# The Subdifferential of Convex Deviation Measures and Risk Functions

Nicole Lorenz *  Gert Wanka †

In this paper we give subdifferential formulas of some convex deviation measures using their conjugate functions and theorems that give the subdifferential of a composition of two functions. Further, we derive other formulas according to a paper of Ruszczyski and Shapiro and the subdifferential of some risk functions.

**Keywords.** Subdifferentials, conjugate functions, deviation measures, risk functions.

**AMS subject classification.** 46N10, 49N15

## 1 Introduction

The subdifferential is a generalization of the classical derivation. It can be used in convex optimization to formulate sufficient and necessary optimality conditions.
*onvex deviation measures* can be used for portfolio optimization problems, where the risk of a portfolio return is measured. They were introduced in 2002 by Rockafellar et al. (cf. [9]) as a consequence of *coherent risk measures* introduced first by Artzner et al. in 1999 (cf. [1]) and *convex risk measures* introduced by Föllmer and Schied in 2002 (cf. [6, 7]).
In order to calculate the subdifferential of deviation measures we consider their conjugate functions and dual representations as given in [3]. Our aim is to give different formulas for the subdifferentials and show how some results given by Ruszczynski and Shapiro in [12] and also some related results can be derived.

The paper is organized as follows. In the following section we introduce some definitions and notations from the convex analysis and stochastic theory we use within the paper. In section 3 we give subdifferential formulas for some basic functions, namely $||\cdot||_p, ||\cdot_-||_p$ and $||\cdot_+||_p$. Further, we consider some deviation measures and calculate their subdifferentials. In the last section we consider risk functions according to a paper by Ruszczynski and Shapiro and show how their results can be derived using conjugate duality.

## 2 Notations and Preliminaries

Let $\mathcal{X}$ be a real Hausdorff locally convex space and $\mathcal{X}^*$ its topological dual space which we endow with the weak* topology $w^* := w(\mathcal{X}^*, \mathcal{X})$. We denote by $\langle x^*, x\rangle := x^*(x)$ the value of the linear continuous functional $x^* \in \mathcal{X}^*$ at $x \in \mathcal{X}$. For $\mathcal{X} = \mathbb{R}^n$ we have $\mathcal{X} = \mathcal{X}^*$ and for $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n, x^* = (x_1^*, \ldots, x_n^*)^T \in \mathbb{R}^n$ it holds $\langle x^*, x\rangle = (x^*)^T x = \sum_{i=1}^n x_i^* x_i$.

*Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: nicole.lorenz@mathematik.tu-chemnitz.de.

†Corresponding author. Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: gert.wanka@mathematik.tu-chemnitz.de, tel.number: +49 371 531 32152, fax number: +49 371 531 22409

For $f : \mathcal{X} \to \overline{\mathbb{R}}$ the *(Fenchel-Moreau) conjugate function of* $f$, $f^* : \mathcal{X}^* \to \overline{\mathbb{R}}$ is defined by

$$f^*(x^*) = \sup_{x\in\mathcal{X}}\{\langle x^*, x\rangle - f(x)\}.$$

Similarly, the *biconjugate function of* $f$, $f^{**} : \mathcal{X} \to \overline{\mathbb{R}}$ is defined by

$$f^{**}(x) = \sup_{x^*\in\mathcal{X}^*}\{\langle x^*, x\rangle - f^*(x^*)\}.$$

Let $K \subseteq \mathcal{X}$ be a nontrivial convex cone. The *dual cone* $K^* \subseteq \mathcal{X}^*$ is defined by

$$K^* := \{x^* \in \mathcal{X}^* : \ \langle x^*, x\rangle \geq 0, \ \forall x \in K\}.$$

For a function $f : \mathcal{X} \to \overline{\mathbb{R}}$ the *effective domain* is defined by $\mathrm{dom}(f) = \{x \in \mathcal{X} : \ f(x) < +\infty\}$. Further $f$ is *proper* if $\mathrm{dom}(f) \neq \emptyset$ and $f(x) > -\infty$, $\forall x \in \mathcal{X}$ and $f$ is called *convex* if for all $x, y \in \mathcal{X}$ and all $\lambda \in [0, 1]$ it holds

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

The function $f$ is *lower semicontinuous at* $\overline{x} \in \mathcal{X}$ if $\liminf_{x\to\overline{x}} f(x) \geq f(\overline{x})$.
The following theorem holds (see [8]):

**Theorem 2.1** (Fenchel-Moreau)**.** *Let $f : \mathcal{X} \to \overline{\mathbb{R}}$ be a proper function. Then it holds $f = f^{**}$ if and only if $f$ is convex and lower semicontinuous. Consequently it follows:*

$$f(x) = \sup_{x^*\in\mathcal{X}^*}\{\langle x^*, x\rangle - f^*(x^*)\}. \tag{1}$$

From formula (1) the so-called *dual representation* can be derived.

**Remark 2.2.** *For $f : \mathcal{X} \to \overline{\mathbb{R}}$ it always holds $f^* = f^{***}$.*

For $f : \mathcal{X} \to \overline{\mathbb{R}}$ proper and $f(x) \in \mathbb{R}$ the *subdifferential of $f$ at $x$*, introduced in [4], is given by

$$\partial f(x) = \{x^* \in \mathcal{X}^* : f(y) - f(x) \geq \langle x^*, y - x\rangle \,, \ \forall y \in \mathcal{X}\}.$$

Otherwise, for $f(x) = +\infty$, we assume by convention that $\partial f(x) = \emptyset$. With the help of conjugate functions we get some necessary characterizations of the subdifferential (cf. [5]). For the function $f : \mathcal{X} \to \overline{\mathbb{R}}$ and $x \in \mathcal{X}, x^* \in \mathcal{X}^*$ it holds (cf. [2, Theorem 3.2.12])

$$x^* \in \partial f(x) \quad \Leftrightarrow \quad f(x) + f^*(x^*) = \langle x^*, x\rangle \quad \Leftrightarrow \quad f(x) + f^*(x) \leq \langle x^*, x\rangle\,. \tag{2}$$

**Remark 2.3.** *Let $f : \mathcal{X} \to \overline{\mathbb{R}}$ be a proper function. As a consequence of the definition of the subdifferential and the Fenchel-Moreau theorem we get the following:*

$$\begin{aligned}
x^* \in \partial f^{**}(x) \quad &\Leftrightarrow \quad f^{**}(x) + f^*(x^*) = \langle x^*, x\rangle \\
&\Leftrightarrow \quad f^{**}(x) = \langle x^*, x\rangle - f^*(x^*) \\
&\Leftrightarrow \quad \sup_{y^*\in\mathcal{X}^*}\{\langle y^*, x\rangle - f^*(y^*)\} = \langle x^*, x\rangle - f^*(x^*) \\
&\Leftrightarrow \quad x^* \in \arg\max\{\langle \cdot, x\rangle - f^*(\cdot)\}.
\end{aligned}$$

*It follows that*

$$\partial f^{**}(x) = \arg\max\{\langle \cdot, x\rangle - f^*(\cdot)\}.$$

*If $f$ is subdifferentiable at $x$ we have $\partial f^{**}(x) = \partial f(x)$ (cf. [2, Theorem 2.3.16]) and this gives another characterization of the subdifferential of a function $f$, namely*

$$\partial f(x) = \arg\max\{\langle \cdot, x\rangle - f^*(\cdot)\}. \tag{3}$$

By $B_{\mathcal{X}}(x,r)$ we denote the *open ball with radius* $r>0$ *and center* $x$ in $\mathcal{X}$ defined by

$$B_{\mathcal{X}}(x,r)=\{y\in\mathcal{X}:d(x,y)<r\},$$

which is a special neighbourhood of $x$, where $d:\mathcal{X}\times\mathcal{X}\to\mathbb{R}$ is the metric induced by the topology in $\mathcal{X}$. The *closed ball* $\overline{B}_{\mathcal{X}}(x,r)$ is defined by $\overline{B}_{\mathcal{X}}(x,r):=\{y\in\mathcal{X}:d(x,y)\leq r\}$.
By $\mathcal{L}(\mathcal{X},\mathcal{Y})$ we denote the set of linear continuous operators mapping from $\mathcal{X}$ into $\mathcal{Y}$. For $A\in\mathcal{L}(\mathcal{X},\mathcal{Y})$ we denote the *adjoint operator* by $A^*\in\mathcal{L}(\mathcal{Y}^*,\mathcal{X}^*)$.
In the following we write min and max instead of inf and sup if we want to express that the infimum/supremum of a scalar optimization problem is attained. By $v(P)$ we denote the *optimal objective value* of the optimization problem $(P)$.

Consider now the atomless *probability space* $(\Omega,\mathfrak{F},\mathbb{P})$, where $\Omega$ is a basic space, $\mathfrak{F}$ a $\sigma$-algebra on $\Omega$ and $\mathbb{P}$ a probability measure on the measurable space $(\Omega,\mathfrak{F})$.
For a random variable $X:\Omega\to\mathbb{R}$ the *expected value* with respect to $\mathbb{P}$ is defined by

$$\mathbb{E}(X)=\int_\Omega X(\omega)d\mathbb{P}(\omega).$$

Furthermore, for $p\in(1,+\infty)$ let $L_p$ be the following space of random variables:

$$L_p:=L_p(\Omega,\mathfrak{F},\mathbb{P},\mathbb{R})=\left\{X:\Omega\to\mathbb{R},\ X\ \text{ measurable,}\int_\Omega|X(\omega)|^p d\mathbb{P}(\omega)<+\infty\right\}.$$

The space $L_p$ equipped with the norm $||X||_p=(\mathbb{E}(|X|^p))^{\frac{1}{p}}$, $X\in L_p$, is a Banach space. It is well-known that the dual space of $L_p$ is $L_q:=L_q(\Omega,\mathfrak{F},\mathbb{P},\mathbb{R})$, where $q\in(1,+\infty)$ fulfills $\frac{1}{p}+\frac{1}{q}=1$.
For $p\in(1,+\infty)$, the cone

$$(L_p)_+=\{X\in L_p:\ X\geq 0 \text{ a.s.}\}$$

is inducing the partial ordering denoted by "$\geq$". The dual cone of $(L_p)_+$ is $(L_q)_+$. The partial ordering induced by $(L_q)_+$ is also denoted by "$\geq$". As these orderings are given in different linear spaces, no confusion is possible.
For $X\in L_p$ and $X^*\in L_q$ we have $\langle X^*,X\rangle:=\int_\Omega X^*(\omega)X(\omega)d\mathbb{P}(\omega)=\mathbb{E}(X^*X)$ as representation of the linear continuous functional.
Equalities and inequalities between random variables are to be viewed in the sense of holding almost surely (a.s.) regarding $\mathbb{P}$. Thus for $X,Y:\Omega\to\mathbb{R}$ when we write "$X=Y$" or "$X\geq Y$" we mean "$X=Y$ *a.s.*" or "$X\geq Y$ *a.s.*", respectively.
For an arbitrary random variable $X:\Omega\to\mathbb{R}$, we also define $X_-,X_+:\Omega\to\mathbb{R}$ as being

$$X_-(\omega):=\max(-X(\omega),0)\qquad\forall\omega\in\Omega,$$
$$X_+(\omega):=\max(X(\omega),0)\qquad\forall\omega\in\Omega.$$

The *essential supremum* of $X$ is $\operatorname{essup}X=\inf\{a\in\mathbb{R}:\mathbb{P}(\omega:X(\omega)>a)=0\}$.

# 3 The Subdifferential of Deviation Measures

In this section we consider some special deviation measures that can be used for portfolio optimization problems and calculate their subdifferential formulas. We give different possibilities to calculate the subdifferential using conjugate functions and formulas for the subdifferential of composed functions, respectively.

## 3.1 Preliminary Facts and Basic Functions

In this first subsection we consider some basic facts and functions we need for the further calculations. In order to get a formula for the subdifferential of the composition between a function $f$ and a linear continuous mapping $T$, $\partial(f \circ T)$, or between two functions $g$ and $h$, $\partial(g \circ h)$, we consider the following theorems. $\mathcal{X}$ and $\mathcal{Y}$ are assumed to be real Hausdorff locally convex spaces.
We consider the following theorem (cf. [2, Theorem 3.5.7]):

**Theorem 3.1.** *Let be $T : \mathcal{X} \to \mathcal{Y}$ a linear continuous mapping and $f : \mathcal{Y} \to \overline{\mathbb{R}}$ a proper and convex function. If the condition*

$$(RC_1) \quad \Big| \; \exists x' \in T^{-1}(\text{dom}(f)) \text{ such that } f \text{ is continuous at } Tx'.$$

*is fulfilled, then for all $x \in \mathcal{X}$ it holds $\partial(f \circ T)(x) = T^* \partial f(Tx)$.*

Further we have, adapted from Theorem 2.8.10 in [13] and using Theorem 3.5.6 (b) in [2], the following:

**Theorem 3.2.** *Let $h : \mathcal{X} \to \mathbb{R}$ be convex and $g : \mathbb{R} \to \overline{\mathbb{R}}$ be proper and convex such that $g$ is increasing on $h(\mathcal{X}) + [0, +\infty)$. Assume that the regularity condition*

$$(RC_2) \quad \Big| \; \exists x' \in \mathcal{X} \text{ such that } h(x') \in \text{dom}(g) \text{ and } g \text{ is continuous at } h(x') \tag{4}$$

*holds. Then for all $x \in \mathcal{X}$ it holds*

$$\partial(g \circ h)(x) = \bigcup_{\lambda \in \partial g(h(x))} \partial(\lambda h)(x).$$

In the following subsections we need formulas for the subdifferentials of some basic functions. These are given in the following examples.

**Corollary 3.3.** *Let be $f_1 : L_p \to \mathbb{R}$ defined by $f_1(X) = ||X||_p$. Then it holds for $X \in L_p$:*

$$\partial f_1(X) = \{X^* \in L_q : \langle X^*, X \rangle = ||X||_p, ||X^*||_q = 1\}, \qquad X \neq 0, \tag{5}$$

$$\partial f_1(0) = \{X^* \in L_q : ||X^*||_q \leq 1\}. \tag{6}$$

This is a classical result that will be proved here in a very short way by using the conjugate function of $f_1$ and using formula (2).

*Proof.* The subdifferential of the norm can be given using the conjugate function $f_1^* : L_q \to \overline{\mathbb{R}}$ taking $X^* \in L_q$ given by (cf. [3])

$$f_1^*(X^*) = \begin{cases} 0, & X^* \in \overline{B}_{L_q}(0,1), \\ +\infty, & \text{otherwise.} \end{cases} \tag{7}$$

Using formula (2) it holds for $X \in L_p, X \neq 0$:

$$\begin{aligned} \partial f_1(X) &= \{X^* \in L_q : f_1(X) + f_1^*(X^*) = \langle X^*, X \rangle\} \\ &= \{X^* \in L_q : \langle X^*, X \rangle = ||X||_p, ||X^*||_q \leq 1\}. \end{aligned}$$

Because of the *Cauchy–Schwarz inequality*,

$$\langle X^*, X \rangle \leq ||X^*||_q \cdot ||X||_p, \quad X \in L_p, X^* \in L_q, \tag{8}$$

follows $||X^*||_q = 1$ from $\langle X^*, X \rangle = ||X||_p, ||X^*||_q \leq 1$. From this formula (5) follows immediately.

Further for $X = 0$ it holds $\partial f_1(0) = \{X^* \in L_q : f_1(0) + f_1^*(X^*) = 0\} = \{X^* \in L_q : ||X^*||_q \leq 1\}$. This follows directly from formula (7). ☐

**Corollary 3.4.** *Let $f_2 : L_p \to \overline{\mathbb{R}}$ be defined by $f_2(X) = ||X_-||_p$. Then it holds for $X \in L_p$*

$$\partial f_2(X) = \{X^* \in L_q : \ ||X_-||_p = \langle X^*, X\rangle , X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}, \quad X \neq 0, \tag{9}$$
$$\partial f_2(0) = \{X^* \in L_q : \ X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}.$$

*Proof.* To get a formula for the subdifferential of $f_2$ we consider the conjugate function $f_2^* : L_q \to \mathbb{R}$ given by (cf. [3])

$$f_2^*(X^*) = \begin{cases} 0, & \text{if } \ ||X^*||_q \leq 1, \ X^* \leq 0, \\ +\infty, & \text{otherwise,} \end{cases} \tag{10}$$
$$= \begin{cases} 0, & \text{if } \ X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+, \\ +\infty, & \text{otherwise.} \end{cases}$$

It follows by means of formula (2) again:

$$\begin{aligned} \partial f_2(X) &= \{X^* \in L_q : f_2(X) + f_2^*(X^*) = \langle X^*, X\rangle\} \\ &= \{X^* \in L_q : ||X_-||_p = \langle X^*, X\rangle , X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}, \quad X \neq 0, \\ \partial f_2(0) &= \{X^* \in L_q : f_2(0) + f_2^*(X^*) = 0\} = \{X^* \in L_q : X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}. \end{aligned}$$

□

**Corollary 3.5.** *Let $f_3 : L_p \to \mathbb{R}$ be given by $f_3(X) = ||X_+||_p$. Then it holds for $X \in L_p$*

$$\partial f_3(X) = \{X^* \in L_q : \ ||X_+||_p = \langle X^*, X\rangle , X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}, \quad X \neq 0,$$
$$\partial f_3(0) = \{X^* \in L_q : \ X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}.$$

*Proof.* In analogy with the proof of Corollary 3.4 the result follows immediately using formula (2) and the conjugate function of $f_3$, $f_3^* : L_q \to \overline{\mathbb{R}}$, (cf. [3])

$$f_3^*(X^*) = \begin{cases} 0, & \text{if } X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+, \\ +\infty, & \text{otherwise.} \end{cases} \tag{11}$$

It holds:

$$\begin{aligned} \partial f_3(X) &= \{X^* \in L_q : f_3(X) + f_3^*(X^*) = \langle X^*, X\rangle\} \\ &= \{X^* \in L_q : ||X_+||_p = \langle X^*, X\rangle , X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}, \quad X \neq 0, \\ \partial f_3(0) &= \{X^* \in L_q : f_3(0) + f_3^*(X^*) = 0\} = \{X^* \in L_q : X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}. \end{aligned}$$

□

Since in the following subsections we will consider some convex deviation measures that were introduced by Rockafellar and his coauthors (cf. [9]), we give here a definition:

**Definition 3.6.** *The function $d : L_p \to \overline{\mathbb{R}}$ is called a convex deviation measure if the following properties are fulfilled:*
*(D1) Translation invariance:* $\quad d(X + b) = d(X), \quad \forall X \in L_p, \ \forall b \in \mathbb{R}$;
*(D2) Non-Negativity:* $\quad d(X) \geq 0, \quad \forall X \in L_p$;
*(D3) Convexity:* $\quad d(\lambda X + (1 - \lambda)Y) \leq \lambda d(X) + (1 - \lambda)d(Y), \quad \forall \lambda \in [0, 1], \forall X, Y \in L_p$.

The following theorem states the connection between convex risk and convex deviation measures (see [9, 10, 11]).

**Definition 3.7.** *The function* $\rho : L_p \to \overline{\mathbb{R}}$ *is a convex risk measure if and only if* $d : L_p \to \overline{\mathbb{R}}$ *given by* $d(X) = \rho(X) + \mathbb{E}(X)$ *for* $X \in L_p$*, is a convex deviation measure.*

Having the subdifferential of any deviation measure $d$ we can easily derive the one of the corresponding risk measure. The following theorem holds.

**Lemma 3.8.** *Assume that* $d : L_p \to \overline{\mathbb{R}}$ *is a convex deviation measure. For* $X \in L_p$ *the subdifferential of the corresponding risk measure* $\rho : L_p \to \overline{\mathbb{R}}$*,* $\rho(X) = d(X) - \mathbb{E}(X)$ *is* $\partial\rho(X) = \partial d(X) - 1$*.*

*Proof.* Since the function $X \mapsto \mathbb{E}(X) = \langle 1, X\rangle, X \in L_p$, is linear and continuous it holds $\partial\rho(X) = \partial d(X) - \partial\langle 1, X\rangle$. Further, since $\partial\langle 1, X\rangle = \nabla\langle 1, X\rangle = 1$, it follows $\partial\rho(X) = \partial d(X) - 1$. □

In the following sections we will only consider convex deviation measures. Here we first show some examples.

**Example 3.9.** *Some deviation measures.*

For $p \in (1, +\infty)$ and $a > 1$ let the deviation measures $d_i : L_p \to \mathbb{R}, i = 1, \ldots, 5$, be given as follows for $X \in L_p$:

$$\begin{aligned}
d_1(X) &= ||X - \mathbb{E}(X)||_p,\\
d_2(X) &= ||(X - \mathbb{E}(X))_-||_p,\\
d_3(X) &= ||(X - \mathbb{E}(X))_+||_p,\\
d_4(X) &= ||X - \mathbb{E}(X)||_p^a,\\
d_5(X) &= ||(X - \mathbb{E}(X))_-||_p^a,
\end{aligned}$$

In case $p = 1$, $d_1, d_2$ and $d_3$ are the so-called *mean absolute deviation*, and the *lower* and *upper semideviation*, respectively.
For $p = 2$ the functions $d_2$ and $d_3$ become the *standard lower* and *upper semideviation*, respectively.
The case $a = p = 2$ leads to the *variance* for $d_4$.

## 3.2 Subdifferentials of Convex Deviation Measures.

In this subsection we consider some deviation measures being the composition of the functions $f_1$ and $f_2$ and the linear continuous functional $A \in \mathcal{L}(L_p, L_p)$ defined by $AX = X - \mathbb{E}(X)$. We calculate the subdifferential formulas using Theorem 3.1 and the conjugate functions, respectively.

The adjoint operator $A^* \in \mathcal{L}(L_q, L_q)$ was given in [3] and for $X^* \in L_q$ it holds $A^*X^* = X^* - \mathbb{E}(X^*)$. First let us consider the following:

**Theorem 3.10.** *Let* $X \in L_p$ *be fixed. Then for* $X^* \in L_q$ *it holds:*

$$\begin{aligned}
&X^* \in \mathcal{D}_1 := \arg \max_{\substack{Y^* \in L_q,\\ ||Y^*||_q = 1}} \langle Y^*, X - \mathbb{E}(X)\rangle\\
\Leftrightarrow\quad &\langle X^*, X - \mathbb{E}(X)\rangle = ||X - \mathbb{E}(X)||_p \text{ and } ||X^*||_q = 1.
\end{aligned}$$

*Proof.* Taking $X \in L_p$ we have $||X||_p = \sup_{X^* \in L_q, ||X^*||_q = 1} \langle X^*, X\rangle$ and hence $||X - \mathbb{E}(X)||_p = \sup_{X^* \in L_q, ||X^*||_q = 1} \langle X^*, X - \mathbb{E}(X)\rangle$ and the result follows directly from the Cauchy–Schwarz inequality given in formula (8). □

It holds:

**Corollary 3.11.** *Let the function $d_1 : L_p \to \mathbb{R}$ be given by $d_1(X) = ||X - \mathbb{E}(X)||_p$. Then it holds for $X \in L_p, X \neq 0$:*

$$\partial d_1(X) = \{X^* \in L_q : ||X - \mathbb{E}(X)||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0, \min_{c\in\mathbb{R}} ||X^* - c||_q \leq 1\} \tag{12}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, ||X - \mathbb{E}(X)||_p = \langle X^* - \mathbb{E}(X^*), X\rangle, ||X^*||_q = 1\} \tag{13}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q,\ X^* \in \mathcal{D}_1\}, \qquad \mathcal{D}_1 = \arg \max_{\substack{Y^*\in L_q,\\ ||Y^*||_q=1}} \langle Y^*, X - \mathbb{E}(X)\rangle, \tag{14}$$

*and further*

$$\partial d_1(0) = \{X^* \in L_q : \mathbb{E}(X^*) = 0, \min_{c\in\mathbb{R}} ||X^* - c||_q \leq 1\} \tag{15}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, ||X^*||_q \leq 1\}. \tag{16}$$

*Proof.* The conjugate function of $d_1$, $d_1^* : L_q \to \overline{\mathbb{R}}$, is given by (cf. [3])

$$d_1^*(X^*) = \begin{cases} 0, & \text{if } \exists Y^* \in L_q : Y^* - \mathbb{E}(Y^*) = X^* \text{ and } ||Y^*||_q \leq 1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{17}$$

$$= \begin{cases} 0, & \text{if } \mathbb{E}(X^*) = 0 \text{ and } \min_{c\in\mathbb{R}} ||X^* - c||_q \leq 1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{18}$$

(a) Using the conjugate function as given in formula (18) and formula (2) we get the subdifferential as in formula (12):

$$\begin{aligned}\partial d_1(X) &= \{X^* \in L_q : d_1(X) + d_1^*(X^*) = \langle X^*, X\rangle\} \\ &= \{X^* \in L_q : ||X - \mathbb{E}(X)||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0 \text{ and } \min_{c\in\mathbb{R}} ||X^* - c||_q \leq 1\}, \quad X \neq 0.\end{aligned}$$

We further have

$$\begin{aligned}\partial d_1(0) &= \{X^* \in L_q : d_1(0) + d_1^*(X^*) = 0\} \\ &= \{X^* \in L_q : \mathbb{E}(X^*) = 0 \text{ and } \min_{c\in\mathbb{R}} ||X^* - c||_q \leq 1\},\end{aligned}$$

which is formula (15).

(b) A formula for the subdifferential of $d_1$ can be given if we consider the composition $d_1 = f_1 \circ A$, where $f_1$ is the norm function, whose subdifferential was given in Corollary 3.3, formula (5), and $A$ is defined as above.
Theorem 3.1 can be applied since $A$ is linear and continuous, $f_1$ is proper and convex and further the composition $f_1 \circ A$ is proper. The regularity condition $(RC_1)$ is fulfilled since $f_1$ is continuous and it holds for $X \in L_p, X \neq 0$:

$$\begin{aligned}\partial d_1(X) &= \partial(f_1 \circ A)(X) = A^*\partial|| \cdot ||_p(AX) \\ &= \{A^*X^* : X^* \in L_q, ||AX||_p = \langle X^*, AX\rangle,\ ||X^*||_q = 1\} \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in L_q,\ ||X - \mathbb{E}(X)||_p = \langle X^*, X - \mathbb{E}(X)\rangle,\ ||X^*||_q = 1\},\end{aligned} \tag{19}$$

which is equal to (13).
Further it holds using formula (6)

$$\partial d_1(0) = \partial(f_1 \circ A)(0) = A^*\partial|| \cdot ||_p(0) = \{X^* - \mathbb{E}(X^*) : X^* \in L_q, ||X^*||_q \leq 1\},$$

which leads to formula (16).

(c) Formula (14) arises by using Theorem 3.10.
Further, we get the same result using the conjugate function given in formula (17) and formula (3) since $d_1$ is subdifferentiable. This can be seen as follows:

$$\begin{aligned}
&\partial d_1(X) = \arg\max_{X^*\in L_q} \{\langle X^*, X\rangle - d_1^*(X^*)\}\\
&= \{Y^* \in L_q : \sup_{X^*\in L_q} \{\langle X^*, X\rangle - d_1^*(X^*)\} = \langle Y^*, X\rangle - d_1^*(Y^*)\}\\
&= \{Y^* \in L_q : \sup_{\substack{Z^*\in L_q: Z^*-\mathbb{E}(Z^*)=X^*,\\ ||Z^*||_q\le 1}} \langle X^*, X\rangle = \langle Y^*, X\rangle, Y^* = W^* - \mathbb{E}(W^*), ||W^*||_q \le 1\}\\
&= \{W^* - \mathbb{E}(W^*) : W^* \in L_q, \sup_{\substack{Z^*\in L_q,\\ ||Z^*||_q\le 1}} \langle Z^* - \mathbb{E}(Z^*), X\rangle = \langle W^* - \mathbb{E}(W^*), X\rangle, ||W^*||_q \le 1\}\\
&= \{W^* - \mathbb{E}(W^*) : W^* \in L_q, W^* \in \arg\max_{\substack{Z^*\in L_q,\\ ||Z^*||_q\le 1}} \langle Z^* - \mathbb{E}(Z^*), X\rangle, ||W^*||_q = 1\}.
\end{aligned}$$

□

In the following we consider the function $d_2 : L_p \to \mathbb{R}$ given by $d_2(X) = ||(X - \mathbb{E}(X))_-||_p$. The conjugate function of $d_2$, $d_2^* : L_q \to \overline{\mathbb{R}}$, is given by (cf. [3])

$$d_2^*(X^*) = \begin{cases} 0, & \text{if } \mathbb{E}(X^*) = 0,\ X^* \le 1,\ ||\operatorname{essup} X^* - X^*||_q \le 1, \\ +\infty, & \text{otherwise,} \end{cases} \tag{20}$$

$$= \begin{cases} 0, & \text{if } \mathbb{E}(X^*) = 0 \text{ and } \exists c \in \mathbb{R} :\ 0 \le c \le 1,\ ||X^* - c||_q \le 1,\ \ X^* \le c, \\ +\infty, & \text{otherwise,} \end{cases} \tag{21}$$

$$= \begin{cases} 0, & \text{if } \ \exists Y^* \in L_q : Y^* - \mathbb{E}(Y^*) = X^*, Y^* \le 0, ||Y^*||_q \le 1, \\ +\infty, & \text{otherwise,} \end{cases} \tag{22}$$

$$= \begin{cases} 0, & \text{if } \ \exists Y^* \in L_q : Y^* - \mathbb{E}(Y^*) = X^*, Y^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+, \\ +\infty, & \text{otherwise.} \end{cases} \tag{23}$$

We can prove the following theorem.

**Theorem 3.12.** *Let $X \in L_p$ be fixed. Then for $X^* \in L_q$ it holds:*

$$\begin{aligned}
& X^* \in \mathcal{D}_2 := \arg\max_{Y^*\in \overline{B}_{L_q}(0,1)\cap -(L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle \\
\Leftrightarrow\quad & \langle X^*, X - \mathbb{E}(X)\rangle = ||(X - \mathbb{E}(X))_-||_p,\ \textit{and } X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+.
\end{aligned}$$

*Proof.* Using Theorem 2.1 and formula (22) for the conjugate function of $d_2$ we get for $X \in L_p$:

$$\begin{aligned}
||(X - \mathbb{E}(X))_-||_p &= d_2(X) = d_2^{**}(X) \\
&= \sup_{X^*\in L_q} \{\langle X^*, X\rangle - d_2^*(X^*)\} = \sup_{\substack{(X^*,Y^*)\in L_q\times L_q:\\ Y^*-\mathbb{E}(Y^*)=X^*,\\ Y^*\le 0, ||Y^*||_q\le 1}} \langle X^*, X\rangle, \\
&= \sup_{Y^*\in \overline{B}_{L_q}(0,1)\cap -(L_q)_+} \langle Y^* - \mathbb{E}(Y^*), X\rangle = \sup_{Y^*\in \overline{B}_{L_q}(0,1)\cap -(L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle .
\end{aligned}$$

Assuming now that $X^* \in \mathcal{D}_2$, which means that $X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+$ and

$$\langle X^*, X - \mathbb{E}(X)\rangle = \sup_{Y^*\in \overline{B}_{L_q}(0,1)\cap -(L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle ,$$

it follows $\langle X^*, X - \mathbb{E}(X)\rangle = ||(X - \mathbb{E}(X))_-||_p$ and $X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+$.
On the other hand having $X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+$ such that $\langle X^*, X - \mathbb{E}(X)\rangle = ||(X - \mathbb{E}(X))_-||_p$.
Using formulas (23) and the fact that for a proper, convex and lower semicontinuous function $f : L_p \to \overline{\mathbb{R}}$ it holds

$$f(X) = f^{**}(X) = \sup_{X^* \in L_q} \{\langle X^*, X\rangle - f^*(X^*)\} = \sup_{X^* \in L_q} \{\mathbb{E}(X^*X) - f^*(X^*)\}, \quad \forall X \in L_p, \tag{24}$$

we get the following dual representation for $d_2$, $X \in L_p$:

$$\begin{aligned} d_2(X) &= \sup\{\mathbb{E}(X^*X) : X^* \in L_q, \exists Y^* \in L_q : Y^* - \mathbb{E}(Y^*) = X^*, Y^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\} \\ &= \sup_{Y^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+} \langle Y^* - \mathbb{E}(Y^*), X\rangle = \sup_{Y^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle . \end{aligned}$$

Using this dual representation we get that $X^* \in \mathcal{D}_2$ and this concludes the proof. □

**Corollary 3.13.** *Let $d_2 : L_p \to \mathbb{R}$ be given by $d_2(X) = ||(X - \mathbb{E}(X))_-||_p$. Then we have for $X \in L_p, X \neq 0$*

$$\begin{aligned} &\partial d_2(X) \\ &= \{X^* \in L_q : ||(X - \mathbb{E}(X))_-||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0, X^* \leq 1, ||\operatorname{essup} X^* - X^*||_q \leq 1\} && (25) \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+, ||(X - \mathbb{E}(X))_-||_p = \langle X^* - \mathbb{E}(X^*), X\rangle\} && (26) \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \mathcal{D}_2\}, \quad \mathcal{D}_2 := \arg \max_{Y^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle . && (27) \end{aligned}$$

*Further it holds:*

$$\begin{aligned} \partial d_2(0) &= \{X^* \in L_q : \mathbb{E}(X^*) = 0, X^* \leq 1, ||\operatorname{essup} X^* - X^*||_q \leq 1\} && (28) \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}. && (29) \end{aligned}$$

*Proof.* (a) Using the conjugate function of $d_2$ given in formula (20) and formula (2) it holds for $X \in L_p, X \neq 0$:

$$\begin{aligned} \partial d_2(X) &= \{X^* \in L_q : d_2(X) + d_2^*(X^*) = \langle X^*, X\rangle\} \\ &= \{X^* \in L_q : ||(X - \mathbb{E}(X))_-||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0,\ X^* \leq 1,\ ||\operatorname{essup} X^* - X^*||_q \leq 1\}. \end{aligned}$$

We get the subdifferential formula (25). Further we get formula (28) by

$$\begin{aligned} \partial d_2(0) &= \{X^* \in L_q : d_2(0) + d_2^*(X^*) = 0\} \\ &= \{X^* \in L_q : \mathbb{E}(X^*) = 0,\ X^* \leq 1,\ ||\operatorname{essup} X^* - X^*||_q \leq 1\}. \end{aligned}$$

(b) Using formula (23) we get the second subdifferential formula (26) for $X \in L_p, X \neq 0$:

$$\begin{aligned} &\partial d_2(X) = \{Y^* \in L_q : d_2(X) + d_2^*(Y^*) = \langle Y^*, X\rangle\} \\ &= \{Y^* \in L_q : ||(X - \mathbb{E}(X))_-||_p = \langle Y^*, X\rangle, \exists X^* \in L_q : X^* - \mathbb{E}(X^*) = Y^*, X^* \leq 0, ||X^*||_q \leq 1\} \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+, ||(X - \mathbb{E}(X))_-||_p = \langle X^* - \mathbb{E}(X^*), X\rangle\}. \end{aligned}$$

We get the same result by considering the composition $f_2 \circ A$, where $f_2$ and its subdifferential is given as in Corollary 3.4. We can apply Theorem 3.1 since the regularity condition $(RC_1)$ is fulfilled ($f_2$ is continuous). Thus for all $X \in L_p$ it holds using formula (9):

$$\begin{aligned} &\partial d_2(X) = \partial(f_2 \circ A)(X) = A^* \partial f_2(AX) \\ &= A^*\{X^* \in L_q : ||AX_-||_p = \langle X^*, AX\rangle, X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\} \\ &= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+, ||(X - \mathbb{E}(X))_-||_p = \langle X^*, X - \mathbb{E}(X)\rangle\}. && (30) \end{aligned}$$

Further we get formula (29) by

$$\begin{aligned}\partial d_2(0) &= \{Y^* \in L_q : d_2(0) + d_2^*(Y^*) = \langle Y^*, 0\rangle\}\\ &= \{Y^* \in L_q : \exists X^* \in L_q : X^* - \mathbb{E}(X^*) = Y^*, X^* \leq 0, ||X^*||_q \leq 1\}\\ &= \{X^* - \mathbb{E}(X^*) : X^* \in \overline{B}_{L_q}(0,1) \cap -(L_q)_+\}.\end{aligned}$$

(c) Using Theorem 3.12 we get from formula (30), where the set $\mathcal{D}_2$ is given as above. This formula can also be given by using the arguments in the proof of Corollary 3.11, part (c). □

The next function we consider is $d_3 : L_p \to \mathbb{R}$ given by $d_3(X) = ||(X - \mathbb{E}(X))_+||_p$. The conjugate function of $d_3$ is given (use $d_3^*(X^*) = d_2^*(-X^*)$, cf. [3])

$$d_3^*(X^*) = \begin{cases} 0, & \mathbb{E}(X^*) = 0, X^* \geq -1, ||\operatorname{essup} X^* - X^*||_q \leq 1,\\ +\infty, & \text{otherwise,}\end{cases} \tag{31}$$

$$= \begin{cases} 0, & \exists Y^* \in L_q : Y^* - \mathbb{E}(Y^*) = X^*, Y^* \in \overline{B}_{L_q} \cap (L_q)_+,\\ +\infty, & \text{otherwise.}\end{cases} \tag{32}$$

In analogy to Theorem 3.12 one can prove the following theorem by making use of the two formulas.

**Theorem 3.14.** *Let $X \in L_p$ be fixed. For all $X^* \in L_q$ it holds:*

$$\begin{aligned} &X^* \in \mathcal{D}_3 := \arg \max_{Y^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle\\ \Leftrightarrow\quad &\langle X^*, X - \mathbb{E}(X)\rangle = ||(X - \mathbb{E}(X))_+||_p, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+.\end{aligned}$$

**Corollary 3.15.** *Let $d_3 : L_p \to \mathbb{R}$ be given by $d_3(X) = ||(X - \mathbb{E}(X))_+||_p$. Then it holds for $X \in L_p, X \neq 0$:*

$$\partial d_3(X)$$

$$= \{X^* \in L_q : ||(X - \mathbb{E}(X))_+||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0, X^* \geq -1, ||\operatorname{essup} X^* - X^*||_q \leq 1\} \tag{33}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+, ||(X - \mathbb{E}(X))_+||_p = \langle X^* - \mathbb{E}(X^*), X\rangle\} \tag{34}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q,\ X^* \in \mathcal{D}_3\}, \mathcal{D}_3 := \arg \max_{Y^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+} \langle Y^*, X - \mathbb{E}(X)\rangle \tag{35}$$

*and further*

$$\partial d_3(0) = \{X^* \in L_q : \mathbb{E}(X^*) = 0, X^* \geq -1, ||\operatorname{essup} X^* - X^*||_q \leq 1\} \tag{36}$$

$$= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}. \tag{37}$$

*Proof.* (a) By using the formula for the conjugate function given in formula (31) and formula (2) we get for $X \in L_p, X \neq 0$ by

$$\begin{aligned}\partial d_3(X) &= \{X^* \in L_q : d_3(X) + d_3^*(X^*) = \langle X^*, X\rangle\}\\ &= \{X^* \in L_q : ||(X - \mathbb{E}(X))_+||_p = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0, X^* \geq -1, ||\operatorname{essup} X^* - X^*||_q \leq 1\}\end{aligned}$$

the subdifferential formula (33). Further it holds:

$$\begin{aligned}\partial d_3(0) &= \{X^* \in L_q : d_3(0) + d_3^*(X^*) = 0\}\\ &= \{X^* \in L_q : \mathbb{E}(X^*) = 0, X^* \geq -1, ||\operatorname{essup} X^* - X^*||_q \leq 1\}.\end{aligned}$$

(b) By formula (32) we get formula (34).

The same formula for the subdifferential of $d_3$ arises in analogy to the one in Corollary 3.13 using the subdifferential of $f_3$ and Theorem 3.1 (condition $(RC_1)$ is fulfilled). It holds for $X \in L_p, X \neq 0$:

$$\begin{aligned}
&\partial d_3(X) = \partial(f_3 \circ A)(X) = A^* \partial f_3(AX) \\
&= A^*\{X^* \in L_q : ||AX_+||_p = \langle X^*, AX\rangle, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\} \\
&= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+, ||(X - \mathbb{E}(X))_+||_p = \langle X^*, X - \mathbb{E}(X)\rangle\}.
\end{aligned}$$

Further it holds:

$$\begin{aligned}
\partial d_3(0) &= \partial(f_3 \circ A)(0) = A^* \partial f_3(0) \\
&= \{X^* - \mathbb{E}(X^*) : X^* \in L_q, X^* \in \overline{B}_{L_q}(0,1) \cap (L_q)_+\}.
\end{aligned}$$

(c) Formula (35) follows from Theorem 3.14. This formula can also be given by using the arguments in the proof of Corollary 3.11, part (c). □

**Corollary 3.16.** *Let $d_4 : L_p \to \mathbb{R}$ be given by $d_4(X) = ||X - \mathbb{E}(X)||_p^a$ ($p \in (1, +\infty), a > 1$). Then it holds for all $X \in L_p, X \neq 0$:*

$$\partial d_4(X) \tag{38}$$

$$= \left\{ X^* \in L_q : ||X - \mathbb{E}(X)||_p^a + \min_{c \in \mathbb{R}} \left\{ (a-1) \left|\left| \frac{1}{a}(X^* - c) \right|\right|_q^{\frac{a}{a-1}} \right\} = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0 \right\}. \tag{39}$$

*For all non-constant $X \in L_p$ it holds:*

$$\begin{aligned}
&\partial d_4(X) \\
&= \{X^* \in L_q : a||X - \mathbb{E}(X)||_p^a = \langle X^*, X\rangle, \mathbb{E}(X^*) = 0, \min_{c \in \mathbb{R}} ||X^* - c||_q \leq a||X - \mathbb{E}(X)||_p^{a-1}\}.
\end{aligned} \tag{40}$$

*Proof.* (a) Using the conjugate function of $d_4$, $d_4^* : L_q \to \overline{\mathbb{R}}$ given by (cf. [3])

$$d_4^*(X^*) = \begin{cases} \min\limits_{c \in \mathbb{R}} \left\{ (a-1) \left|\left| \frac{1}{a}(X^* - c) \right|\right|_q^{\frac{a}{a-1}} \right\}, & \text{if } \mathbb{E}(X^*) = 0, \\ +\infty, & \text{otherwise,} \end{cases} \tag{41}$$

we get the subdifferential formula (38).

(b) The function can be written as $d_4 = g \circ d_1$, where $g : \mathbb{R} \to \overline{\mathbb{R}}$ is given by

$$g(x) = \begin{cases} x^a, & \text{if } x \geq 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{42}$$

Especially, the application of Theorem 3.2 is possible since $g$ is convex, continuous, and increasing on $d_1(L_p) + [0, +\infty) = [0, +\infty)$ and $d_1$ is convex (especially $(RC_2)$ in formula (4) is fulfilled).
For $X^* \in L_q$ and $\lambda > 0$ we have

$$(\lambda d_1)^*(X^*) = \lambda d_1^* \left( \frac{X^*}{\lambda} \right) = \begin{cases} 0, & \text{if } \mathbb{E}(X^*) = 0 \text{ and } \min\limits_{c \in \mathbb{R}} ||X^* - c||_q \leq \lambda, \\ +\infty, & \text{otherwise.} \end{cases}$$

Since the function $g$ is differentiable for all $x > 0$ it holds $\partial g(x) = \{ax^{a-1}\}$. Further, using the conjugate function $g^* : \mathbb{R} \to \mathbb{R}$ given by

$$g^*(\beta) = \begin{cases} (a-1) \left( \frac{\beta}{a} \right)^{\frac{a}{a-1}}, & \text{if } \beta \geq 0, \\ 0, & \text{otherwise,} \end{cases} \tag{43}$$

we have $\partial g(0) = -\mathbb{R}_+$, which follows by

$$x^* \in \partial g(0) \quad \Leftrightarrow \quad g(0) + g^*(x^*) = \langle x^*, 0 \rangle \quad \Leftrightarrow \quad g^*(x^*) = 0 \quad \Leftrightarrow \quad x^* \leq 0,$$

and for $x < 0$ we have $\partial g(x) = \emptyset$ since $g(x) = +\infty$. It finally holds

$$\partial g(x) = \begin{cases} ax^{a-1}, & \text{if } x > 0, \\ -\mathbb{R}_+, & \text{if } x = 0, \\ \emptyset, & \text{if } x < 0. \end{cases} \tag{44}$$

We get the following formula:

$$\partial d_4(X) = \partial(g \circ d_1)(X) = \bigcup_{\lambda \in \partial g(d_1(X))} \partial(\lambda d_1)(X).$$

Let us assume that $X$ is non-constant. Then it holds $d_1(X) > 0$ and hence $\partial g(d_1(X)) = a(d_1(X))^{a-1} = a||X - \mathbb{E}(X)||_p^{a-1}$.
It follows now:

$$\begin{aligned}
&\partial d_4(X) \\
&= \bigcup_{\lambda = a||X-\mathbb{E}(X)||_p^{a-1}} \partial(\lambda d_1)(X) \\
&= \{X^* \in L_q : (\lambda d_1)(X) + (\lambda d_1)^*(X^*) = \langle X^*, X \rangle, \lambda = a||X - \mathbb{E}(X)||_p^{a-1}\} \\
&= \{X^* \in L_q : \lambda ||X - \mathbb{E}(X)||_p = \langle X^*, X \rangle, \mathbb{E}(X^*) = 0, \\
&\qquad \min_{c \in \mathbb{R}} ||X^* - c||_q \leq \lambda, \lambda = a||X - \mathbb{E}(X)||_p^{a-1}\} \\
&= \{X^* \in L_q : a||X - \mathbb{E}(X)||_p^{a-1} ||X - \mathbb{E}(X)||_p = \langle X^*, X \rangle, \mathbb{E}(X^*) = 0, \\
&\qquad \min_{c \in \mathbb{R}} ||X^* - c||_q \leq a||X - \mathbb{E}(X)||_p^{a-1}\} \\
&= \{X^* \in L_q : a||X - \mathbb{E}(X)||_p^{a} = \langle X^*, X \rangle, \mathbb{E}(X^*) = 0, \min_{c \in \mathbb{R}} ||X^* - c||_q \leq a||X - \mathbb{E}(X)||_p^{a-1}\},
\end{aligned}$$

which is formula (40). □

**Corollary 3.17.** *Let $d_5 : L_p \to \overline{\mathbb{R}}$ be given by $d_5(X) = ||(X - \mathbb{E}(X))_-||_p^a$ ($p \in (1, +\infty), a > 1$). Then it holds for all $X \in L_p$:*

$$\partial d_5(X) = \left\{ X^* \in L_q : ||X - \mathbb{E}(X)||_p^a + (a-1) \Big|\Big| \frac{1}{a}(\operatorname{essup} x^* - x^*) \Big|\Big|_q^{\frac{a}{a-1}} = \langle X^*, X \rangle, \mathbb{E}(X^*) = 0 \right\}. \tag{45}$$

*Proof.* Using the conjugate function of $d_5$, $d_5^* : L_p \to \overline{\mathbb{R}}$ given by (cf. [3])

$$d_5^*(X^*) = \begin{cases} (a-1) \big|\big| \frac{1}{a}(\operatorname{essup} X^* - X^*) \big|\big|_q^{\frac{a}{a-1}}, & \text{if } \mathbb{E}(X^*) = 0, \\ +\infty, & \text{otherwise}, \end{cases} \tag{46}$$

we get the subdifferential formula (45). □

# 4 Risk Functions in Portfolio Optimization Problems

In a paper of Ruszczyński and Shapiro [12] some risk functions and their subdifferentials are considered. Our aim is to rediscover these results by using conjugate duality techniques as done in the above section and to extend them to other functions.

An uncertain outcome is represented by the random variable $X \in L_p$. Ruszczyński and Shapiro assume that the smaller the values of $X$ the better. Thus, their aim is to minimize a so-called *risk function* of the form

$$\mathbb{E}(X) + cd(X),$$

where $d : L_p \to \overline{\mathbb{R}}$ is any deviation measure and $c \in \mathbb{R}_+$.

As we will see in the next chapter in portfolio optimization problems we minimize the risk of the portfolio and maximize its expected return. Thus risk functions $r : L_p \to \overline{\mathbb{R}}$ we consider in our framework have the form

$$r(X) = \lambda d(X) - \mathbb{E}(X),$$

where $\lambda > 0$. This results in a different sign and other weights ($c$ and $\lambda$) in the subdifferential formulas in contrast to the one in [12]. For some risk functions the subdifferential formulas are given in the following corollaries. The notation is used according to [12].

**Corollary 4.1** (Mean-deviation risk function)**.** *Let $r_1 : L_p \to \mathbb{R}$ be the risk function given by $r_1(X) = \lambda d_1(X) - \mathbb{E}(X) = \lambda||X - \mathbb{E}(X)||_p - \mathbb{E}(X)$ for $\lambda > 0$. Then it holds for $X \in L_p, X \neq 0$*

$$\partial r_1(X) = \{X^* - \mathbb{E}(X^*) - 1 : X^* \in L_q,\ X^* \in \mathcal{D}_{1'}\}, \tag{47}$$

$$\mathcal{D}_{1'} = \arg \max_{\substack{X^* \in L_q, \\ ||X^*||_q = \lambda}} \langle X^*, X - \mathbb{E}(X) \rangle \tag{48}$$

*and further*

$$\partial r_1(0) = \{X^* - \mathbb{E}(X^*) - 1 : X^* \in L_q, ||X^*||_q \leq \lambda\} \tag{49}$$

*Proof.* It holds $\partial r_1(X) = \lambda \partial d_1(X) - 1$ (see Lemma 3.8). We use Corollary 3.11, formula (14) to get the following for $X \in L_p, X \neq 0$:

$$\begin{aligned}
&\partial r_1(X) = \{\lambda(X^* - \mathbb{E}(X^*)) - 1 : X^* \in L_q,\ X^* \in \mathcal{D}_1\}, \\
&\mathcal{D}_1 := \arg \max_{\substack{X^* \in L_q, \\ ||X^*||_q = 1}} \langle X^*, X - \mathbb{E}(X) \rangle .
\end{aligned}$$

By setting $\lambda X^* =: X^*$ we get formula (47).

Further it holds:

$$\begin{aligned}
&\partial r_1(0) = \lambda \partial d_1(0) - 1 \\
&= \lambda\{X^* - \mathbb{E}(X^*) : X^* \in L_q, ||X^*||_q \leq 1\} - 1 \\
&= \{\lambda(X^* - \mathbb{E}(X^*)) - 1 : X^* \in L_q, ||X^*||_q \leq 1\} \\
&= \{X^* - \mathbb{E}(X^*) - 1 : X^* \in L_q, ||X^*||_q \leq \lambda\}
\end{aligned}$$

□

**Remark 4.2.** *Ruszczyński and Shapiro give a formula for the subdifferential of the risk function $\mathbb{E}(X) + c||X - \mathbb{E}(X)||_p$ at zero, called $\mathcal{A}_p$ (cf. [12, formula (4.7)]), namely (using our notations)*

$$\mathcal{A}_p = \{X^* - \mathbb{E}(X^*) + 1 : Z^* \in L_q, ||X^*||_q \leq c\}.$$

*We get the same result from formula (49).*

**Corollary 4.3** ((Lower) Mean-semideviation risk measure)**.** *Let the risk function $r_2 : L_p \to \mathbb{R}$ be given by $r_2(X) = \lambda d_2(X) - \mathbb{E}(X) = \lambda||(X - \mathbb{E}(X))_-||_p - \mathbb{E}(X)$ for $\lambda > 0$. Then it holds for $X \in L_p$:*

$$\partial r_2(X) = \{X^* - \mathbb{E}(X^*) - 1 : X^* \in L_q,\, X^* \in \mathcal{D}_{2'}\}\,, \tag{50}$$
$$\mathcal{D}_{2'} = \arg \max_{\substack{X^* \in \overline{B}_{L_q}(0,\lambda) \\ \cap -(L_q)_+}} \langle X^*, X - \mathbb{E}(X)\rangle\,. \tag{51}$$

*Proof.* As in the latter example we have $\partial r_2(X) = \lambda \partial d_2(X) - 1$ (see Lemma 3.8) and using Corollary 3.13, formula (27) and get

$$\partial r_2(X) = \{\lambda(X^* - \mathbb{E}(X^*)) - 1 : X^* \in L_q,\, X^* \in \mathcal{D}_2\}\,,$$
$$\mathcal{D}_2 = \arg \max_{\substack{X^* \in \overline{B}_{L_q}(0,1) \\ \cap -(L_q)_+}} \langle X^*, X - \mathbb{E}(X)\rangle\,.$$

By setting $\lambda X^* =: X^*$ we get formula (50) and this concludes the proof. □

**Corollary 4.4** ((Upper) Mean-semideviation risk measure)**.** *Let the risk function $r_3 : L_p \to \mathbb{R}$ be given by $r_3(X) = \lambda d_3(X) - \mathbb{E}(X) = \lambda||(X - \mathbb{E}(X))_+||_p - \mathbb{E}(X)$ for $\lambda > 0$. Then it holds for $X \in L_p$:*

$$\partial r_3(X) = \{X^* - \mathbb{E}(X^*) - 1 : X^* \in L_q,\, X^* \in \mathcal{D}_{3'}\}\,, \tag{52}$$
$$\mathcal{D}_{3'} = \arg \max_{\substack{X^* \in \overline{B}_{L_q}(0,\lambda) \\ \cap (L_q)_+}} \langle X^*, X - \mathbb{E}(X)\rangle\,. \tag{53}$$

*Proof.* In analogy with Corollary 4.3 we get formula (52) by using Corollary 3.15, formula (35). □

**Remark 4.5.** *Formula (52) is identical to the formula that was given in [12, formula (4.12)].*

# References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] R.I. Boţ, S.-M. Grad, and G. Wanka. *Duality in Vector Optimization.* Springer-Verlag, Berlin Heidelberg, 2009.

[3] R.I. Boţ, N. Lorenz, and G. Wanka. Dual representations for convex risk measures via conjugate duality. *Journal of Optimization Theory and Applications*, 144(2):185–203, 2010.

[4] A. Brønsted and R.T. Rockafellar. On the subdifferentiability of convex functions. *Proceedings of the American Society*, 16:605–611, 1965.

[5] I. Ekeland and R. Temam. *Convex analysis and variational problems.* North-Holland Publishing Company, Amsterdam, 1976.

[6] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[7] H. Föllmer and A. Schied. *Stochastic finance. An introduction in discrete time.* Walter de Gruyter, Berlin, 2002.

[8] R.T. Rockafellar. Conjugate duality and optimization. Regional Conference Series in Applied Mathematics 16. Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[9] R.T. Rockafellar, S. Uryasev, and M. Zabarankin. Deviation measures in risk analysis and optimization. *Research Report 2002-7*, Department of Industrial and Systems Engineering, University of Florida, 2002.

[10] R.T. Rockafellar, S. Uryasev, and M. Zabarankin. Generalized deviations in risk analysis. *Finance and Stochastics*, 10:51–74, 2006.

[11] R.T. Rockafellar, S. Uryasev, and M. Zabarankin. Optimality conditions in portfolio analysis with general deviation measures. *Mathematical Programming*, Ser. B 108:515–540, 2006.

[12] A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. *Mathematics of Operations Research*, 31(3):433–452, 2006.

[13] C. Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific Publishing, Singapore, 2002.