European Journal of Operational Research 249 (2016) 200–208

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

CVaR (superquantile) norm: Stochastic case[☆]

Alexander Mafusalov, Stan Uryasev[*]

Risk Management and Financial Engineering Lab Department of Industrial and Systems Engineering, University of Florida, 303 Weil Hall, Gainesville, FL 32611, US

ARTICLE INFO

Article history:
Received 17 January 2014
Accepted 26 September 2015
Available online 5 October 2015

Keywords:
CVaR norm
L-p norm
Superquantile
Risk quadrangle
Linear regression

ABSTRACT

The concept of Conditional Value-at-Risk (CVaR) is used in various applications in uncertain environment. This paper introduces CVaR (superquantile) norm for a random variable, which is by definition CVaR of absolute value of this random variable. It is proved that CVaR norm is indeed a norm in the space of random variables. CVaR norm is defined in two variations: scaled and non-scaled. L-1 and L-infinity norms are limiting cases of the CVaR norm. In continuous case, scaled CVaR norm is a conditional expectation of the random variable. A similar representation of CVaR norm is valid for discrete random variables. Several properties for scaled and non-scaled CVaR norm, as a function of confidence level, were proved. Dual norm for CVaR norm is proved to be the maximum of L-1 and scaled L-infinity norms. CVaR norm, as a Measure of Error, is related to a Regular Risk Quadrangle. Trimmed L1-norm, which is a non-convex extension for CVaR norm, is introduced analogously to function L-p for $p < 1$. Linear regression problems were solved by minimizing CVaR norm of regression residuals.

Published by Elsevier B.V.

1. Introduction

The concept of Conditional Value-at-Risk (CVaR) is widely used in risk management and various applications in uncertain environment. This paper introduces a concept of *CVaR norm* in the space of random variables. CVaR norm in $\mathbb{R}^n$ was introduced in (Pavlikov & Uryasev, 2014) and developed in (Gotoh & Uryasev, 2013), and is a particular case of general error measures introduced and developed by Rockafellar, Uryasev, and Zabarankin (2008). The term "superquantile", free from dependence on financial terminology, can be used as a neutral alternative name for "CVaR", like it was done in Rockafellar and Royset (2010); Rockafellar and Uryasev (2013). For the similar reason the alternative name "superquantile norm" is proposed to replace "CVaR norm" when desired. For the sake of consistency within the paper and with the earlier study (Pavlikov & Uryasev, 2014), this paper will use mostly the "CVaR norm" term.

This section provides a short introduction in the CVaR norm in $\mathbb{R}^n$ and shows the relation with the CVaR norm in the space of random variables. This paper is motivated by applications of norms in optimization. We consider norms in $\mathbb{R}^n$ and in the space of random variables. We use symbols $\mathbf{x}$ and x_i for a vector and an ith vector component in $\mathbb{R}^n$, i.e. $\mathbf{x} = (x_1, \ldots, x_n)$. We use symbol X for a random variable.

l_p norms are broadly used in $\mathbb{R}^n$, and L_p norms are considered in the space of random variables. For $p \in [1, \infty]$, the norms l_p and L_p are defined as follows[1]:

$$l_p(\mathbf{x}) = \left(\frac{1}{n} \sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \quad L_p(X) = (E|X|^p)^{1/p},$$

where E is the expectation sign. The most popular cases are $p = 1, 2, \infty$, i.e.,

- $l_1(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n |x_i|, \quad L_1(X) = E|X|;$
- $l_\infty(\mathbf{x}) = \max_{i=1,\ldots,n} |x_i|, \quad L_\infty(X) = \sup|X|;$
- $l_2(\mathbf{x}) = \left(\frac{1}{n}\sum_{i=1}^n x_i^2\right)^{1/2}, \quad L_2(X) = (EX^2)^{1/2}.$

It is known that $l_1(\mathbf{x}) \le l_2(\mathbf{x}) \le l_\infty(\mathbf{x})$ and $L_1(X) \le L_2(X) \le L_\infty(X)$, which follow from $l_p(\mathbf{x}) \le l_q(\mathbf{x})$ and $L_p(X) \le L_q(X)$ for $p < q$, see (e.g. Brezis, 2010, page 118).

The other family of norms, CVaR norm for $\mathbb{R}^n$ was defined in Pavlikov and Uryasev (2014) and studied in Gotoh and Uryasev (2013). According to Pavlikov and Uryasev (2014, Definition 3), the *non-scaled*

[1] Note that the classic definition for l_p norm is $l_p(\mathbf{x}) = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ does not satisfy inequality $l_p(\mathbf{x}) \le l_q(\mathbf{x})$ for $p < q$. This paper uses an equivalent scaled version of this norm $l_p(\mathbf{x}) = \left(\frac{1}{n}\sum_{i=1}^n |x_i|^p\right)^{1/p}$, which satisfies that inequality. L_p norm is commonly defined as $L_p(f) = \|f\|_p \equiv \left(\int_S |f|^p \, d\mu\right)^{1/p}$, where S is a considered space. It is known (e.g. Brezis, 2010, page 118) that for $\|\cdot\|_p$ and $\|\cdot\|_q$ norms inequality $\|f\|_p \le \mu(S)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q$ holds for $1 \le p \le q \le \infty$, where S is a considered space and $\mu(S)$ is the measure of the space S. When S is a probability space, $\mu(S) = 1$ and inequality $L_p(X) \le L_q(X)$ holds for $1 \le p \le q \le \infty$, where $L_p(X) = (E|X|^p)^{1/p}$.

[☆] This paper was processed by Guest-Editor prof. Ruediger Schultz for an uncompleted EJOR Feature Cluster on Stochastic Optimization.

[*] Corresponding author. Tel.: +1 352 213 3457.
E-mail addresses: mafusalov@ufl.edu (A. Mafusalov), uryasev@ufl.edu (S. Uryasev).

http://dx.doi.org/10.1016/j.ejor.2015.09.058
0377-2217/Published by Elsevier B.V.

CVaR norm with parameter α, or $\langle\langle \mathbf{x} \rangle\rangle_\alpha$, is defined on $\mathbf{x} \in \mathbb{R}^n$ as a sum of absolute values of biggest $n(1-\alpha)$ components. If $n(1-\alpha)$ is not an integer, then $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ is defined as a weighted average of two norms $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_1}$ and $\langle\langle \mathbf{x} \rangle\rangle_{\alpha_2}$ for closest values α_1, α_2 such that $n(1-\alpha_1)$ and $n(1-\alpha_2)$ are integers.

A similar norm, called D-norm, was introduced in Bertsimas, Pachamanova, and Sim (2004, Section 3) in a different way. D-norm is defined as a maximum of a sum of weighted absolute values of vector components. The maximization is performed over all sets of indexes for components in the sum, with a constraint on cardinality. For $\alpha \in \left[0, \frac{n-1}{n}\right]$, the CVaR norm $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ coincides with the D-norm $|||\mathbf{x}|||_p$ with parameter p defined by $p = n(1-\alpha)$, see Pavlikov and Uryasev (2014, Proposition 3.4). (Bertsimas et al., 2004, Proposition 2) find a dual norm to D-norm; this result was generalized with Item 2 of Proposition 2.1 of this paper for the stochastic case.

Both CVaR norm and D-norm can be viewed as important special cases of Ordered Weighting Averaging (OWA) operators, see Merigó and Yager (2013); Torra and Narukawa (2007); Yager (2010). A subfamily of OWA operators with monotonically non-increasing weights, when implied to the absolute values of the vector, were formalized as norms in Yager (2010). The worst-case averages, corresponding to CVaR, were also studied, e.g., in Ogryczak and Zawadzki (2002); Romeijn, Ahuja, Dempsey, and Kumar (2005).

The paper (Pavlikov & Uryasev, 2014, Definition 1) has also defined *scaled* CVaR norm $\langle\langle \mathbf{x} \rangle\rangle^S_\alpha$. Scaled version calculates average value of components instead of sum: $n(1-\alpha)\langle\langle \mathbf{x} \rangle\rangle^S_\alpha = \langle\langle \mathbf{x} \rangle\rangle_\alpha$. This paper defines (scaled) CVaR norm of a random variable X as an expectation of $|X|$ in its right $(1-\alpha)$-tail. It can be shown that proposed norm is a generalization of $\langle\langle \mathbf{x} \rangle\rangle^S_\alpha$ in a following way. Consider mapping $X(\mathbf{x}) : \mathbb{R}^n \to \mathcal{L}^1(\Omega)$ from Eucledian space of dimension n to the space of L_1-finite random variables. Denote $\mathbf{x} = (x_1, \dots, x_n)$. Let $X(\mathbf{x})$ be discretely distributed over atoms $x_1, \dots, x_n$ with equal probabilities $\frac{1}{n}$. Then it is easy to see that $\langle\langle \mathbf{x} \rangle\rangle^S_\alpha = \mathrm{CVaR}_\alpha(|X(\mathbf{x})|) \equiv \langle\langle X(\mathbf{x}) \rangle\rangle^S_\alpha$, see Pavlikov and Uryasev (2014, Definition 2).

This paper also defines *non-scaled* CVaR norm $\langle\langle X \rangle\rangle_\alpha = (1-\alpha)\langle\langle X \rangle\rangle^S_\alpha$, which corresponds to $\langle\langle \mathbf{x} \rangle\rangle_\alpha$ from Pavlikov and Uryasev (2014). Non-scaled version has attractive properties with respect to parameter α, see Items 6, 7 from Section 2.

Risk Quadrangle considers risk $\mathcal{R}(X)$, deviation $\mathcal{D}(X)$, regret $\mathcal{V}(X)$, error $\mathcal{E}(X)$ and statistic $\mathcal{S}(X)$, related with a set of equations, called The General Relationships, see Rockafellar and Uryasev (2013, Diagram 3). If a functional satisfies a corresponding set of axioms, it is called regular, see Rockafellar and Uryasev (2013, Section 3). It can be proved that if $\mathcal{R}(X)$ is a coherent and regular Measure of Risk, then $\mathcal{R}(|X|)$ is both a norm and a regular Measure of Error, however, this proof is beyond the scope of this paper. This paper proves that $\langle\langle X \rangle\rangle_\alpha$ is a regular Measure of Error and finds the corresponding functions $\mathcal{R}(X)$, $\mathcal{D}(X)$, $\mathcal{V}(X)$ and $\mathcal{S}(X)$ in risk quadrangle related to the Measure of Error $\mathcal{E}(X) = \langle\langle X \rangle\rangle_\alpha$, see Section 2.

Item 3 from Proposition 2.1 can be viewed as a stochastic generalization of Hall and Tymoczko (2012, Lemma 1). Paper (Hall & Tymoczko, 2012) considers functions $\Sigma_j(\mathbf{x})$ on nonnegative orthant $\mathbb{R}^n_+$, corresponding to $\mathbb{R}^n_+$ reduction of special cases of CVaR norm. Paper (Hall & Tymoczko, 2012) relies on majorization theory, see (Marshall, Olkin, & Arnold, 2011), which is generalized for the stochastic case with a concept of stochastic dominance, see Ogryczak and Ruszczynski (2002); Dentcheva and Ruszczynski (2003, e.g.).

This paper considers also non-convex functions closely related to CVaR norm. In deterministic case, by definition, CVaR norm is the average of biggest $(1-\alpha)n$ absolute values of components of a vector. The *trimmed L1-norm* is also known as a trimmed sum of absolute deviations estimator for LTA regression, see (Hawkins & Olive, 1999; Bassett Jr, 1991; Hössjer, 1994, e.g.). L1-based LTA regression is introduced similarly to the more common L2-based Least-Trimmed-Squares (LTS) regression, see Zabarankin and Uryasev (2014b, e.g.). Trimmed L1-norm is denoted here by t_α, it is the average of smallest αn absolute values of components of a vector. Calculation formulas and mathematical properties for $t_\alpha(\mathbf{x})$ in Eucledian space and $T_\alpha(X)$ in the space of random variables are considered in Section 3. Trimmed L1-norm is also related to the sparse optimization, similar to functions l_p for $0 < p < 1$, see Ge, Jiang, and Ye (2011). Note that the constraint on trimmed L1-norm directly specifies sparsity of the solution vector (see Item 6 of Section 3), compared to “indirect” sparsity specification with l_p function. This paper also provides an illustration for t_α in Euclidean space.

Paper (Krzemienowski, 2009, Definition 2) defines conditional average CAVG function. Both average quantile and CVaR are subfamilies of CAVG family, therefore, both $\langle\langle X \rangle\rangle^S_\alpha$ and $T_\alpha(X)$ are subfamilies of $\mathrm{CAVG}_{\beta,\gamma}(|X|)$ function family. Unfortunately, these functions are not convex or concave in general, and are out of the scope of this study, although robust regression applications based on these functions is a promising research direction.

The paper is organized as follows. Section 2 gives a formal definition of CVaR norm in stochastic case and enlists various mathematical of CVaR norm, including that it is indeed a norm and a regular measure of error. CVaR norm is a parametric family of norms with respect to the confidence parameter α, properties of CVaR norm as a function of α are proved. Dual representation of the CVaR norm is derived, and a dual norm to the CVaR norm is defined. A short introduction to the concept of Risk Quadrangle is given. We derive the quadrangle related to the CVaR norm as a measure of error and we prove that this quadrangle is regular. Section 3 defines the *trimmed L1-norm*, both in $\mathbb{R}^n$ and in the space of random variables, and enlists several basic properties. The trimmed L1-norm is an extension of CVaR norm, but it is not actually a norm. Section 4 illustrates properties of CVaR norm with a case study. Section 5 provides concluding remarks and acknowledgements.

2. CVaR (superquantile) norm properties and connection to risk quadrangle

This section gives a formal definition of CVaR norm in stochastic case and proves various properties of the norm. Let us denote $[x]^+ = \max\{0, x\}$, $[x]^- = \max\{0, -x\}$. Consider cumulative distribution function $F_X(x) = P(X \le x)$. If, for a probability level $\alpha \in (0, 1)$, there is a unique x such that $F_X(x) = \alpha$, then this x is called the α-quantile $q_\alpha(X)$. In general, however, the value x is not unique, or may not even exist. There are two boundary values:

$$q^+_\alpha(X) = \inf\{x | F_X(x) > \alpha\}, \quad q^-_\alpha(X) = \sup\{x | F_X(x) < \alpha\}.$$

We will call by the *quantile* the entire interval between the two boundary values,

$$q_\alpha(X) = [q^-_\alpha(X), q^+_\alpha(X)]. \tag{1}$$

We will use notation $\int q_p(X)dp \equiv \int q^-_p(X)dp$, which is reasonable since $\int q^+_p(X)dp = \int q^-_p(X)dp$.

The CVaR norm is defined as follows.

Definition 1. Let X be a random variable with $E|X| < \infty$. Then CVaR (superquantile) norm of X with parameter $\alpha \in [0, 1]$ is defined by

$$\langle\langle X \rangle\rangle^S_\alpha = \mathrm{CVaR}_\alpha(|X|) = \bar{q}_\alpha(|X|).$$

Following the logic of (Pavlikov & Uryasev, 2014), $\langle\langle X \rangle\rangle^S_\alpha$ is called *scaled* CVaR norm, while Definition 2 introduces $\langle\langle X \rangle\rangle_\alpha$, corresponding to *non-scaled* CVaR norm for $\mathbb{R}^n$ in (Pavlikov & Uryasev, 2014, Definition 3). By default we call by CVaR norm the function $\langle\langle X \rangle\rangle^S_\alpha$. The second version of the norm, *non-scaled* CVaR norm, is defined as follows.

Definition 2. Let X be a random variable with $E|X| < \infty$. Then non-scaled CVaR (superquantile) norm of X with parameter $\alpha \in [0, 1)$ is defined as follows:

$$\langle\langle X \rangle\rangle_\alpha = (1-\alpha)\langle\langle X \rangle\rangle^S_\alpha.$$