theory and applications in decision analysis, *Operations Research*, 43, 807-825, 1995.