Optim Lett
DOI 10.1007/s11590-013-0713-7

ORIGINAL PAPER

CVaR norm and applications in optimization

Konstantin Pavlikov · Stan Uryasev

Received: 17 May 2013 / Accepted: 21 November 2013

Abstract This paper introduces the family of *CVaR norms* in $\mathbb{R}^n$, based on the CVaR concept. The CVaR norm is defined in two variations: scaled and non-scaled. The well-known L_1 and L_∞ norms are limiting cases of the new family of norms. The D-norm, used in robust optimization, is equivalent to the non-scaled CVaR norm. We present two relatively simple definitions of the CVaR norm: (i) as the average or the sum of some percentage of largest absolute values of components of vector; (ii) as an optimal solution of a CVaR minimization problem suggested by Rockafellar and Uryasev. CVaR norms are piece-wise linear functions on $\mathbb{R}^n$ and can be used in various applications where the Euclidean norm is typically used. To illustrate, in the computational experiments we consider the problem of projecting a point onto a polyhedral set. The CVaR norm allows formulating this problem as a convex or linear program for any level of conservativeness.

Keywords CVaR norm · L_p norm · Projection

1 Introduction

The notion of norm is used in various mathematical applications. The great variety of norms arises from the existence of various applications, requiring the choice of

This research has been supported by the AFOSR grant FA9550-11-1-0258, "New Developments in Uncertainty: Linking Risk Management, Reliability, Statistics and Stochastic Optimization".

K. Pavlikov · S. Uryasev (✉)
Risk Management and Financial Engineering Lab, Department of Industrial and Systems Engineering, University of Florida,
303 Weil Hall, Gainesville, FL 32611, USA
e-mail: uryasev@ufl.edu

K. Pavlikov
e-mail: kpavlikov@ufl.edu

Published online: 22 January 2014

a specific norm. A well-known example of norms in $\mathbb{R}^n$ is the family of L_p norms, with popular cases $p = 1, 2, \infty$. Other values of p, especially $p > 2$, are rarely used because of the computational effort involving the calculation of norm values.

This paper introduces the family of CVaR norms in $\mathbb{R}^n$ space, which is based on the notion of CVaR, thoroughly discussed in [7,8]. The CVaR norm has a parameter $\alpha \in [0, 1]$ controlling the conservativeness of the norm. We consider scaled and non-scaled CVaR norms, different by a multiplying coefficient, and denoted by C_α^S and C_α, respectively. Despite this simple difference between the two families of CVaR norms, they exhibit different properties with respect to parameter α. The non-scaled CVaR norm spans the L_1 and L_∞ norms as the parameter α varies from 0 to 1. For every value of α, optimization involving the CVaR norm is a convex problem, which can be reduced to linear programming. In contrast, optimization of an L_p norm, especially with a large value p, can not easily be done using standard optimization solvers. For instance, the numerical experiments in Sect. 4 address the problem of projection of a point onto a polyhedral set using the CVaR norm, and show that the solution times of this problem do not significantly depend on the value of α.

This paper shows that the CVaR norm is equivalent to the D-norm introduced in [1] for robust optimization [1,2]. Robust optimization is a methodology dealing with optimization problems where problem parameters are not known exactly. For instance, a robust counterpart of a linear problem is a problem where column vectors in the matrix of constraints belong to some convex set. Various norms can be used to define this convex set. The robust counterpart of a linear problem with the uncertainty set defined by the D-norm is a linear problem [1]. While the D-norm is defined in combinatorial terms as a maximum over some subset of indices, the equivalent CVaR norm has a simple and intuitive definition that can be widely used in various mathematical and engineering areas.

The motivation for the definition of CVaR norms comes from the concept of the Fundamental Risk Quadrangle, developed in [9], which deals with stochastic random variables and relates Risk, Deviation, Error and Regret. Whereas Risk provides a single numerical surrogate for a random variable (usually evaluating large outcomes) and Regret has a strong connection to the utility function of outcomes of a random variable, Deviation and Error are used to assess the stochastic nature of a random value, namely “nonconstancy” and “nonzeroness”. For example, in regression analysis, the error, measured by the L_2 norm (corresponding to the Mean-Based Risk Quadrangle [9]) leads to the well-known simple least-squares linear regression. In statistics, the error, measured by the supremum of absolute difference between two cumulative distribution functions, or L_∞ (corresponding to the Range-Based Quadrangle), leads to the Kolmogorov-Smirnov distance between probability distributions and the corresponding goodness of fit test. In the process of evaluating the distances between random variables, we have observed that the CVaR of the absolute value of a random variable can serve as another example of Error measure and generates a norm in the space of random variables. However, discussion of CVaR norms in the stochastic case is beyond the scope of this paper, since the goal here is to explain the concept in the simplest deterministic $\mathbb{R}^n$ case.

The paper is organized as follows. Section 2 defines the family of scaled CVaR norms and discusses its properties. Section 3 introduces the family of non-scaled

CVaR norms. Section 4 considers a case study with the projection problem of a point onto a polyhedron using the scaled CVaR norm. Section 5 summarizes our results.

2 The scaled CVaR norm

By definition, a *norm* on $\mathbb{R}^n$ is a function $\rho : \mathbb{R}^n \to \mathbb{R}$ that satisfies the following properties:

1. $\rho(\lambda \mathbf{x}) = |\lambda| \mathbf{x},\ \forall \mathbf{x} \in \mathbb{R}^n,\ \forall \lambda \in \mathbb{R}$.
2. $\rho(\mathbf{x} + \mathbf{y}) \leq \rho(\mathbf{x}) + \rho(\mathbf{y}),\ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
3. $\rho(\mathbf{x}) = 0 \Rightarrow \mathbf{x} = 0$.

This section introduces the *scaled CVaR norm* C_α^S on $\mathbb{R}^n$ space and establishes connection to the scaled L_p^S norm of vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, defined by

$$\|\mathbf{x}\|_p^S = \left(\frac{1}{n} \sum_{i=1}^{n} |x|_i^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

The class of scaled L_p^S norms includes the least conservative L_1^S norm (scaled "Manhattan" norm):

$$\|\mathbf{x}\|_1^S = \frac{1}{n} \sum_{i=1}^{n} |x_i|,$$

and the most conservative L_∞ norm:

$$\|\mathbf{x}\|_\infty^S = \max_i |x_i|. \tag{1}$$

We will denote the scaled CVaR norm for the vector $\mathbf{x} = (x_1, \dots, x_n)$ by $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S$, where α is parameter, in the range $0 \leq \alpha \leq 1$. The following definition of C_α^S is inspired by the CVaR definition in Portfolio Safeguard Package [5].

Definition 1 Let us order absolute values of components of vector $\mathbf{x} \in \mathbb{R}^n$, as follows $|x_{(1)}| \leq |x_{(2)}| \leq \cdots \leq |x_{(n)}|$.
For $\alpha_j = \frac{j}{n}$, $j = 0, \dots, n-1$, the scaled CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}^S$ of vector $\mathbf{x}$ with parameter α_j equals:

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}^S = \frac{1}{n-j} \sum_{i=j+1}^{n} |x_{(i)}|. \tag{2}$$

For α, such that $\alpha_j < \alpha < \alpha_{j+1}$, $j = 0, \dots, n-2$, the scaled CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S$ equals the weighted average of $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}^S$ and $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}^S$:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \mu \langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}^S + (1-\mu) \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}^S, \tag{3}$$

where

$$\mu = \frac{(\alpha_{j+1} - \alpha)(1 - \alpha_j)}{(\alpha_{j+1} - \alpha_j)(1 - \alpha)}.$$

For α, such that $\frac{n-1}{n} < \alpha \leq 1$,

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \max_i |x_i|. \tag{4}$$

The fact that Definition 1 defines a norm follows from Propositions 2.2 and 2.3. To build some intuition about C_α^S, we provide the following examples.

Example 1 Figure 1 compares unit disks (balls) for C_α^S and L_p^S norms in $\mathbb{R}^2$ space. A unit disk (ball) for a norm is a set of vectors with norm less or equal than 1, i.e., $U_\alpha^S = \{\mathbf{x} = (x_1, x_2) \mid \langle\langle \mathbf{x} \rangle\rangle_\alpha^S \leq 1\}$ for the scaled CVaR norm with parameter α. The following values of parameter α are considered: $0, 0.1, 1 - \frac{1}{\sqrt{2}}, 0.4, 1$. The corresponding values of the parameter p of L_p^S norm are obtained using transformation $p(\alpha) = \frac{1}{(1-\alpha)^2}$.

The following example illustrates the C_α^S norm numerically.

Example 2 Let $\mathbf{x} \in \mathbb{R}^5$, $\mathbf{x} = (2, 1, 7, 10, -12)$. The vector of ordered absolute values of components, $(|x_{(1)}|, |x_{(2)}|, \ldots, |x_{(5)}|)$, equals $(1, 2, 7, 10, 12)$. Then, the C_α^S norm equals

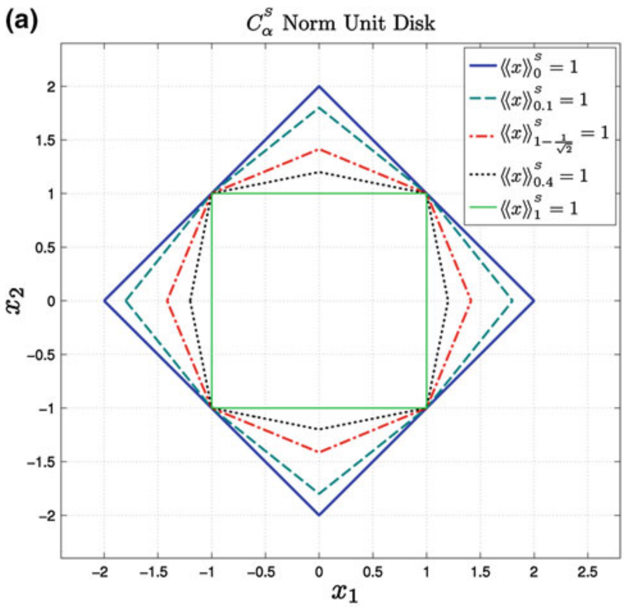
$$\begin{aligned}
\langle\langle \mathbf{x} \rangle\rangle_0^S &= \frac{1+2+7+10+12}{5} = 6.4, \quad \text{for } \alpha = \frac{0}{5} = 0, \\
\langle\langle \mathbf{x} \rangle\rangle_{0.2}^S &= \frac{2+7+10+12}{4} = 7.75, \quad \text{for } \alpha = \frac{1}{5} = 0.2, \\
\langle\langle \mathbf{x} \rangle\rangle_{0.4}^S &= \frac{7+10+12}{3} \cong 9.67, \quad \text{for } \alpha = \frac{2}{5} = 0.4, \\
\langle\langle \mathbf{x} \rangle\rangle_{0.6}^S &= \frac{10+12}{2} = 11, \quad \text{for } \alpha = \frac{3}{5} = 0.6, \\
\langle\langle \mathbf{x} \rangle\rangle_{0.8}^S &= \max_i(|x_i|) = 12, \quad \text{for } \alpha = \frac{4}{5} = 0.8.
\end{aligned}$$

For $\alpha = 0.05$, which is between $\alpha_0 = 0$ and $\alpha_1 = 0.2$, $\langle\langle \mathbf{x} \rangle\rangle_{0.05}^S$ equals the weighted average of $\langle\langle \mathbf{x} \rangle\rangle_0^S$ and $\langle\langle \mathbf{x} \rangle\rangle_{0.2}^S$,

$$\begin{aligned}
&\mu = \frac{0.15 \cdot 1}{0.2 \cdot 0.95} \cong 0.79, \\
&\langle\langle \mathbf{x} \rangle\rangle_{0.05}^S = \mu \langle\langle \mathbf{x} \rangle\rangle_0^S + (1-\mu)\langle\langle \mathbf{x} \rangle\rangle_{0.2}^S \cong 0.79 \cdot 6.4 + 0.21 \cdot 7.75 \cong 6.68.
\end{aligned}$$

For $\alpha > 0.8$,

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \langle\langle \mathbf{x} \rangle\rangle_{0.8}^S = 12.$$



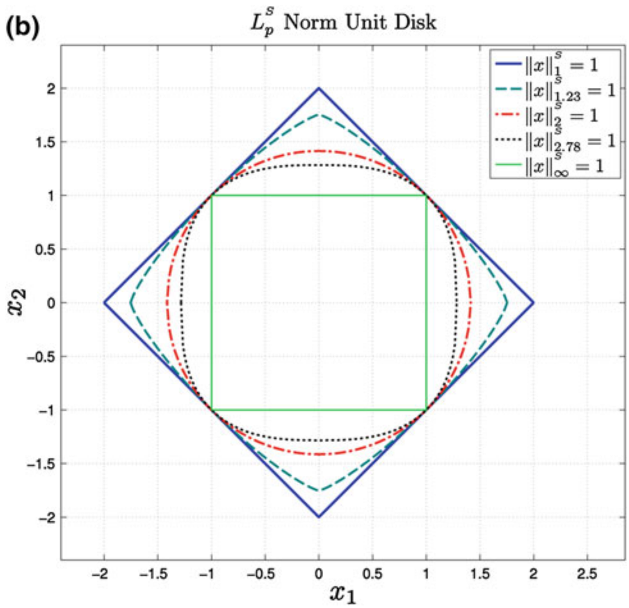


Fig. 1 **a** Unit disks for C_α^S in $\mathbb{R}^2$ with $\alpha = 0, 0.1, 1 - \frac{1}{\sqrt{2}}, 0.4, 1$. A unit disk is a set of vectors $\mathbf{x} = (x_1, x_2)$ in $\mathbb{R}^2$, such that $\langle\langle \mathbf{x}\rangle\rangle_\alpha^S \leq 1$. The scaled CVaR norm with $\alpha = 0$ corresponds to the scaled L_1^S, and the scaled CVaR norm with $\alpha \geq 0.5$ corresponds to L_∞, according to (4). Observe that $U_1^S \subset U_{0.4}^S \subset U_{1-\frac{1}{\sqrt{2}}}^S \subset U_{0.1}^S \subset U_0^S$, which is a corollary of Proposition 2.1. **b** Unit disks for the L_p^S norm in $\mathbb{R}^2$ with parameter $p = 1, 1.23, 2, 2.78, \infty$, where p values are obtained from α using transformation $p(\alpha) = \frac{1}{(1-\alpha)^2}$

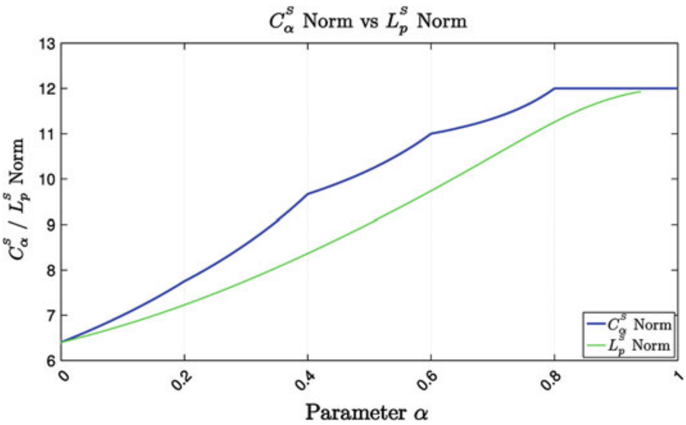


Fig. 2 The graph of C_α^S for the vector $\mathbf{x} = (2, 1, 7, 10, -12)$. The graph of the L_p^S norm $\|\mathbf{x}\|_p^S$ with $p(\alpha) = \frac{1}{(1-\alpha)^2}$ is included for comparison purpose. Parameter p is set to make $\langle\langle x \rangle\rangle_\alpha^S$ and $\|x\|_p^S$ visible on the same scale. For instance, when $\alpha = 0.6$, $p(\alpha) = 6.25$; when $\alpha = 0.8$, $p(\alpha) = 25$. All levels $p > 25$ of the L_p^S norm correspond to the most conservative parameter $\alpha = 1$ of C_α^S norm

Figure 2 compares C_α^S and L_p^S norms numerically. Parameter α of C_α^S is in the range $0 \leq \alpha \leq 1$ and parameter p of L_p^S is in the range $1 \leq p \leq \infty$. We set parameter p so that $\alpha = 0$ corresponds to $p = 1$ and $\alpha = 1$ corresponds to $p = \infty$ and both norms can be presented on the same figure. We use the following mapping function, $p(\alpha) = \frac{1}{(1-\alpha)^2}$, even though it is not the only possible choice.

Below, we provide an alternative definition of the scaled CVaR norm. Let us denote $[t]^+ = \max(t, 0)$. This alternative definition comes from the concept of CVaR for random variables, discussed in [7,8]. The CVaR of a random variable X with confidence level α is defined as follows:

$$\text{CVaR}_\alpha(X) = \min_c \left(c + \frac{1}{1-\alpha} E[X - c]^+ \right), \tag{5}$$

where $E(\cdot)$ denotes the expectation of a random variable. If X is a discrete random variable, distributed on $\{x_1, \ldots, x_n\}$, such that the probability $\mathbb{P}(X = x_i) = p_i$, the CVaR of X is expressed as follows:

$$\text{CVaR}_\alpha(X) = \min_c \left(c + \frac{1}{1-\alpha} \sum_{i=1}^{n} p_i [x_i - c]^+ \right). \tag{6}$$

The scaled CVaR norm of vector $\mathbf{x} = (x_1, \ldots, x_n)$ defined by Definition 1, is the CVaR of a random variable X, uniformly distributed on the set of outcomes $\{|x_1|, \ldots, |x_n|\}$, i.e., the probability $\mathbb{P}(X = |x_i|) = \frac{1}{n}$, if every outcome $|x_i|$ is unique. If some outcomes coincide, say $|x_i| = |x_j|$, for some $i \neq j$, then probability $\mathbb{P}(X = |x_i|) = \frac{2}{n}$. Thus, the following alternative definition of the C_α^S norm is valid.

Definition 2 For $\mathbf{x} \in \mathbb{R}^n$, the scaled CVaR norm C_α^S is defined as follows:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \min_c \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [|x_i| - c]^+ \right), \quad \text{for } 0 \le \alpha < 1, \tag{7}$$

$$\langle\langle \mathbf{x} \rangle\rangle_1^S = \max_i |x_i|. \tag{8}$$

We proceed as follows. First, we show that C_α^S defined by (7), (8) is a nondecreasing function of parameter α. Second, we prove that Definition 2 defines a norm. Then, we prove that Definitions 1 and 2 are equivalent.

Proposition 2.1 *For* $\mathbf{x} \in \mathbb{R}^n$, $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S$ *is a nondecreasing function of the parameter* α*, i.e.,*

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_1}^S \le \langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}^S, \quad \textit{for } 0 \le \alpha_1 \le \alpha_2 \le 1, \ \mathbf{x} \in \mathbb{R}^n.$$

Proof For $\alpha_1 \le \alpha_2 < 1$,

$$\begin{aligned}
\langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}^S &= \min_c \left(c + \frac{1}{n(1-\alpha_2)} \sum_{i=1}^n [|x_i| - c]^+ \right) \\
&= \min_c \left(c + \frac{1}{n(1-\alpha_1)} \sum_{i=1}^n [|x_i| - c]^+ + \left(\frac{1}{n(1-\alpha_2)} - \frac{1}{n(1-\alpha_1)} \right) \sum_{i=1}^n [|x_i| - c]^+ \right) \\
&\ge \min_c \left(c + \frac{1}{n(1-\alpha_1)} \sum_{i=1}^n [|x_i| - c]^+ \right) + \min_c \frac{1}{n} \left(\frac{1}{1-\alpha_2} - \frac{1}{1-\alpha_1} \right) \sum_{i=1}^n [|x_i| - c]^+ .
\end{aligned}$$

Since $\frac{1}{1-\alpha_2} - \frac{1}{1-\alpha_1} = \frac{1-\alpha_1-1+\alpha_2}{(1-\alpha_2)(1-\alpha_1)} = \frac{\alpha_2-\alpha_1}{(1-\alpha_2)(1-\alpha_1)} \ge 0$, then $\min_c \frac{1}{n} \left(\frac{1}{1-\alpha_2} - \frac{1}{1-\alpha_1} \right) \sum_{i=1}^n [|x_i| - c]^+ = \frac{1}{n} \left(\frac{1}{1-\alpha_2} - \frac{1}{1-\alpha_1} \right) \min_c \sum_{i=1}^n [|x_i| - c]^+ = 0$.

Therefore,

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}^S \ge \min_c \left(c + \frac{1}{n(1-\alpha_1)} \sum_{i=1}^n [|x_i| - c]^+ \right) = \langle\langle \mathbf{x} \rangle\rangle_{\alpha_1}^S .$$

For $\alpha_1 < \alpha_2 = 1$,

$$\begin{aligned}
\langle\langle \mathbf{x} \rangle\rangle_{\alpha_1}^S &= \min_c \left(c + \frac{1}{n(1-\alpha_1)} \sum_{i=1}^n [|x_i| - c]^+ \right) \le |x_{(n)}| + \frac{1}{n(1-\alpha_1)} \sum_{i=1}^n \left[|x_i| - |x_{(n)}| \right]^+ \\
&= |x_{(n)}| = \langle\langle \mathbf{x} \rangle\rangle_1^S = \langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}^S .
\end{aligned}$$

Thus, for any α_1, α_2, $0 \le \alpha_1 \le \alpha_2 \le 1$, we have $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_1}^S \le \langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}^S$, which completes the proof. □

Proposition 2.2 *Definition 2 defines a norm, i.e., the following properties of a norm are satisfied:*

1. $\langle\langle \lambda \cdot \mathbf{x} \rangle\rangle_\alpha^S = |\lambda| \cdot \langle\langle \mathbf{x} \rangle\rangle_\alpha^S,\ \lambda \in \mathbb{R},\ \mathbf{x} \in \mathbb{R}^n.$
2. $\langle\langle \mathbf{x} + \mathbf{y} \rangle\rangle_\alpha^S \le \langle\langle \mathbf{x} \rangle\rangle_\alpha^S + \langle\langle \mathbf{y} \rangle\rangle_\alpha^S,\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^n.$
3. $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = 0 \Rightarrow \mathbf{x} = \mathbf{0},\ \mathbf{x} \in \mathbb{R}^n.$

Proof **Property 1**. *If* $\lambda = 0$, *then* $\langle\langle 0 \cdot \mathbf{x} \rangle\rangle_\alpha^S = 0 = 0 \cdot \langle\langle \mathbf{x} \rangle\rangle_\alpha^S$. If $\lambda \neq 0$,

$$\begin{aligned}\langle\langle \lambda \cdot \mathbf{x} \rangle\rangle_\alpha^S &= \min_c \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|\lambda||x_i| - c\right]^+ \right) \\ &= \min_c |\lambda| \left(\frac{c}{|\lambda|} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i| - \frac{c}{|\lambda|}\right]^+ \right) \\ &= \min_c |\lambda| \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i| - c\right]^+ \right) = |\lambda| \cdot \langle\langle \mathbf{x} \rangle\rangle_\alpha^S.\end{aligned}$$

Property 2. A similar property is proved for a more general case in [7, Theorem 2], and in [4]. Here we provide the proof for this special case for the convenience of the reader. The scaled CVaR norm of $\mathbf{x}$ equals:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \min_c \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [|x_i| - c]^+ \right) = c_{\mathbf{x}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [|x_i| - c_{\mathbf{x}}]^+ ,$$

where $c_{\mathbf{x}}$ denotes argminimum of c in expression (7). Similarly,

$$\langle\langle \mathbf{y} \rangle\rangle_\alpha^S = \min_c \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [|y_i| - c]^+ \right) = c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|y_i| - c_{\mathbf{y}}\right]^+ .$$

Thus,

$$\begin{aligned}\langle\langle \mathbf{x} + \mathbf{y} \rangle\rangle_\alpha^S &= \min_c \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [|x_i + y_i| - c]^+ \right) \\ &\le c_{\mathbf{x}} + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i + y_i| - c_{\mathbf{x}} - c_{\mathbf{y}}\right]^+ .\end{aligned} \tag{9}$$

Since $[\cdot]^+$ is a nondecreasing function, we can write

$$\begin{aligned}&c_{\mathbf{x}} + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i + y_i| - c_{\mathbf{x}} - c_{\mathbf{y}}\right]^+ \\ &\quad\le c_{\mathbf{x}} + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i| + |y_i| - c_{\mathbf{x}} - c_{\mathbf{y}}\right]^+ .\end{aligned} \tag{10}$$

The function $[\cdot]^+$ satisfies the inequality:

$$[a+b]^+ \le [a]^+ + [b]^+, \quad a, b \in \mathbb{R}. \tag{11}$$

Therefore, (10), (11) imply

$$c_{\mathbf{x}} + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left[|x_i| + |y_i| - c_{\mathbf{x}} - c_{\mathbf{y}}\right]^+ \tag{12}$$

$$\le c_{\mathbf{x}} + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left([|x_i| - c_{\mathbf{x}}]^+ + \left[|y_i| - c_{\mathbf{y}}\right]^+\right) \tag{13}$$

$$= c_{\mathbf{x}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} [|x_i| - c_{\mathbf{x}}]^+ + c_{\mathbf{y}} + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left[|y_i| - c_{\mathbf{y}}\right]^+$$

$$= \langle\langle \mathbf{x} \rangle\rangle_\alpha^S + \langle\langle \mathbf{y} \rangle\rangle_\alpha^S. \tag{14}$$

Thus, $\langle\langle \mathbf{x}+\mathbf{y} \rangle\rangle_\alpha^S \le \langle\langle \mathbf{x} \rangle\rangle_\alpha^S + \langle\langle \mathbf{y} \rangle\rangle_\alpha^S$, which completes the proof of Property 2.

Property 3. Since $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = 0$ and parameter $\alpha \ge 0$, then $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S \ge \langle\langle \mathbf{x} \rangle\rangle_0^S$ by the monotonicity property in Proposition 2.1. Thus

$$0 = \langle\langle \mathbf{x} \rangle\rangle_\alpha^S \ge \langle\langle \mathbf{x} \rangle\rangle_0^S = \frac{1}{n} \sum_{i=1}^{n} |x_i| \ge 0.$$

Consequently, $\sum_{i=1}^{n} |x_i| = 0$, and $x_i = 0$, $i = 1, \ldots, n$. □

Proposition 2.3 *Definitions 1 and 2 are equivalent.*

Proof Let us consider an $\alpha < 1$ and corresponding α_j, α_{j+1}, such that

$$\alpha_j = \frac{j}{n}, \quad j \in \{0, \ldots, n\}, \tag{15}$$

$$\alpha_j \le \alpha < \alpha_{j+1}. \tag{16}$$

Let us denote, $f(c) = c + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} [|x_i| - c]^+$. First, we show that $|x_{(j+1)}| \in \operatorname{argmin}_c f(c)$, i.e.,

$$\min_c f(c) = |x_{(j+1)}| + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left[|x_i| - |x_{(j+1)}|\right]^+.$$

The function $f(c)$ is a convex function because it is a linear combination of convex functions c and $[|x_i| - c]^+$ [6]. To show that $|x_{(j+1)}| \in \operatorname{argmin}_c f(c)$, it is sufficient to prove that zero belongs to the generalized gradient $\partial f(|x_{(j+1)}|)$ of the convex function $f(\cdot)$ at point $|x_{(j+1)}|$,

$$0 \in \partial f\left(|x_{(j+1)}|\right).$$

Let us find $\partial f\left(|x_{(j+1)}|\right)$,

$$\begin{aligned}\partial f\left(|x_{(j+1)}|\right) &= \{1\} + \frac{1}{n(1-\alpha)}\partial\left(\sum_{i=1}^{n}\left([|x_i|-c]^+\right)\right)\Bigg|_{c=|x_{(j+1)}|}\\ &= \{1\} + \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\partial\left(\left[|x_{(i)}|-c\right]^+\right)\Big|_{c=|x_{(j+1)}|}.\end{aligned}$$

The generalized gradient of the function $[-t]^+$ can be expressed as follows:

$$\partial\left[-t\right]^+ = \begin{cases}\{0\}, & \text{if } t>0,\\ [-1,0], & \text{if } t=0,\\ \{-1\}, & \text{otherwise.}\end{cases} \tag{17}$$

If $j=n-1$, then

$$\begin{aligned}\partial f\left(|x_{(n)}|\right) &= \{1\} + \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\partial\left(\left[|x_{(i)}|-c\right]^+\right)\Big|_{c=|x_{(n)}|}\\ &= \{1\} + \frac{1}{n(1-\alpha)}\sum_{i=1}^{n-1}\{0\} + \left[-\frac{1}{n(1-\alpha)},0\right]\\ &= \left[1-\frac{1}{n(1-\alpha)},1\right] = \left[\frac{n-n\alpha-1}{n(1-\alpha)},1\right].\end{aligned}$$

Since $j=n-1$ and (16), we know that $n\alpha \geq n-1$, therefore $n-n\alpha-1 \leq 0$. Thus,

$$0 \in \partial f\left(|x_{(n)}|\right).$$

If $1 \leq j \leq n-2$, then

$$\begin{aligned}\partial f\left(|x_{(j+1)}|\right) &= \{1\} + \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\partial\left(\left[|x_{(i)}|-c\right]^+\right)\Big|_{c=|x_{(j+1)}|}\\ &= \{1\} + \frac{1}{n(1-\alpha)}\sum_{i=1}^{j}\{0\} + \left[-\frac{1}{n(1-\alpha)},0\right] + \frac{1}{n(1-\alpha)}\sum_{i=j+2}^{n}\{-1\}\\ &= \{1\} + \left[-\frac{1}{n(1-\alpha)},0\right] + \left\{-\frac{n-j-1}{n(1-\alpha)}\right\}\\ &= \{1\} + \left[-\frac{n-j}{n(1-\alpha)},-\frac{n-j-1}{n(1-\alpha)}\right]\\ &= \left[1-\frac{n-j}{n(1-\alpha)},1-\frac{n-j-1}{n(1-\alpha)}\right]\\ &= \left[\frac{n-n\alpha-n+j}{n(1-\alpha)},\frac{n-n\alpha-n+j+1}{n(1-\alpha)}\right]\\ &= \left[\frac{-n\alpha+j}{n(1-\alpha)},\frac{-n\alpha+j+1}{n(1-\alpha)}\right].\end{aligned}$$

Since $\alpha_j = \frac{j}{n} \le \alpha$ by (16), then $-n\alpha + j \le 0$. Since $\alpha < \frac{j+1}{n}$ by (16), then $j + 1 - n\alpha > 0$. Therefore,

$$0 \in \left[\frac{-n\alpha + j}{n(1-\alpha)}, \frac{-n\alpha + j + 1}{n(1-\alpha)}\right].$$

If $j = 0$, then

$$\begin{aligned}
\partial f\left(|x_{(1)}|\right) &= \{1\} + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \partial\left(\left[|x_{(i)}| - c\right]^+\right)\Big|_{c=|x_{(1)}|} \\
&= \{1\} + \left[-\frac{1}{n(1-\alpha)}, 0\right] + \frac{1}{n(1-\alpha)} \sum_{i=2}^{n} \{-1\} \\
&= \left[1 - \frac{n}{n(1-\alpha)}, 1 - \frac{n-1}{n(1-\alpha)}\right] \\
&= \left[-\frac{n\alpha}{n(1-\alpha)}, \frac{1 - n\alpha}{n(1-\alpha)}\right].
\end{aligned} \tag{18}$$

Since $0 \le \alpha < \frac{1}{n}$ by (16), then $1 - n\alpha > 0$. Therefore,

$$0 \in \partial f\left(|x_{(1)}|\right).$$

Thus, we have proven that

$$|x_{(j+1)}| \in \operatorname*{argmin}_c \ f(c).$$

Therefore, Definition 2 can be rewritten as follows:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = |x_{(j+1)}| + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left[|x_i| - |x_{(j+1)}|\right]^+. \tag{19}$$

Let $\alpha = \alpha_j = \frac{j}{n}$, $j = 0, \ldots, n-1$, then

$$\begin{aligned}
\langle\langle \mathbf{x} \rangle\rangle_\alpha^S &= |x_{(j+1)}| + \frac{1}{n-j} \sum_{i=j+1}^{n} \left(|x_{(i)}| - |x_{(j+1)}|\right) \\
&= |x_{(j+1)}| + \frac{1}{n-j} \sum_{i=j+1}^{n} |x_{(i)}| - \frac{1}{n-j} \sum_{i=j+1}^{n} |x_{(j+1)}| = \frac{1}{n-j} \sum_{i=j+1}^{n} |x_{(i)}|.
\end{aligned} \tag{20}$$

Therefore, equality (20) is identical to (2), which proves the equivalence of Definitions 1 and 2 for the case $\alpha = \alpha_j$. Let us consider α, such that $\frac{j}{n} = \alpha_j < \alpha < \alpha_{j+1} =$

$\frac{j+1}{n}$, $j = 0, \ldots, n-2$. Since $\frac{1}{1-\alpha_j} < \frac{1}{1-\alpha} < \frac{1}{1-\alpha_{j+1}}$, we can represent $\frac{1}{1-\alpha}$ as a convex combination of $\frac{1}{1-\alpha_j}$ and $\frac{1}{1-\alpha_{j+1}}$:

$$\frac{1}{1-\alpha} = \frac{\mu}{1-\alpha_j} + \frac{1-\mu}{1-\alpha_{j+1}},$$

where

$$\begin{aligned}\mu &= \frac{\frac{1}{1-\alpha} - \frac{1}{1-\alpha_{j+1}}}{\frac{1}{1-\alpha_j} - \frac{1}{1-\alpha_{j+1}}} = \frac{\frac{(1-\alpha_j)(1-\alpha_{j+1})}{1-\alpha} - (1-\alpha_j)}{\alpha_j - \alpha_{j+1}} \\ &= \frac{(1-\alpha_j)(\alpha - \alpha_{j+1})}{(\alpha_j - \alpha_{j+1})(1-\alpha)} = \frac{(1-\alpha_j)(\alpha_{j+1} - \alpha)}{(\alpha_{j+1} - \alpha_j)(1-\alpha)}.\end{aligned}$$

According to Definition 2, C_α^S is represented as follows:

$$\begin{aligned}\langle\langle \mathbf{x} \rangle\rangle_\alpha^S &= |x_{(j+1)}| + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_i| - |x_{(j+1)}|\right]^+ \\ &= |x_{(j+1)}| + \frac{1}{n(1-\alpha)} \sum_{i=1}^n \left[|x_{(i)}| - |x_{(j+1)}|\right]^+ \\ &= |x_{(j+1)}| + \left(\frac{\mu}{n(1-\alpha_j)} + \frac{1-\mu}{n(1-\alpha_{j+1})}\right) \sum_{i=1}^n \left[|x_{(i)}| - |x_{(j+1)}|\right]^+ \\ &= |x_{(j+1)}| + \frac{\mu}{n(1-\alpha_j)} \sum_{i=j+1}^n \left(|x_{(i)}| - |x_{(j+1)}|\right) \\ &\quad + \frac{1-\mu}{n(1-\alpha_{j+1})} \sum_{i=j+2}^n \left(|x_{(i)}| - |x_{(j+1)}|\right) \\ &= |x_{(j+1)}| + \frac{\mu}{n(1-\alpha_j)} \sum_{i=j+1}^n |x_{(i)}| - \frac{\mu(n-j)}{n(1-\alpha_j)} |x_{(j+1)}| \\ &\quad + \frac{1-\mu}{n(1-\alpha_{j+1})} \sum_{i=j+2}^n |x_{(i)}| - \frac{(1-\mu)(n-j-1)}{n(1-\alpha_{j+1})} |x_{(j+1)}|.\end{aligned}$$

Since $n - j = n(1-\alpha_j)$ and $n - j - 1 = n(1-\alpha_{j+1})$, we obtain

$$\begin{aligned}\langle\langle \mathbf{x} \rangle\rangle_\alpha^S &= \frac{\mu}{n(1-\alpha_j)} \sum_{i=j+1}^n |x_{(i)}| + \frac{1-\mu}{n(1-\alpha_{j+1})} \sum_{i=j+2}^n |x_{(i)}| \\ &= \mu \langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}^S + (1-\mu) \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}^S.\end{aligned}$$

Finally, let us consider α, such that $\frac{n-1}{n} < \alpha < 1$. With (19), C_α^S can be expressed as

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = |x_{(n)}| + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} \left[|x_i| - |x_{(n)}| \right]^+ = |x_{(n)}|,$$

which corresponds to (4), and the proof is complete. □

Remark 1 The scaled L_1^S norm is a special case of the scaled CVaR norm defined by Definition 2, when $\alpha = 0$, i.e.,

$$\|\mathbf{x}\|_1^S = \langle\langle \mathbf{x} \rangle\rangle_0^S.$$

Proof From representation (19) of C_α^S,

$$\langle\langle \mathbf{x} \rangle\rangle_0^S = |x_{(1)}| + \frac{1}{n} \sum_{i=1}^{n} \left(|x_{(i)}| - |x_{(1)}| \right) = \frac{1}{n} \sum_{i=1}^{n} |x_{(i)}| = \|\mathbf{x}\|_1^S.$$

□

Remark 2 The L_∞ norm is a special case of the CVaR norm defined by Definition 2, when $\frac{n-1}{n} \leq \alpha \leq 1$, i.e.,

$$\|\mathbf{x}\|_\infty^S = \langle\langle \mathbf{x} \rangle\rangle_\alpha^S, \quad \text{for } \frac{n-1}{n} \leq \alpha \leq 1.$$

Proof Follows from definitions (1) and (4). □

Finally, note that $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S$, $\mathbf{x} \in \mathbb{R}^\mathbf{n}$ can be obtained by solving the following optimization problem:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \min_{z_i, c} \left(c + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} z_i \right) \tag{21}$$

subject to

$$z_i \geq |x_i| - c, i = 1, \ldots, n, \tag{22}$$

$$z_i \geq 0, i = 1, \ldots, n. \tag{23}$$

The set of constraints (22)–(23) follows from the definition of $[|x_i| - c]^+$.

3 The non-scaled CVaR norm

This section defines the non-scaled CVaR norm in $\mathbb{R}^n$ and studies its properties. We will call the non-scaled CVaR norm by just CVaR norm.

Definition 3 The CVaR norm $\langle\langle \cdot \rangle\rangle_\alpha$ with parameter α, such that $0 \leq \alpha < 1$, is defined by:

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = n(1-\alpha) \cdot \langle\langle \mathbf{x} \rangle\rangle_\alpha^S. \tag{24}$$

The family of standard L_p norms, $\|\cdot\|_p$, for a vector $\mathbf{x}$ is defined by

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x|_i^p\right)^{\frac{1}{p}}.$$

Notice that the scaled L_p^S norm $\|\cdot\|_p^S$, considered in the previous section, is obtained by multiplying $\|\cdot\|_p$ by $n^{-\frac{1}{p}}$:

$$\|\mathbf{x}\|_p^S = n^{-\frac{1}{p}} \cdot \|\mathbf{x}\|_p.$$

The boundary cases of L_p norms are the L_1 and L_∞ norms:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|, \tag{25}$$

$$\|\mathbf{x}\|_\infty = \max_i |x_i|. \tag{26}$$

Remark 3 The L_1 norm is a special case of C_α for $\alpha = 0$, i.e.,

$$\|\mathbf{x}\|_1 = \langle\langle \mathbf{x} \rangle\rangle_0.$$

Proof

$$\langle\langle \mathbf{x} \rangle\rangle_0 = n(1-0)\langle\langle \mathbf{x} \rangle\rangle_{\alpha=0}^S = n\frac{1}{n}\sum_{i=1}^{n} |x_i| = \|\mathbf{x}\|_\mathbf{1}.$$

□

Remark 4 The L_∞ norm is a special case of C_α for $\alpha = \frac{n-1}{n}$, i.e.,

$$\|\mathbf{x}\|_\infty = \langle\langle \mathbf{x} \rangle\rangle_{\frac{n-1}{n}}.$$

Proof Follows from definitions (24) and the fact that $\langle\langle \mathbf{x} \rangle\rangle_\alpha^S = \max_i |x_i|$ and $n(1-\alpha) = 1$ for $\alpha = \frac{n-1}{n}$. □

Let us order the absolute values of components of the vector $\mathbf{x} \in \mathbb{R}^n$, as follows $|x_{(1)}| \le |x_{(2)}| \le \cdots \le |x_{(n)}|$. The CVaR norm C_α can be written similar to Definition 1 for the scaled CVaR norm C_α^S.

Proposition 3.1 *For* $\alpha_j = \frac{j}{n},\ j = 0, \ldots, n-1$, *the CVaR norm* $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}$ *of the vector* $\mathbf{x}$ *with parameter* α_j *equals:*

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j} = \sum_{i=j+1}^{n} |x_{(i)}|. \tag{27}$$

For α, *such that* $\alpha_j < \alpha < \alpha_{j+1},\ j = 0, \ldots, n-2$, *the CVaR norm* $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ *equals the weighted average of* $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}$ *and* $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}$*:*

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = \lambda \langle\langle \mathbf{x} \rangle\rangle_{\alpha_j} + (1-\lambda)\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}, \tag{28}$$

where

$$\lambda = \frac{\alpha_{j+1} - \alpha}{\alpha_{j+1} - \alpha_j}.$$

For α*, such that* $\frac{n-1}{n} < \alpha < 1$*, the CVaR norm* $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ *equals*

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = n(1-\alpha)\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{n-1}} = n(1-\alpha)\max_i |x_i|.$$

Proof Representation (27) directly follows from the representation (2) by noticing that $n(1-\alpha) = n(1-\alpha_j) = n - j$. Then, (28) follows from the multiplication of (3) by $n(1-\alpha)$. □

The following proposition was proved for a more general stochastic case in [3]. Here we provide a simplified proof in the deterministic case.

Proposition 3.2 *For* $\mathbf{x} \in \mathbb{R}^n$*, the CVaR norm* $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ *is a non-increasing, concave, piecewise-linear function of the parameter* α*.*

Proof First, we show that C_α is a non-increasing function at discrete points $\alpha_j = \frac{j}{n}$, $j = 0, \ldots, n-1$. Let $\alpha_{j_1} < \alpha_{j_2}$, then

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j_1}} = \sum_{i=j_1+1}^{n} |x_{(i)}| = \sum_{i=j_1+1}^{j_2} |x_{(i)}| + \sum_{i=j_2+1}^{n} |x_{(i)}| \geq \sum_{i=j_2+1}^{n} |x_{(i)}| = \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j_2}}. \quad (29)$$

Using (28) for $\alpha \neq \frac{j}{n}$, $j = 0, \ldots, n-2$ and $\alpha < \frac{n-1}{n}$, we observe that $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ is a convex combination of corresponding $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}$ and $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}$ with $\lambda = \frac{\alpha_{j+1}-\alpha}{\alpha_{j+1}-\alpha_j}$, linear with respect to α. Since $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j} \geq \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j_1+1}}$ by (29), the proposition is proved for $\alpha \in [0, \frac{n-1}{n}]$. To finish the proof, it suffices to note that for $\alpha \in (\frac{n-1}{n}, 1)$, the CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_\alpha = n(1-\alpha) \cdot \max_i |x_i|$, which is a linear non-increasing function of parameter α. To prove concavity, we need to show that for every α_{j-1}, α_j, α_{j+1}, $j = 1, \ldots, n-2$

$$\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j} \geq \frac{1}{2}\left(\langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j-1}} + \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}}\right).$$

Indeed, using (27), we obtain

$$\begin{aligned} \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j-1}} + \langle\langle \mathbf{x} \rangle\rangle_{\alpha_{j+1}} &= \sum_{i=j}^{n} |x_{(i)}| + \sum_{i=j+2}^{n} |x_{(i)}| \\ &= 2\sum_{i=j+1}^{n} |x_{(i)}| + |x_{(j)}| - |x_{(j+1)}| \leq 2\sum_{i=j+1}^{n} |x_{(i)}| = 2\langle\langle \mathbf{x} \rangle\rangle_{\alpha_j}. \end{aligned} \quad (30)$$

□

Similar to the scaled CVaR norm C_α^S, we provide the minimization formula for C_α.

Proposition 3.3 *The CVaR norm C_α can be obtained by solving the following minimization problem:*

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = \min_c \left(n(1-\alpha)c + \sum_{i=1}^{n} [|x_i| - c]^+ \right), \quad \textit{for } 0 \le \alpha < 1. \tag{31}$$

Proof The statement follows from Definitions 2 and 3. □

Definition 4 The D-norm with parameter $p \in [1, n]$, introduced in [1] and denoted by $|||\mathbf{x}|||_p$, is defined as follows:

$$|||\mathbf{x}|||_p = \max_{S,t} \left(\sum_{i \in S} |x_i| + (p - \lfloor p \rfloor)|x_t| \right), \tag{32}$$

where

$$\lfloor p \rfloor = \max_{i \le p,\, i \in \{1,\ldots,n\}} i, \tag{33}$$

$$S \subseteq \{1, \ldots, n\}, \tag{34}$$

$$|S| \le \lfloor p \rfloor, \tag{35}$$

$$t \in \{1, \ldots, n\}. \tag{36}$$

Proposition 3.4 *For $\mathbf{x} \in \mathbb{R}^n$, the CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ with parameter $\alpha \in [0, \frac{n-1}{n}]$ coincides with the D-norm $|||\mathbf{x}|||_p$ with parameter $p = n(1-\alpha)$:*

$$\langle\langle \mathbf{x} \rangle\rangle_\alpha = |||\mathbf{x}|||_p. \tag{37}$$

Proof The proposition follows from the dual representation of the D-norm presented in the proof of Proposition 2 in [1]. It is the linearized version of optimization problem (31), which is the alternative definition of the CVaR norm. □

Remark 5 The CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ with parameter α is in correspondence with the D-norm only when $\alpha \in [0, \frac{n-1}{n}]$; in this case $\langle\langle \mathbf{x} \rangle\rangle_\alpha \ge \max_i |x_i|$. For $\frac{n-1}{n} < \alpha \le 1$ the parameter $p = n(1-\alpha) > n$ and the D-norm is not defined, since $p \in [1, n]$ according to the definition of D-norm. When parameter α varies from $\frac{n-1}{n}$ to 1, the norm C_α linearly decreases from $\max_i |x_i|$ to zero.

To provide some intuition about C_α, we consider the following examples:

Example 3 Figure 3 compares unit disks (balls) for the C_α and L_p norms in $\mathbb{R}^2$ space. A unit disk for a norm is a set of vectors with norm less or equal than 1; i.e., $U_\alpha = \{\mathbf{x} = (x_1, x_2) \mid \langle\langle \mathbf{x} \rangle\rangle_\alpha \le 1\}$ for the CVaR norm with parameter α. The following values of α are considered: $0, 0.1, 1 - \frac{1}{\sqrt{2}}, 0.4, 0.5$. The corresponding values of parameter p for the L_p^S norm are obtained by the transformation $p(\alpha) = \frac{1}{(1-\alpha)^2}$.

The following example illustrates the C_α norm numerically.

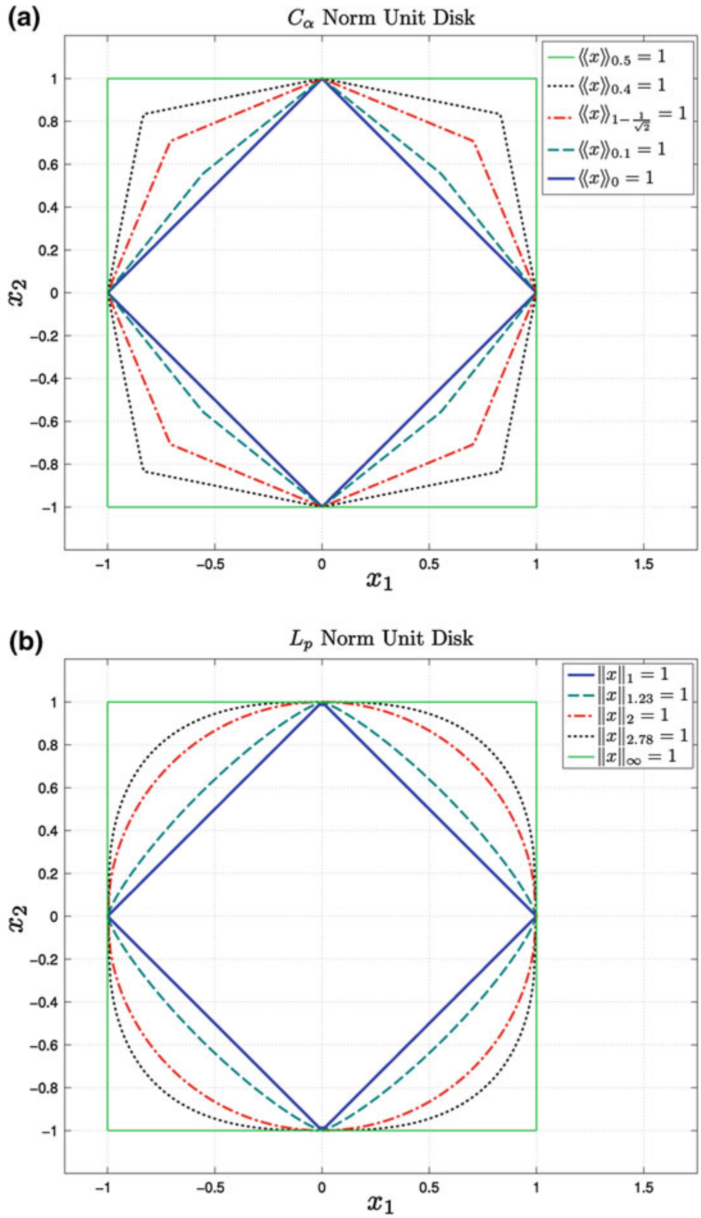


Fig. 3 **a** Unit disks for C_α in $\mathbb{R}^2$ with parameter $\alpha = 0, 0.1, 1-\frac{1}{\sqrt{2}}, 0.4, 0.5$. A unit disk is a set of vectors $\mathbf{x} = (x_1, x_2)$ in $\mathbb{R}^2$, such that $\langle\langle \mathbf{x}\rangle\rangle_\alpha \leq 1$. The CVaR norm with $\alpha = 0$ corresponds to L_1, and the CVaR norm with $\alpha = 0.5$ corresponds to L_∞, according to Remark 4. Observe that $U_0 \subset U_{0.1} \subset U_{1-\frac{1}{\sqrt{2}}} \subset U_{0.4} \subset U_{0.5}$, which is a corollary of Proposition 3.2. **b** Unit disks for the L_p norm in $\mathbb{R}^2$ with parameter $p = 1, 1.23, 2, 2.78, \infty$, where parameter p values were obtained from parameter α using the mapping function $p(\alpha) = \frac{1}{(1-\alpha)^2}$ for $\alpha < 0.5$ and $p(0.5) = \infty$ according to the Remark 4

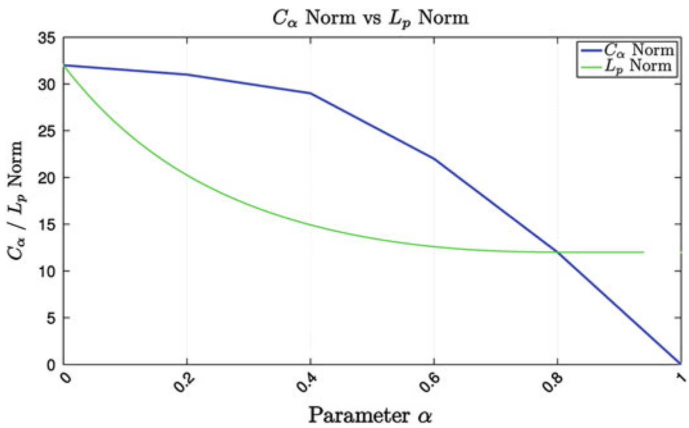


Fig. 4 The graph of C_α for the vector $\mathbf{x} = (2, 1, 7, 10, -12)$. The graph of the L_p norm $\|\mathbf{x}\|_p$ with $p(\alpha) = \frac{1}{(1-\alpha)^2}$ is included for comparison purpose. Parameter p is set to make $\langle\langle x \rangle\rangle_\alpha$ and $\|x\|_p$ visible on the same scale. For instance, when $\alpha = 0.6$, $p(\alpha) = 6.25$; when $\alpha = 0.8$, $p(\alpha) = 25$. C_α of $\mathbf{x}$ equals L_∞ norm of $\mathbf{x}$ when $\alpha = 0.8$

Example 4 Let $\mathbf{x} \in \mathbb{R}^5$, $\mathbf{x} = (2, 1, 7, 10, -12)$. The vector of ordered absolute values of components, $(|x_{(1)}|, |x_{(2)}|, \ldots, |x_{(5)}|)$, equals $(1, 2, 7, 10, 12)$. Then, C_α equals

$$\langle\langle \mathbf{x} \rangle\rangle_0 = 1 + 2 + 7 + 10 + 12 = 32, \quad \text{for } \alpha = \frac{0}{5} = 0,$$
$$\langle\langle \mathbf{x} \rangle\rangle_{0.2} = 2 + 7 + 10 + 12 = 31, \quad \text{for } \alpha = \frac{1}{5} = 0.2,$$
$$\langle\langle \mathbf{x} \rangle\rangle_{0.4} = 7 + 10 + 12 = 29, \quad \text{for } \alpha = \frac{2}{5} = 0.4,$$
$$\langle\langle \mathbf{x} \rangle\rangle_{0.6} = 10 + 12 = 22, \quad \text{for } \alpha = \frac{3}{5} = 0.6,$$
$$\langle\langle \mathbf{x} \rangle\rangle_{0.8} = \max_i(|x_i|) = 12, \quad \text{for } \alpha = \frac{4}{5} = 0.8.$$

For $\alpha = 0.05$, which is between $\alpha_0 = 0$ and $\alpha_1 = 0.2$, $\langle\langle \mathbf{x} \rangle\rangle_{0.05}$ equals the weighted average of $\langle\langle \mathbf{x} \rangle\rangle_0$ and $\langle\langle \mathbf{x} \rangle\rangle_{0.2}$,

$$\lambda = \frac{0.2 - 0.05}{0.2} = 0.75,$$
$$\langle\langle \mathbf{x} \rangle\rangle_{0.05} = \lambda \langle\langle \mathbf{x} \rangle\rangle_0 + (1 - \lambda) \langle\langle \mathbf{x} \rangle\rangle_{0.2} = 0.75 \cdot 32 + 0.25 \cdot 31 = 31.75.$$

Figure 4 compares the C_α and L_p norms numerically. Parameter α of C_α is in the range $0 \le \alpha < 1$ and parameter p of L_p is in the range $1 \le p \le \infty$. In this graph, $\alpha = 0$ corresponds to $p = 1$ and $\alpha = \frac{n-1}{n} = \frac{4}{5}$ corresponds to $p = 25$, and both norms can be presented on the same figure. We use the following mapping function, $p(\alpha) = \frac{1}{(1-\alpha)^2}$.

4 Numerical experiments

This section considers a projection problem using the C_α^S norm. Let P be a convex polyhedron $P \subset \mathbb{R}^n$ and let $\mathbf{w} \notin P$. The projection problem finds a point $\mathbf{w}_p \in P$ (called a projection of $\mathbf{w}$ on P), such that the distance between $\mathbf{w}$ and $\mathbf{w}_p$ is minimized over all points in P. The projection problem using C_α^S can be done efficiently with convex and linear programming techniques.

Let $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$ and P be defined by a set of m hyperplanes. Each hyperplane j, $j = 1, \ldots, m$ is defined by a vector $(a_1^j, \ldots, a_n^j, b^j)$, $j = 1, \ldots, m$. By introducing the matrix $\mathbf{A} = (a_i^j)$, $j = 1, \ldots, m$; $i = 1, \ldots, n$ and vector $\mathbf{b}^T = (b^1, \ldots, b^m)$, the polyhedron P is represented by the set of m linear constraints: $\mathbf{Ax} \leq \mathbf{b}$ and $\mathbf{x} \geq 0$. The problem of projecting a point $\mathbf{w}$ onto P with the scaled CVaR norm is formulated as follows:

$$\min_{\mathbf{x}} \langle\langle \mathbf{w} - \mathbf{x} \rangle\rangle_\alpha^{\mathbf{S}} \tag{38}$$

subject to

$$\mathbf{Ax} \leq \mathbf{b}, \tag{39}$$

$$\mathbf{x} \geq 0. \tag{40}$$

The problem (38)–(40) with representation (21)–(23) can be reformulated as the following convex optimization problem.

$$\min_{x_1,\ldots,x_n,c} c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n z_i \tag{41}$$

subject to

$$z_i \geq |x_i - w_i| - c, \quad i = 1, \ldots, n, \tag{42}$$

$$\sum_{i=1}^n a_i^j x_i \leq b^j, \quad j = 1, \ldots, m, \tag{43}$$

$$z_i \geq 0,\ x_i \geq 0, \quad i = 1, \ldots, n. \tag{44}$$

Constraint (42) can be represented by two linear constraints:

$$x_i - w_i \leq z_i + c, \quad i = 1, \ldots, n, \tag{45}$$

$$x_i - w_i \geq -z_i - c, \quad i = 1, \ldots, n, \tag{46}$$

which gives the linear problem formulation:

$$\min_{x_1,\ldots,x_n,c} c + \frac{1}{n(1-\alpha)} \sum_{i=1}^n z_i \tag{47}$$

subject to

$$x_i - z_i - c \leq w_i, \quad i = 1, \ldots, n, \tag{48}$$

$$x_i + z_i + c \geq w_i, \quad i = 1, \ldots, n, \tag{49}$$

$$\sum_{i=1}^{n} a_i^j x_i \le b^j, \qquad j = 1, \dots, m, \tag{50}$$

$$z_i \ge 0, \; x_i \ge 0, \qquad i = 1, \dots, n. \tag{51}$$

For the numerical experiments, we generated the matrix $\mathbf{A} = (a_i^j), \; j = 1, \dots, m\,; \; i = 1, \dots, n$ and the vector $\mathbf{b}^T = (b^1, \dots, b^m)$ randomly. We sampled the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ assuming that a_i^j and b_j are realizations of the uniformly distributed random variable on [0, 1]. Therefore, $0 \notin P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge 0\}$. We consider the problem of projecting $\mathbf{w} = 0$ on P using the scaled CVaR norm with parameter $\alpha = 0.1, \dots, 0.9$.

Computational results presented in Tables 1 and 2 were performed on a machine equipped with Windows 7x64 operating system, AMD Opteron(tm) Processor 6128 (CPU 2.0GHz), RAM 12GB, using XPress-MP software [10], XPRS_BAR solver with 1 thread. Table 2 shows that the projection problem can be solved efficiently for various values of parameter α, dimension n, and the number of constraints m defining the polyhedron. In particular, we want to emphasize that setting various conservativeness levels α for the CVaR norm does not dramatically impact the efficiency of the optimization. This is in contrast to using the L_p norm, which leads to computational difficulties, especially for large $p > 2$.

Another advantage of the CVaR norm is that it is a convex piece-wise linear function, which can be precoded in software packages. In particular, Portfolio Safeguard [5] has a precoded CVaR norm function, which can be easily used in mathematical programming. A case study solving projection problems with Portfolio Safeguard is pre-

Table 1 Projection in $\mathbb{R}^n$ space

i	n	m	Objective function ($\times 10^{-2}$)								
			$\alpha =$								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	100	10,000	2.23	2.26	2.28	2.29	2.30	2.30	2.30	2.30	2.30
2		20,000	2.28	2.33	2.38	2.43	2.49	2.55	2.55	2.55	2.55
3		50,000	2.32	2.36	2.39	2.40	2.42	2.43	2.43	2.43	2.43
4		100,000	2.33	2.35	2.37	2.38	2.40	2.41	2.41	2.41	2.41
5	500	10,000	0.410	0.414	0.416	0.417	0.418	0.418	0.418	0.418	0.418
6		20,000	0.414	0.418	0.420	0.421	0.422	0.422	0.422	0.422	0.422
7		50,000	0.418	0.422	0.424	0.426	0.427	0.427	0.427	0.427	0.427
8		100,000	0.424	0.429	0.433	0.437	0.442	0.445	0.445	0.445	0.445
9	1,000	10,000	0.203	0.205	0.206	0.207	0.208	0.208	0.208	0.208	0.208
10		20,000	0.204	0.206	0.208	0.209	0.210	0.210	0.210	0.210	0.210
11		50,000	0.204	0.207	0.208	0.208	0.209	0.209	0.209	0.209	0.209
12		100,000	0.206	0.208	0.209	0.210	0.210	0.210	0.210	0.210	0.210

Polyhedron P_i in $\mathbb{R}^n$ is defined by m hyperplanes, specified in the third column of the table. Formulation (47)–(51) solves the problem of projection of 0 onto P_i using the scaled CVaR norm with parameter $\alpha = 0.1, \dots, 0.9$. The table shows the optimal objective values of the optimization problems

Table 2 Projection in $\mathbb{R}^n$ space

i	n	m	Execution Time (sec)								
			$\alpha =$								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	100	10, 000	12.65	9.32	9.71	8.32	8.48	6.07	5.66	5.67	5.17
2		20,000	18.33	17.64	15.26	20.28	16.55	12.37	12.26	11.73	10.5
3		50,000	37.46	52.4	36.44	39.05	37.14	32.93	34.68	27.82	26.6
4		100,000	100.03	95.28	72.57	70.06	75.46	62.87	60.2	56.63	57.53
5	500	10,000	100.84	100.56	89.36	116.84	94.76	76.43	75.24	79.44	80
6		20,000	194.67	158.73	229.71	212.53	203.76	149.07	150.73	160.49	144.93
7		50,000	951.7	615.68	430.15	470.23	451.52	386.16	409.39	359.92	382.16
8		100,000	1190.73	1575.88	1493.01	945.55	1201.15	724.19	816.63	731.4	728.04
9	1,000	10, 000	420.35	367.77	365.21	287.88	368	238.91	235.53	240.5	272.08
10		20,000	920.61	1086.76	576.32	575.91	956.38	478.15	479.35	508.94	542.36
11		50,000	4234.25	2484.64	1771.73	1525.61	1370.4	1211	1214.4	1130.85	1287.61
12		100,000	1713.53	3136.04	2876.51	1539.45	1791.79	1287.55	1353.42	1200.97	1374.89

Polyhedron P_i in $\mathbb{R}^n$ is defined by m hyperplanes, specified in the third column of the table. Formulation (47)–(51) solves the problem of projection of 0 onto P_i using the scaled CVaR norm with parameter $\alpha = 0.1, \ldots, 0.9$. The table shows the execution times of the optimization problems sented in the following link: http://www.ise.ufl.edu/uryasev/research/testproblems/advanced-statistics/case-study-projection-on-polyhedron-with-cvar-absolute-norm-4/. The case study at this link investigates projection problems with various norms.

5 Conclusion

This paper introduces the two families (scaled and non-scaled) of CVaR norms with parameter α controlling the level of conservativeness. The family of non-scaled CVaR norms spans the L_1 and L_∞ norms. We provide two equivalent definitions of the CVaR norm: as the sum or the average of the specified percentage of the largest absolute values of the vector components, and as the solution of the CVaR minimization problem. The non-scaled CVaR norm is equivalent to the D-norm employed in robust optimization. We show that the CVaR norm can be efficiently used in optimization problems. In the numerical experiments, we consider the projection problems with various levels of the conservativeness parameter α and observe that the solution time is not significantly affected by the value of this parameter.

Acknowledgments Authors would like to thank the referees for their comments and suggestions, which helped to improve the quality of the paper. Authors are also grateful to Prof. Donald W. Hearn, University of Florida, for valuable general comments and suggestions.

References

1. Bertsimas, D., Pachamanova, D., Sim, M.: Robust linear optimization under general norms. Oper. Res. Lett. **32**(6), 510–516 (2004)
2. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
3. Mafusalov, A., Uryasev, S.: CVaR norm: stochastic case. University of Florida, Research Report (2013, in preparation)
4. Pflug, G.C.: Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. Methodology and Application. Probabilistic Constrained Optimization. Kluwer, Dordrecht (2000)
5. Portfolio Safeguard version 2.1, 2009. http://www.aorda.com/aod/welcome.action
6. Rockafellar, R.T.: Convex Analysis, vol. 28. Princeton University Press, Princeton (1996)
7. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. J. Risk **2**(3), 21–41 (2000)
8. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. J. Banking Fin. **26**(7), 1443–1471 (2002)
9. Rockafellar, R.T., Uryasev, S.: The fundamental risk quadrangle in risk management, optimization and statistical estimation. Surv. Oper. Res. Manage. Sci. **18**, 33–53 (2013)
10. Xpress. FicoTM xpress optimization suite 7.4 (2012). http://www.fico.com