TÜBİTAK

Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2021) 45: 695 – 717
© TÜBİTAK
doi:10.3906/mat-2012-45

Portfolio optimization with two quasiconvex risk measures

Çağın ARARAT*
Department of Industrial Engineering, Faculty of Engineering, Bilkent University, Ankara, Turkey

Received: 10.12.2020 • **Accepted/Published Online:** 14.01.2021 • **Final Version:** 26.03.2021

Abstract: We study a static portfolio optimization problem with two risk measures: a principle risk measure in the objective function and a secondary risk measure whose value is controlled in the constraints. This problem is of interest when it is necessary to consider the risk preferences of two parties, such as a portfolio manager and a regulator, at the same time. A special case of this problem where the risk measures are assumed to be coherent (positively homogeneous) is studied recently in a joint work of the author. The present paper extends the analysis to a more general setting by assuming that the two risk measures are only quasiconvex. First, we study the case where the principal risk measure is convex. We introduce a dual problem, show that there is zero duality gap between the portfolio optimization problem and the dual problem, and finally identify a condition under which the Lagrange multiplier associated to the dual problem at optimality gives an optimal portfolio. Next, we study the general case without the convexity assumption and show that an approximately optimal solution with prescribed optimality gap can be found by using the well-known bisection algorithm combined with a duality result that we prove.

Key words: Portfolio optimization, quasiconvex risk measure, minimal penalty function, maximal risk function, Lagrange duality, bisection method

1. Introduction

Risk measures are functionals that are defined on a linear space of real-valued random variables hosted by a common probability space. In the context of financial mathematics, each random variable can denote the uncertain future worth of an investor's position, and a risk measure assigns a (deterministic) extended real number to the random variable; this number quantifies the initial capital that is needed for compensating the risk of the position.

Introduced in [3], risk measures have been studied extensively in the financial mathematics literature over the last two decades. In [3], the so-called coherent risk measures are studied within an axiomatic framework; we provide the mathematical formulations of these axioms in Section 2 for the convenience of the reader. Among the properties of a coherent risk measure, positive homogeneity imposes that the risk of a financial position is scalable by the size of the position. While some classical risk measures such as negative expected value and average value-at-risk (see [7, Example 4.40]) enjoy this property, positive homogeneity can be found restrictive from a financial point of view. To this end, convex risk measures provide a richer class of risk measures in which positive homogeneity is not taken for granted. A classical example of a convex but not coherent risk measure is the entropic risk measure, which has a simple expression of the log-sum-exp form (see Example 6.1). The reader is referred to [7, Chapter 4] for a detailed discussion on convex risk measures.

*Correspondence: cararat@bilkent.edu.tr

2010 *AMS Mathematics Subject Classification:* 90C11, 90C20, 90C90, 91B30, 91G10

The convexity property of a risk measure is often motivated by the statement "Diversification does not increase risk", which formulates the role of allocating one's capital into a variety of investment opportunities. More recently, it has been argued that quasiconvexity can be used as a relaxed alternative for convexity as it still captures the idea behind diversification. Therefore, quasiconvex risk measures, as argued in [5] and [6], cover a wider range of functionals that can be used for risk measurement purposes; these include certainty equivalents (see [6, Example 8] and Section 6.2) and economic indices of riskiness (see [6, Example 3]) in addition to convex (and coherent) risk measures described above.

A rich class of problems where risk measures appear naturally is that of portfolio optimization problems. In these problems, one wishes to minimize or control the risk associated to a portfolio that consists of multiple risky assets. Assuming special families of asset return distributions, the works [10, 11, 13] study static portfolio optimization problems with a single coherent risk measure that appears in the objective function. More recently, in the previous work [2] of the author, joint with Aktürk, a static portfolio optimization problem with two coherent risk measures is formulated. In this problem, the decision maker aims to minimize the value of a principle risk measure, e.g., the risk measure of the portfolio manager, while keeping the value of a secondary risk measure, e.g., the risk measure declared by a regulatory authority, below a critical threshold. In [2], a complete analysis of this problem is provided for the general case of arbitrary asset return distributions and arbitrary coherent risk measures that satisfy certain regularity conditions.

On the other hand, the use of quasiconvex risk measures in portfolio optimization is relatively new. In [12], a static portfolio optimization problem is studied, where the objective function is the composition of a quasiconvex risk measure and a concave functional that is defined on a space of portfolios, and this composition is to be minimized over a convex compact set of portfolios. The main result [12, Theorem 4] provides a sufficient condition for a portfolio to be optimal in terms of a set relation between normal cones and generalized subdifferentials for quasiconvex functions. In particular, the derivations rely on the dual representations for quasiconvex risk measures developed in [6] as well as the general duality theory for quasiconvex functions initiated earlier in [14]. In [9], quasiconvex risk measures are used in a dynamic portfolio optimization problem in continuous time in order to model ambiguity-averse preferences. The work [9] also makes use of the dual representation results of [6].

The aim of the present paper is to extend the static portfolio optimization problem in [2] by assuming that both the principle and the secondary risk measures are quasiconvex. In particular, we cover the case where the two risk measures are convex. It should be noted that the extension from the coherent case to the quasiconvex case requires entirely different duality arguments, explaining the mathematical originality of the present paper. On the other hand, compared to the portfolio optimization problem in [12] with a single quasiconvex risk measure and a general convex set constraint, our problem assumes that the constraint has a special structure induced by the secondary risk measure. This structure makes it possible to formulate a more explicit dual problem with a linear inequality constraint.

The rest of the paper is organized as follows. After reviewing some basic notions about quasiconvex risk measures in Section 2 and introducing the primal problem in Section 3, we break down the analysis of the problem into two steps. First, in Section 4, we work under the assumption that the principle risk measure is a convex functional (not necessarily translative though). We formulate the dual problem in Section 4.1 and prove that (Theorem 4.8) there is zero duality gap between the primal and dual problems. In Section 4.2, we impose further structural properties on the principle risk measure and prove that (Theorem 4.11), at optimality, a Lagrange multiplier attached to the linear inequality constraint of the dual problem yields an optimal portfolio

vector for the primal problem. Next, in Section 5, we remove the convexity assumption and reformulate the quasiconvex portfolio optimization problem via a family of convex feasibility problems parametrized by a decision variable of the quasiconvex problem. Similar to the results of Section 4, we provide a duality-based method to solve each of these feasibility problems. Then, we employ the well-known bisection method that iterates through different values of the parameter of the feasibility problems and stops with prescribed suboptimality in finitely many iterations. Hence, combining the duality result with the bisection method provides a way to find an approximately optimal solution for the portfolio optimization problem with two quasiconvex risk measures. In Section 6, we consider convex risk measures and certainty equivalents in order to illustrate the use of the dual problems, and we discuss the validity of some technical assumptions stated in Section 4 and Section 5. Finally, Section 7 concludes the paper.

2. Quasiconvex risk measures

In this section, we fix the notation for the rest of the paper and review some preliminary notions related to risk measures. For the latter, we focus on the more recent quasiconvex framework studied in [5, 6].

Let $n \in \mathbb{N} := \{1, 2, \ldots\}$. We assume that the standard Euclidean space $\mathbb{R}^n$ is equipped with an arbitrary norm $|\cdot|$ and the usual inner product defined by $x^\mathsf{T} z := \sum_{i=1}^n x_i z_i$ for $x, z \in \mathbb{R}^n$. We denote by $\mathbb{R}^n_+$ the positive orthant in $\mathbb{R}^n$, that is, the set of all $x = (x_1, \ldots, x_n)^\mathsf{T} \in \mathbb{R}^n$ with $x_i \geq 0$ for each $i \in \{1, \ldots, n\}$.

To introduce the probabilistic setup, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and denote by L^0_n the set of all $\mathcal{F}$-measurable random variables taking values in $\mathbb{R}^n$, where two elements are considered identical if they are equal $\mathbb{P}$-almost surely. For each $X \in L^0_n$, we define

$$\|X\|_p := \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}}$$

for $p \in [1, +\infty)$ and

$$\|X\|_p := \inf\{c \geq 0 \mid \mathbb{P}\{|X| \leq c\} = 1\}$$

for $p = +\infty$. Let $p \in [1, +\infty]$. The space $L^p_n := \{X \in L^0_n \mid \|X\|_p < +\infty\}$ is a Banach space equipped with the norm $\|\cdot\|_p$. For brevity, let $L^p := L^p_1$; for $Y_1, Y_2 \in L^p$, we write $Y_1 \leq Y_2$ if $\mathbb{P}\{Y_1 \leq Y_2\} = 1$, which yields the cone $L^p_+ := \{Y \in L^p \mid 0 \leq Y\}$.

Let $\mathcal{Y} = L^p$, where $p \in [1, +\infty]$. We assume that $\mathcal{Y}$ is equipped with the strong topology induced by the norm $\|\cdot\|_p$ if $p < +\infty$ and with the weak* topology $\sigma(L^\infty, L^1)$ if $p = +\infty$. Under this topology, we denote by $\mathcal{Y}^*$ the topological dual space of $\mathcal{Y}$ with the bilinear duality mapping $\langle \cdot, \cdot \rangle : \mathcal{Y}^* \times \mathcal{Y} \to \mathbb{R}$. Hence, $\mathcal{Y}^* = L^q$ with $\langle V, Y \rangle = \mathbb{E}[VY]$ for every $V \in \mathcal{Y}^*, Y \in \mathcal{Y}$, where $q \in [1, +\infty]$ is the conjugate exponent of p, that is, $\frac{1}{p} + \frac{1}{q} = 1$. In all cases, we consider $\mathcal{Y}^*$ with the weak topology $\sigma(\mathcal{Y}^*, \mathcal{Y}) = \sigma(L^q, L^p)$.

For a functional $\rho \colon \mathcal{Y} \to \bar{\mathbb{R}} := [-\infty, +\infty]$, let us consider the following properties.

(i) Monotonicity: $Y_1 \leq Y_2$ implies $\rho(Y_1) \geq \rho(Y_2)$ for every $Y_1, Y_2 \in \mathcal{Y}$.

(ii) Quasiconvexity: It holds $\rho(\lambda Y_1 + (1-\lambda)Y_2) \leq \max\{\rho(Y_1), \rho(Y_2)\}$ for every $Y_1, Y_2 \in \mathcal{Y}$ and $\lambda \in (0, 1)$.

(iii) Convexity: It holds $\rho(\lambda Y_1 + (1-\lambda)Y_2) \leq \lambda\rho(Y_1) + (1-\lambda)\rho(Y_2)$ for every $Y_1, Y_2 \in \mathcal{Y}$ and $\lambda \in (0, 1)$ (with the inf-addition convention $(+\infty) + (-\infty) = +\infty$ for the right hand side of the inequality).

(iv) Translativity: It holds $\rho(Y+y)=\rho(Y)-y$ for every $Y\in\mathcal{Y}$ and $y\in\mathbb{R}$.

(v) Positive homogeneity: It holds $\rho(\lambda Y)=\lambda\rho(Y)$ for every $Y\in\mathcal{Y}$ and $\lambda>0$.

For each $t\in\mathbb{R}$, let us define

$$\mathscr{A}^t:=\{Y\in\mathcal{Y}\mid\rho(Y)\leq t\},$$

which is called the *acceptance set* of ρ at level t. It is easy to check that

$$\rho(Y)=\inf\left\{t\in\mathbb{R}\mid Y\in\mathscr{A}^t\right\},\quad Y\in\mathcal{Y},\tag{2.1}$$

where $\inf\emptyset:=+\infty$ as a convention. Note that ρ satisfies quasiconvexity if and only if $\mathscr{A}^t$ is convex for each $t\in\mathbb{R}$. The functional ρ is called a quasiconvex risk measure if it satisfies (i) and (ii), a convex risk measure if it satisfies (i), (ii), (iii), (iv), and a coherent risk measure if it satisfies (i), (ii), (iii), (iv), (v). Hence, a coherent risk measure is necessarily a convex risk measure, and a convex risk measure is necessarily a quasiconvex risk measure. In addition, convexity implies quasiconvexity; and, under monotonicity, quasiconvexity and translativity imply convexity; see [7, Exercise 4.1.1]. Hence, when working within the general framework of quasiconvex risk measures, one does not assume translativity.

In the current paper, we study a static portfolio optimization problem with two quasiconvex risk measures. In order to formulate certain dual problems associated to the portfolio optimization problem, the dual representations of these risk measures will have a crucial role. We review the dual representation result of [6] next. To that end, let $\mathcal{Y}_+^*:=L_+^q$.

Definition 2.1 *[6, Definition 9] A function $\beta\colon(\mathcal{Y}_+^*\setminus\{0\})\times\mathbb{R}\to\bar{\mathbb{R}}$ is called a maximal risk function if it satisfies the following properties.*

(a) β is increasing (nondecreasing) and left-continuous in the second argument.

(b) β is jointly quasiconcave.

(c) It holds $\beta(\lambda V,\lambda s)=\beta(V,s)$ for every $V\in\mathcal{Y}_+^*, s\in\mathbb{R},\lambda>0$.

(d) It holds $\lim_{s\to-\infty}\beta(V_1,s)=\lim_{s\to-\infty}\beta(V_2,s)$ for every $V_1,V_2\in\mathcal{Y}_+^*$.

(e) The right-continuous version $(V,s)\mapsto\beta^+(V,s):=\inf_{s'>s}\beta(V,s')$ of β is (weakly) upper semicontinuous in the first argument.

Thanks to property (c) in Definition 2.1, one can simply work with the restriction of β on the set $\mathcal{Y}_+^{*,1}\times\mathbb{R}$, where

$$\mathcal{Y}_+^{*,1}:=L_+^{q,1}:=\left\{Y\in L_+^q\mid\mathbb{E}[Y]=1\right\}.\tag{2.2}$$

Let ρ be a lower semicontinuous quasiconvex risk measure. We define the minimal penalty function $\alpha\colon\mathcal{Y}_+^*\times\mathbb{R}\to\bar{\mathbb{R}}$ of ρ by

$$\alpha(V,t):=\sup_{Y\in\mathscr{A}^t}\mathbb{E}\left[-VY\right],\quad V\in\mathcal{Y}_+^*, t\in\mathbb{R}.$$

For each $t\in\mathbb{R}$, the acceptance set $\mathscr{A}^t$ is a closed convex subset of $\mathcal{Y}$, hence it is characterized by its support function $V\mapsto\alpha(V,t)$. Consequently, the risk measure ρ is uniquely determined by its minimal penalty function

α. In quasiconvex analysis, it is sometimes useful to work with the left-continuous version α^- of α (with respect to the second variable) defined by

$$\alpha^-(V,t) \coloneqq \sup_{t'<t} \alpha(V,t'), \quad V \in \mathcal{Y}_+^*, t \in \mathbb{R}. \tag{2.3}$$

From the proof of [6, Theorem 3], it follows that

$$\alpha^-(V,t) = \sup_{Y \in \mathscr{A}^{t-}} \mathbb{E}\left[-VY\right], \quad V \in \mathcal{Y}_+^*, t \in \mathbb{R},$$

where, for each $t \in \mathbb{R}$, $\mathscr{A}^{t-}$ is the strict acceptance set at level t defined by

$$\mathscr{A}^{t-} \coloneqq \{Y \in \mathcal{Y} \mid \rho(Y) < t\}.$$

For convenience, we recall the following result from [6], which provides a precise formula to recover ρ from α or α^-.

Proposition 2.2 *[6, Theorem 3] Let $\rho\colon \mathcal{Y} \to \bar{\mathbb{R}}$ be a lower semicontinuous quasiconvex risk measure with minimal penalty function α. Then, there exists a unique maximal risk function β such that*

$$\rho(Y) = \sup_{V \in \mathcal{Y}_+^* \setminus \{0\}} \beta(V, \mathbb{E}[-VY]), \quad Y \in \mathcal{Y}.$$

Moreover, such β is given as the left-continuous pseudo-inverse of α or that of α^- with respect to the second variable, that is,

$$\beta(V,s) = \inf\{t \in \mathbb{R} \mid s \le \alpha(V,t)\} = \inf\left\{t \in \mathbb{R} \mid s \le \alpha^-(V,t)\right\}$$

for every $V \in \mathcal{Y}_+^*$ and $s \in \mathbb{R}$.

3. Portfolio optimization problem

In this section, we formulate and study a risk-averse portfolio optimization problem with two quasiconvex risk measures.

We suppose that there are n risky assets with possibly correlated returns in a static model as in [2]. To that end, let $X = (X_1, \dots, X_n)^\mathsf{T} \in L_n^p$ be a random vector, where $p \in [1, +\infty]$. For each $i \in \{1, \dots, n\}$, the component X_i denotes the return of the i^{th} asset as a multiple of the initial price of that asset. In this static model, a portfolio is naturally defined as a vector $w \in \mathbb{R}^n$ with $\sum_{i=1}^n w_i = 1$, where, for each $i \in \{1, \dots, n\}$, w_i denotes the weight of the corresponding asset in the portfolio, which is based on the asset prices at the beginning of the investment period. Hence, prohibiting short-selling, the set of all portfolios is defined as

$$\mathcal{W} \coloneqq \left\{w \in \mathbb{R}_+^n \mid \mathbf{1}^\mathsf{T} w = 1\right\}, \tag{3.1}$$

where $\mathbf{1} = (1, \dots, 1)^\mathsf{T} \in \mathbb{R}^n$.

Given a portfolio $w \in \mathcal{W}$, note that the corresponding return $w^\mathsf{T} X$ is in $\mathcal{Y} = L^p$.

To model risk-aversion, suppose that we have two proper lower semicontinuous quasiconvex risk measures $\rho_1, \rho_2\colon L^p \to \bar{\mathbb{R}}$. The portfolio manager aims to choose a portfolio $w \in \mathcal{W}$ that minimizes the type 1 risk $\rho_1(w^\mathsf{T} X)$ while satisfying the type 2 risk constraint

$$\rho_2(w^\mathsf{T} X) \le r,$$

where $r \in \mathbb{R}$ is a fixed threshold for this type of risk. Hence, the porfolio optimization problem of interest is formulated as

$$\begin{aligned} &\text{minimize} \quad \rho_1(w^{\mathsf{T}}X) && (\mathscr{P}(r)) \\ &\text{subject to} \quad \rho_2(w^{\mathsf{T}}X) \leq r \\ &\qquad\qquad w \in \mathcal{W}. \end{aligned}$$

We denote by $p(r)$ the optimal value of $(\mathscr{P}(r))$.

As in [2], the motivation to use two risk measures comes from decision making under two risk perceptions. For instance, the portfolio manager may choose the portfolio by using ρ_1 as a suitable risk measure for her own risk perception; however, there might be an obligation to consider the opinion of a regulatory authority whose risk perception is encoded by ρ_2. In another setting, the portfolio manager may wish to work with two risk measures but the principle risk measure ρ_1 may have higher seniority than ρ_2. In particular, this framework covers as special cases the problem of maximizing expected return subject to a quasiconvex risk constraint (arbitrary ρ_2) if we take $\rho_1(Y) = \mathbb{E}[-Y]$ for each $Y \in L^p$, as well as the problem of minimizing a quasiconvex risk measure (arbitrary ρ_1) while maintaining a sufficiently high expected return if we take $\rho_2(Y) = \mathbb{E}[-Y]$ for each $Y \in L^p$. In the general case, while the previous work [2] is restricted to coherent risk measures (e.g., negative expected value, average value-at-risk), the current work covers a much larger class of risk measures as it allows for both convex (translative) risk measures (e.g., entropic risk measure, utility-based shortfall risk measures) as well as quasiconvex and non-convex (non-translative) risk measures (e.g., certainty equivalent, economic index of riskiness). Some detailed examples will be studied in Section 6.

For each $j \in \{1,2\}$, let us denote by $\mathscr{A}_j^t$ the acceptance set of ρ_j at level $t \in \mathbb{R}$ and by α_j the minimal penalty function of ρ_j. In view of Proposition 2.2, there exists a unique maximal risk function β_j such that ρ_j admits the dual representation

$$\rho_j(Y) = \sup_{V \in L^q_+ \setminus \{0\}} \beta_j(V, \mathbb{E}[-VY]), \quad Y \in L^p, \tag{3.2}$$

and we have $\beta_j(V,s) = \inf\{t \in \mathbb{R} \mid s \leq \alpha_j(V,t)\}$ for every $V \in L^q_+$, $s \in \mathbb{R}$. Let us also define $g_j \colon \mathbb{R}^n \to \bar{\mathbb{R}}$ by

$$g_j(w) = \rho_j(w^{\mathsf{T}}X), \quad w \in \mathbb{R}^n, \tag{3.3}$$

which is a quasiconvex function.

We present the analysis of $(\mathscr{P}(r))$ in the next two sections.

4. Analysis of the problem when the principle risk measure is convex

In this section, we analyze the porfolio optimization problem $(\mathscr{P}(r))$ under the following assumption.

Assumption 4.1 *ρ_1 satisfies convexity.*

Note that Assumption 4.1 does not ensure that ρ_1 is a convex risk measure in general as the latter terminology assumes translativity in addition to convexity (see Section 2).

4.1. First main result: establishing strong duality

As a preparation for Theorem 4.8, we first introduce an auxiliary problem that will show up in the derivation of dual problem. For each $V_1, V_2 \in L^q_+$, let us define

$$h^P(V_1,V_2) \coloneqq \inf\left\{t + \mathbb{E}[-V_2X]^{\mathsf{T}}w \mid \alpha_1^-(V_1,t) \geq \mathbb{E}[-V_1X]^{\mathsf{T}}w,\ w \in \mathcal{W},\ t \in \mathbb{R}\right\}. \tag{4.1}$$

For the first result of this section (Proposition 4.4 below), we need the following continuity assumption on the maximal risk function of ρ_1; note that this assumption also appears as part of [9, Assumption 3.2] in the context of a dynamic portfolio optimization problem. Recall (2.2) for the definition of $L^{q,1}_+$.

Assumption 4.2 *β_1 is jointly (weakly) upper semicontinuous on $L^{q,1}_+ \times \mathbb{R}$.*

Before stating Proposition 4.4, we introduce a restrictive finiteness assumption which will be removed later in Section 4.2; see Remark 4.5 for a discussion on this assumption.

Assumption 4.3 *For every $w \in \mathcal{W}$, $V_1 \in L^{q,1}_+$ and $V_2 \in L^q_+$, it holds $\beta_1(V_1, \mathbb{E}[-V_1w^{\mathsf{T}}X]) \in \mathbb{R}$ and $\alpha_2(V_2,r) \in \mathbb{R}$.*

The next result establishes the connection between $(\mathscr{P}(r))$ and h^P.

Proposition 4.4 *Suppose that Assumption 4.1, Assumption 4.2, Assumption 4.3 hold. Then,*

$$p(r) = \sup_{V_1,V_2 \in L^q_+} \left(h^P(V_1,V_2) - \alpha_2(V_2,r)\right).$$

Proof Let $I_{\mathscr{A}_2^r}$ be the convex analytic indicator function of $\mathscr{A}_2^r$, that is, $I_{\mathscr{A}_2^r}(Y) = 0$ whenever $Y \in \mathscr{A}_2^r$ and $I_{\mathscr{A}_2^r}(Y) = +\infty$ whenever $Y \in L^p \setminus \mathscr{A}_2^r$. Then, using [17, Theorem 2.3.3, (2.33)], Proposition 2.2 and property (c) in Definition 2.1, we obtain

$$\begin{aligned}
p(r) &= \inf\left\{\rho_1(w^{\mathsf{T}}X) \mid \rho_2(w^{\mathsf{T}}X) \leq r,\ w \in \mathcal{W}\right\} \\
&= \inf\left\{\rho_1(w^{\mathsf{T}}X) \mid w^{\mathsf{T}}X \in \mathscr{A}_2^r,\ w \in \mathcal{W}\right\} \\
&= \inf_{w\in\mathcal{W}} \left(\rho_1(w^{\mathsf{T}}X) + I_{\mathscr{A}_2^r}(w^{\mathsf{T}}X)\right) \\
&= \inf_{w\in\mathcal{W}} \left(\sup_{V_1 \in L^{q,1}_+} \beta_1\left(V_1, \mathbb{E}[-V_1w^{\mathsf{T}}X]\right) + \sup_{V_2 \in L^q_+} \left(\mathbb{E}[-V_2w^{\mathsf{T}}X] - \alpha_2(V_2,r)\right)\right) \\
&= \inf_{w\in\mathcal{W}} \sup_{V_1 \in L^{q,1}_+, V_2 \in L^q_+} b(w,V_1,V_2),
\end{aligned}$$

where

$$b(w,V_1,V_2) \coloneqq \beta_1\left(V_1, \mathbb{E}[-V_1w^{\mathsf{T}}X]\right) + \mathbb{E}[-V_2w^{\mathsf{T}}X] - \alpha_2(V_2,r), \quad w \in \mathcal{W}, V_1 \in L^{q,1}_+, V_2 \in L^q_+.$$

For fixed $w \in \mathcal{W}, V_1 \in L_+^{q,1}$, clearly $V_2 \mapsto \mathbb{E}[-V_2 w^{\mathsf{T}} X] - \alpha_2(V_2, r)$ is concave and upper semicontinuous by the properties of support function. Let $w \in \mathcal{W}, V_2 \in L_+^q$. For each $V_1, V_1' \in L_+^{q,1}$ and $\lambda \in (0,1)$, we have

$$\begin{aligned}
&\beta_1\left(\lambda V_1 + (1-\lambda)V_1', \mathbb{E}[-(\lambda V_1 + (1-\lambda)V_1')w^{\mathsf{T}} X]\right) \\
&= \beta_1\left(\lambda V_1 + (1-\lambda)V_1', \lambda \mathbb{E}[-V_1 w^{\mathsf{T}} X] + (1-\lambda)\mathbb{E}[-V_1' w^{\mathsf{T}} X]\right) \\
&\geq \min\left\{\beta_1\left(V_1, \mathbb{E}[-V_1 w^{\mathsf{T}} X]\right), \beta_1\left(V_1', \mathbb{E}[-V_1' w^{\mathsf{T}} X]\right)\right\}
\end{aligned}$$

by the joint quasiconcavity of β_1. Hence, the function $V_1 \mapsto \beta_1\left(V_1, \mathbb{E}[-V_1 w^{\mathsf{T}} X]\right)$ is quasiconcave. We claim that this function is also weakly upper semicontinuous. To that end, let $(V_\alpha)_{\alpha \in I}$ be a weakly convergent net in $L_+^{q,1}$ with some index set I and limit V. Hence, $(\mathbb{E}[-V_\alpha w^{\mathsf{T}} X])_{\alpha \in I}$ converges to $\mathbb{E}[-V w^{\mathsf{T}} X]$. So $(V_\alpha, \mathbb{E}[-V_\alpha w^{\mathsf{T}} X])_{\alpha \in I}$ is a weakly convergent net in $L_+^{q,1} \times \mathbb{R}$ with limit $(V, \mathbb{E}[-V w^{\mathsf{T}} X])$. By Assumption 4.2, we get

$$\limsup_{\alpha \in I} \beta_1(V_\alpha, \mathbb{E}[-V_\alpha w^{\mathsf{T}} X]) \leq \beta_1(V, \mathbb{E}[-V w^{\mathsf{T}} X]).$$

Hence, the claim follows. Since β_1 is increasing and left-continuous in the second argument, it is also lower semicontinuous (indeed continuous by Assumption 4.2). Hence, for fixed $V_1 \in L_1^{q,1}, V_2 \in L_+^q$, the function $w \mapsto \beta_1\left(V_1, \mathbb{E}[-V_1 w^{\mathsf{T}} X]\right)$ is lower semicontinuous (indeed continuous) and quasiconvex on $\mathcal{W}$. Note that $\mathcal{W}$ is a convex and compact set. Finally, by Assumption 4.3, b is real-valued on $\mathcal{W} \times L_+^{q,1} \times L_+^q$. Therefore, by Sion's minimax theorem [16, Corollary 3.3] and property (c) in Definition 2.1, we may write

$$\begin{aligned}
p(r) &= \sup_{V_1 \in L_+^{q,1}, V_2 \in L_+^q} \inf_{w \in \mathcal{W}} b(w, V_1, V_2) \\
&= \sup_{V_1 \in L_+^{q,1}, V_2 \in L_+^q} \left(\inf_{w \in \mathcal{W}} \left(\beta_1(V_1, \mathbb{E}[-V_1 X]^{\mathsf{T}} w) + \mathbb{E}[-V_2 X]^{\mathsf{T}} w\right) - \alpha_2(V_2, r)\right) \\
&= \sup_{V_1, V_2 \in L_+^q} \left(\inf_{w \in \mathcal{W}} \left(\beta_1(V_1, \mathbb{E}[-V_1 X]^{\mathsf{T}} w) + \mathbb{E}[-V_2 X]^{\mathsf{T}} w\right) - \alpha_2(V_2, r)\right) \\
&= \sup_{V_1, V_2 \in L_+^q} \left(\inf_{w \in \mathcal{W}} \left(\inf\left\{t \in \mathbb{R} \mid \alpha_1^-(V_1, t) \geq \mathbb{E}[-V_1 X]^{\mathsf{T}} w\right\} + \mathbb{E}[-V_2 X]^{\mathsf{T}} w\right) - \alpha_2(V_2, r)\right) \\
&= \sup_{V_1, V_2 \in L_+^q} \left(h^P(V_1, V_2) - \alpha_2(V_2, r)\right),
\end{aligned}$$

where $h^P(V_1, V_2)$ is defined by (4.1). □

Remark 4.5 *In the proof of Proposition 4.4, the only role of Assumption 4.3 is to ensure that the function b has finite values so that Sion's minimax theorem can be applied. In cases where Assumption 4.3 is not valid, the proof of Proposition 4.4 can be seen as a heuristic argument to come up with a dual formulation of $(p(r))$. Thanks to Theorem 4.11 below, it turns out that the conclusion of Proposition 4.4 is still valid without Assumption 4.3.*

Let $V_1, V_2 \in L_+^q$. Note that $h^P(V_1, V_2)$ is the optimal value of a finite-dimensional optimization problem which is in general nonconvex due to the inequality constraint when ρ_1 is only assumed to be a quasiconvex risk

measure. However, under Assumption 4.1, as argued in the proof of [6, Proposition 6], the function $t \mapsto \alpha_1^-(V_1, t)$ is concave and the optimization problem in $h^P(V_1, V_2)$ becomes convex. Before proceeding further, we introduce an additional assumption related to the asymptotic behavior of α_1^-.

Assumption 4.6 *For each $V_1 \in L_+^q$, it holds*

$$\lim_{t\to\infty} \alpha_1^-(V_1, t) = +\infty.$$

To formulate the next proposition, let us define

$$\alpha_1^*(V_1, z) := \inf_{t\in\mathbb{R}} \left(tz - \alpha_1^-(V_1, t)\right) = \inf_{t\in\mathbb{R}} \left(tz - \alpha_1(V_1, t)\right), \quad V_1 \in L_+^q, z \in \mathbb{R}.$$

Note that $z \mapsto -\alpha_1^*(V_1, -z)$ is the conjugate function of α_1^- (and also of α_1) with respect to the second variable. Since $t \mapsto \alpha_1^-(V_1, t)$ is an increasing function, it is easy to check that $\alpha_1^*(V_1, z) = -\infty$ whenever $z < 0$. Moreover, under Assumption 4.1, for each $V_1 \in L_+^q$, the function $t \mapsto \alpha_1^-(V_1, t)$ is concave and upper semicontinuous so that Fenchel–Moreau theorem [17, Theorem 2.3.3] gives

$$\alpha_1^-(V_1, t) = \inf_{z\geq 0} \left(tz - \alpha_1^*(V_1, z)\right), \quad V_1 \in L_+^q, t \in \mathbb{R}. \tag{4.2}$$

The following special value of α_1^* will play an important role in the dual problem of $(\mathscr{P}(r))$:

$$\tilde{\alpha}_1(V_1) := \alpha_1^*(V_1, 1) = \inf_{t\in\mathbb{R}} \left(t - \alpha_1^-(V_1, t)\right), \quad V_1 \in L_+^q. \tag{4.3}$$

Using this quantity, let us define, for each $V_1, V_2 \in L_+^q$,

$$h^D(V_1, V_2) := \sup\left\{\tilde{\alpha}_1(xV_1) - y \mid x\mathbb{E}[V_1X] + \mathbb{E}[V_2X] \leq y\mathbf{1},\ x > 0,\ y \in \mathbb{R}\right\},$$

where the inequality constraint is understood in the componentwise manner. As the proof of the next result shows, the dual of the problem in $h^P(V_1, V_2)$ gives rise to $h^D(V_1, V_2)$.

Proposition 4.7 *Suppose that Assumption 4.1 and Assumption 4.6 hold, and let $V_1, V_2 \in L_+^q$. Then,*

$$h^P(V_1, V_2) = h^D(V_1, V_2).$$

Proof As noted above, the problem defining $h^P(V_1, V_2)$ is a convex optimization problem under Assumption 4.1. By Assumption 4.6, there exists $w^0 \in \mathcal{W}$ and $t^0 \in \mathbb{R}$ such that $\alpha_1^-(V_1, t^0) > \mathbb{E}[-V_1X]^\intercal w^0$ (indeed, for every $w^0 \in \mathcal{W}$ such $t^0 \in \mathbb{R}$ exists). Hence, Slater's condition holds for this convex optimization problem

and we have

$$\begin{aligned}
h^P(V_1,V_2) &= \sup_{\substack{x\geq 0,\\ y\in\mathbb{R}}} \inf_{\substack{w\in\mathbb{R}^n_+,t\in\mathbb{R}:\\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t+\mathbb{E}[-V_2X]^{\mathsf{T}}w + x(-\alpha_1^-(V_1,t)+\mathbb{E}[-V_1X]^{\mathsf{T}}w) + y(\mathbf{1}^{\mathsf{T}}w-1)\right)\\
&= \sup_{\substack{x\geq 0,\\ y\in\mathbb{R}}} \inf_{\substack{w\in\mathbb{R}^n_+,t\in\mathbb{R}:\\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t-x\alpha_1^-(V_1,t)+(x\mathbb{E}[-V_1X]+\mathbb{E}[-V_2X]+y\mathbf{1})^{\mathsf{T}}w - y\right)\\
&= \sup_{\substack{x>0,\\ y\in\mathbb{R}}} \inf_{\substack{w\in\mathbb{R}^n_+,t\in\mathbb{R}:\\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t-\alpha_1^-(xV_1,t)+(x\mathbb{E}[-V_1X]+\mathbb{E}[-V_2X]+y\mathbf{1})^{\mathsf{T}}w - y\right)\\
&= \sup_{\substack{x>0,\\ y\in\mathbb{R}}} \left\{\tilde{\alpha}_1(xV_1)-y \mid x\mathbb{E}[V_1X]+\mathbb{E}[V_2X]\leq y\mathbf{1}\right\}.
\end{aligned}$$

Let us justify each passage in the above derivation: the first equality is by strong duality for convex optimization, the second is by simple manipulations, the third excludes the case $x=0$ from further consideration since in this case the infimum yields $-\infty$, the fourth is by evaluating the infimum with respect to $w\in\mathbb{R}^n_+$ and $t\in\mathbb{R}$. Therefore, $h^P(V_1,V_2)=h^D(V_1,V_2)$. □

We are ready to prove the first main result of the paper, which establishes strong duality between $(\mathscr{P}(r))$ and a new problem, which we refer to as the dual problem of $(\mathscr{P}(r))$.

Theorem 4.8 *Suppose that Assumption 4.1, Assumption 4.2, Assumption 4.3, Assumption 4.6 hold, and consider the problem*

$$\begin{aligned}
&\text{maximize } \tilde{\alpha}_1(V_1)-\alpha_2(V_2,r)-y && (\mathscr{D}(r))\\
&\text{subject to } \mathbb{E}[V_1X]+\mathbb{E}[V_2X]\leq y\mathbf{1}\\
&V_1,V_2\in L^q_+,\ y\in\mathbb{R},
\end{aligned}$$

where the inequality constraint is understood in the componentwise sense. Then, $(\mathscr{P}(r))$ and $(\mathscr{D}(r))$ have the same optimal value $p(r)$.

Proof Combining Proposition 4.4 and Proposition 4.7, we obtain

$$\begin{aligned}
p(r) &= \sup_{V_1,V_2\in L^q_+} \left(h^D(V_1,V_2)-\alpha_2(V_2,r)\right)\\
&= \sup_{V_1,V_2\in L^q_+} \left(\sup\left\{\tilde{\alpha}_1(xV_1)-y \mid \mathbb{E}[xV_1X]+\mathbb{E}[V_2X]\leq y\mathbf{1},\ x>0,\ y\in\mathbb{R}\right\}-\alpha_2(V_2,r)\right)\\
&= \sup\left\{\tilde{\alpha}_1(xV_1)-\alpha_2(V_2,r)-y \mid \mathbb{E}[xV_1X]+\mathbb{E}[V_2X]\leq y\mathbf{1}, V_1,V_2\in L^q_+, x>0, y\in\mathbb{R}\right\}\\
&= \sup\left\{\tilde{\alpha}_1(V_1)-\alpha_2(V_2,r)-y \mid \mathbb{E}[V_1X]+\mathbb{E}[V_2X]\leq y\mathbf{1},\ V_1,V_2\in L^q_+,\ y\in\mathbb{R}\right\},
\end{aligned}$$

which coincides with the optimal value of $(\mathscr{D}(r))$. □

It is worth noting that the dual problem $(\mathscr{D}(r))$ is a convex optimization problem thanks to Assumption 4.1.

4.2. Second main result: establishing optimality

While Theorem 4.8 provides strong duality between $(\mathscr{P}(r))$ and $(\mathscr{D}(r))$, it does not make a statement on how to find an optimal portfolio $w^* \in \mathcal{W}$ for $(\mathscr{P}(r))$. The aim of Theorem 4.11, the second main result of the paper, is to find such w^* in relation to $(\mathscr{D}(r))$.

As a preparation, we recall some well-known concepts and facts from convex analysis. To that end, let us fix an arbitrary Hausdorff locally convex topological real vector space $\mathcal{X}$ with topological dual $\mathcal{X}^*$ and bilinear duality mapping $\langle \cdot, \cdot \rangle : \mathcal{X}^* \times \mathcal{X} \to \mathbb{R}$. For our purposes, the following special cases of $\mathcal{X}$ are particularly important:

(i) $\mathcal{X} = \mathbb{R}^n$ with the usual topology, which yields $\mathcal{X}^* = \mathbb{R}^n$ together with $\langle z, x \rangle = z^\mathsf{T} x$ for every $x, z \in \mathbb{R}^n$.

(ii) $\mathcal{X} = L^q$ with $q \in [1, +\infty)$ with the weak topology $\sigma(L^q, L^p)$, which yields $\mathcal{X}^* = L^p$ together with $\langle Y, U \rangle = \mathbb{E}[UY]$ for every $U \in L^q$, $Y \in L^p$. (Recall that p, q are conjugate exponents.)

(iii) $\mathcal{X} = L^\infty$ with the weak* topology $\sigma(L^\infty, L^1)$, which yields $\mathcal{X}^* = L^1$ together with $\langle Y, U \rangle = \mathbb{E}[UY]$ for every $U \in L^\infty$, $Y \in L^1$.

Consider a set $A \subseteq \mathcal{X}$. The function $I_A \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$I_A(x) = \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \in \mathcal{X} \setminus A, \end{cases}$$

is called the indicator function of A; note that A is a convex set if and only if I_A is a convex function. For a point $x \in A$, the convex cone

$$\mathcal{N}(A, x) \coloneqq \{z \in \mathcal{X}^* \mid \forall x' \in A \colon \langle z, x \rangle \geq \langle z, x' \rangle\}$$

is called the normal cone of A at x. Let $g \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be a function. Given $x \in \mathcal{X}$, the set

$$\partial g(x) \coloneqq \{z \in \mathcal{X}^* \mid \forall x' \in \mathcal{X} \colon g(x') \geq g(x) + \langle z, x' - x \rangle\}$$

is called the subdifferential of g at x. If A is a nonempty convex set, then by [17, Section 2.4], $\partial I_A(x) = \mathcal{N}_A(x)$ for every $x \in A$, and $\partial I_A(x) = \emptyset$ for every $x \in \mathcal{X} \setminus A$. The function $g^* \colon \mathcal{X}^* \to \bar{\mathbb{R}}$ defined by

$$g^*(z) \coloneqq \sup_{x \in \mathcal{X}} \left(\langle z, x \rangle - g(x) \right), \quad z \in \mathcal{X}^*$$

is called the conjugate function of g. We have

$$z \in \partial g(x) \quad \Leftrightarrow \quad x \in \partial g^*(z) \tag{4.4}$$

for every $x \in \mathcal{X}, z \in \mathcal{X}^*$ such that g is lower semicontinuous at x. If A is a nonempty closed convex set, then it is well-known that $\sigma_A \coloneqq (I_A)^*$ is the support function of A defined by

$$\sigma_A(z) \coloneqq \sup_{x \in A} \langle z, x \rangle, \quad z \in \mathcal{X}^*.$$

From the above definitions, it is clear that, for a point $x \in A$, we have

$$\mathcal{N}(A, x) = \{z \in \mathcal{X}^* \mid \sigma_A(z) = \langle z, x \rangle\}. \tag{4.5}$$

Consider the problem of minimizing g over A. Suppose that g is proper convex and let $x \in A$ with $g(x) < +\infty$. By Pshenichnyi–Rockafellar theorem [17, Theorem 2.9.1], if

$$\partial g(x) \cap -\mathcal{N}(A, x) \neq \emptyset, \tag{4.6}$$

then x is a minimizer of g over A, that is, $g(x) = \inf_{x' \in A} g(x')$; the converse holds if g is continuous at x.

The next lemma is devoted to the calculation of a certain subdifferential that is relevant to $(\mathscr{P}(r))$.

Lemma 4.9 *Suppose that Assumption 4.1 holds. Let $w \in \mathbb{R}^n$ be such that $g_1(w) = \rho_1(w^\mathsf{T} X) < +\infty$. Then,*

$$\left\{\mathbb{E}\left[-VX\right] \mid \beta_1(V, -w^\mathsf{T}\mathbb{E}[VX]) = g_1(w),\ V \in L^q_+\right\} \subseteq \partial g_1(w).$$

Proof Thanks to Assumption 4.1, the function g_1 defined by (3.3) is convex. Let us define a continuous linear operator $L\colon \mathbb{R}^n \to L^p$ by

$$Lw' \coloneqq X^\mathsf{T} w', \quad w' \in \mathbb{R}^n.$$

Then, its adjoint operator $L^*\colon L^q \to \mathbb{R}^n$ is given by

$$L^* V = \mathbb{E}\left[VX\right], \quad V \in L^q.$$

Since $g_1 = \rho_1 \circ L$ and ρ_1 is finite at $w^\mathsf{T} X$, by [17, Theorem 2.8.3(iii)], we have

$$\left\{L^* V \mid V \in \partial\rho_1(w^\mathsf{T} X)\right\} = \left\{\mathbb{E}\left[VX\right] \mid V \in \partial\rho_1(w^\mathsf{T} X)\right\} \subseteq \partial g_1(w). \tag{4.7}$$

On the other hand, the subdifferential of the convex lower semicontinuous function ρ_1 at a point $Y \in L^p$ is given by

$$\partial\rho_1(Y) = \left\{-V \mid \beta_1(V, \mathbb{E}\left[-VY\right]) = \rho_1(Y),\ V \in L^q_+\right\}, \tag{4.8}$$

that is, it is the set of maximizers in the dual representation (3.2). Taking $Y = w^\mathsf{T} X$ in (4.8) and combining it with (4.7) yield the claim of the lemma. □

Let $L^q_{++} \coloneqq \{V \in L^q_+ \mid \mathbb{P}\{V = 0\} = 0\}$. The next assumption is a constraint qualification for $(\mathscr{D}(r))$.

Assumption 4.10 *There exist $V_1, V_2 \in L^q_{++}$ such that $\tilde{\alpha}_1(V_1) \in \mathbb{R}$, $\alpha_2(V_2, r) \in \mathbb{R}$.*

We are ready to prove the second main theorem of the paper, which establishes for $(\mathscr{P}(r))$ the optimality of a Lagrange multiplier associated to $(\mathscr{D}(r))$.

Theorem 4.11 *Suppose that Assumption 4.1, Assumption 4.2, Assumption 4.6, Assumption 4.10 hold, and there exists an optimal solution $(V_1^*, V_2^*, y^*) \in L^q_+ \times L^q_+ \times \mathbb{R}$ for $(\mathscr{D}(r))$. Then, there exists an optimal Lagrange multiplier $w^* \in \mathbb{R}^n$ associated to the inequality constraint of $(\mathscr{D}(r))$. Moreover, every $w^* \in \mathbb{R}^n$ that is the Lagrange multiplier of the equality constraint of $(\mathscr{D}(r))$ at optimality is an optimal solution for $(\mathscr{P}(r))$, and $(\mathscr{P}(r))$ and $(\mathscr{D}(r))$ have the same optimal value $p(r)$.*

Proof Let $(V_1^*, V_2^*, y^*) \in L^q_+ \times L^q_+ \times \mathbb{R}$ be an optimal solution for $(\mathscr{D}(r))$. Let us denote by $d(r)$ the optimal value of $(\mathscr{D}(r))$. By Assumption 4.10, Slater's condition holds, that is, there exist $V_1, V_2 \in L^q_{++}$, $y \in \mathbb{R}$ such that $\tilde{\alpha}_1(V_1) \in \mathbb{R}$, $\alpha_2(V_2, r) \in \mathbb{R}$ and

$$\max_{i \in \{1,\dots,n\}} \left(\mathbb{E}\left[V_1 X_i\right] + \mathbb{E}\left[V_2 X_i\right]\right) < y$$

since we may simply take $y := \max_{i\in\{1,\dots,n\}} \left(\mathbb{E}\left[V_1X_i\right] + \mathbb{E}\left[V_2X_i\right]\right) + 1$. Hence, by [4, Corollary 4.8], there is zero duality gap between $(\mathscr{D}(r))$ and its Lagrange dual problem, and we may write

$$\begin{aligned} d(r) &= \inf_{w\in\mathbb{R}^n_+} \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(\tilde{\alpha}_1(V_1) - \alpha_2(V_2,r) - y - w^{\mathsf{T}}\left(\mathbb{E}\left[V_1X\right] + \mathbb{E}\left[V_2X\right] - y\mathbf{1}\right)\right) \\ &= \inf_{w\in\mathbb{R}^n_+} \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(\tilde{\alpha}_1(V_1) - \alpha_2(V_2,r) - y + \mathbb{E}[-V_1w^{\mathsf{T}}X] + \mathbb{E}[-V_2w^{\mathsf{T}}X] + yw^{\mathsf{T}}\mathbf{1}\right). \end{aligned} \tag{4.9}$$

Moreover, [4, Corollary 4.8] also ensures that there exists an optimal Lagrange multiplier $w^* \in \mathbb{R}^n$ so that

$$d(r) = \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(\tilde{\alpha}_1(V_1) - \alpha_2(V_2,r) - y + \mathbb{E}[-V_1(w^*)^{\mathsf{T}}X] + \mathbb{E}[-V_2(w^*)^{\mathsf{T}}X] + y(w^*)^{\mathsf{T}}\mathbf{1}\right),$$

and (V_1^*, V_2^*, y^*) is an optimal solution of the above concave maximization problem.

Let $w \in \mathbb{R}^n_+$. Note that the inner (maximization) problem in (4.9) is easily separated into three terms as

$$\sup_{V_1\in L^q_+} \left(\tilde{\alpha}_1(V_1) + \mathbb{E}[-V_1w^{\mathsf{T}}X]\right) + \sup_{V_2\in L^q_+} \left(-\alpha_2(V_2,r) + \mathbb{E}[-V_2w^{\mathsf{T}}X]\right) + \sup_{y\in\mathbb{R}} y(w^{\mathsf{T}}\mathbf{1} - 1). \tag{4.10}$$

From the last term in (4.10), it follows immediately that

$$\sup_{y\in\mathbb{R}} y(w^{\mathsf{T}}\mathbf{1} - 1) = I_{\mathcal{W}}(w).$$

In particular, if $w = w^*$, then we must have $w^* \in \mathcal{W}$. It is also easy to check that

$$y^*\mathbf{1} \in \mathcal{N}(\mathcal{W}, w^*). \tag{4.11}$$

For the first term in (4.10), note that

$$\begin{aligned} \rho_1(w^{\mathsf{T}}X) &= \sup_{V_1\in L^q_+} \beta_1\left(V_1, \mathbb{E}[-V_1w^{\mathsf{T}}X]\right) \\ &= \sup_{V_1\in L^q_+} \inf\left\{t\in\mathbb{R} \mid \alpha_1^-(V_1,t) \geq \mathbb{E}[-V_1w^{\mathsf{T}}X]\right\} \\ &= \sup_{V_1\in L^q_+, x\geq 0} \inf_{\substack{t\in\mathbb{R}: \\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t + x\mathbb{E}[-V_1w^{\mathsf{T}}X] - x\alpha_1^-(V_1,t)\right) \\ &= \sup_{V_1\in L^q_+, x\geq 0} \inf_{\substack{t\in\mathbb{R}: \\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t + \mathbb{E}[-xV_1w^{\mathsf{T}}X] - \alpha_1^-(xV_1,t)\right) \\ &= \sup_{V_1\in L^q_+} \inf_{\substack{t\in\mathbb{R}: \\ \alpha_1^-(V_1,t)\in\mathbb{R}}} \left(t + \mathbb{E}[-V_1w^{\mathsf{T}}X] - \alpha_1^-(V_1,t)\right) \\ &= \sup_{V_1\in L^q_+} \left(\tilde{\alpha}_1(V_1) + \mathbb{E}[-V_1w^{\mathsf{T}}X]\right). \end{aligned}$$

The steps of this calculation are justified by following the same arguments as in the proof of Proposition 4.7, hence we omit this justification for brevity. In particular, when $w = w^*$, by the optimality of V_1^* and property (c) of the definition of maximal risk function, we have

$$g_1(w^*) = \rho_1((w^*)^{\mathsf{T}}X) = \tilde{\alpha}_1(V_1^*) + \mathbb{E}[-V_1^*(w^*)^{\mathsf{T}}X] = \beta_1\left(V_1^*, \mathbb{E}[-V_1^*(w^*)^{\mathsf{T}}X]\right).$$

By Lemma 4.9, it follows that

$$\mathbb{E}[-V_1^* X] \in \partial g_1(w^*). \tag{4.12}$$

For the second term in (4.10), we first note that $V_2 \mapsto \alpha_2(V_2, r)$ is closely related to the support function of the closed convex set $\mathscr{A}_2^r$; indeed, we have

$$\alpha_2(V_2, r) = \sup_{Y \in \mathscr{A}_2^r} \mathbb{E}[-V_2 Y] = \sigma_{\mathscr{A}_2^r}(-V_2), \quad V_2 \in L_+^q.$$

Hence, by the conjugate duality between indicator function and support function, we have

$$\sup_{V_2 \in L_+^q} \left(-\alpha_2(V_2, r) + \mathbb{E}[-V_2 w^{\mathsf{T}} X]\right) = I_{\mathscr{A}_2^r}(w^{\mathsf{T}} X) = I_{\{g_2 \le r\}}(w),$$

where $\{g_2 \le r\} \coloneqq \{w \in \mathbb{R}^n \mid g_2(w) \le r\}$. In particular, when $w = w^*$, by the first-order condition, we have

$$-(w^*)^{\mathsf{T}} X \in \partial \alpha_2(V_2^*, r), \tag{4.13}$$

where the subdifferential is with respect to the first variable. Hence, by (4.4), (4.13) is equivalent to

$$-(w^*)^{\mathsf{T}} X \in -\partial \sigma_{\mathscr{A}_2^r}(-V_2^*)$$

as well as to

$$-V_2^* \in \partial I_{\mathscr{A}_2^r}((w^*)^{\mathsf{T}} X).$$

In particular, $\partial I_{\mathscr{A}_2^r}((w^*)^{\mathsf{T}} X) \neq \emptyset$ so that $(w^*)^{\mathsf{T}} X \in \mathscr{A}_2^r$, that is, $g_2(w^*) \le r$, and

$$-V_2^* \in \partial I_{\mathscr{A}_2^r}((w^*)^{\mathsf{T}} X) = \mathcal{N}(\mathscr{A}_2^r, (w^*)^{\mathsf{T}} X).$$

So

$$\sup_{w \in \mathbb{R}^n \,:\, g_2(w) \le r} w^{\mathsf{T}} \mathbb{E}[-V_2^* X] = \sup_{w \in \mathbb{R}^n \,:\, g_2(w) \le r} \mathbb{E}[-V_2^* w^{\mathsf{T}} X] \le \sup_{Y \in \mathscr{A}_2^r} \mathbb{E}[-V_2^* Y] = \mathbb{E}[-V_2^* (w^*)^{\mathsf{T}} X],$$

which implies that the inequality in the middle is indeed an equality. Therefore, the equality of the first and last quantities yields

$$\mathbb{E}[-V_2^* X] \in \mathcal{N}(\{g_2 \le r\}, w^*). \tag{4.14}$$

Combining the results for the three terms of (4.10) for a generic $w \in \mathbb{R}_+^n$, we see that

$$d(r) = \inf_{w \in \mathbb{R}_+^n} \left(\rho_1(w^{\mathsf{T}} X) + I_{\mathscr{A}_2^r}(w^{\mathsf{T}} X) + I_{\mathcal{W}}(w)\right) = \inf\left\{\rho_1(w^{\mathsf{T}} X) \mid \rho_2(w^{\mathsf{T}} X) \le r,\ w \in \mathcal{W}\right\} = p(r),$$

establishing the strong duality between $(\mathscr{P}(r))$ and $(\mathscr{D}(r))$.

Note that the inequality constraint in $(\mathscr{D}(r))$ ensures that

$$w^{\mathsf{T}} \mathbb{E}[V_1^* X] \le w^{\mathsf{T}} (\mathbb{E}[-V_2^* X] + y^* \mathbf{1}), \quad w \in \mathcal{W}. \tag{4.15}$$

Moreover, the complementary slackness condition for this constraint yields

$$(w^*)^{\mathsf{T}} \mathbb{E}[V_1^* X] = (w^*)^{\mathsf{T}} (\mathbb{E}[-V_2^* X] + y^* \mathbf{1}). \tag{4.16}$$

On the other hand, bringing together (4.11) and (4.14) gives

$$\begin{aligned}\mathbb{E}[-V_2^*X]+y^*\mathbf{1}\in\mathcal{N}(\{g_2\le r\},w^*)+\mathcal{N}(\mathcal{W},w^*) &= \partial I_{\{g_2\le r\}}(w^*)+\partial I_{\mathcal{W}}(w^*)\\ &\subseteq \partial(I_{\{g_2\le r\}}+I_{\mathcal{W}})(w^*)\\ &= \partial I_{\{g_2\le r\}\cap\mathcal{W}}(w^*)\\ &= \mathcal{N}(\{g_2\le r\}\cap\mathcal{W},w^*).\end{aligned}$$

In the above calculation, only the passage to the second line is nontrivial and it is justified by the rules of subdifferential calculus; see, for instance, [15, Theorem 23.8]. Hence, by (4.15) and (4.16), we have

$$\sigma_{\{g_2\le r\}\cap\mathcal{W}}(\mathbb{E}[V_1^*X])\le\sigma_{\{g_2\le r\}\cap\mathcal{W}}(\mathbb{E}[-V_2^*X]+y^*\mathbf{1})=(w^*)^{\mathsf{T}}\left(\mathbb{E}[-V_2^*X]+y^*\mathbf{1}\right)=(w^*)^{\mathsf{T}}\mathbb{E}[V_1^*X].$$

Hence, by (4.5), we conclude that $\mathbb{E}[V_1^*X]\in\mathcal{N}(\{g_2\le r\}\cap\mathcal{W},w^*)$, that is,

$$\mathbb{E}[-V_1^*X]\in-\mathcal{N}(\{g_2\le r\}\cap\mathcal{W},w^*).$$

Combining this with (4.12), we obtain

$$\mathbb{E}[-V_1^*X]\in\partial g_1(w^*)\cap-\mathcal{N}(\{g_2\le r\}\cap\mathcal{W},w^*).$$

By (4.6), this implies that w^* is a minimizer of g_1 over $\{g_2\le r\}\cap\mathcal{W}$, that is, w^* is an optimal solution of $(\mathscr{P}(r))$. □

Remark 4.12 *It should be noted that we do not work under Assumption 4.3 in Theorem 4.11, hence the strong duality established by Theorem 4.8 is not taken for granted; instead we re-establish strong duality in Theorem 4.11. Then, the reader might naturally question the need for Theorem 4.8. As noted in Remark 4.5, in the absence of Assumption 4.3, the arguments in the proof of Proposition 4.4 (and Theorem 4.8) can be seen as a heuristic way to derive the dual problem $(\mathscr{D}(r))$. Theorem 4.11 provides the formal justification of this heuristic approach without using Sion's minimax theorem.*

5. Analysis of the problem when the principle risk measure is quasiconvex

In this section, we remove Assumption 4.1 and study $(\mathscr{P}(r))$ with ρ_1 being a quasiconvex risk measure.

Recalling (2.1), we may write

$$\begin{aligned}p(r)&=\inf\left\{\rho_1(w^{\mathsf{T}}X)\mid\rho_2(w^{\mathsf{T}}X)\le r,\ w\in\mathcal{W}\right\}\\ &=\inf\left\{t\in\mathbb{R}\mid w^{\mathsf{T}}X\in\mathscr{A}_1^t,\ w^{\mathsf{T}}X\in\mathscr{A}_2^r,\ w\in\mathcal{W}\right\}\\ &=\inf_{t\in\mathbb{R}}\left(t+f(t,r)\right)=\inf_{t\in\mathbb{R}:f(t,r)<+\infty}\left(t+f(t,r)\right),\end{aligned}\tag{5.1}$$

where

$$f(t,r):=\inf_{w\in\mathcal{W}}\left(I_{\mathscr{A}_1^t}(w^{\mathsf{T}}X)+I_{\mathscr{A}_2^r}(w^{\mathsf{T}}X)\right),\quad t\in\mathbb{R}.\tag{5.2}$$

Note that, for each $t\in\mathbb{R}$, $f(t,r)$ is closely related to the feasibility problem

$$\text{Find } w\in\mathcal{W} \text{ such that } w^{\mathsf{T}}X\in\mathscr{A}_1^t\cap\mathscr{A}_2^r.\tag{$\mathscr{F}^P(t,r)$}$$

Indeed, if there exists $w \in \mathbb{R}^n$ solving $(\mathscr{F}^P(t,r))$, then $f(t,r) = 0$; otherwise, $f(t,r) = +\infty$. The expression in (5.2) formulates $(\mathscr{F}^P(t,r))$ as an optimization problem whose optimal value is $f(t,r)$; this problem is convex because $\mathscr{A}_1^t, \mathscr{A}_2^r$ are convex sets. Hence, in view of (5.1), the quasiconvex portfolio optimization problem $(\mathscr{P}(r))$ is characterized by a family of convex optimization problems; this is a well-known paradigm in quasiconvex programming as discussed recently in [1, Section 2.1].

We introduce an analogue of Assumption 4.10 that will be needed in recovering a solution for $(\mathscr{F}^P(t,r))$.

Assumption 5.1 *Given $t \in \mathbb{R}$, there exist $V_1, V_2 \in L^q_{++}$ such that $\alpha_1(V_1,t) \in \mathbb{R}$, $\alpha_2(V_2,r) \in \mathbb{R}$.*

In the next theorem, following a similar path as in Section 4 (see Theorem 4.8, Theorem 4.11), we provide a dual formulation of $f(t,r)$ and a method to calculate a solution for $(\mathscr{F}^P(t,r))$.

Theorem 5.2 *Let $t \in \mathbb{R}$ and consider the problem*

$$\begin{aligned} &\text{maximize} \quad -\alpha_1(V_1,t) - \alpha_2(V_2,r) - y && (\mathscr{F}^D(t,r)) \\ &\text{subject to} \quad \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \leq y\mathbf{1} \\ &\qquad V_1, V_2 \in L^q_+, \ y \in \mathbb{R}. \end{aligned}$$

(i) Then, $(\mathscr{F}^P(t,r))$ and $(\mathscr{F}^D(t,r))$ have the same optimal value $f(t,r)$.

(ii) Suppose that Assumption 5.1 holds for t and there exists an optimal solution (V_1^t, V_2^t, y^t) for $(\mathscr{F}^D(t,r))$. Then, there exists an optimal Lagrange multiplier $w^t \in \mathbb{R}^n$ associated to the inequality constraint of $(\mathscr{F}^D(t,r))$. Moreover, every $w^t \in \mathbb{R}^n$ that is the Lagrange multiplier of the inequality constraint of $(\mathscr{F}^D(t,r))$ at optimality is an optimal solution for $(\mathscr{F}^P(t,r))$.

Proof We first prove (i) under the additional assumption that $\alpha_1(V_1,t) \in \mathbb{R}, \alpha_2(V_2,r) \in \mathbb{R}$ for all $V_1, V_2 \in L^q_+$. Since $\mathscr{A}_1^t, \mathscr{A}_2^r$ are closed convex subsets of L^p, similar to the proof of Proposition 4.4, we have

$$\begin{aligned} f(t,r) &= \inf_{w \in \mathcal{W}} \left(I_{\mathscr{A}_1^t}(w^{\mathsf{T}} X) + I_{\mathscr{A}_2^r}(w^{\mathsf{T}} X) \right) \\ &= \inf_{w \in \mathcal{W}} \left(\sup_{V_1 \in L^q_+} \left(\mathbb{E}[-V_1 w^{\mathsf{T}} X] - \alpha_1(V_1,t) \right) + \sup_{V_2 \in L^q_+} \left(\mathbb{E}[-V_2 w^{\mathsf{T}} X] - \alpha_2(V_1,r) \right) \right) \\ &= \inf_{w \in \mathcal{W}} \sup_{V_1 \in L^q_+, V_2 \in L^q_+} v_{t,r}(w, V_1, V_2), \end{aligned}$$

where

$$v_{t,r}(w, V_1, V_2) := \mathbb{E}[-V_1 w^{\mathsf{T}} X] + \mathbb{E}[-V_2 w^{\mathsf{T}} X] - \alpha_1(V_1,t) - \alpha_2(V_2,r), \quad w \in \mathcal{W}, V_1 \in L^q_+, V_2 \in L^q_+.$$

By the properties of support function, it is clear that $(V_1, V_2) \mapsto v_{t,r}(w, V_1, V_2)$ is concave and upper semicontinuous for fixed $w \in \mathcal{W}$. On the other hand, for fixed $(V_1, V_2) \in L^q_+ \times L^q_+$, the function $w \mapsto v_{t,r}(w, V_1, V_2)$ is continuous and affine, hence lower semicontinuous and convex. Since $\mathcal{W}$ is a convex compact set and $v_{t,r}$ has finite values thanks to our additional assumption, we may apply the standard minimax theorem [16, Corollary 3.3]

and get

$$\begin{aligned} f(t,r) &= \sup_{V_1\in L^q_+, V_2\in L^q_+} \inf_{w\in\mathcal{W}} v_{t,r}(w,V_1,V_2) \\ &= \sup_{V_1\in L^q_+, V_2\in L^q_+} \left(\inf\left\{(\mathbb{E}[-V_1X]+\mathbb{E}[-V_2X])^{\mathsf{T}} w \mid \mathbf{1}^{\mathsf{T}}w=1,\ w\in\mathbb{R}^n_+\right\} - \alpha_1(V_1,t)-\alpha_2(V_2,r)\right). \end{aligned}$$

Note that the inner minimization problem is a finite-dimensional linear optimization problem with nonempty feasible region. Hence, by linear programming duality, we may pass to its dual formulation, which yields

$$\begin{aligned} f(t,r) &= \sup_{V_1\in L^q_+, V_2\in L^q_+} \left(\sup\left\{-y \mid \mathbb{E}[V_1X]+\mathbb{E}[V_2X]\le y\mathbf{1},\ y\in\mathbb{R}\right\} - \alpha_1(V_1,t)-\alpha_2(V_2,r)\right) \\ &= \sup\left\{-\alpha_1(V_1,t)-\alpha_2(V_2,r)-y \mid \mathbb{E}[V_1X]+\mathbb{E}[V_2X]\le y\mathbf{1},\ V_1,V_2\in L^q_+,\ y\in\mathbb{R}\right\}, \end{aligned}$$

which coincides with the optimal value of $(\mathscr{F}^D(t,r))$.

Next, we prove (i) without the additional assumption and we also prove (ii). Let $\tilde{f}(t,r)$ be the optimal value of $(\mathscr{F}^D(t,r))$. Let $(V_1^t,V_2^t,y^t)\in L^q_+\times L^q_+\times\mathbb{R}$ be an optimal solution for $(\mathscr{F}^D(t,r))$. By Assumption 5.1, there exist $V_1,V_2\in L^q_{++}$ such that $\alpha_1(V_1,t)\in\mathbb{R}$ and $\alpha_2(V_2,r)\in\mathbb{R}$. Similar to the proof of Theorem 4.11, it follows that Slater's condition holds for $(\mathscr{F}^D(t,r))$; hence, by strong duality for convex optimization, there exists an optimal Lagrange multiplier $w^t\in\mathbb{R}^n$ such that

$$\begin{aligned} \tilde{f}(t,r) &= \inf_{w\in\mathbb{R}^n} \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(-\alpha_1(V_1,t)-\alpha_2(V_2,r)-y-w^{\mathsf{T}}\left(\mathbb{E}[V_1X]+\mathbb{E}[V_2X]-y\mathbf{1}\right)\right) \\ &= \inf_{w\in\mathbb{R}^n} \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(-\alpha_1(V_1,t)-\alpha_2(V_2,r)-y-\mathbb{E}[V_1w^{\mathsf{T}}X]+\mathbb{E}[V_2w^{\mathsf{T}}X]+yw^{\mathsf{T}}\mathbf{1}\right) \\ &= \sup_{V_1,V_2\in L^q_+, y\in\mathbb{R}} \left(-\alpha_1(V_1,t)-\alpha_2(V_2,r)-y-\mathbb{E}[V_1(w^t)^{\mathsf{T}}X]-\mathbb{E}[V_2(w^t)^{\mathsf{T}}X]+y(w^t)^{\mathsf{T}}\mathbf{1}\right) \\ &= \sup_{V_1\in L^q_+}\left(-\alpha_1(V_1,t)-\mathbb{E}[V_1(w^t)^{\mathsf{T}}X]\right)+\sup_{V_2\in L^q_+}\left(-\alpha_2(V_2,r)-\mathbb{E}[V_2(w^t)^{\mathsf{T}}X]\right)+\sup_{t\in\mathbb{R}} y\left((w^t)^{\mathsf{T}}\mathbf{1}-1\right). \end{aligned}$$

Then, following similar arguments as in the proof of Theorem 4.11, it can be checked that $\tilde{f}(t,r)=f(t,r)$ so that (i) holds without the additional assumption. Moreover, it can be checked that

$$w^t\in\mathcal{W},\quad y^t\mathbf{1}\in\mathcal{N}(\mathcal{W},w^t),\quad -(w^t)^{\mathsf{T}}X\in\partial\alpha_1(V_1^t,t),\quad (w^t)^{\mathsf{T}}X\in\partial\alpha_2(V_2^t,r).$$

These imply that

$$\mathbb{E}[-V_1^tX]\in-\mathcal{N}(\{g_1\le t\},w^t),\quad \mathbb{E}[-V_2^tX]\in-\mathcal{N}(\{g_2\le r\},w^t),$$

and finally we obtain

$$\mathbb{E}[-V_1^tX]\in-\mathcal{N}(\{g_1\le t\}\cap\{g_2\le r\}\cap\mathcal{W},w^t).$$

Hence, we conclude that w^t solves the feasibility problem $(\mathscr{F}^P(t,r))$ and (ii) holds. □

Let $\operatorname{dom} g_1 := \{w\in\mathbb{R}^n \mid g_1(w)\in\mathbb{R}\}$. The next assumption will be useful when devising a method to find an approximately optimal solution for $(\mathscr{P}(r))$.

Assumption 5.3 *It holds* $\mathbf{1} \in \operatorname{dom} g_1$ *and* $\operatorname{dom} g_1 \cap \{g_2 \leq r\} \cap \mathcal{W} \neq \emptyset$. *In other words,* $g_1(\mathbf{1}) = \rho_1(\mathbf{1}^{\mathsf{T}} X) \in \mathbb{R}$ *and there exists* $w^0 \in \mathcal{W}$ *such that* $\rho_1((w^0)^{\mathsf{T}} X) \in \mathbb{R}$ *and* $\rho_2((w^0)^{\mathsf{T}} X) \leq r$.

Finally, we discuss a simple method to solve $(\mathscr{P}(r))$ with the help of Theorem 5.2. To that end, note that

$$p(r) = \inf \{g_1(w) \mid g_2(w) \leq r, w \in \mathcal{W}\}.$$

Since $\{g_2 \leq r\} \cap \mathcal{W}$ is a convex compact set and g_1 is a lower semicontinuous function, $(\mathscr{P}(r))$ has an optimal solution, that is, there exists $w^* \in \mathcal{W}$ such that $g_2(w^*) \leq r$ and

$$g_1(w^*) = p(r).$$

Moreover, under Assumption 5.3, we also have

$$p(r) = g_1(w^*) \leq g_1(w^0) < +\infty.$$

On the other hand, by the monotonicity of ρ_1 and Assumption 5.3,

$$-\infty < g_1(\mathbf{1}) = \rho_1(\mathbf{1}^{\mathsf{T}} X) \leq \rho_1((w^*)^{\mathsf{T}} X) = g_1(w^*) = p(r).$$

Hence, $p(r) \in \mathbb{R}$ with finite upper bound $u_1 \coloneqq g_1(w^0)$ and finite lower bound $\ell_1 \coloneqq g_1(\mathbf{1})$. Let $\varepsilon > 0$. Using these bounds, the well-known bisection algorithm (see [1, Section 3]) can be employed to find an ε-optimal solution for $(\mathscr{P}(r))$ as follows. At each iteration $k \in \mathbb{N}$, we start with $\ell_k, u_k \in \mathbb{R}$ such that $\ell_k \leq p(r) \leq u_k$ and we let

$$t_k \coloneqq \frac{\ell_k + u_k}{2}.$$

Then, under Assumption 4.10, we solve the feasibility problem $(\mathscr{F}^P(t_k, r))$, that is, we calculate $f(t_k, r)$. If $f(t_k, r) = 0$, then we have $\ell_k \leq p(r) \leq t_k$, in which case we proceed to the next iteration using $\ell_{k+1} \coloneqq \ell_k$ and $u_{k+1} \coloneqq t_k$. Otherwise, we have $f(t_k, r) = +\infty$ and $t_k \leq p(r) \leq u_k$, in which case we proceed to the next iteration using $\ell_{k+1} \coloneqq t_k$ and $u_{k+1} \coloneqq u_k$. We stop this procedure at the first iteration number K for which $u_K - \ell_K \leq \varepsilon$. It can be checked that

$$K \leq \left\lceil \log_2 \left(\frac{g_1(w^0) - g_1(\mathbf{1})}{\varepsilon} \right) \right\rceil$$

so that the algorithm stops after finitely many iterations. Then, we may apply Theorem 5.2 and find an optimal Lagrange multiplier w^{t_K}, which also solves $(\mathscr{F}^P(t_K, r))$. Hence, $w^{t_K} \in \mathcal{W}$, $g_2(w^{t_K}) \leq r$ and

$$p(r) \leq g_1(w^{t_K}) \leq t_K \leq p(r) + \varepsilon,$$

which shows that w^{t_K} is an ε-optimal solution for $(\mathscr{P}(r))$.

6. Examples

In this section, we consider some well-known classes of quasiconvex risk measures as special cases of ρ_1 and ρ_2.

6.1. Convex risk measures

Let ρ be a lower semicontinuous convex risk measure on L^p with $\rho(0) \in \mathbb{R}$ and acceptance set family $(\mathscr{A}^t)_{t\in\mathbb{R}}$. Since ρ is translative, its acceptance set $\mathscr{A} := \mathscr{A}^0$ at level 0 determines ρ completely. As a result, the dual representation in Proposition 2.2 reduces to a simpler form which we derive here for the convenience of the reader. Let $V \in L^q_+ \setminus \{0\}$ and $t \in \mathbb{R}$. Then, by the translativity of ρ,

$$\alpha(V,t) = \sup_{Y\in L^p\colon \rho(Y)\le t} \mathbb{E}[-VY] = \sup_{Y\in L^p\colon \rho(Y+t)\le 0} \mathbb{E}[-VY] = \sup_{Y\in L^p\colon Y+t\in\mathscr{A}} \mathbb{E}[-VY] = \gamma(V) + t\mathbb{E}[V],$$

where

$$\gamma(V) := \sup_{Y\in\mathscr{A}} \mathbb{E}[-VY].$$

The function $\gamma\colon L^q_+\setminus\{0\} \to \bar{\mathbb{R}}$ is called the minimal penalty function of ρ in the sense of convex risk measures (not to be confused with the minimal penalty function α in the sense of quasiconvex risk measures); it follows from the definition that γ is convex and lower semicontinuous. Since $V \neq 0$, we have $\mathbb{E}[V] > 0$. From this and the above expression for α, it is evident that $t \mapsto \alpha(V,t)$ is concave (indeed affine) and continuous with

$$\lim_{t\to\infty} \alpha^-(V,t) = \lim_{t\to\infty} \alpha(V,t) = +\infty$$

(see Assumption 4.6); and $t \mapsto \alpha(V,t)$ has its true inverse given by

$$\beta(V,s) = \frac{s-\gamma(V)}{\mathbb{E}[V]}, \quad s \in \mathbb{R}.$$

Moreover, for each $a \in \mathbb{R}$, the set

$$\Big\{(V,s) \in L^{q,1}_+ \times \mathbb{R} \mid \beta(V,s) \ge a\Big\} = \Big\{(V,s) \in L^{q,1}_+ \times \mathbb{R} \mid \gamma(V) + a\mathbb{E}[V] - s \le 0\Big\}$$

is closed by the lower semicontinuity of γ; therefore, β is jointly upper semicontinuous on $L^{q,1}_+ \times \mathbb{R}$ (see Assumption 4.2). On the other hand, we have

$$\tilde{\alpha}(V) := \inf_{t\in\mathbb{R}} (t - \alpha(V,t)) = \inf_{t\in\mathbb{R}} ((1-\mathbb{E}[V])t - \gamma(V)) = \begin{cases} -\gamma(V) & \text{if } \mathbb{E}[V] = 1, \\ -\infty & \text{else.} \end{cases}$$

Now suppose that, in $(\mathscr{P}(r))$, both ρ_1 and ρ_2 are convex risk measures with respective minimal penalty functions γ_1 and γ_2 in the sense of convex risk measures. Then, by the above discussion, Assumption 4.1, Assumption 4.2 and Assumption 4.6 hold, and the dual problem $(\mathscr{D}(r))$ can be rewritten as

$$\begin{aligned} \text{maximize} \quad & -\gamma_1(V_1) - \gamma_2(V_2) - r\mathbb{E}[V_2] - y \\ \text{subject to} \quad & \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \le y\mathbf{1} \\ & \mathbb{E}[V_1] = 1 \\ & V_1, V_2 \in L^q_+, \ y \in \mathbb{R}. \end{aligned}$$

Example 6.1 *For this example, we assume that $p = +\infty$ (and $q = 1$). For each $j \in \{1,2\}$, let us suppose that ρ_j is the* entropic risk measure *with risk aversion parameter $r_j > 0$ (see, for instance, [7, Example 4.34]), that is,*

$$\rho_j(Y) = \frac{1}{r_j} \log \mathbb{E}\left[e^{-r_j Y}\right], \quad Y \in L^\infty.$$

In this case, it is well-known that γ_j is the relative entropy function given by

$$\gamma_j(V) = \frac{1}{r_j}\mathbb{E}\left[V \log\left(\frac{V}{\mathbb{E}[V]}\right)\right] = \frac{1}{r_j}\left(\mathbb{E}[V\log(V)] - \mathbb{E}[V]\log(\mathbb{E}[V])\right), \quad V \in L^1_+ \setminus \{0\}.$$

Note that Assumption 4.10 is satisfied here: by taking $V_1 = V_2 \equiv 1$, we have $\tilde{\alpha}_1(V_1) = 0 \in \mathbb{R}$ and $\alpha_2(V_2, r) = r \in \mathbb{R}$. Moreover, $(\mathscr{D}(r))$ takes the form

$$\begin{aligned} \textit{maximize} \quad & -\frac{1}{r_1}\mathbb{E}[V_1\log(V_1)] - \frac{1}{r_2}\mathbb{E}[V_2\log(V_2)] + \frac{1}{r_2}\mathbb{E}[V_2]\log(\mathbb{E}[V_2]) - r\mathbb{E}[V_2] - y \\ \textit{subject to} \quad & \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \leq y\mathbf{1} \\ & \mathbb{E}[V_1] = 1 \\ & V_1, V_2 \in L^1_+, \ y \in \mathbb{R}. \end{aligned}$$

In the case of a finite probability space, this problem can be solved numerically using, for instance, the convex optimization package CVX (see [8]) as it is able to work with the convex function $z \mapsto z\log(z)$ on $\mathbb{R}_+$. As is standard in convex optimization, these packages also provide the value of the Lagrange multiplier w^* that corresponds to the inequality constraint in $(\mathscr{D}(r))$ at (approximate) optimality. By Theorem 4.11, such w^* is an (approximately) optimal solution of $(\mathscr{P}(r))$.*

6.2. Certainty equivalents

Certainty equivalents form an important class of quasiconvex risk measures. We briefly recall their definitions and properties; see [6, Example 8] for more details. To avoid integrability issues, we assume that $p = +\infty$ (hence $q = 1$) in all examples although, in each example, a larger L^p space can be considered depending on the nature of the loss function.

Let $\ell\colon \mathbb{R} \to (-\infty, +\infty]$ be a convex lower semicontinuous increasing function that is differentiable on $\operatorname{dom}\ell := \{y \in \mathbb{R} \mid \ell(y) < +\infty\}$, we call ℓ a loss function. Let ρ be the certainty equivalent corresponding to ℓ, that is,

$$\rho(Y) = \ell^{-1}\left(\mathbb{E}[\ell(-Y)]\right), \quad Y \in L^\infty,$$

where ℓ^{-1} is the left-continuous inverse of ℓ. The minimal penalty function of ρ is given by

$$\alpha(V,t) = \mathbb{E}\left[Vh\left(\lambda(V,t)\frac{V}{\mathbb{E}[V]}\right)\right], \quad V \in L^1_+ \setminus \{0\}, \ t \in \mathbb{R}, \tag{6.1}$$

where h is the right-continuous inverse of the derivative ℓ' of ℓ, and $\lambda(V,t) > 0$ is a multiplier that is determined by the equation

$$\mathbb{E}\left[\ell\left(h\left(\lambda(V,m)\frac{V}{\mathbb{E}[V]}\right)\right)\right] = \ell^+(t), \tag{6.2}$$

*Grant M, Boyd SP (2020). CVX: Matlab Software for Disciplined Convex Programming [online]. Version 2.2. Website http://cvxr.com/cvx [accessed 27 November 2020].

where ℓ^+ is the right-continuous version of ℓ.

Next, we consider some special cases of ℓ for which more explicit forms of α can be obtained.

Example 6.2 *Suppose that ℓ is the quadratic loss function given by*

$$\ell(y) = \begin{cases} \frac{1}{2}y^2 + y & \text{if } y \geq -1, \\ -\frac{1}{2} & \text{else.} \end{cases}$$

Let $V \in L^1_+ \setminus \{0\}$, $t, s \in \mathbb{R}$. After elementary calculations, we may solve (6.2) for $\lambda(V,t)$ and use the resulting expression in (6.1) to get

$$\alpha(V,t) = \begin{cases} (1+t)\left\|V\right\|_2 - \mathbb{E}[V] & \text{if } t > -1, \\ -\mathbb{E}[V] & \text{else,} \end{cases} \qquad \beta(V,s) = \begin{cases} \frac{s+\mathbb{E}[V]}{\left\|V\right\|_2} - 1 & \text{if } s > -\mathbb{E}[V], \\ -\infty & \text{else.} \end{cases}$$

In particular, as in Section 6.1, it is easy to verify that Assumption 4.2 and Assumption 4.6 hold for $\alpha_1 = \alpha$.

Example 6.3 *Suppose that ℓ is the logarithmic loss function given by*

$$\ell(y) = \begin{cases} -\log(-y) & \text{if } y < 0, \\ +\infty & \text{else.} \end{cases}$$

Let $V \in L^1_+ \setminus \{0\}$, $t, s \in \mathbb{R}$. Then, after some elementary calculations, we obtain

$$\alpha(V,t) = \begin{cases} te^{\mathbb{E}[\log(V)]} & \text{if } t < 0, \\ +\infty & \text{else,} \end{cases} \qquad \beta(V,s) = \begin{cases} se^{-\mathbb{E}[\log(V)]} & \text{if } s < 0, \\ 0 & \text{else.} \end{cases}$$

It follows that Assumption 4.2 and Assumption 4.6 hold for $\alpha_1 = \alpha$.

We conclude this section with an example where ρ_1, ρ_2 are assumed to be certainty equivalents whose respective loss functions ℓ_1, ℓ_2 are among the two examples described above, which illustrates a possible special form of $(\mathscr{F}^D(t,r))$.

Example 6.4 *Suppose that ℓ_1 is the quadratic loss function in Example 6.2 and ℓ_2 is the logarithmic loss function in Example 6.3. Assume that $r < 0$ and take $t \in \mathbb{R}$. If $t \leq -1$, then $(\mathscr{F}^D(t,r))$ can be rewritten as*

$$\begin{aligned} &\textit{maximize} && -\mathbb{E}[V_1] - re^{\mathbb{E}[\log(V_2)]} - y \\ &\textit{subject to} && \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \leq y\mathbf{1} \\ & && V_1, V_2 \in L^1_+,\ y \in \mathbb{R}. \end{aligned}$$

If $t \in (-1,0)$, then $(\mathscr{F}^D(t,r))$ becomes

$$\begin{aligned} &\textit{maximize} && (1+t)\left\|V_1\right\|_2 - \mathbb{E}[V_1] - re^{\mathbb{E}[\log(V_2)]} - y \\ &\textit{subject to} && \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \leq y\mathbf{1} \\ & && V_1, V_2 \in L^1_+,\ y \in \mathbb{R}. \end{aligned}$$

Finally, if $t \geq 0$*, then* $(\mathscr{F}^D(t,r))$ *becomes*

$$\begin{aligned} &\textit{maximize} \ \ (1+t)\left\|V_1\right\|_2 - \mathbb{E}[V_1] - y \\ &\textit{subject to} \ \ \mathbb{E}[V_1 X] + \mathbb{E}[V_2 X] \leq y\mathbf{1} \\ &\qquad V_1, V_2 \in L^1_+, \ y \in \mathbb{R}. \end{aligned}$$

As noted in Example 6.2, Assumption 4.2 and Assumption 4.6 hold in this case. Moreover, Assumption 5.1 holds for every $t \in \mathbb{R}$ *trivially since* α_1 *and* α_2 *are real-valued. Then, Theorem 5.2 is applicable and we may apply the procedure described at the end of Section 5 to find an approximately optimal portfolio for* $(\mathscr{P}(r))$*.*

7. Conclusion

In this paper, we study a static portfolio optimization problem with two quasiconvex risk measures: one in the objective function and one in the constraints. When the former risk measure is a convex functional, the portfolio optimization problem is convex and we are able to identify an optimal portfolio as a Lagrange multiplier of the corresponding dual problem at optimality. In the general quasiconvex case, the portfolio optimization problem is characterized by a family of convex feasibility problems each of which can be solved with the help of its own dual problem. Our analysis shows that it is possible to find an approximately optimal solution for the portfolio optimization problem by solving a finite number of these convex feasibility problems.

Acknowledgment

We would like to thank an anonymous referee whose comments improved the readibility of the paper.

References

[1] Agrawal A, Boyd SP. Disciplined quasiconvex programming. Optimization Letters (2020); 14: 1643-1657.

[2] Aktürk TD, Ararat Ç. Portfolio optimization with two coherent risk measures. Journal of Global Optimization 2020; 78 (3): 597-626.

[3] Artzner P, Delbaen F, Eber JM, Heath D. Coherent measures of risk. Mathematical Finance 1999; 9 (3): 203-228.

[4] Borwein JM, Lewis AS. Partially finite convex programming, Part I: Quasi relative interiors and duality theory. Mathematical Programming 1992; 57 (1): 15-48.

[5] Cerreia-Vioglio S, Maccheroni F, Marinacci M, Montrucchio L. Risk measures: rationality and diversification. Mathematical Finance 2011; 21 (4): 743-774.

[6] Drapeau S, Kupper M. Risk preferences and their robust representation. Mathematics of Operations Research 2013; 38 (1): 28-62.

[7] Föllmer H, Schied A. Stochastic Finance: An Introduction in Discrete Time. 4th revised edition. Berlin, Germany: De Gruyter, 2016.

[8] Grant M, Boyd SP. Graph implementations for nonsmooth convex programs. In: Blondel VD, Boyd SP, Kimura H (editors). Recent Advances in Learning and Control. London, UK: Springer, 2008, pp. 95-110.

[9] Källblad S. Risk- and ambiguity-averse portfolio optimization with quasiconvex utility functionals. Finance and Stochastics 2017; 21 (2): 397-425.

[10] Landsman Z. Minimization of the root of a quadratic functional under an affine equality constraint. Journal of Computational and Applied Mathematics 2008; 216: 319-327.

[11] Landsman Z, Makov U. Minimization of a function of a quadratic functional with application to optimal portfolio selection. Journal of Optimization Theory and Applications 2016; 170: 308-322.

[12] Mastrogiacomo E, Rosazza Gianin E. Portfolio optimization with quasiconvex risk measures, Mathematics of Operations Research 2015; 40 (4): 1042-1059.

[13] Owadally I. An improved closed-form solution for the constrained minimization of the root of a quadratic functional. Journal of Computational and Applied Mathematics 2011; 236: 4428-4435.

[14] Penot JP, Volle M. On quasiconvex duality. Mathematics of Operations Research 1990; 15 (4): 597-625.

[15] Rockafellar RT. Convex Analysis. Princeton, NJ, USA: Princeton University Press, 1970.

[16] Sion M. On general minimax theorems. Pacific Journal of Mathematics 1958; 8 (1): 171-176.

[17] Zălinescu C. Convex Analysis in General Vector Spaces. Singapore: World Scientific, 2002.