On Linear Programming Duality and Necessary and Sufficient Conditions in Minimax Theory

J.B.G. Frenk · P. Kas · G. Kassay

Abstract In this paper we discuss necessary and sufficient conditions for different minimax results to hold using only linear programming duality and the finite intersection property for compact sets. It turns out that these necessary and sufficient conditions have a clear interpretation within zero-sum game theory. We apply these results to derive necessary and sufficient conditions for strong duality for a general class of optimization problems.

Keywords Minimax theory · Game theory · Lagrangian duality · Linear programming duality

1 Introduction

Let A and B be nonempty sets and $f : A \times B \to \mathbb{R}$ a given function. In this paper we consider Borel probability measures on A and B and we assume without loss of generality that A and B are topological spaces with Borel σ-algebras $\mathfrak{A}$ and $\mathfrak{B}$. A minimax result for the function f defined on $A \times B$ is a theorem which asserts that

$$\inf_{b \in B} \sup_{a \in A} f(a, b) = \sup_{a \in A} \inf_{b \in B} f(a, b). \tag{1}$$

J.B.G. Frenk (✉)
Econometric Institute, Erasmus University, Rotterdam, Netherlands
e-mail: frenk@few.eur.nl

P. Kas
Department of Mathematics, Eastern Mediterranean University, TRNC, Gazimagusa, Turkey

G. Kassay
Faculty of Mathematics and Computer Science, Babes-Bolyai University, Cluj, Romania

Since relation (1) is important in game theory and optimization a lot of papers have appeared in the literature (for a survey see Refs. [1, 2]) introducing different proof techniques in combination with sufficient conditions on the function f and the sets A and B to verify relation (1). Among the used techniques are fixed point theorems, techniques from topology (cf. Refs. [3, 4]), and versions of the Hahn–Banach theorem in finite/infinite dimensional topological vector spaces. The purpose of this paper is to derive necessary and sufficient conditions on the function f and the sets A and B for the above and related minimax results. It turns out for the proof of these necessary and sufficient conditions that we only need the strong duality theorem of linear programming (cf. Ref. [5]) and some standard result on compact sets and lower semicontinuous functions. To introduce the other minimax results we first define the notion of a mixed strategy. For the set A let $\mathcal{P}_F(A)$ denote the convex set of all probability measures on A with finite support. If ϵ_a represents the one-point probability measure concentrated on the point $a \in A$, this means that λ belongs to $\mathcal{P}_F(A)$ if and only if there exists some finite set $\{a_1, \ldots, a_m\} \subseteq A$ and a vector $s(\lambda) := (s_1(\lambda), \ldots, s_m(\lambda))$ satisfying

$$\lambda = \sum_{i=1}^{m} s_i(\lambda)\epsilon_{a_i}, \quad \sum_{i=1}^{m} s_i(\lambda) = 1 \text{ and } s_i(\lambda) > 0, \ 1 \le i \le m. \tag{2}$$

Within game theory (cf. Ref. [6]) the set $\mathcal{P}_F(A)$ is known as the set of mixed strategies available to a player having set A as its set of pure strategies. A larger set of strategies (cf. Ref. [7]) is given by the convex set $\mathcal{P}(A)$ of Borel probability measures on A. To extend the minimax result involving the pure strategy sets A and B to the strategy sets $\mathcal{P}(A)$ and $\mathcal{P}(B)$ we first extend the function f to the larger domain $\mathcal{P}_F(A) \times \mathcal{P}_F(B)$. Therefore, if $\mu \in \mathcal{P}_F(B)$ is represented by

$$\mu = \sum_{j=1}^{p} s_j(\mu)\epsilon_{b_j}, \quad \sum_{j=1}^{p} s_j(\mu) = 1 \text{ and } s_j(\mu) > 0, \ 1 \le j \le p \tag{3}$$

with $\{b_1, \ldots, b_p\} \subseteq B$, then the extension $f_e : \mathcal{P}_F(A) \times \mathcal{P}_F(B) \to \mathbb{R}$ of the function f is given by

$$f_e(\lambda, \mu) := \sum_{i=1}^{m} \sum_{j=1}^{p} s_i(\lambda) s_j(\mu) f(a_i, b_j) \tag{4}$$

with $\lambda \in \mathcal{P}_F(A)$ and $\mu \in \mathcal{P}_F(B)$ listed in relations (2) and (3). To extend the function f to the larger domain $\mathcal{P}_F(A) \times \mathcal{P}(B)$ we always assume that the function $f(a, .) : B \to \mathbb{R}$ belongs for every $\mu \in \mathcal{P}(B)$ and $a \in A$ to the set $\mathcal{L}^1_\mu(B)$ of Borel measurable functions on B, which are Lebesgue absolutely integrable with respect to μ. The function $f_e : \mathcal{P}_F(A) \times \mathcal{P}(B) \to \mathbb{R}$ is now defined by

$$f_e(\lambda, \mu) := \sum_{i=1}^{m} s_i(\lambda) \int_B f(a_i, b) d\mu(b) \tag{5}$$

with λ represented by relation (2). Finally, if we extend the function f to the largest domain $\mathcal{P}(A) \times \mathcal{P}(B)$, we assume that the function f belongs for every $\mu \in \mathcal{P}(B)$

and $\lambda \in \mathcal{P}(A)$ to the set $\mathcal{L}^1_{\lambda \otimes \mu}(A \times B)$ of Borel measurable functions on $A \times B$ (measurable with respect to the product σ-algebra $\mathfrak{A} \otimes \mathfrak{B}$), which are Lebesgue absolutely integrable with respect to the Borel product measure $\lambda \otimes \mu$. In this case it is well-known for f belonging to $\mathcal{L}^1_{\lambda \otimes \mu}(A \times B)$ that the Fubini theorem holds (cf. Refs. [8, 9]) and so it follows that

$$\int_{A \times B} f \, d(\lambda \otimes \mu) = \int_A \int_B f \, d\mu \, d\lambda = \int_B \int_A f \, d\lambda \, d\mu. \tag{6}$$

The function $f_e : \mathcal{P}(A) \times \mathcal{P}(B) \to \mathbb{R}$ is now defined by

$$f_e(\lambda, \mu) := \int_{A \times B} f \, d(\lambda \otimes \mu). \tag{7}$$

By relation (6) it follows that the function f_e is convex and concave in both arguments. Consider now the following different minimax results given by:

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{\lambda \in \mathcal{P}(A)} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}(A)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu), \tag{8}$$

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu), \tag{9}$$

$$\inf_{\mu \in \mathcal{P}_F(B)} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu), \tag{10}$$

$$\inf_{b \in B} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \epsilon_b) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b), \tag{11}$$

$$\inf_{b \in B} \sup_{a \in A} f(a, b) = \sup_{a \in A} \inf_{b \in B} f(a, b). \tag{12}$$

In the next section it will be verified that the minimax results considered in the above relations satisfy the strict inclusions (12) $\Rightarrow$ (11) $\Rightarrow$ (10) $\Rightarrow$ (9) $\Rightarrow$ (8). In this paper we derive in Sect. 2 for relations (9) up to (12) a necessary and sufficient condition on the function f and the sets A and B. In Sect. 3 we apply the minimax results of Sect. 2 to derive results for the special case of Lagrangian duality in optimization.

2 On Minimax Results and LP Duality

Let $\mathcal{F}(A_0)$ denote the set of all finite subsets of the set $A_0 \subseteq A$ and $\mathcal{P}(J)$ with $J \in \mathcal{F}(A_0)$ the set of all probability measures on J. Clearly the set $\mathcal{P}(J)$ is convex and in particular $\mathcal{P}(J) = \mathrm{co}(\{\epsilon_a\}_{a \in J})$, with $\mathrm{co}(C)$ the convex hull of a set C. Also for $A_0 \subseteq A$ we obtain

$$\mathcal{P}_F(A_0) = \mathrm{co}(\{\epsilon_a\}_{a \in A_0}) = \bigcup_{J \in \mathcal{F}(A_0)} \mathcal{P}(J), \tag{13}$$

and this set is again convex. Hence by the convexity of $\lambda \mapsto f_e(\lambda, \mu)$ it follows that

$$\max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) = \max_{a \in J} f_e(\epsilon_a, \mu) \tag{14}$$

for every $\mu \in \mathcal{P}_F(B)$ and by the concavity of $\mu \mapsto f_e(\lambda, \mu)$ that

$$\inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu) = \inf_{b \in B} f_e(\lambda, \epsilon_b) \tag{15}$$

for every $\lambda \in \mathcal{P}(J)$. To derive necessary and sufficient conditions for the different minimax results we use the well-known minimax theorem of Wald (cf. Ref. [10]). In this paper an alternative proof of this result is given by means of the strong duality theorem of linear programming and a well-known result on so-called inf-compact functions. We recall that a function $k : B \to \mathbb{R}$ is called inf-compact if all its lower level sets $\{b \in B : k(b) \leq r\}, r \in \mathbb{R}$ are compact and it is called sup-compact if the function $-k$ is inf-compact (cf. Ref. [11]). For inf-compact functions the next result is well-known (cf. Ref. [11]).

Lemma 2.1 *If the function $b \mapsto f(a, b)$ is lower semicontinuous for every $a \in A$ and there exists some set $J_0 \in F(A)$ such that the function $b \mapsto \max_{a \in J_0} f(a, b)$ is inf-compact, then*

$$\sup_{J \in \mathcal{F}(A)} \inf_{b \in B} \max_{a \in J} f(a, b) = \inf_{b \in B} \sup_{a \in A} f(a, b)$$

and in both expressions we may replace inf by min.

A symmetrical version of Lemma 2.1 is given by

$$\inf_{I \in \mathcal{F}(B)} \sup_{a \in A} \min_{b \in I} f(a, b) = \sup_{a \in A} \inf_{b \in B} f(a, b) \tag{16}$$

and this holds if the function $a \mapsto f(a, b)$ is upper semicontinuous for every $b \in B$ and there exists some set $I_0 \in \mathcal{F}(B)$ such that the function $a \mapsto \min_{b \in I_0} f(a, b)$ is sup-compact. Also in relation (16) we may replace sup by max. It follows (cf. Ref. [12]) that the conditions of Lemma 2.1 are satisfied if the set B is compact and the function $b \mapsto f(a, b)$ is lower semicontinuous for every $a \in A$. Since we need Lemma 2.1 we also introduce on the set $\mathcal{P}(A)$ the weak*-topology (cf. Ref. [13]). Using relation (16) and the strong duality theorem of linear programming the following proof of Wald's minimax theorem can be given.

Theorem 2.1 (Wald's Minimax Theorem) *For every set J belonging to $\mathcal{F}(A)$*

$$\inf_{\mu \in \mathcal{P}_F(B)} \max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) = \max_{\lambda \in \mathcal{P}(J)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu).$$

Proof By relations (13) (replace A_0 by B) and (14) it follows that

$$\inf_{\mu \in \mathcal{P}_F(B)} \max_{\lambda \in \mathcal{P}(J)} f_e(\lambda, \mu) = \inf_{I \in \mathcal{F}(B)} \min_{\mu \in \mathcal{P}(I)} \max_{a \in J} f_e(\epsilon_a, \mu). \tag{17}$$

The optimization problem $\min_{\mu \in \mathcal{P}(I)} \max_{a \in J} f_e(\epsilon_a, \mu)$ is a linear programming problem and by the strong duality theorem for linear programming (cf. Ref. [5]) we obtain von Neumann's minimax result (cf. Refs. [14, 15]) given by

$\min_{\mu\in\mathcal{P}(I)}\max_{a\in J} f_e(\epsilon_a,\mu)=\max_{\lambda\in\mathcal{P}(J)}\min_{b\in I} f_e(\lambda,\epsilon_b)$. Applying now relation (17) yields

$$\inf_{\mu\in\mathcal{P}_F(B)}\max_{\lambda\in\mathcal{P}(J)} f_e(\lambda,\mu)=\inf_{I\in\mathcal{F}(B)}\max_{\lambda\in\mathcal{P}(J)}\min_{b\in I} f_e(\lambda,\epsilon_b). \tag{18}$$

Moreover, since the set J is finite the set $\mathcal{P}(J)$ is compact in the weak*-topology (isomorphic with the $|J|$-dimensional unit simplex $\Delta_{|J|}$ and $\Delta_{|J|}\subseteq\mathbb{R}^{|J|}$ is compact in the Euclidean topology) and in the same topology $\lambda\mapsto f_e(\lambda,\epsilon_b)$ is continuous on $\mathcal{P}(J)$ for every $b\in B$. Hence we may apply relation (16) (replace A by $\mathcal{P}(J)$ and $f(a,b)$ by $f_e(\lambda,\epsilon_b)$) to the last expression in relation (18) and using now (15) shows the result. $\qquad\square$

If we impose in Wald's minimax theorem an additional integrability assumption on the function f the next corollary holds. For a more detailed proof see Lemma 5 of Ref. [16].

Corollary 2.1 *If the function $b\mapsto f(a,b)$ belongs to the set $\bigcap_{\mu\in\mathcal{P}(B)} L^1_\mu(B)$ for every $a\in A$, then one may replace in Theorem 2.1 everywhere the set $\mathcal{P}_F(B)$ by the set $\mathcal{P}(B)$ without changing any values.*

Proof Since $f_e(\lambda,\mu)\geq\inf_{b\in B} f_e(\lambda,\epsilon_b)$ the result can be derived using Theorem 2.1. $\qquad\square$

For further reference we list for any f satisfying the conditions of Corollary 2.1 that

$$\inf_{\mu\in\mathcal{P}_F(B)}\max_{a\in J} f_e(\epsilon_a,\mu)=\inf_{\mu\in\mathcal{P}(B)}\max_{a\in J} f_e(\epsilon_a,\mu) \tag{19}$$

for any $J\in\mathcal{F}(A)$. An immediate consequence of Theorem 2.1 and Corollary 2.1 is given by the next result.

Corollary 2.2 *For any function $f:A\times B\to\mathbb{R}$,*

$$\sup_{J\in\mathcal{F}(A)}\inf_{\mu\in\mathcal{P}_F(B)}\max_{a\in J} f_e(\epsilon_a,\mu)=\sup_{\lambda\in\mathcal{P}_F(A)}\inf_{\mu\in\mathcal{P}_F(B)} f_e(\lambda,\mu).$$

Moreover, if the function $b\mapsto f(a,b)$ belong to the set $\bigcap_{\mu\in\mathcal{P}(B)} L^1_\mu(B)$ for every $a\in A$, then we may replace in the above equality without changing any values the set $\mathcal{P}_F(B)$ by the set $\mathcal{P}(B)$.

Proof Apply relation (13) to Theorem 2.1, respectively Corollary 2.1. $\qquad\square$

We will now derive the different necessary and sufficient conditions for the minimax results in relation (9) up to (12). Starting with relation (9) we first introduce the following class of functions.

Definition 2.1 The function $f : A \times B \to R$ belongs to the set $\mathcal{A}$ if

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}(B)} \max_{a \in J} f_e(\epsilon_a, \mu) = \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) \qquad (20)$$

and the above expressions are well defined.

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{A}$ is given by the observation that for player 2 using strategy set $\mathcal{P}(B)$ and the minimax approach it does not make any difference whether his opponent given by player 1 selects a pure strategy from the set A or first considers all finite subsets of A and then selects from one of these finite subsets a pure strategy. In the following proposition it is shown that this property holds if both the set B and the function f satisfy some topological condition. For a detailed proof see Lemma 9 of Ref. [16].

Proposition 2.1 *If the set B is a compact Hausdorff space and the function $b \mapsto f(a, b)$ is both lower semicontinuous and belongs to the set $\bigcap_{\mu \in \mathcal{P}(B)} L^1_\mu(B)$ for every $a \in A$, then the function f belongs to the set $\mathcal{A}$.*

It is now easy to verify the following equivalent characterization of the minimax result in relation (9). Note that iff is an abbreviation of if and only if.

Theorem 2.2 *If the function $b \mapsto f(a, b)$ belongs to the set $\bigcap_{\mu \in \mathcal{P}(B)} L^1_\mu(B)$ for every $a \in A$, then relation (9) holds iff the function f belongs to the set $\mathcal{A}$.*

Proof Since the equality in relation (9) is the same as

$$\inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}(B)} f_e(\lambda, \mu),$$

the result follows by the second part of Corollary 2.2. □

To derive a natural necessary and sufficient condition for the minimax result in relation (10) we introduce the following class of functions.

Definition 2.2 The function $f : A \times B \to R$ belongs to the set $\mathcal{B}$ if

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{a \in J} f_e(\epsilon_a, \mu) = \inf_{\mu \in \mathcal{P}_F(B)} \sup_{a \in A} f_e(\epsilon_a, \mu). \qquad (21)$$

One can give a similar game theoretic interpretation as for the set $\mathcal{A}$ (replacing $\mathcal{P}(B)$ by $\mathcal{P}_F(B)$) for functions belonging to the set $\mathcal{B}$. If we know additionally that the set B is a compact Hausdorff space and the function $b \mapsto f(a, b)$ is both lower semicontinuous and belongs to $\bigcap_{\mu \in \mathcal{P}(B)} \mathcal{L}^1_\mu(B)$ for every $a \in A$, then the game theoretic interpretation of the set $\mathcal{B}$ can be simplified. If this holds we obtain by relation (19) and Proposition 2.1 that

$$f \in \mathcal{B} \Longleftrightarrow \inf_{\mu \in \mathcal{P}(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) = \inf_{\mu \in \mathcal{P}_F(B)} \sup_{a \in A} f_e(\epsilon_a, \mu). \qquad (22)$$

For this case the game theoretic interpretation of the set $\mathcal{B}$ is given by the observation that player 2 using strategy set $\mathcal{P}(B)$ can restrict himself to the strategy set $\mathcal{P}_F(B)$. However, it might be possible that the value for player 2 cannot be achieved if he restricts himself to the set $\mathcal{P}_F(B)$ of mixed strategies.

Theorem 2.3 *Relation* (10) *holds iff the function* f *belongs to the set* $\mathcal{B}$.

Proof Apply a similar proof as in Theorem 2.2 and use the first part of Corollary 2.2. □

The minimax result listed in relation (10) is of importance in the theory of zero-sum games. It states that both players should use the set of mixed strategies to achieve the (maybe not attainable) value of a zero-sum game. If the function f is continuous on $A \times B$ and the sets A and B are compact sets in a metric space Ville (cf. Refs. [2, 17]), showed that relation (10) holds. Applying the result that any continuous function on a compact set in a metric space is uniformly continuous (cf. Ref. [18]) it is easy to verify that the function f belongs to the set $\mathcal{B}$ and so Ville's minimax result follows from Theorem 2.3. To derive a necessary and sufficient condition for the minimax result in relation (11) we introduce the following class of functions.

Definition 2.3 The function $f : A \times B \to R$ belongs to the set $\mathcal{C}$ if

$$\sup_{J \in \mathcal{F}(A)} \inf_{\mu \in \mathcal{P}_F(B)} \max_{a \in J} f_e(\epsilon_a, \mu) = \inf_{b \in B} \sup_{a \in A} f(a, b). \tag{23}$$

The game theoretic interpretation of a function belonging to the set $\mathcal{C}$ is the same as the one for a function belonging to the set $\mathcal{B}$ with the additional observation that the mixed strategy set $\mathcal{P}_F(B)$ for player 2 is always dominated by his pure strategy set B. Again by relation (22) one can give an easier game theoretic interpretation of the set $\mathcal{C}$ for B a compact Hausdorff space and the function $b \mapsto f(a, b)$ both lower semicontinuous and belonging to $\bigcap_{\mu \in \mathcal{P}(B)} \mathcal{L}^1_\mu(B)$ for every $a \in A$. Under these topological conditions the game theoretic interpretation of the set $\mathcal{C}$ is given by the observation that player 2 using the strategy set $\mathcal{P}(B)$ can restrict himself to the pure strategy set B and this means that the function $b \mapsto f(a, b)$ satisfies some convexity-type condition for every $a \in A$. Hence a topological condition in combination with a convexity-type condition imply the property listed in Definition 2.3. Moreover, to determine the optimal strategy $b_{\mathrm{opt}} \in B$ player 2 must solve the optimization problem $\min_{b \in B} \sup_{a \in A} f(a, b)$. One can now show the following result. A sufficient condition for the listed minimax result was given in Ref. [19].

Theorem 2.4 *Relation* (11) *holds iff the function* f *belongs to the set* $\mathcal{C}$.

Proof Relation (11) is the same as

$$\inf_{b \in B} \sup_{a \in A} f(a, b) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{\mu \in \mathcal{P}_F(B)} f_e(\lambda, \mu).$$

Applying the first part of Corollary 2.2 yields the desired result. □

In the section on Lagrangian duality we need the following equivalent formulation of Theorem 2.4 given by

$$f \in \mathcal{C} \iff \inf_{b \in B} \sup_{a \in A} f(a, b) = \sup_{\lambda \in \mathcal{P}_F(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b). \tag{24}$$

Finally we derive a necessary and sufficient condition for a minimax result involving the pure strategy sets A and B.

Definition 2.4 The function $f : A \times B \to R$ belongs to the set $\mathcal{D}$ if

$$\sup_{\lambda \in \mathcal{P}_F(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) = \sup_{a \in A} \inf_{b \in B} f(a, b).$$

A game theoretic interpretation of the payoff function f belonging to the set $\mathcal{D}$ is given by the observation that for player 1 the mixed strategy set $\mathcal{P}_F(A)$ is always dominated by his pure strategy set. This means that player 1 can restrict himself to the set of pure strategies instead of using the set of mixed strategies and this means that the function $a \mapsto f(a, b)$ satisfy some concavity-type condition for every $b \in B$. One can now show the most useful minimax result.

Theorem 2.5 *Relation* (12) *holds iff the function* f *belongs to the set* $\mathcal{C} \cap \mathcal{D}$.

Proof If the function f belongs to the set $\mathcal{C} \cap \mathcal{D}$, then by Theorem 2.4 we obtain that

$$\inf_{b \in B} \sup_{a \in A} f(a, b) = \inf_{b \in B} \sup_{\lambda \in \mathcal{P}_F(A)} f_e(\lambda, \epsilon_b) = \sup_{a \in A} \inf_{b \in B} f(a, b).$$

To show the reverse implication consider an arbitrary λ belonging to $\mathcal{P}_F(A)$. By relation (13) there exists some $J_0 \in \mathcal{F}(A)$ such that $\lambda \in \mathcal{P}(J_0)$ and this implies $\inf_{b \in B} f_e(\lambda, \epsilon_b) \leq \inf_{b \in B} \sup_{a \in A} f(a, b)$. Applying this inequality and the minimax equality (12) (the reverse inequality trivially holds) we obtain

$$\sup_{\lambda \in \mathcal{P}_F(A)} \inf_{b \in B} f_e(\lambda, \epsilon_b) = \sup_{a \in A} \inf_{b \in B} f(a, b) \tag{25}$$

or the function f belongs to $\mathcal{D}$. Again by the minimax equality (12), relation (25) and Theorem 2.4 we obtain that the function f belongs to $\mathcal{C}$.   □

By Theorem 2.5 it follows immediately that

$$-f \in \mathcal{C} \cap \mathcal{D} \iff \sup_{b \in B} \inf_{a \in A} f(a, b) = \inf_{a \in A} \sup_{b \in B} f(a, b) \tag{26}$$

and this result will be needed in the section on Lagrangian duality. In the next subsection we will investigate in more detail the relations between the different minimax results.

2.1 On relations between different minimax results

In this subsection we investigate in more detail the relations between the different minimax results given by relations (8) up to (12). Introducing the notation L_i and R_i for the left and right-hand side of relation (i) for $8 \leq i \leq 12$ we obviously obtain that

$$L_{12} = L_{11} \geq L_{10} \geq L_9 = L_8 \geq R_8 \geq R_9 = R_{10} = R_{11} \geq R_{12}. \qquad (27)$$

This implies that $(12) \Rightarrow (11) \Rightarrow (10) \Rightarrow (9) \Rightarrow (8)$. Below we show by means of some counterexamples that none of the arrows in relation (27) can be reversed. In the first counterexample we show an instance for which (11) holds and (12) does not hold. For more details on the counter examples the reader should verify Ref. [16].

Example 2.1 Let $A = [0, 1] \subseteq \mathbb{R}$, $B = \{b_1, b_2, b_3\} \subset \mathbb{R}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by

$$f(a, b) = \begin{cases} a^2 & \text{if } b = b_1, \\ (a-1)^2 & \text{if } b = b_2, \\ 2^{-1} & \text{if } b = b_3. \end{cases}$$

It is easy to check that $R_{11} = L_{11} = L_{12} = 2^{-1}$, while $R_{12} = 4^{-1}$, thus (11) holds and (12) does not hold.

The next counterexample shows an instance for which (10) holds and (11) does not.

Example 2.2 Take $A = [0, 1] \subseteq \mathbb{R}$, $B = \{b_1, b_2\} \subset \mathbb{R}$ and introduce the function $f : A \times B \to \mathbb{R}$ given by

$$f(a, b) = \begin{cases} a^2 & \text{if } b = b_1, \\ (a-1)^2 & \text{if } b = b_2. \end{cases}$$

Consider the probability measure $\lambda^* = 2^{-1}\epsilon_{a_1} + 2^{-1}\epsilon_{a_2}$ with $a_1 = 0$ and $a_2 = 1$. It is easy to check that $\min_{b \in B} f_e(\lambda^*, \epsilon_b) = 2^{-1}$, and so it follows that $R_{11} \geq 2^{-1}$. Moreover, we observe by the definition of the sets A and B that $L_{10} = 2^{-1}$ and since we already know that $L_{10} \geq R_{11} = R_{10}$ and $R_{11} \geq 2^{-1}$ we obtain $L_{10} = R_{11} = R_{10} = 2^{-1}$. It is now easy to check that $L_{11} = 1$ and hence we have found an instance for which (10) holds and (11) does not hold.

In order to construct an instance for which (9) holds and (10) does not hold consider the set c_0 of all (real valued) sequences converging to 0. It is well-known that the space c_0 endowed with the norm $\|a\|_{c_0} = \sup_{k \in \mathbb{N}} |a_k|$ is a Banach space. Let $A = \{a = (a_k) \in C_0 : a_1 = 0\}$, $B = [0, 1] \subset \mathbb{R}$ and take the function $f : A \times B \to \mathbb{R}$ given by

$$f(a, b) = f((a_k), b) = \begin{cases} 1 & \text{if there exist some } k \in \mathbb{N} \text{ such that } b = a_k, \\ 0 & \text{otherwise.} \end{cases} \qquad (28)$$

One can now show the following lemma. For its proof see Ref. [16], Lemma 18.

Lemma 2.2 *The function f in relation* (28) *belongs to $L^1_{\lambda \otimes \mu}$ for every $\lambda \in \mathcal{P}(A)$ and* $\mu \in \mathcal{P}(B)$.

We will now list the counterexample for which (9) holds and (10) does not hold.

Example 2.3 Let $f : A \times B \to \mathbb{R}$ be the function defined in relation (28) and consider some $\lambda \in \mathcal{P}_F(A)$. Hence there exists a finite number of sequences $a^i = (a^i_k)_{k \in \mathbb{N}}$, $1 \le i \le m$, belonging to A and some vector $s(\lambda) = (s_1(\lambda), \dots, s_m(\lambda))$, $s_i(\lambda) > 0$ and $\sum_{i=1}^m s_i(\lambda) = 1$ such that $\lambda = \sum_{i=1}^m s_i(\lambda) \epsilon_{a^i}$. Since the set $[0, 1]$ contains more than a countable number of elements one can now choose a number $b \in [0, 1]$ such that none of the above sequences a^i, $1 \le i \le m$, contain this number. Using this number and the definition of f it can be easily seen that $\inf_{b \in [0,1]} f_e(\lambda, \epsilon_b) = \inf_{b \in [0,1]} \sum_{i=1}^m s_i(\lambda) f(a^i, b) = 0$, and so $R_{10} = 0$. On the other hand, consider some $\mu \in \mathcal{P}_F(B)$. By definition one can find some finite set $\{b_1, \dots, b_p\} \subseteq [0, 1]$ and a vector $s(\mu) = (s_1(\mu), \dots, s_p(\mu))$, $s_j(\mu) > 0$ and $\sum_{j=1}^p s_j(\mu) = 1$ such that $\mu = \sum_{j=1}^p s_j(\mu) \epsilon_{b_j}$. Taking the element $a_0 := (0, b_1, \dots, b_p, 0, 0, \dots) \in c_0$ it is obvious by the definition of f that $\sup_{a \in A} f_e(\epsilon_a, \mu) \ge \sum_{j=1}^n s_j(\mu) f(a_0, b_j) = 1$. Since f is bounded by 1 this shows that $L_{10} := \inf_{\mu \in \mathcal{P}_F(B)} \sup_{a \in A} f_e(\epsilon_a, \mu) = 1$, and so we have verified that (10) does not hold. To see that (9) holds, observe that $R_9 = R_{10} = 0$ and let μ_0 be the Lebesgue measure on $[0, 1]$. Clearly the measure μ_0 belongs to $P(B)$. Also, since for every $a \in A$ the function $f(a, .)$ takes the value 1 on a countable set and zero elsewhere and by Lemma 2.2 the function f belongs to $L^1_{\lambda \otimes \mu}$ for every $\lambda \in P(A)$ and $\mu \in P(B)$ this implies $\int_0^1 f(a, b) d\mu_0(b) = 0$ for every $a \in A$. Hence it follows that $L_{10} = 0$ and so (10) holds.

We now list an instance for which (8) holds and (9) does not hold.

Example 2.4 Let $A := [0, 1] \subseteq \mathbb{R}$ and $B := \{(b_k)_{k \in \mathbb{N}} \in c_0 : b_1 = 0\}$ and take the function $f : A \times B \to \mathbb{R}$ given by

$$f(a, b) = \begin{cases} 0 & \text{if there exist some } k \in \mathbb{N} \text{ such that } a = b_k, \\ 1 & \text{otherwise.} \end{cases}$$

As in Example 2.3 one can verify for every $\lambda \in \mathcal{P}_F(A)$ that $\inf_{b \in B} f_e(\lambda, \epsilon_b) = 0$ and so $R_9 = 0$. On the other hand, by Lemma 2.2 the function f is Borel measurable and if λ_0 is the Lebesgue measure on $[0, 1]$ we obtain as before that $\int_0^1 f(a, b) d\lambda_0(a) = 1$ for every $b \in B$. Also it is easy to verify by a similar argument as used in Example 2.3 that

$$\sup_{a \in A} f_e(\epsilon_a, \mu) = \sup_{a \in [0,1]} \sum_{j=1}^p s_j(\mu) f(a, b^j) = 1 \qquad (29)$$

for every $\mu \in \mathcal{P}_F(B)$. Using now the last two equalities we obtain that $1 = L_8 \ge R_8 \ge 1$ and so (8) holds. Moreover, since $R_9 = 0$ and $L_9 = L_8 = 1$ it follows that (9) does not hold.

The above examples showed that none of the implications in relation (27) can be reversed. To conclude this section we give an example which shows that (8) can also fail.

Example 2.5 Let $A = B := [0, \infty) \subset \mathbb{R}$ and consider the function $f : A \times B \to \mathbb{R}$ given by

$$f(a, b) = \begin{cases} 1 & \text{if } a \geq b, \\ 0 & \text{otherwise.} \end{cases}$$

For any $\lambda \in \mathcal{P}(A)$ it follows $\int_0^\infty f(a, b) d\lambda(a) = \lambda([b, \infty)) = 1 - \lambda([0, b))$ for every $b \geq 0$, and so $R_8 = 0$. On the other hand, for any $\mu \in \mathcal{P}(B)$ we observe that $\int_0^\infty f(a, b) d\mu(b) = \mu([0, a))$ for every $a \geq 0$, and so it follows that $L_8 = 1$. Hence (8) does not hold.

In the next section we apply the minimax results of Sect. 2 to derive Lagrangian duality results.

3 Necessary and Sufficient Conditions for Lagrangian Duality

We first recall some well-known notions. Let Y be a Hausdorff locally convex vector space and $K \subseteq Y$ some closed convex cone. Introduce on Y the ordering $\leq_K$ defined by $y_1 \leq_K y_2 \iff y_2 - y_1 \in K$ (cf. Ref. [20]). If Y^* denotes the topological dual space of Y, let $K^* \subseteq Y^*$ be the so-called dual cone given by $K^* := \{y^* \in Y^* : \langle y^*, y \rangle \geq 0 \text{ for every } y \in K\}$ with $\langle y^*, y \rangle := y^*(y)$. If X is a topological space and $p : X \to \mathbb{R}$ and $q : X \to Y$ are some given mappings, consider for the nonempty feasible region $F := \{x \in X : q(x) \geq_K 0\}$ the optimization problem

$$v(\mathrm{P}) := \sup\{p(x) : x \in F\}. \tag{P}$$

Introducing now the Lagrangian function $\theta : K^* \to (-\infty, \infty]$ given by $\theta(y^*) := \sup_{x \in X}\{p(x) + \langle y^*, q(x) \rangle\}$ the (Lagrangian) dual of optimization problem (P) is represented by

$$v(\mathrm{D}) := \inf_{y^* \in K^*} \theta(y^*). \tag{D}$$

It is well known and easy to verify that $v(\mathrm{D}) \geq v(\mathrm{P})$. In this section we are now interested under which necessary and sufficient conditions we actually have an equality. Whether or not an equality holds (no duality gap) plays a central role in the theory of optimization and this question has been extensively studied within the literature. Using the minimax approach and imposing for a noncompact set X the well-known Slater-type regularity condition, we derive in this section a necessary and sufficient condition for equality of the optimal objective values of both problems. The same can be done for a compact set X without Slater's condition. Moreover, we show that the Slater-type condition is actually equivalent to the inf-compactness of the Lagrangian bifunction and hence this regularity condition is nothing else than a compactness-type condition. To start with our analysis we first give an alternative expression for the value $v(\mathrm{P})$.

Lemma 3.1 *If the function $f : X \times K^* \to R$ is given by $f(x, y^*) := p(x) + \langle y^*, q(x) \rangle$, then $v(\mathrm{P}) = \sup_{\mathbf{x} \in X} \inf_{\mathbf{y}^* \in K^*} f(x, y^*)$.*

Proof Copy the proof for a finite dimensional optimization problem to be found in Ref. [21] and use the bipolar theorem $K^{**} = K$ (cf. Ref. [22]) in a Hausdorff locally convex vector space.   □

If the function f is given as in Lemma 3.1 and in Theorem 2.5 the set A is replaced by X and B by K^*, then we obtain by Lemma 3.1 that

$$v(\mathrm{D}) = v(\mathrm{P}) \Longleftrightarrow f \in \mathcal{C} \cap \mathcal{D}. \tag{30}$$

For the function f in Lemma 3.1 we will show under some regularity condition that f actually belongs to $\mathcal{C}$. To verify this we first observe, using the fact that K^* is a convex cone, that $\max_{x \in J} f_e(\epsilon_x, \mu) \geq \inf_{y^* \in K^*} \max_{x \in J} f(x, y^*)$ for every $J \in \mathcal{F}(X)$ and $\mu \in \mathcal{P}_F(K^*)$. This implies

$$\inf_{\mu \in \mathcal{P}_F(K^*)} \max_{x \in J} f_e(\epsilon_x, \mu) = \inf_{y^* \in K^*} \max_{x \in J} f(x, y^*), \tag{31}$$

and by relation (31) it follows that

$$f \in \mathcal{C} \Longleftrightarrow \sup_{J \in \mathcal{F}(X)} \inf_{y^* \in K^*} \max_{x \in J} f(x, y^*) = \inf_{y^* \in K^*} \sup_{x \in X} f(x, y^*). \tag{32}$$

To check that f belongs to $\mathcal{C}$ we therefore need to prove the last equality in relation (32). To do so we verify whether the conditions of Lemma 2.1 with $A = X$ and $B = K^*$ holds. Hence we introduce a convenient topology on the set Y^* and this will be given by the weak*-topology. Recall in the weak*-topology on Y^* the neighborhood base of zero is given by sets of the form

$$\mathcal{N}(y_1, \ldots, y_k, \epsilon) := \{y^* \in Y^* : |\langle y^*, y_i \rangle| < \epsilon, 1 \leq i \leq k\}$$

with $\epsilon > 0$ and $\{y_1, \ldots, y_k\}$ some finite subset of Y. It is well-known that the net $\{y_i^*\}_{i \in I} \subseteq Y^*$ converges in the weak*-topology to $y^* \in Y^*$ (notation $y_i^* \to^* y^*$) if and only if $\lim_{i \in I} \langle y_i^*, y \rangle = \langle y^*, y \rangle$ for every $y \in Y$. This shows that the function $y \longmapsto f(x, y^*)$ listed in Lemma 3.1 is continuous (in the weak*-topology) for every $x \in X$ and so all its lower level sets are weak*-closed. Hence to apply Lemma 2.1 and show by relation (32) that f belongs to $\mathcal{C}$ we still need to verify the inf-compactness property. Recall now the following well-known regularity condition.

Definition 3.1 A point x_0 is called a Slater point of the feasible region F if $x_0 \in X$ and $q(x_0) \in \mathrm{int}(K)$.

It is now possible to show the following equivalent characterization of a Slater point.

Lemma 3.2 *The point x_0 is a Slater point of the set F iff the function $y^* \mapsto f(x_0, y^*)$ listed in Lemma 3.1 is inf-compact (in the weak*-topology).*

Proof To show that the function $y^* \mapsto f(x_0, y^*)$ is inf-compact we need to verify for every $r \in \mathbb{R}$ that the lower level set $L(r) := \{y^* \in K^* : f(x_0, y^*) \leq r\}$ is weak*-compact. Since $q(x_0)$ belongs to $\mathrm{int}(K)$ there exists some balanced neighborhood

V of 0 (cf. Ref. [23]) such that $q(x_0) - V \subseteq K$. Consider now some y^* belonging to $L(r)$. Since $y^* \in K^*$ this implies for every $y \in V$ that $r \geq p(x_0) + \langle y^*, q(x_0) - y \rangle + \langle y^*, y \rangle \geq p(x_0) + \langle y^*, y \rangle$ and so it follows

$$L(r) \subseteq \{y^* \in Y^* : \langle y^*, y \rangle \leq r - p(x_0) \text{ for every } y \in V\}. \tag{33}$$

Since V is a balanced neighborhood of 0 implying $y \in V$ if and only if $-y \in V$ we obtain that the last set in relation (33) equals

$$\{y^* \in Y^* : |\langle y^*, y \rangle| \leq r - p(y_0) \text{ for every } y \in V\}. \tag{34}$$

Hence by the Banach–Alaoglu theorem (cf. Ref. [23]) the set in relation (34) is weak*-compact and by relation (33), using the fact that $L(r)$ is weak*-closed, we obtain that $L(r)$ is weak*-compact. To show the reverse implication, let $f(x_0, .)$ be inf-compact (with respect to the weak*-topology) and consider some neighborhood V of 0. Since $\alpha \longmapsto \alpha q(x_0)$ is continuous, there exists some $\beta > 0$ such that $q(x_0) \in \beta V$ and so $q(x_0) = \beta y_0$ for some $y_0 \in V$. Take now $r := p(x_0) + \beta$ and assume by contradiction that there exists some nonzero $y_0^* \in L(r)$ satisfying $\langle y_0^*, q(x_0) \rangle \leq 0$. This implies that αy_0^* belongs to $L(r)$ for every $\alpha \geq 1$ and since $L(r)$ is weak*-compact we obtain by part (b) of Theorem 1.15 in Ref. [23] that $L(r)$ is bounded. This contradicts the fact that the nonzero αy_0^* belongs to $L(r)$ for every $\alpha \geq 1$ and so $\langle y^*, q(x_0) \rangle > 0$ for every nonzero $y^* \in L(r)$. Since $q(x_0) = \beta y_0$ for some $y_0 \in V$ we obtain $\{y^* \in K^* : \langle y^*, y \rangle \leq 1 \text{ for every } y \in V\} \subseteq L(r)$ and hence $\langle y^*, q(x_0) \rangle > 0$ for every nonzero $y^* \in K^*$. This shows (cf. [20]) that $q(x_0)$ belongs to $\text{int}(K)$ and so x_0 is a Slater point of the set F. □

Using Lemma 3.2 and relation (30) one can verify the following important duality result for optimization problem P.

Theorem 3.1 *If the set F contains a Slater point, then the function f listed in Lemma* 3.1 *belongs to* $\mathcal{D}$, *i.e.*

$$\sup_{\lambda \in \mathcal{P}_F(X)} \inf_{y^* \in K^*} f_e(\lambda, \epsilon_{y^*}) = \sup_{x \in X} \inf_{y^* \in K^*} f(x, y^*)$$

iff $v(\mathrm{P}) = v(\mathrm{D})$. *Moreover, the dual problem* D *has an optimal solution.*

Proof By Lemma 3.2 and $y^* \mapsto p(x) + \langle y^*, q(x) \rangle$ is continuous in the weak*-topology for every $x \in X$, the conditions of Lemma 2.1 are satisfied. Hence by relation (32) the function f in Lemma 3.1 belongs to the set $\mathcal{C}$ and so by relation (30) $v(\mathrm{D}) = v(\mathrm{P})$ iff $f \in \mathcal{D}$. Actually, by the inf-compactness of $y^* \mapsto p(x_0) + \langle y^*, q(x_0) \rangle$ with x_0 the Slater point, it holds by Lemma 2.1 that the dual problem has an optimal solution. □

Irrespective of whether or not the function f belongs to $\mathcal{D}$, one can give a geometrical interpretation of the value $v(\mathrm{D})$. A similar interpretation was given by means of a completely different proof in Ref. [24] for $v(\mathrm{D})$ associated with a finite dimensional primal problem with a continuous objective function and continuous inequality restrictions.

Theorem 3.2 *If the set F contains a Slater point and the set $H \subseteq R \times Y$ is given by*

$$H := \{(p(x), q(x)) : x \in X\},$$

then $v(D) = \sup\{\alpha : (\alpha, \beta) \in \mathrm{co}(H), \beta \geq_K 0\}$ and the dual problem (D) has an optimal solution.

Proof Since f belongs to $\mathcal{C}$ we obtain by relation (24) (take $A = X$ and $B = K^*$)

$$
\begin{aligned}
v(D) &= \sup_{\lambda \in \mathcal{P}_F(X)} \inf_{y^* \in K^*} f_e(\lambda, \epsilon_{y*}) \\
&= \sup_{(s(\lambda), x) \in \Delta_m \times X^m, m \in \mathbb{N}} \left\{ \sum_{i=1}^{m} s_i(\lambda) p(x_i) + \inf_{y^* \in K^*} \left\langle y^*, \sum_{i=1}^{m} s_i(\lambda) q(x_i) \right\rangle \right\},
\end{aligned}
\tag{35}
$$

where $x = (x_1, \ldots, x_m)$. Applying the bipolar theorem to the last expression in relation (35) and using $v(D) > -\infty$ and relation (35) we obtain

$$
v(D) = \sup_{(s(\lambda), x) \in \Delta_m \times X^m, m \in \mathbb{N}} \left\{ \sum_{i=1}^{m} s_i(\lambda) p(x_i) : \sum_{i=1}^{m} s_i(\lambda) q(x_i) \in K \right\}.
\tag{36}
$$

Hence by relation (36) the desired result follows.   □

In the next example we will consider an important class of optimization problems for which the Lagrangian dual can be simplified. For a detailed derivation see Ref. [16].

Example 3.1 Let Y be a normed linear space with $L \subseteq Y$ some closed linear subspace, $b \in Y$ and $C \subseteq Y$ some closed convex cone, and consider the conic convex programming problem $v(CP) := \sup\{\langle y_0^*, x \rangle : x \in C \cap (L + b)\}$ with y_0^* some element of the topological dual space Y^*. The Lagrangian dual is now given by $v(D) = \inf\{\langle y^*, b \rangle : y^* \in L^{\perp} \cap (C^* + y_0^*)\}$. If the space Y is a real Hilbert space the dual Y^* of Y can be identified with Y and the dual decision variables belong to Y and we obtain in this case $\langle y^*, x \rangle = (y^*, x)$ with $(., .)$ denoting the inner product on the real Hilbert space (cf. Ref. [18]).

We will now investigate what happens if the feasible set $F = \{x \in X : q(x) \geq_K 0\}$ does not necessarily contain a Slater point. Consider in this case the function $g : K^* \times X \to \mathbb{R}$ given by

$$g(y^*, x) := p(x) + \langle y^*, q(x) \rangle. \tag{37}$$

It is now possible to show the following result and this should be compared with Theorem 3.1. Observe that, for X and g satisfying some topological condition, one can give a simpler necessary and sufficient condition.

Theorem 3.3 *The function* $-g$, *given by relation* (37), *belongs to* $\mathcal{C}$, *i.e.*

$$\inf_{J \in \mathcal{F}(K^*)} \sup_{\mu \in \mathcal{P}_F(X)} \inf_{y^* \in J} g_e(\epsilon_{y^*}, \mu) = \sup_{x \in X} \inf_{y^* \in K^*} g(y^*, x)$$

iff $v(\mathrm{D}) = v(\mathrm{P})$. *Additionally, if* X *is a compact Hausdorff space and the function* $x \mapsto g(y^*, x)$ *listed in relation* (37) *is both upper semicontinuous and belongs to* $\bigcap_{\mu \in \mathcal{P}(X)} L^1_\mu(X)$ *for every* $y^* \in K^*$, *the primal problem* (P) *has an optimal solution and* $v(\mathrm{D}) = v(\mathrm{P})$ *iff*

$$\sup_{\mu \in \mathcal{P}(X)} \inf_{y^* \in K^*} g_e(\epsilon_{y^*}, \mu) = \sup_{x \in X} \inf_{y^* \in K^*} g(y^*, x).$$

Proof By Lemma 3.1 we observe $v(P) = \sup_{x \in X} \inf_{y^* \in K^*} g(y^*, x)$. This shows replacing in relation (26) the set A by K^* and B by X that

$$v(\mathrm{P}) = v(\mathrm{D}) \Longleftrightarrow -g \in \mathcal{C} \cap \mathcal{D}. \tag{38}$$

For the bifunction g one can easily verify that $\inf_{\lambda \in \mathcal{P}_F(K^*)} \sup_{x \in X} g_e(\lambda, \epsilon_x) = \inf_{y^* \in K^*} \sup_{x \in X} g(y^*, x)$ and so $-g$ belongs to $\mathcal{D}$ (with $A = K^*$ and $B = X$). Applying relation (38) yields the first part. To show the second part we observe by relation (19) that $\sup_{\mu \in \mathcal{P}_F(X)} \inf_{y^* \in J} g_e(\epsilon_{y^*}, \mu) = \sup_{\mu \in \mathcal{P}(X)} \inf_{y^* \in J} g_e(\epsilon_{y^*}, \mu)$ for every $J \in \mathcal{F}(X)$. This implies by Proposition 2.1 with f_e replaced by $-g_e$, $A = K^*$ and $B = X$ that

$$\inf_{J \in \mathcal{F}(K^*)} \sup_{\mu \in \mathcal{P}_F(X)} \inf_{y^* \in J} g_e(\epsilon_{y^*}, \mu) = \sup_{\mu \in \mathcal{P}(X)} \inf_{y^* \in K^*} g_e(\epsilon_{y^*}, \mu) \tag{39}$$

and by the first part the result follows. Clearly under the above conditions the primal problem has an optimal solution and so the result is proved. $\square$

In Remark 3.1 a characterization is given of the value $v(\mathrm{D})$ irrespective whether the function $-g$ belongs to $\mathcal{C}$ or not.

Remark 3.1 By the definition of $v(\mathrm{D})$, relation (13) and the specific form of the function g it is easy to show that

$$v(\mathrm{D}) = \inf_{\lambda \in \mathcal{P}_F(K^*)} \sup_{x \in X} g_e(\lambda, \epsilon_x) = \inf_{J \in \mathcal{F}(K^*)} \inf_{\lambda \in \mathcal{P}(J)} \sup_{x \in X} g_e(\lambda, \epsilon_x). \tag{40}$$

Applying Wald's minimax result (see Theorem 2.1) to the last part of relation (40) we obtain

$$v(\mathrm{D}) = \inf_{J \in \mathcal{F}(K^*)} \sup_{\mu \in \mathcal{P}_F(X)} \inf_{y^* \in J} g_e(\epsilon_{y^*}, \mu). \tag{41}$$

If we additionally know that the set X is a compact Hausdorff space and the function $x \mapsto g(y^*, x)$ with g listed in relation (37) is both upper semicontinuous and belongs to $\bigcap_{\mu \in \mathcal{P}(X)} L^1_\mu(X)$ for every $y^* \in K^*$, then it is possible to simplify the previous expression for $v(\mathrm{D})$. Applying relation (39), we obtain

$$v(\mathrm{D}) = \sup_{\mu \in \mathcal{P}(X)} \inf_{y^* \in K^*} g_e(\epsilon_{y^*}, \mu). \tag{42}$$

Finally, if the set X is a compact Hausdorff space and additionally the functions $p : X \to R$ and $q : X \to Y$ are continuous (implying $x \mapsto g(y^*, x)$ is continuous for every $y^* \in Y^*$ and hence automatically belongs to $\bigcap_{\mu \in \mathcal{P}(X)} L^1_\mu(X)$), one can give a geometrical interpretation of the value $v(\mathrm{D})$. One can show in this case (for more details see Ref. [16]) that

$$v(\mathrm{D}) = \sup \left\{ \alpha : \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathrm{cl}(\mathrm{co}(H)), \beta \geq_K 0 \right\}$$

with H listed in Theorem 3.2. Since the functions p and q are continuous and X is a compact Hausdorff space it follows that H is compact in $R \times Y$. This shows, if Y is a Fréchet space, that $\mathrm{cl}(\mathrm{co}(H))$ is compact (cf. Ref. [23]). Moreover, for $Y = R^m$ we obtain that $\mathrm{co}(H)$ is closed and hence $\mathrm{cl}(\mathrm{co}(H)) = \mathrm{co}(H)$.

Finally we note that the different convexity-type conditions on the functions p and q for strong duality (for a recent overview see for example Ref. [21]) are easy to verify to belong to the sets $\mathcal{C}$ or $\mathcal{D}$. Also the results in this section can be more easily proved for finite dimensional optimization problems, i.e. $X \subseteq \mathbb{R}^n$ and $Y = \mathbb{R}^m$. In this case the set $Y^* = \mathbb{R}^m$ is finite dimensional and instead of the Banach–Alaoglu theorem and the weak*-topology on Y^* we use the ordinary Euclidean topology on R^m and the well-known result that a set $C \subseteq \mathbb{R}^m$ is compact if and only if the set C is closed and bounded.

References

1. Simons, S.: Minimax theorems and their proofs. In: Du, D.Z., Pardalos, P.M. (eds.) Minimax and Applications, pp. 1–23. Kluwer Academic, Dordrecht (1995)
2. Frenk, J.B.G., Kassay, G., Kolumbán, J.: On equivalent results in minimax theory. Eur. J. Oper. Res. **157**, 46–58 (2004)
3. Frenk, J.B.G., Kassay, G.: The level set method of Joó and its use in minimax theory. Math. Program. Ser. A **105**, 145–155 (2006)
4. Joó, I.: On some convexities. Acta Math. Hung. **54**, 163–172 (1989)
5. Chvátal, V.: Linear Programming. Freeman, New York (1999)
6. Jones, A.N.: Game Theory: Mathematical Models of Conflict. Horwood, Chichester (2000)
7. Frenk, J.B.G., Kassay, G., Protassov, V.: On Borel probability measures and noncooperative game theory. Optimization **54**, 81–101 (2005)
8. Ash, R.B.: Real Analysis and Probability. Academic, New York (1972)
9. Neveu, J.: Mathematical Foundations of the Calculus of Probability. Holden Day, San Francisco (1965)
10. Wald, A.: Generalization of a theorem by von Neumann concerning zero-sum two-person games. Ann. Math. **46**, 281–286 (1945)
11. Aubin, J.B.: Optima and Equilibria: An Introduction to Nonlinear Analysis. Graduate Texts in Mathematics. Springer, Berlin (1993)
12. Royden, H.L.: Real Analysis. Macmillan, New York (1988)
13. Choquet, G.: Lectures on Analysis 1: Integration and Topological Vector Spaces. Mathematics Lecture Note Series. Benjamin, Reading (1976)
14. von Neumann, J.: Zur Theorie der Gesellschaftsspiele. Math. Ann. **100**, 295–320 (1928)
15. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1994)
16. Frenk, J.B.G., Kas, P., Kassay, G.: On linear programming duality and a necessary and sufficient condition in minimax theory. Econometric Institute Report E.I. 2004-14, https://ep.eur.nl/handle/1765/1219
17. Ville, J.: Sur la théorie générale des jeux ou intervient l'habilité des jouers. In: Borel, E., et al. (eds.) Traité du Calcul des Probabilités et de Ses Applications, pp. 105–113. Gauthier-Villars, Paris (1938)

18. Kolmogorov, A.N., Fomin, S.V.: Introductory Real Analysis. Dover Books on Mathematics. Dover, New York (1975)
19. Kassay, G., Kolumbán, J.: On a generalized sup–inf problem. J. Optim. Theory Appl. **91**, 651–670 (1996)
20. Jahn, J.: Mathematical Vector Optimization in Partially Ordered Spaces. Lang, Frankfurt am Main (1986)
21. Frenk, J.B.G., Kassay, G.: On classes of generalized convex functions Gordan–Farkas type theorems and Lagrangian duality. J Optim. Theory Appl. **102**, 315–343 (1999)
22. Aubin, J.B., Ekeland, I.: Applied Nonlinear Analysis. Wiley, New York (1984)
23. Rudin, W.: Functional Analysis. International Series in Pure and Applied Mathematics. McGraw–Hill, New York (1991)
24. Faigle, U., Kern, W., Still, G.: Algorithmic Principles of Mathematical Programming. Kluwer Texts in Mathematical Sciences, vol. 24. Kluwer Academic, Dordrecht (2002)